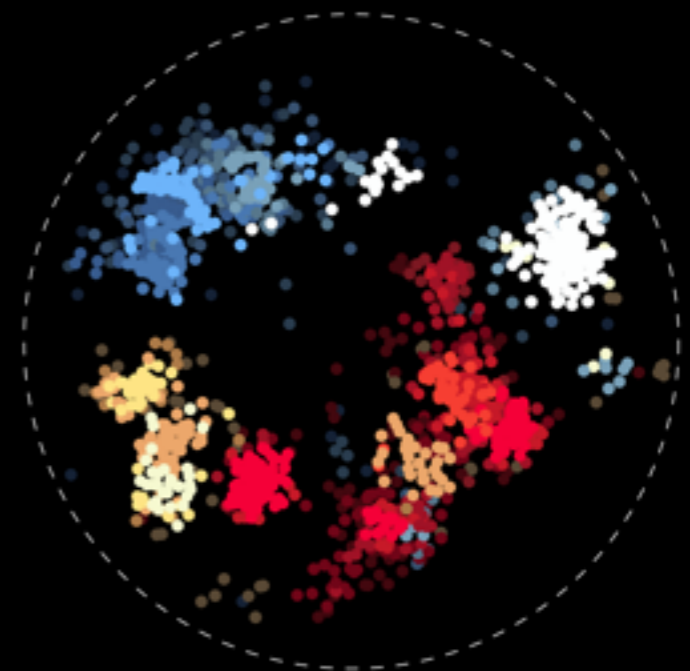
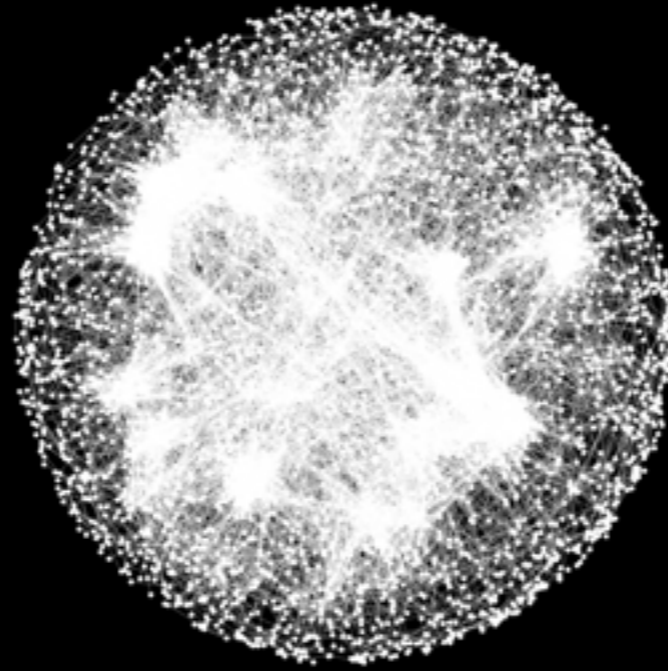
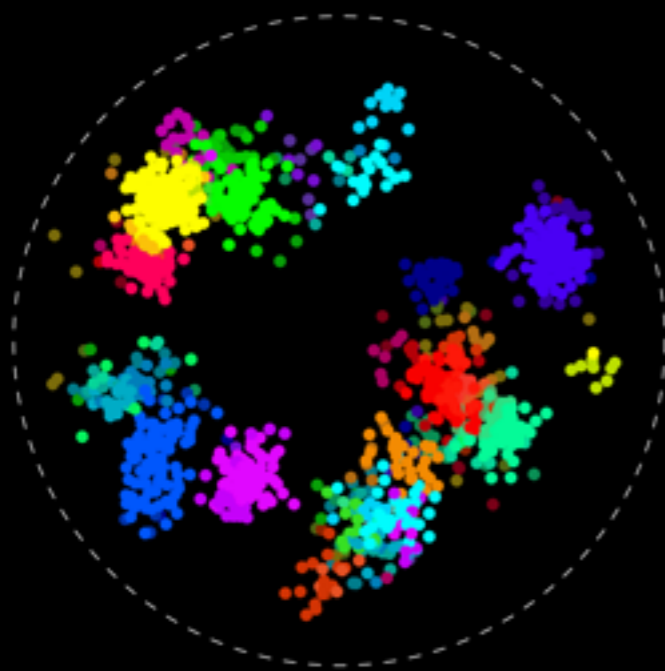


# Systematic Functional Annotation of Large-Scale Biological Networks

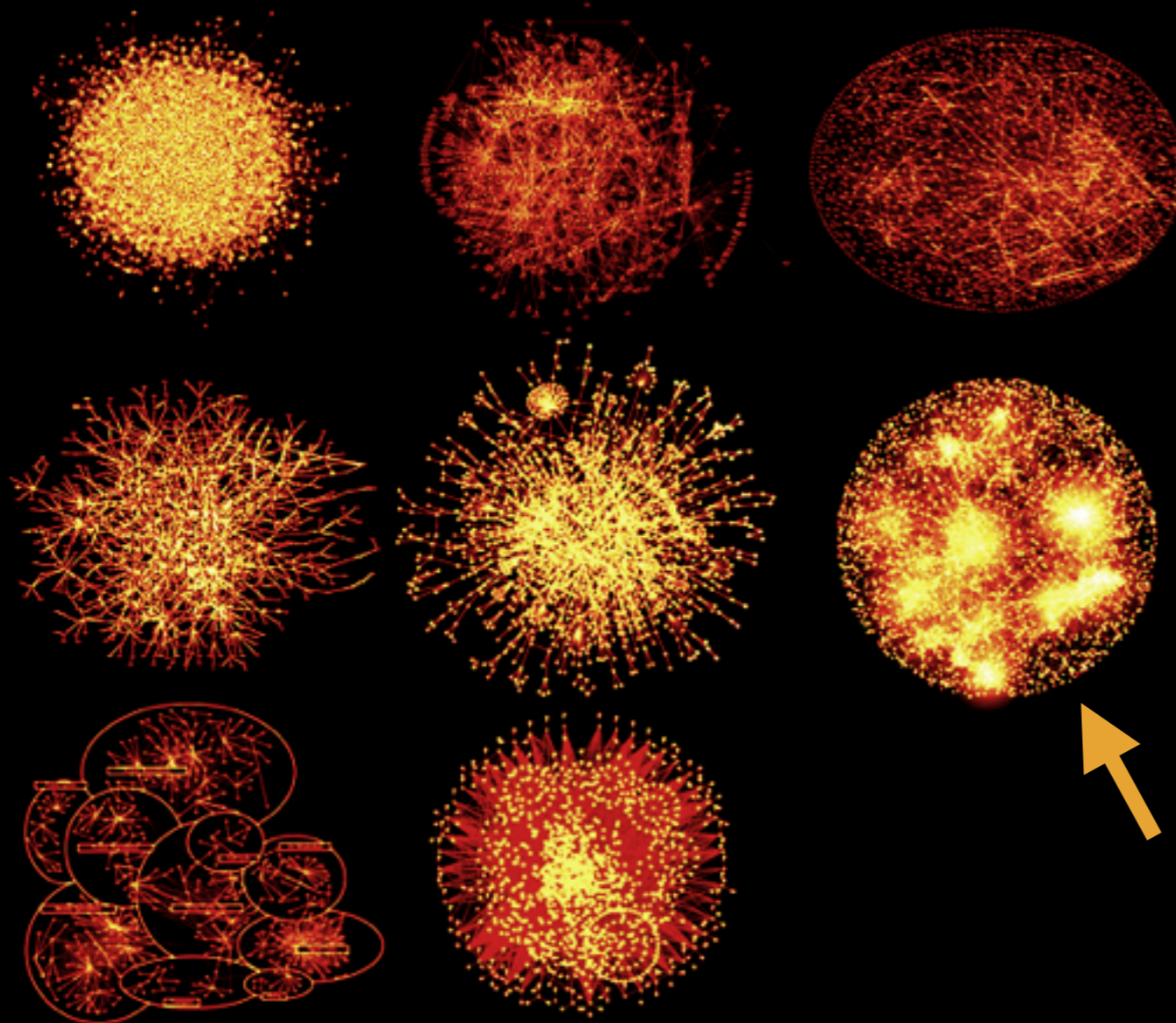
Anastasia Baryshnikova

Lewis-Sigler Fellow  
Princeton University

[www.baryshnikova-lab.org](http://www.baryshnikova-lab.org)



# Networks as maps of biological systems



Good biological interpretation for:

- individual interactions
- local structures/modules

Much worse understanding of:

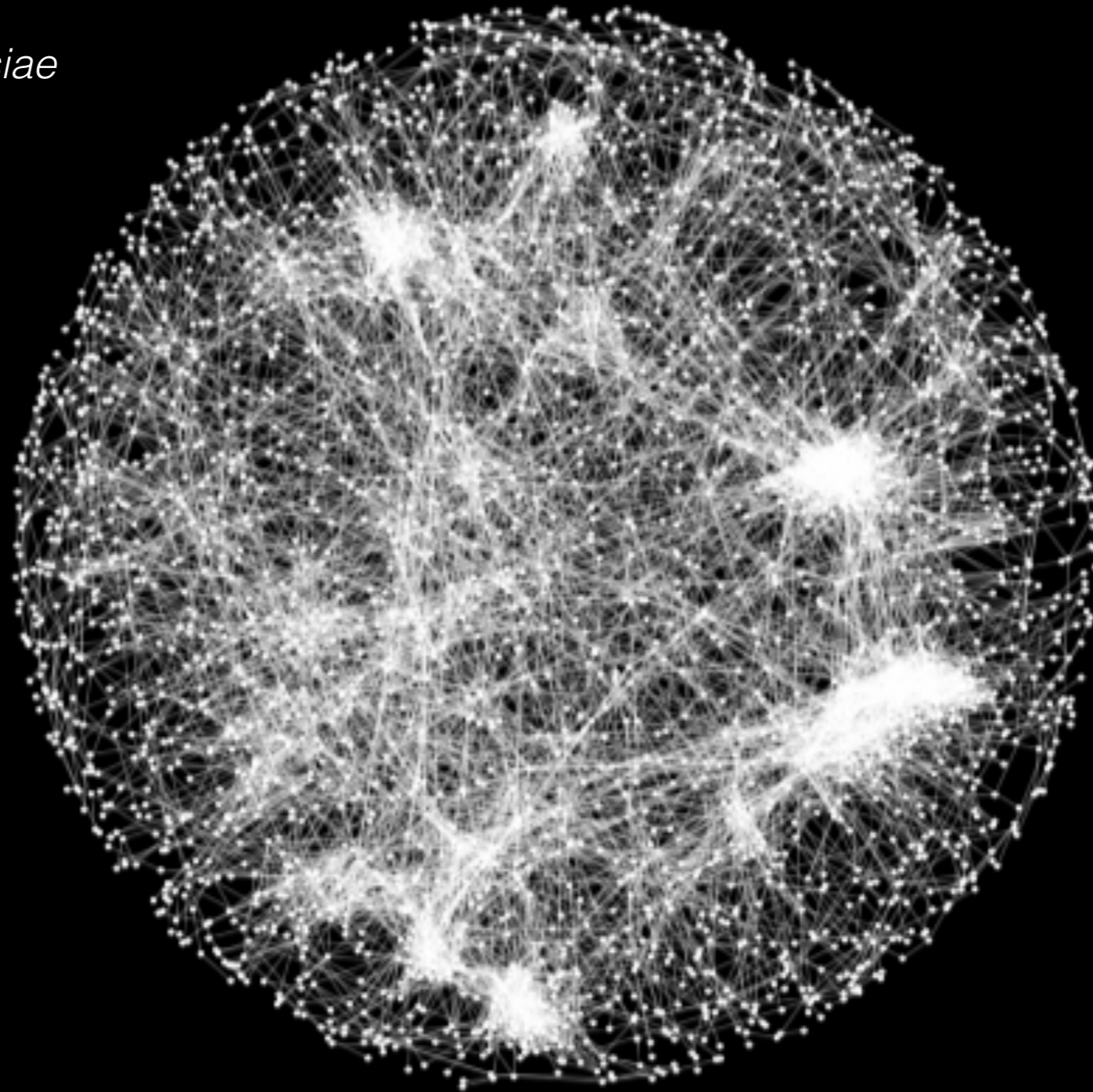
- global organization

# Genetic interaction similarity network (year 2010)

*Saccharomyces cerevisiae*

2,838 nodes

10,016 edges



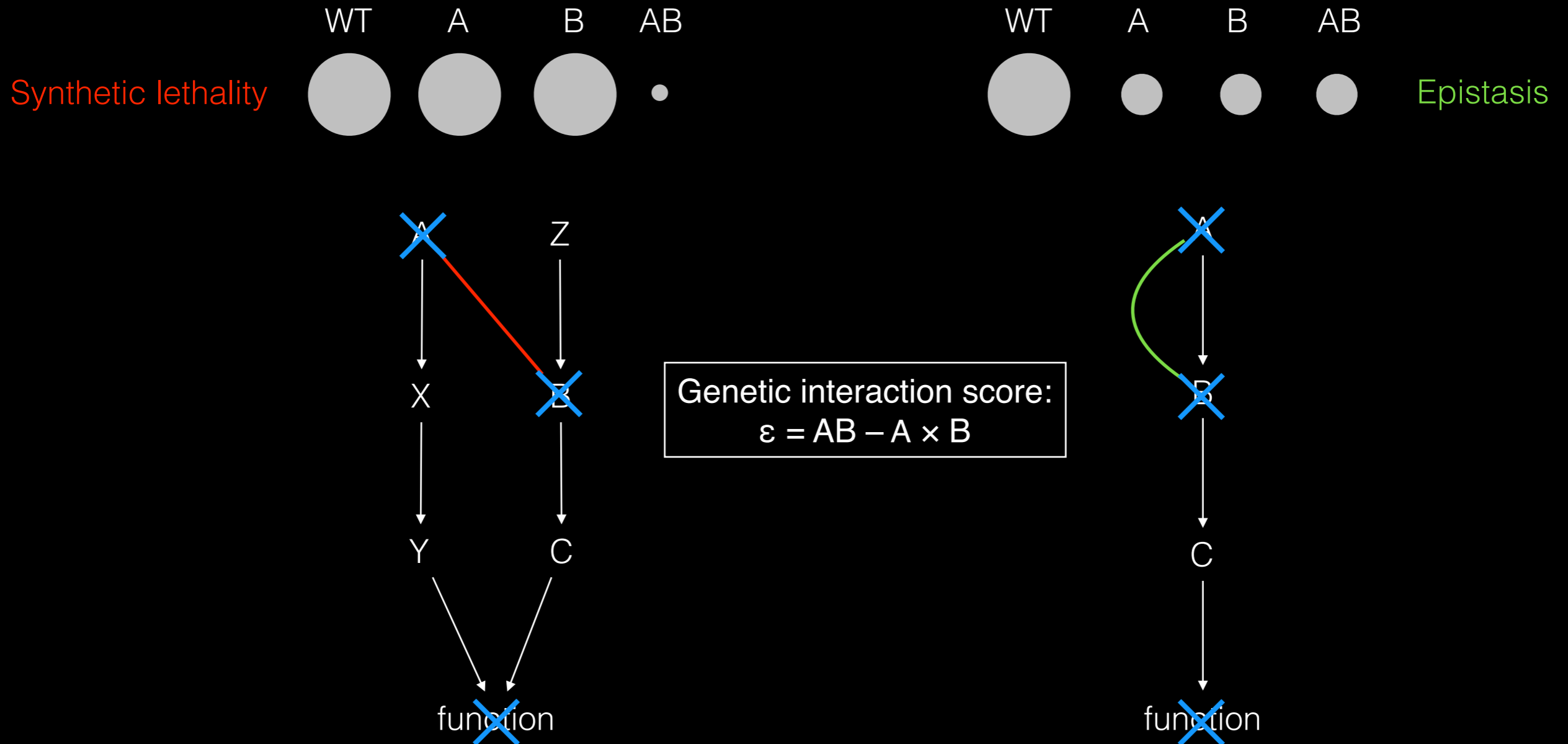
Gene A      Gene B



Genetic interaction profile  
correlation  $> 0.2$

# Genetic interactions

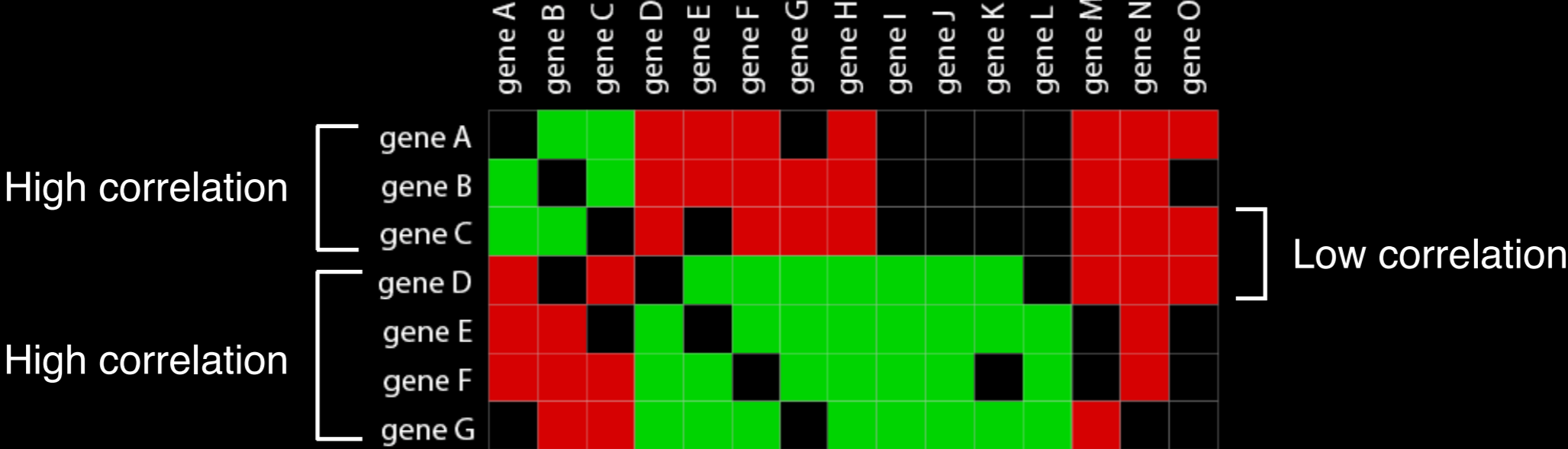
Double mutants with unexpected phenotypes.



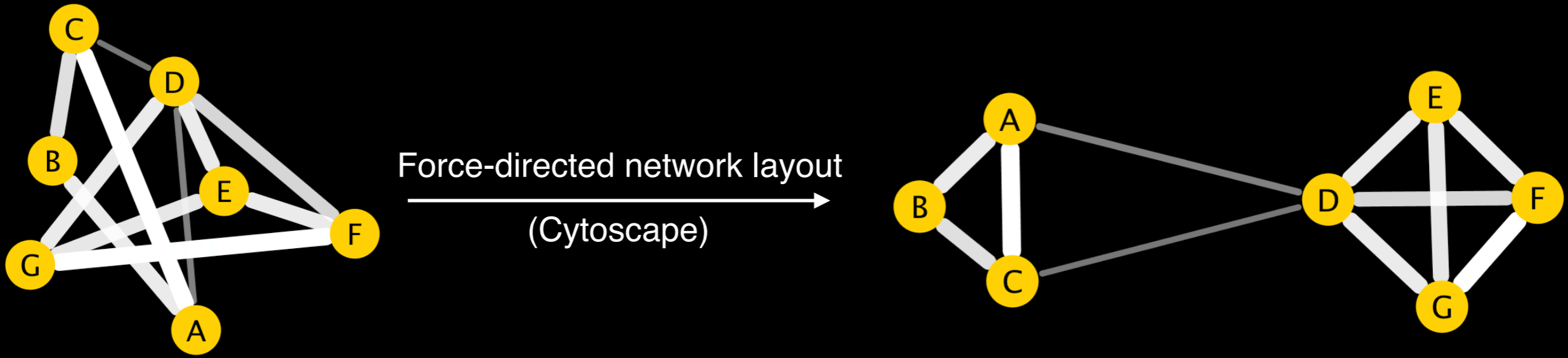
Genome-scale analysis:

- Construct & analyze all 36 000 000 combinations of double mutants in yeast
- Synthetic Genetic Array (SGA) by Charlie Boone & team (University of Toronto)

# Similarity of genetic interaction profiles



■ Negative genetic interaction (e.g., synthetic lethality)  
■ Positive genetic interaction (e.g., epistasis)

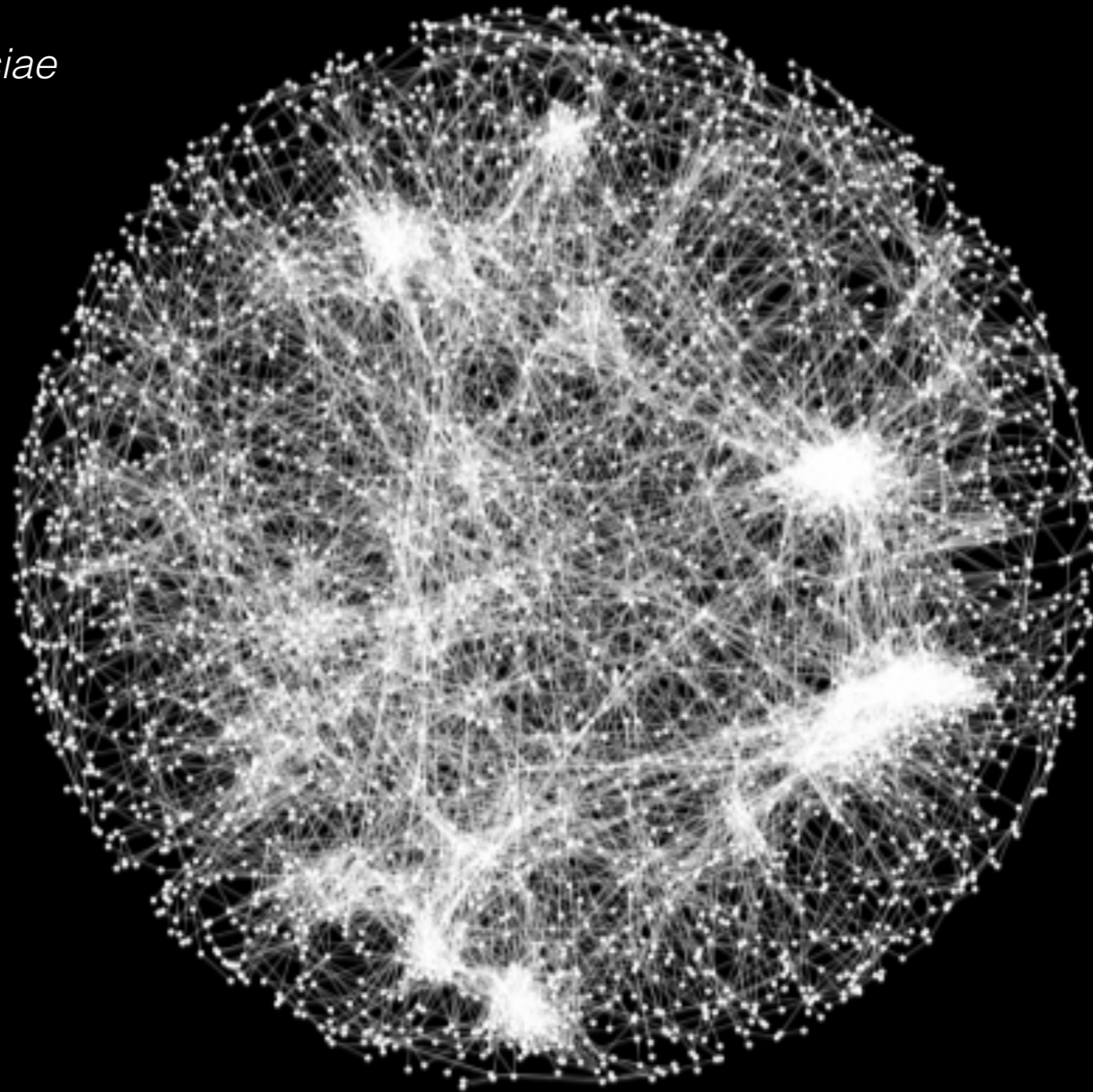


# Genetic interaction similarity network (year 2010)

*Saccharomyces cerevisiae*

2,838 nodes

10,016 edges



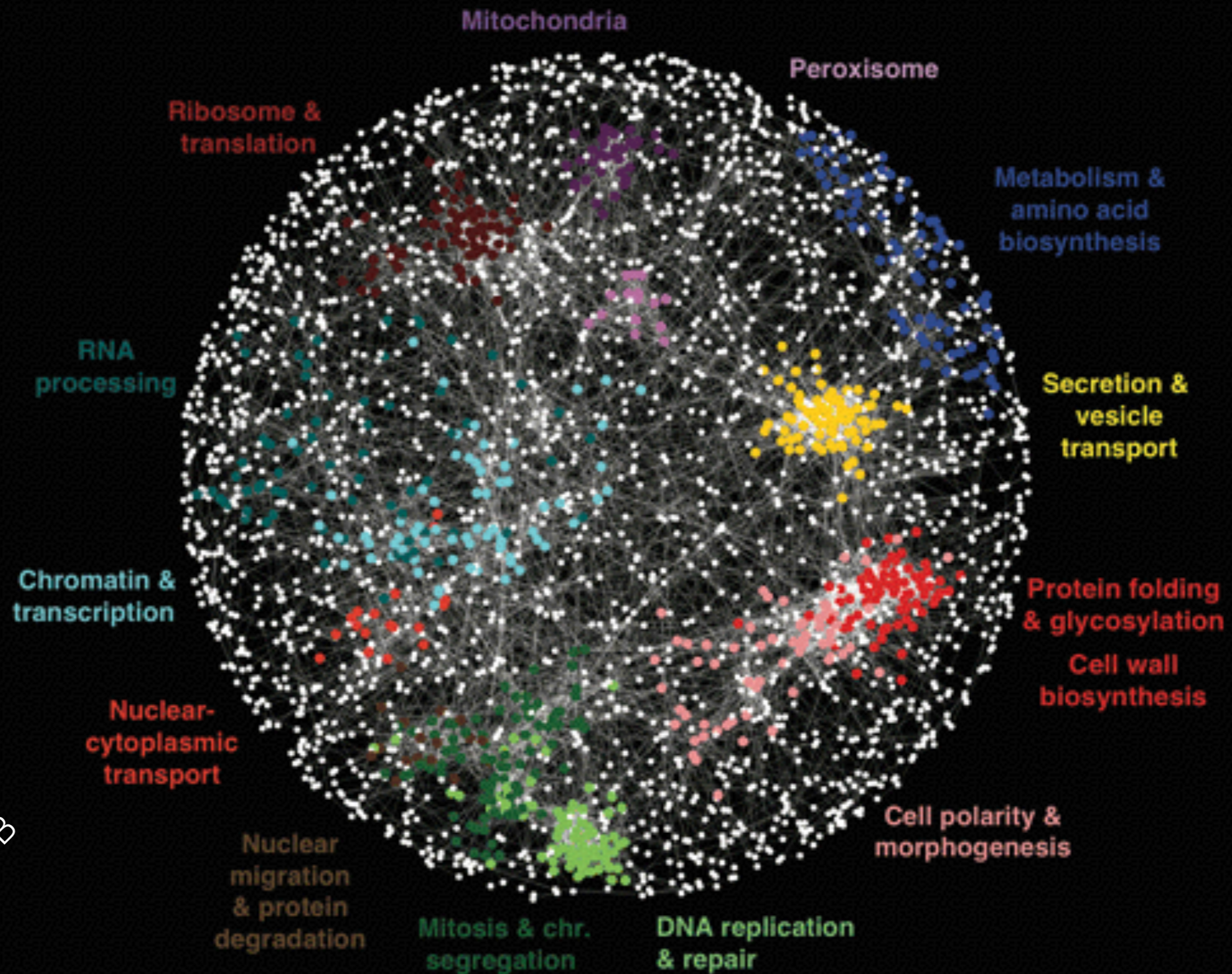
Gene A      Gene B



Genetic interaction profile  
correlation  $> 0.2$

# Genetic interaction similarity network (year 2010)

2,838 nodes  
10,016 edges



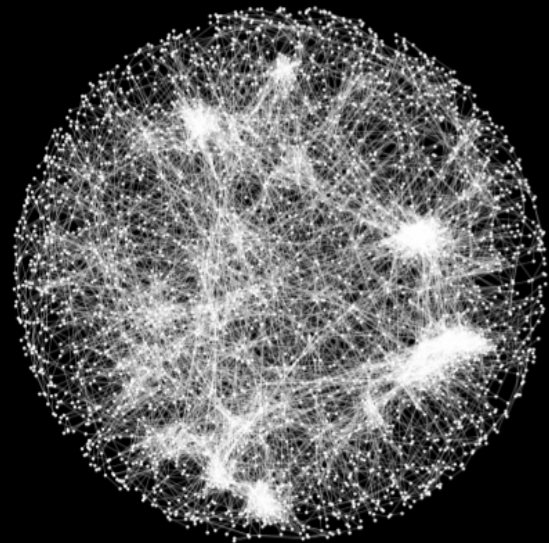
Gene A

Gene B

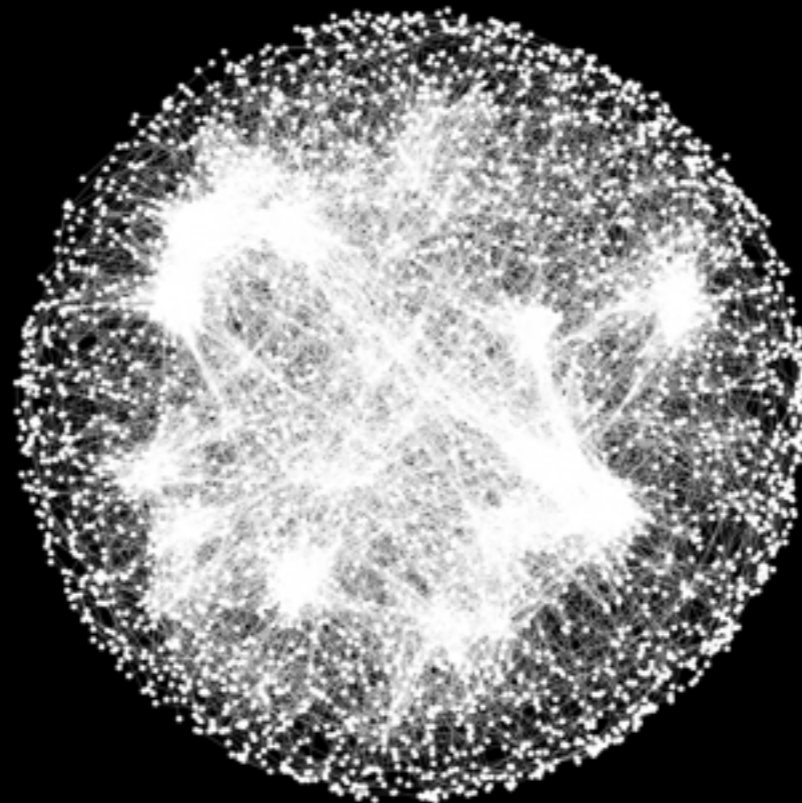


Genetic interaction profile  
correlation  $> 0.2$

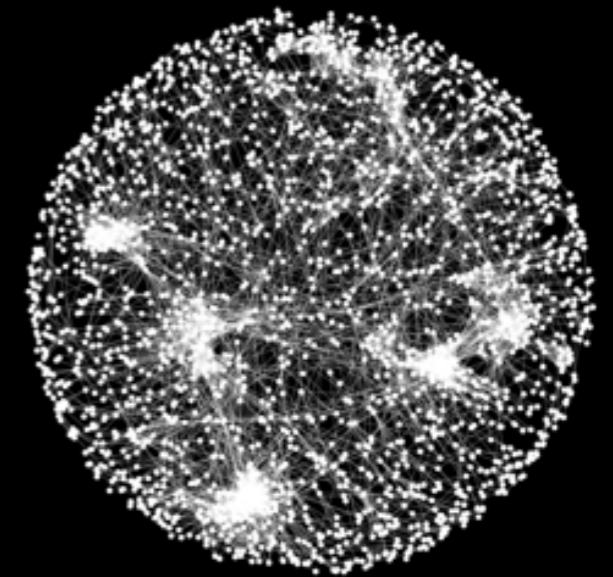
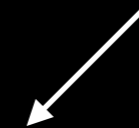
# The Yeast Genetic Interaction Similarity Network (Year 2016)



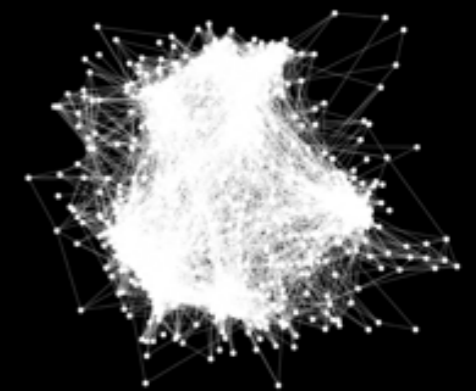
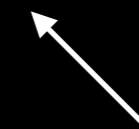
**2010 network**  
**2,838** nodes, **10,016** edges  
Average degree = **7.1**



**2016 network**  
**3,996** nodes, **28,688** edges  
Average degree = **14.4**



Non-essential network

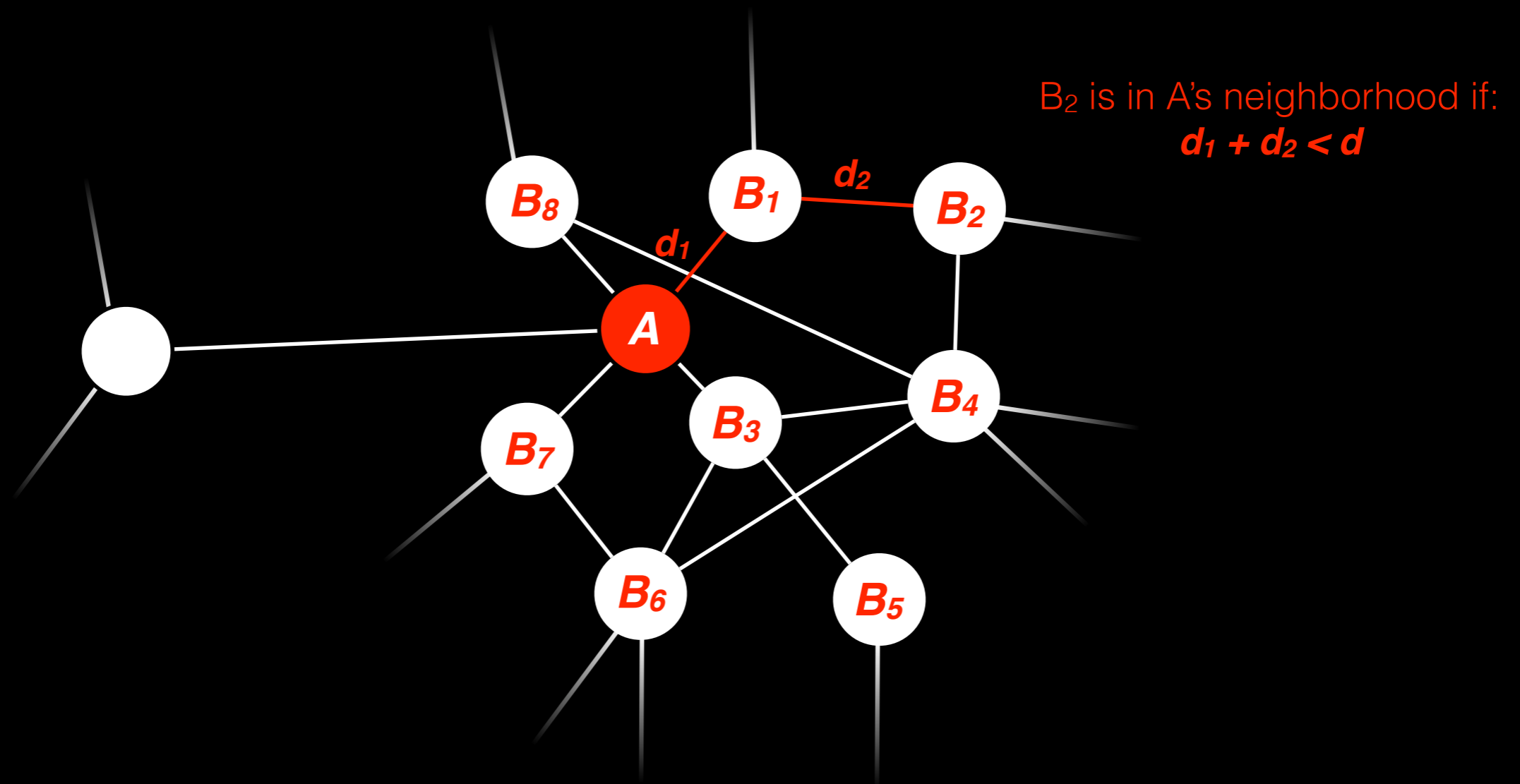


Essential network

Many large networks = Need for an automated method for functional annotation.

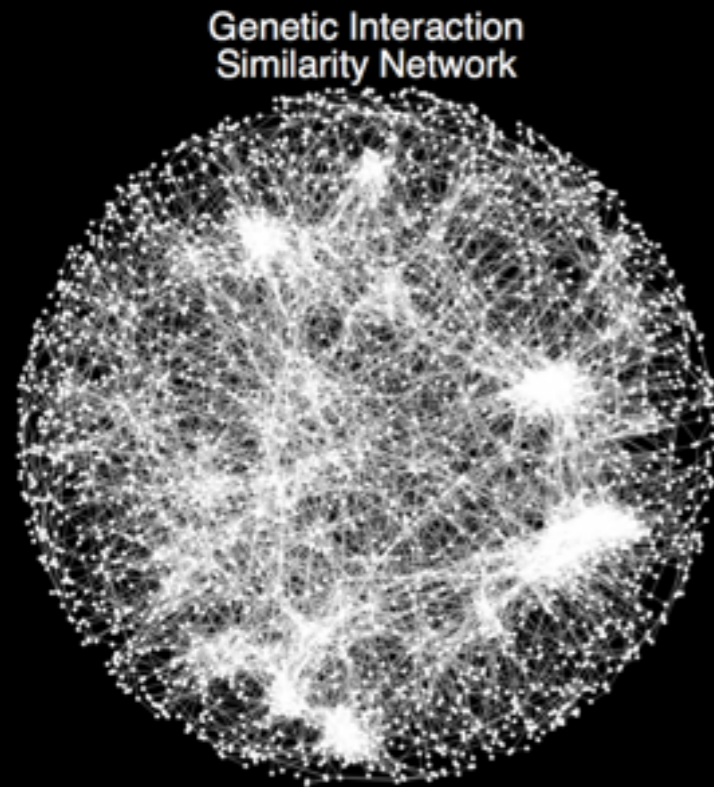


# Spatial Analysis of Functional Enrichment (SAFE)

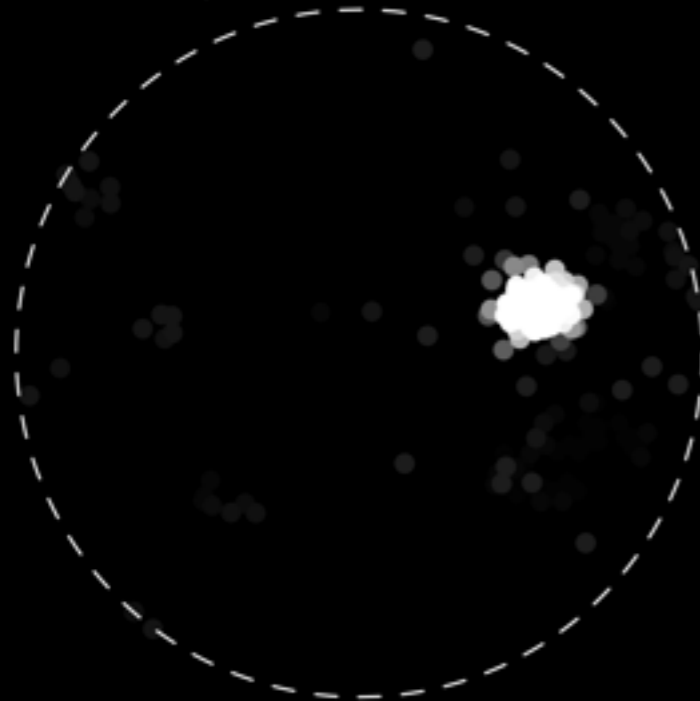


1. Take a node **A** in the network
2. Find all other nodes **B** that can be reached from **A** by traveling no more than  $d$
3. Determine whether or not nodes **B** are statistically enriched for a functional group (e.g., a GO term)
4. Associate **A** with the  $-\log_{10}$  of the enrichment  $p$ -value (normalized to  $[0,1]$  range).

# Different GO terms show different patterns of enrichment

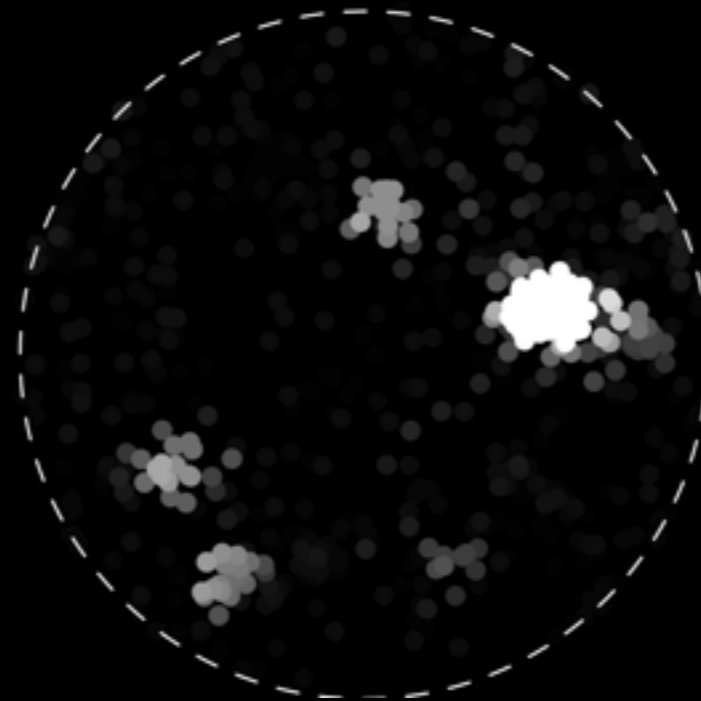


GO:0006888  
ER to Golgi vesicle-mediated transport



12%  
Region-specific

GO:0051649  
Establishment of localization in cell



4%  
Multi-regional

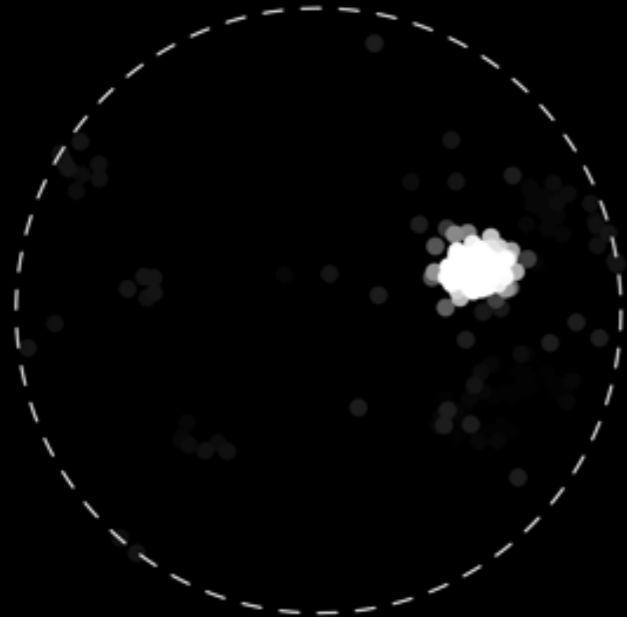
GO:0006820  
Anion transport



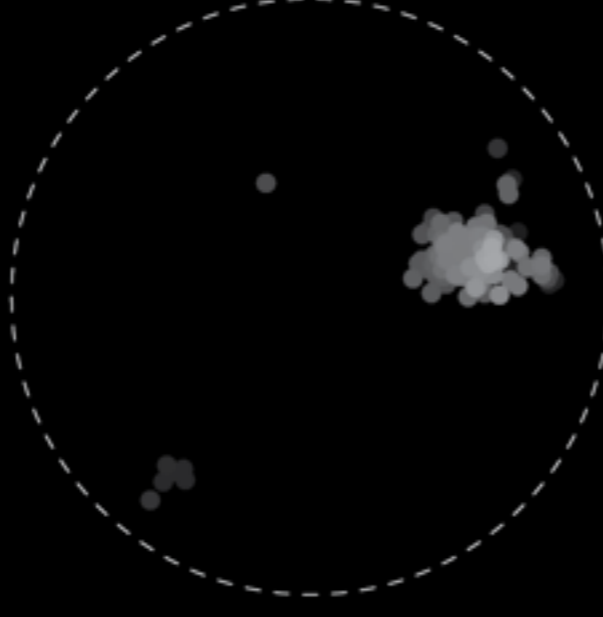
84%  
Sparse/small

# Related processes = similar patterns of enrichment

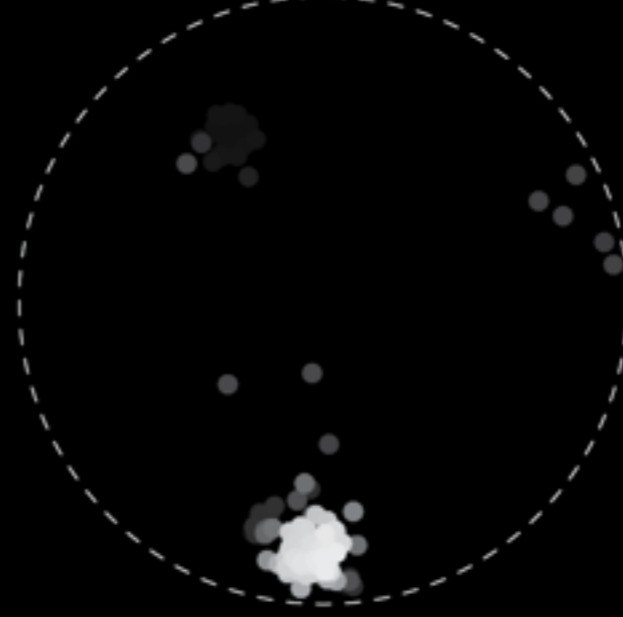
GO:0006888  
ER to Golgi vesicle-mediated transport



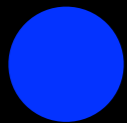
GO:0006895  
Golgi to endosome transport



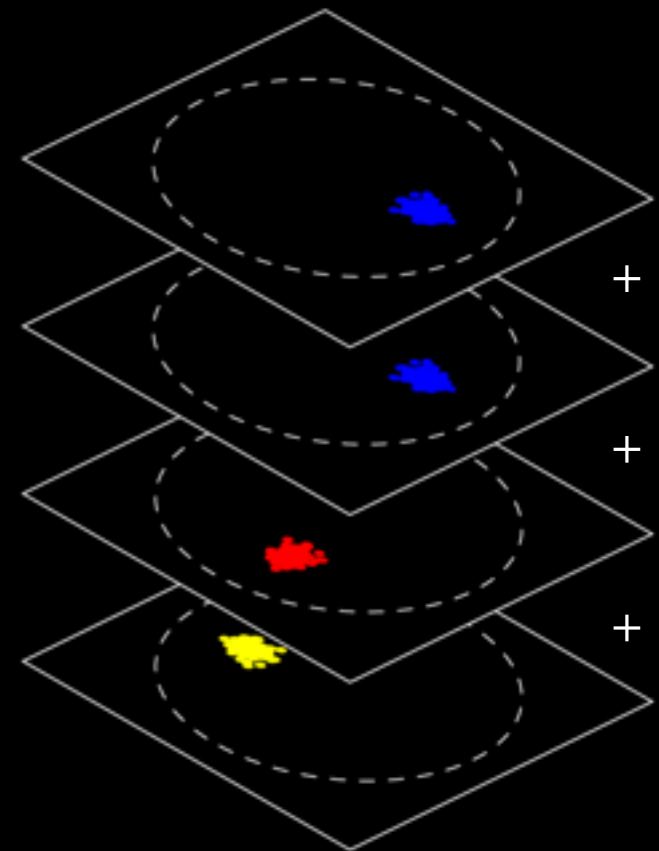
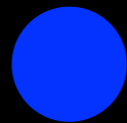
GO:0022616  
DNA strand elongation



Similar  
enrichment  
landscapes



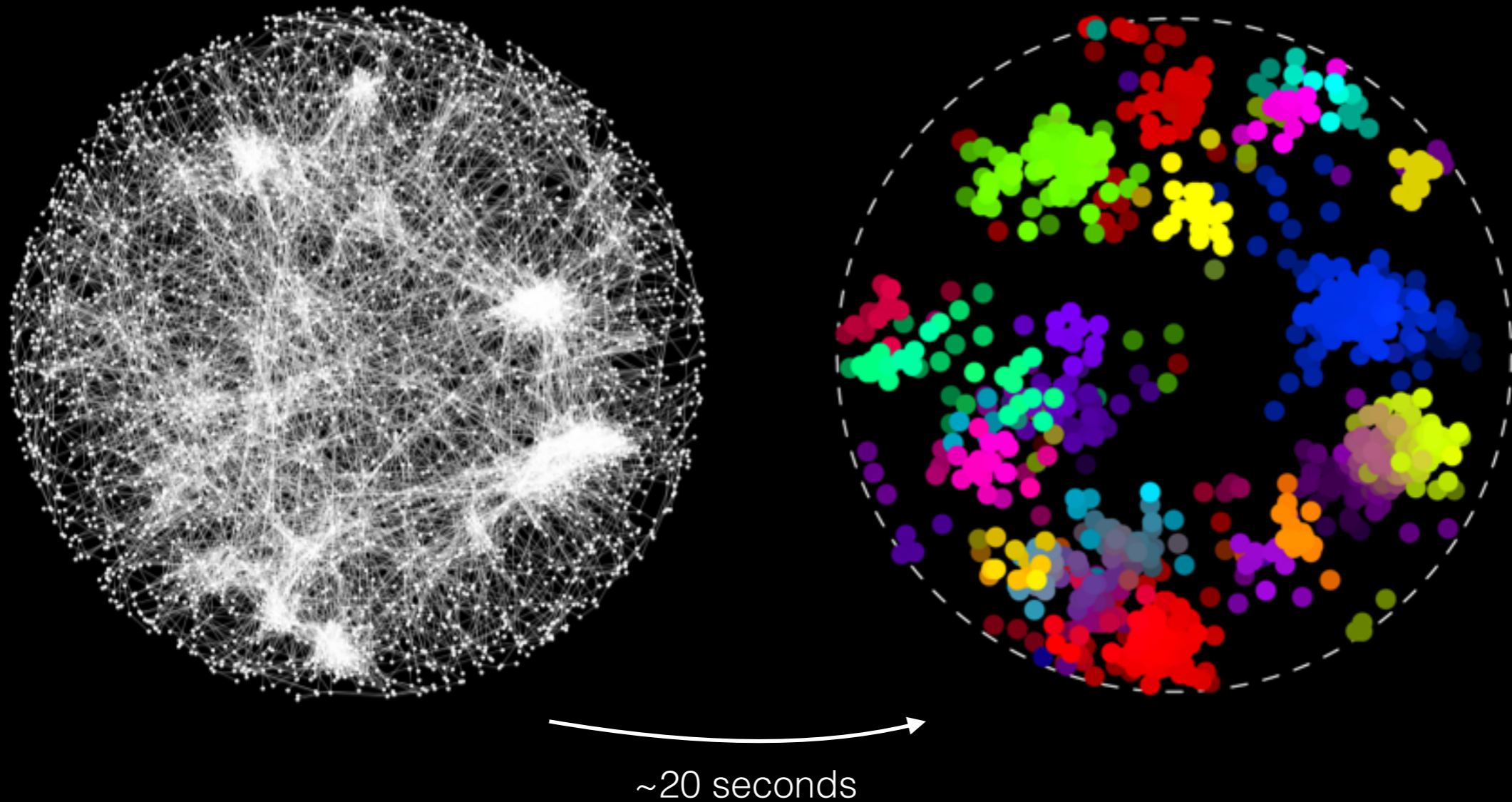
Different  
enrichment  
landscapes



Combine all enrichment  
landscapes  
(proportionally to their  
intensities)

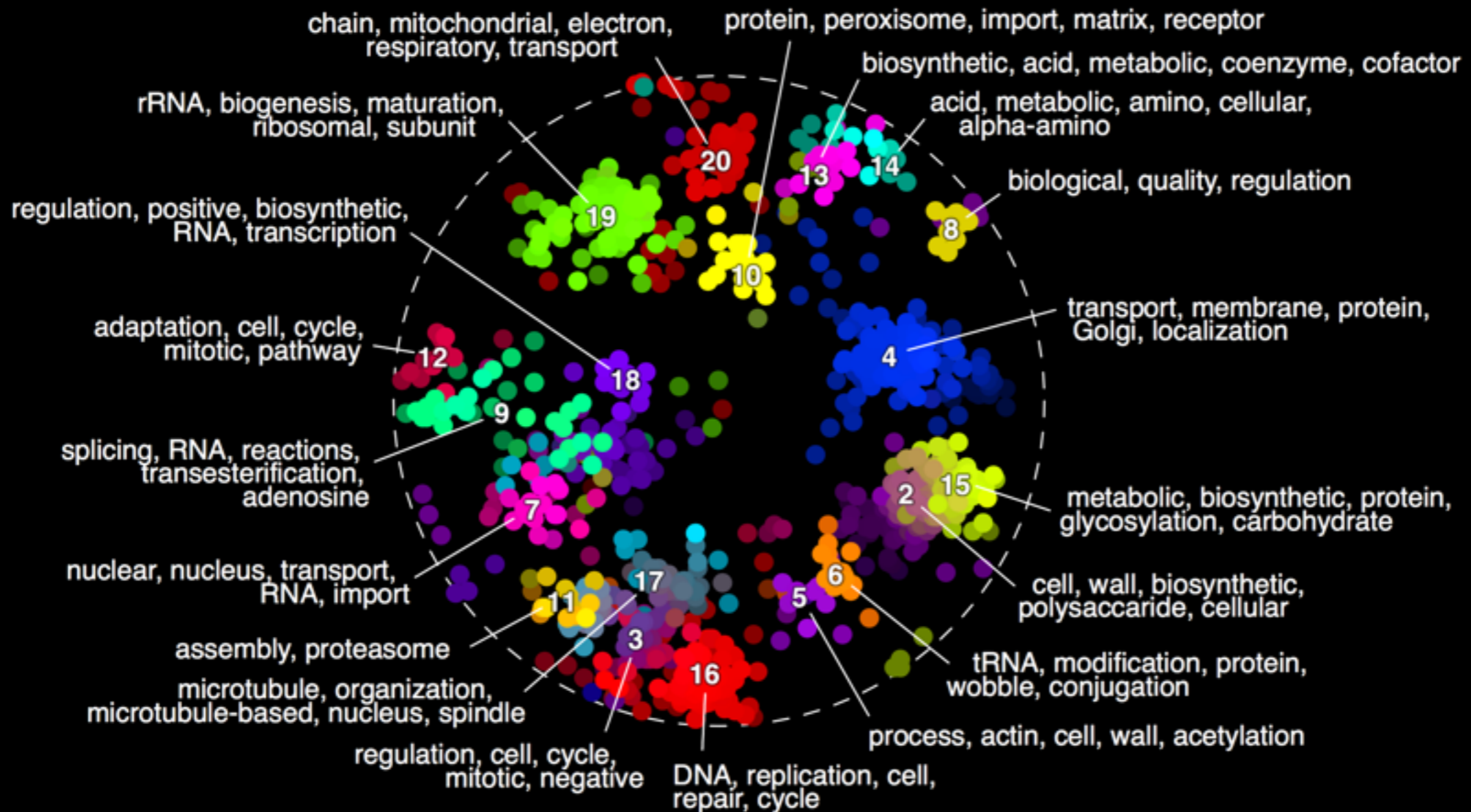
# The automated functional map of the yeast genetic interaction similarity network

Every color = a group of GO terms enriched in that region



# The Automated Functional Map of the Yeast Genetic Interaction Similarity Network

Every color = a group of GO terms enriched in that region



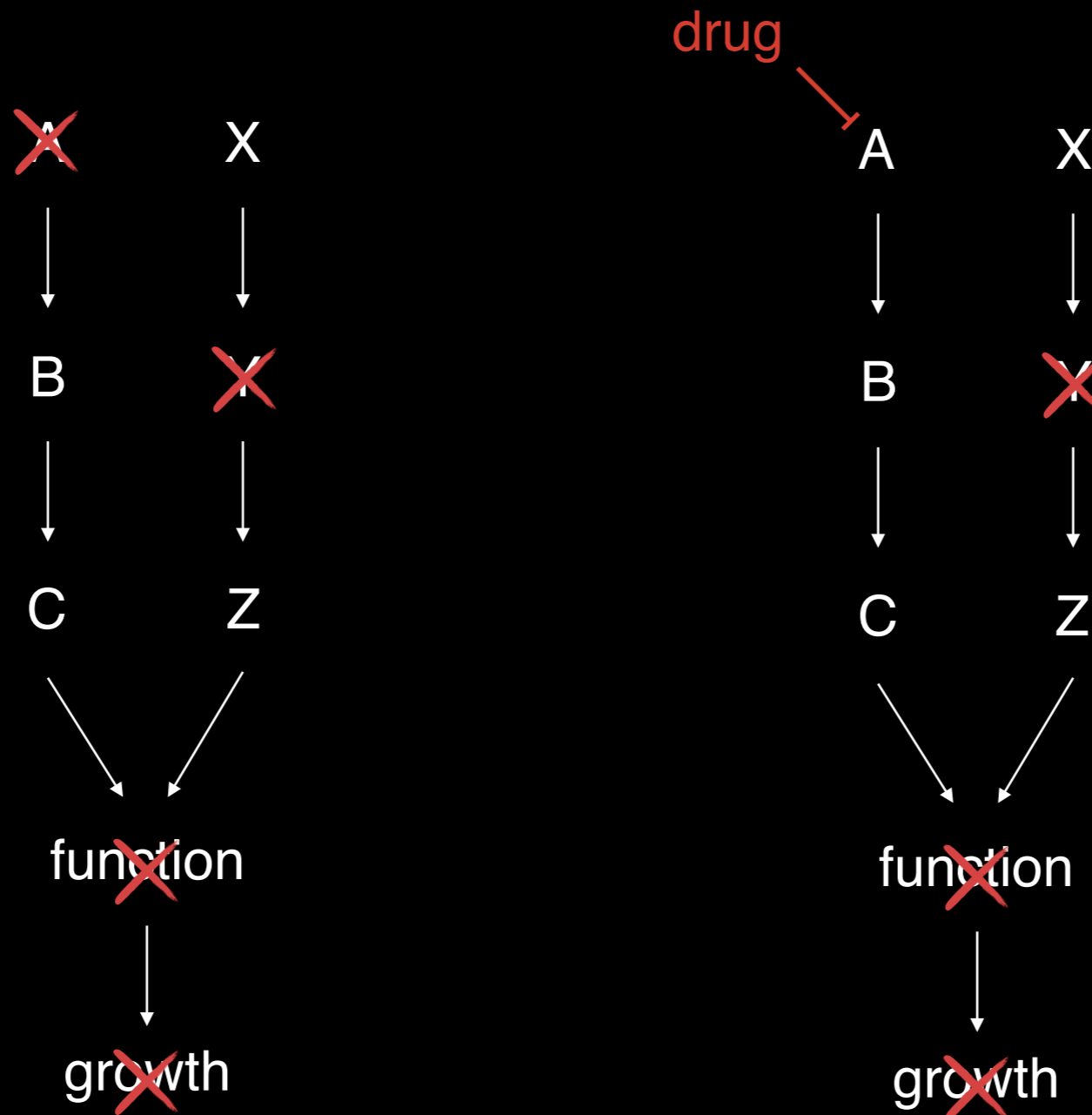
Every label = the top 5 most frequent words in the names of the GO terms

# SAFE is sensitive & robust to biological signal

- It identifies all manually annotated regions + 3 more.
- It is robust to numerous sources of variation (independent layout runs, distance metrics, neighborhood radius, annotation errors).
- It is fast & automated → use multiple independent functional standards to annotate the same network.
  - E.g., yeast: 1000s of phenotypic screens, including chemical genomics.

# The chemical genomic advantage

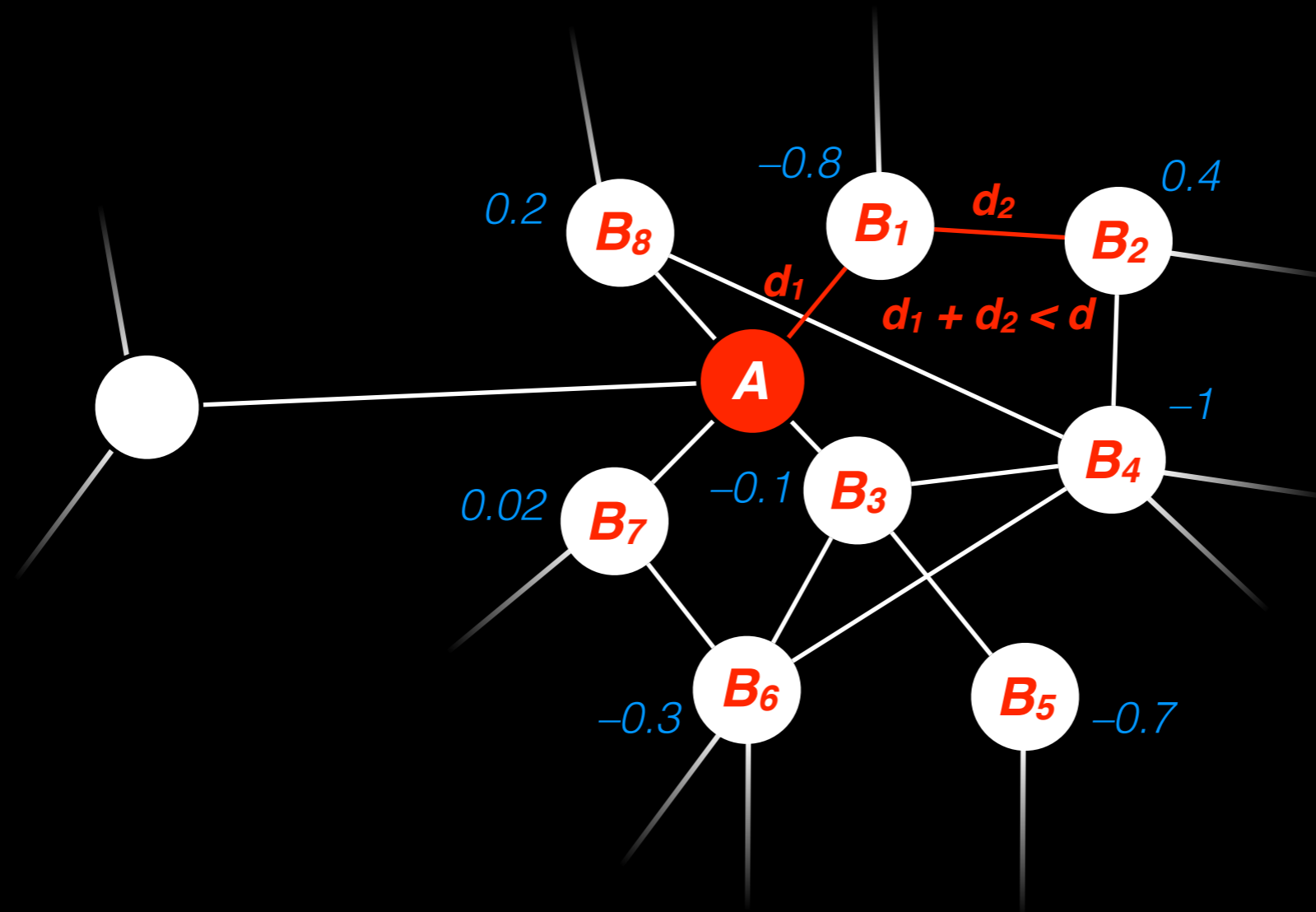
Drugs can mimic the phenotypes of their mutated targets.



## Test case:

132 chemical-genomics screens for drugs with known modes-of-action  
Hoepfner *et al.*, *Microb. Res.*, 2014

# Spatial Analysis of Functional Enrichment (SAFE)



$$\text{Score} = 0.4 + 0.2 + 0.02 - 0.1 - 0.3 - 0.7 - 0.8 - 1 = -2.28$$

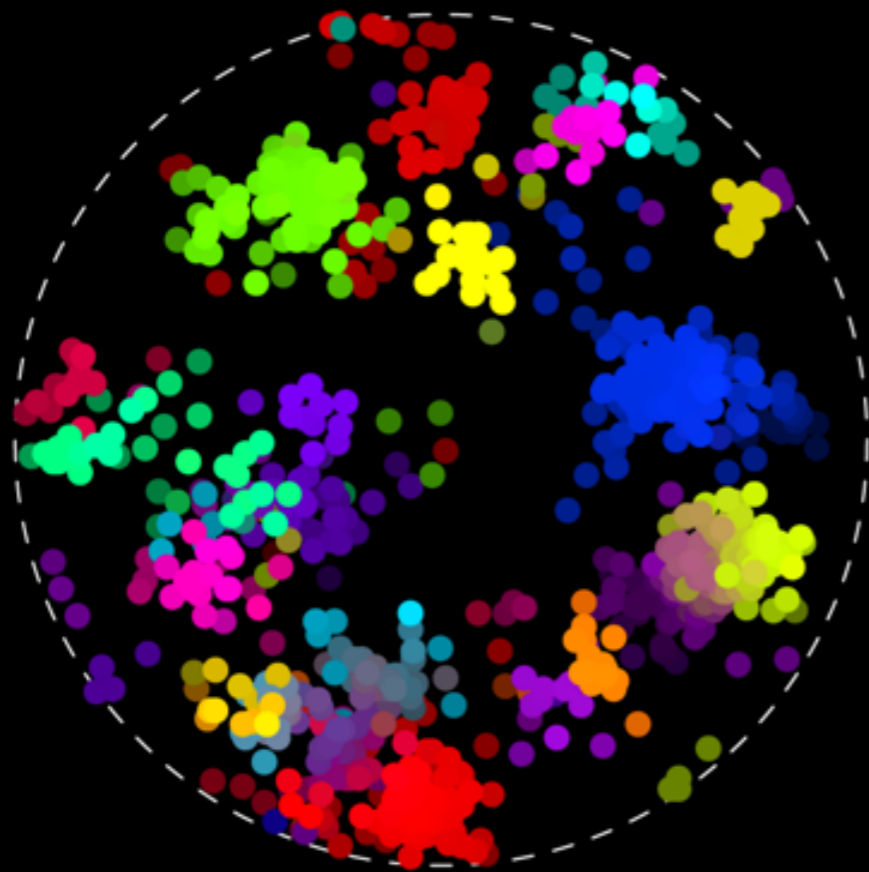
Is this higher or lower than you would expect by random chance?

1. Take a node **A** in the network
2. Find all other nodes **B** that can be reached from **A** by traveling no more than **d**
3. Determine whether or not nodes **B** are statistically enriched for a ~~functional group~~ (e.g., a GO term) *quantitative phenotype*
4. Associate **A** with the  $-\log_{10}$  of the enrichment *p*-value (normalized to [0,1] range).

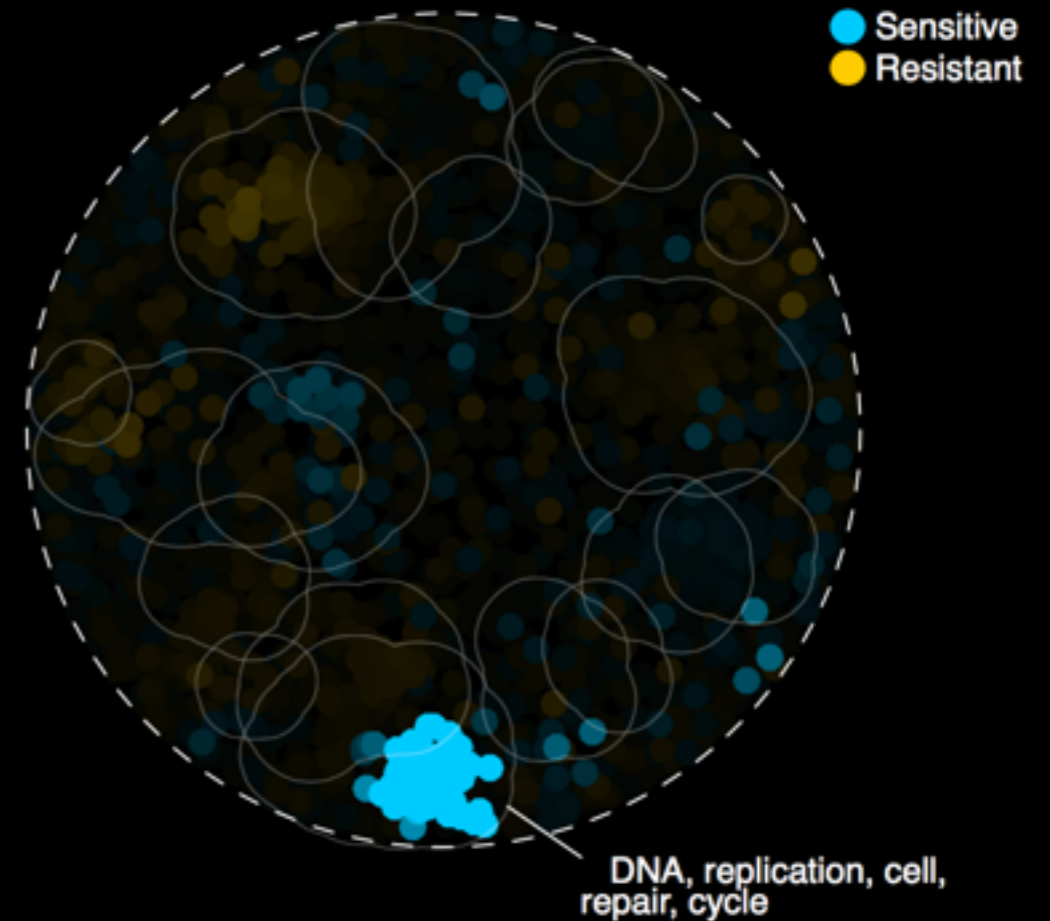


# SAFE recapitulates the known modes-of-action of chemical compounds

Gene Ontology Annotation



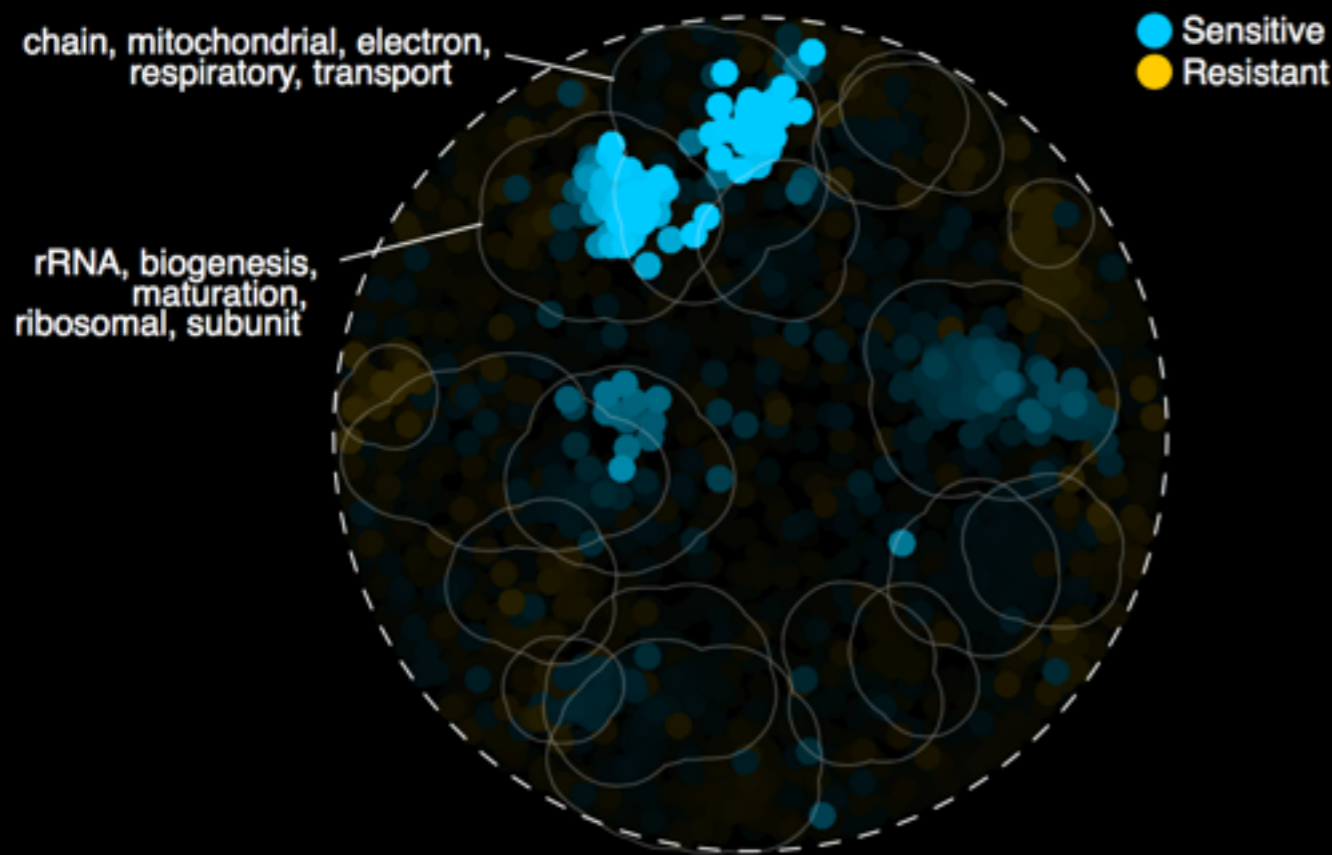
Doxorubicin  
DNA intercalator, blocks replication



# SAFE recapitulates the known modes-of-action of chemical compounds

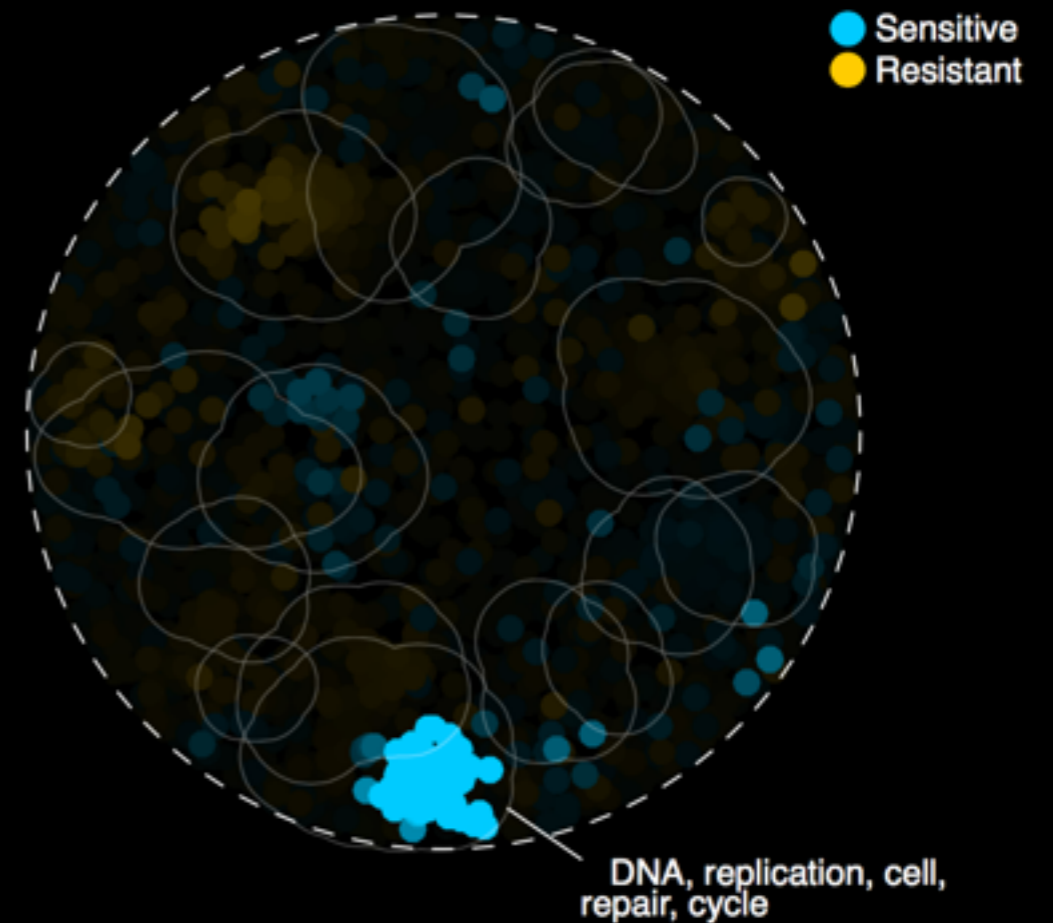
## Verrucarin A

Inhibits protein synthesis and mitochondrial function



## Doxorubicin

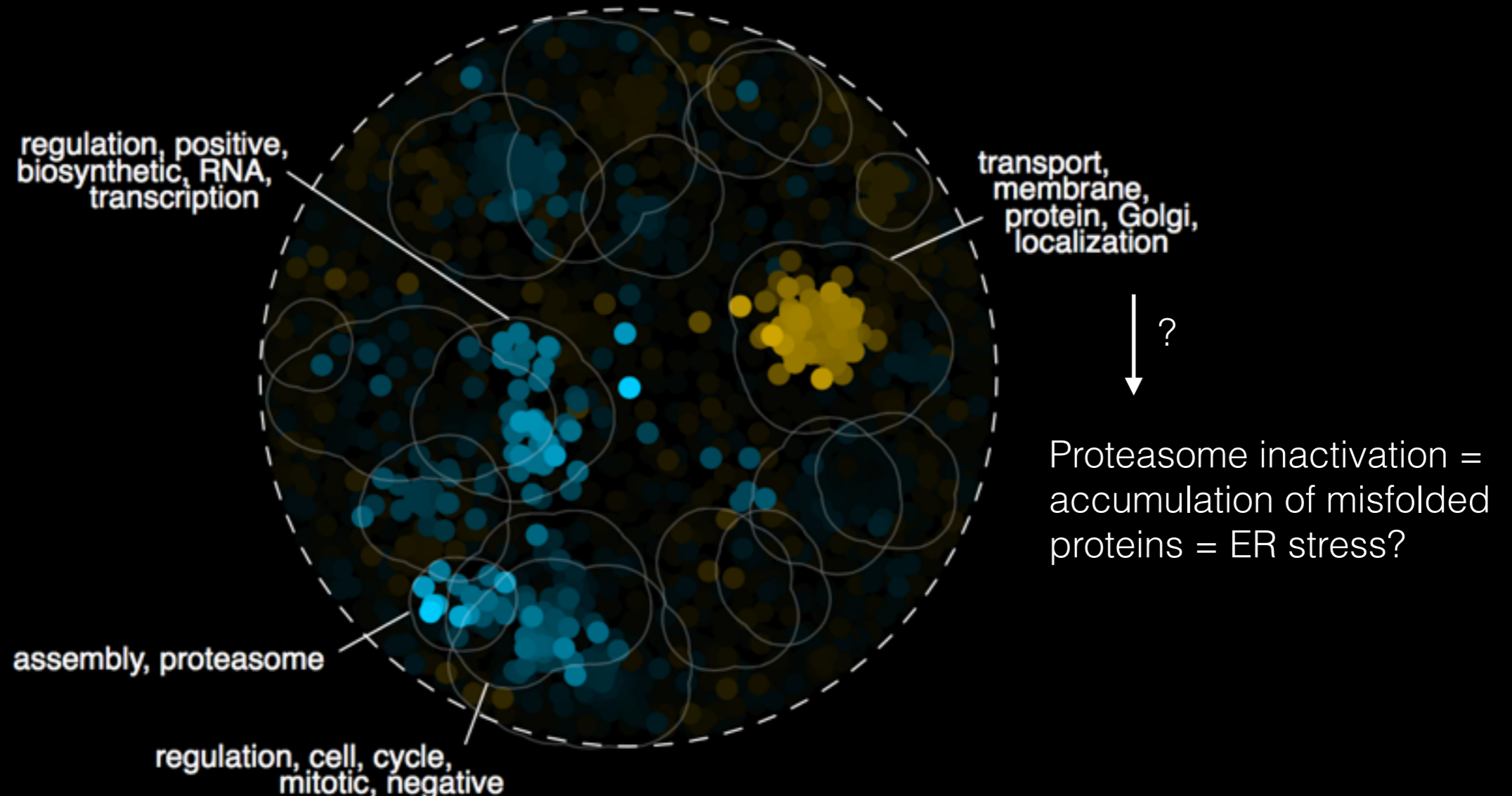
DNA intercalator, blocks replication



# SAFE uncovers potentially novel mechanisms of drug activity

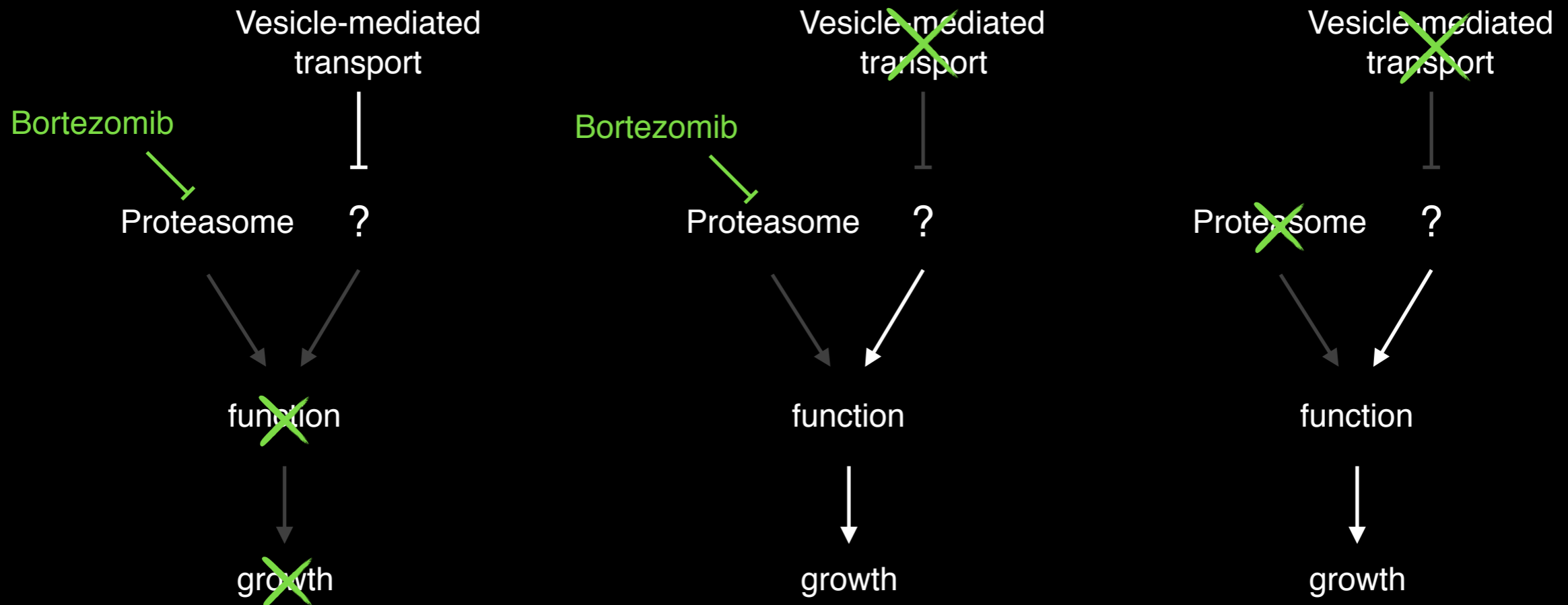
## Bortezomib a.k.a. Velcade and Ps-341

- The only FDA-approved proteasome inhibitor.
- Used for treating multiple myeloma and mantle cell lymphoma.
- Promotes programmed cell death.
- Synergistic with HDAC inhibitors.



# Could it be a side effect of bortezomib?

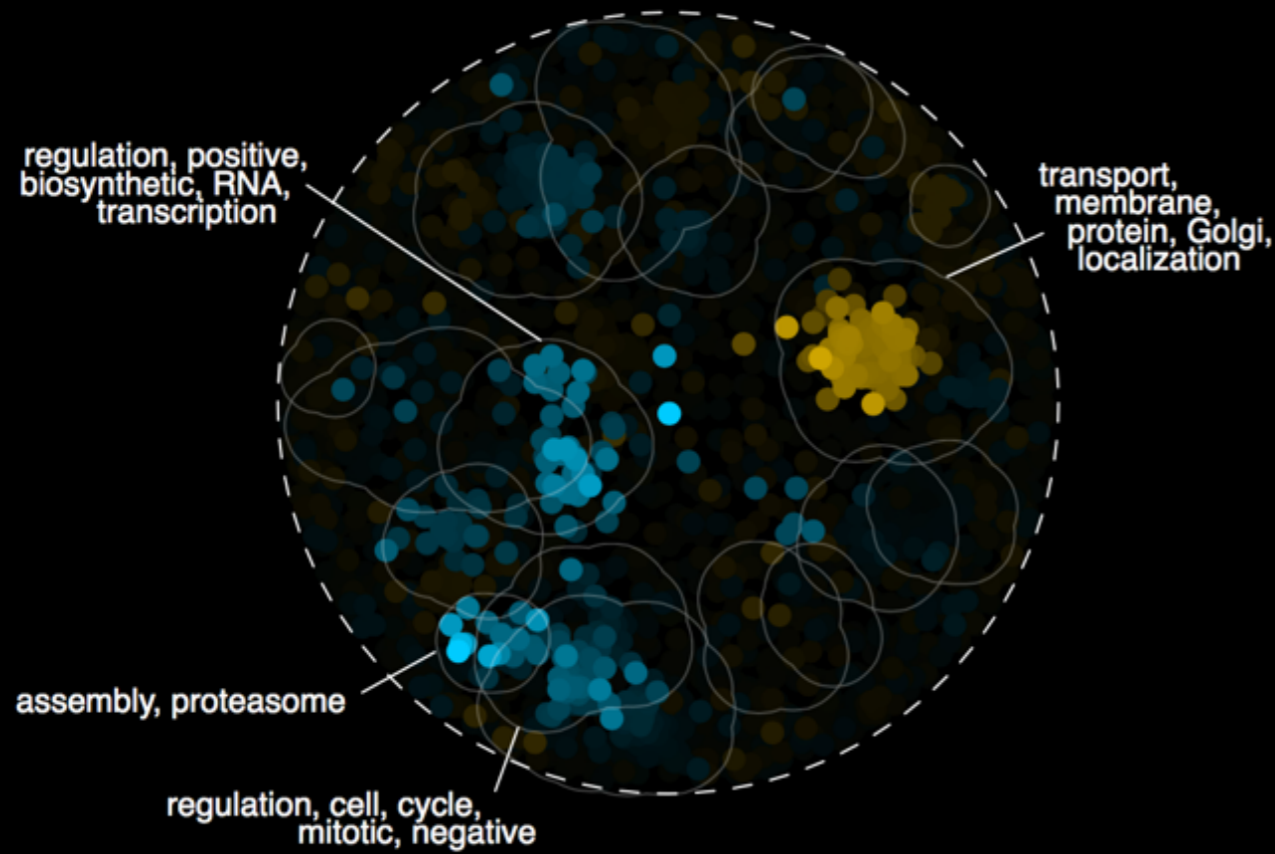
Mutating drug targets can recapitulate their chemical-genetic interactions.



# Proteasome mutants recapitulate the effect of bortezomib treatment

Bortezomib

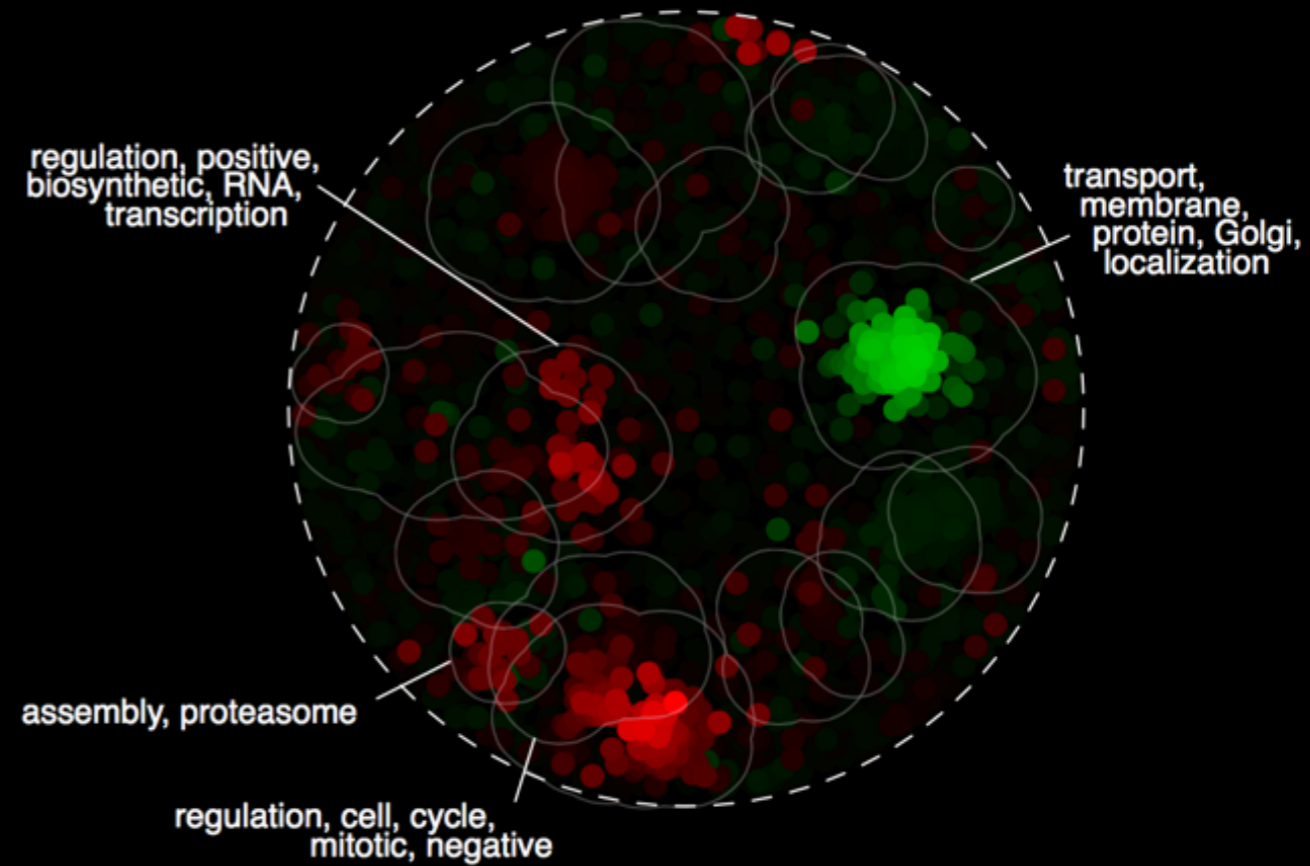
Proteasome inhibitor



● Sensitive  
● Resistant

*rpt4-145*

ts mutant in an ATPase of the regulatory subunit of the proteasome



● Negative genetic interaction  
● Positive genetic interaction

# Spatial Analysis of Functional Enrichment (SAFE)

- An automated & systematic method for annotating biological networks.
- Answers 3 fundamental questions:
  1. Are any regions of the network specifically associated with a given function or phenotype?
  2. Where in the network are these regions localized?
  3. How does their localization compare to that of other functions or phenotypes?
- Builds a functional map of the network and enables the investigation of inter-process relationships.
- Allows to integrate multiple functional annotations of the same network and gain insight into the molecular mechanisms of drug response.

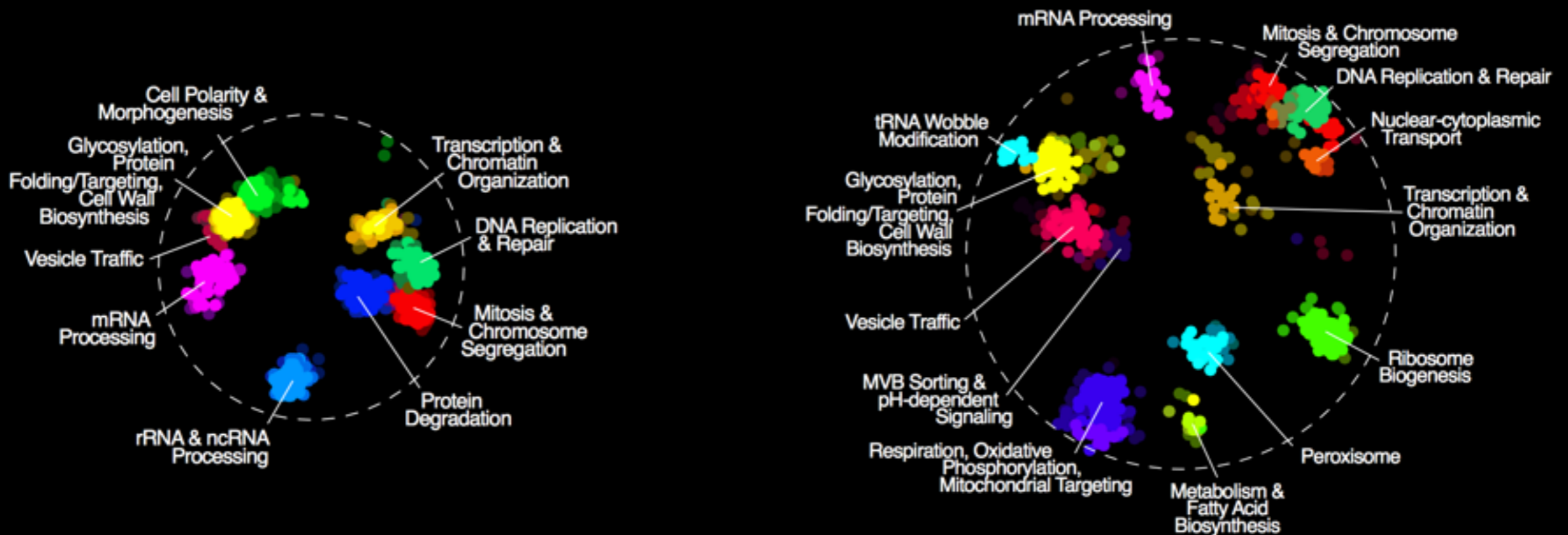
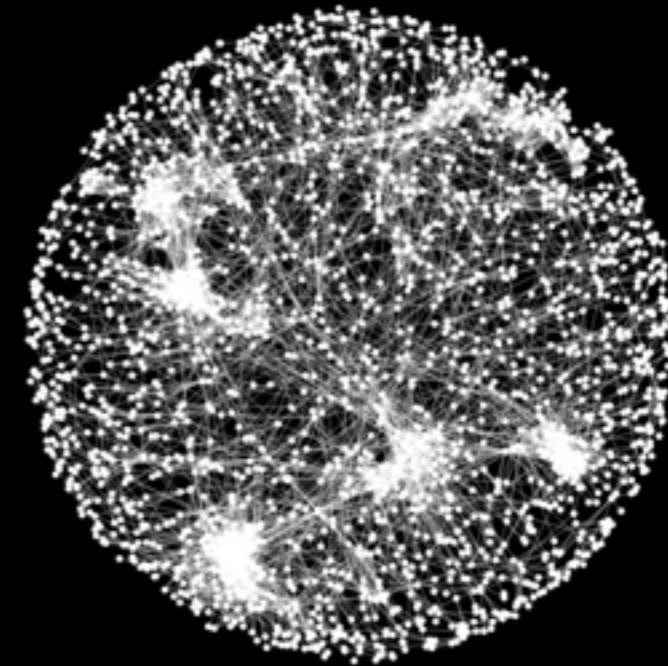
# 2016 Yeast Genetic Interaction Network

*The essential and non-essential networks are both similar and different.*

Essential Network

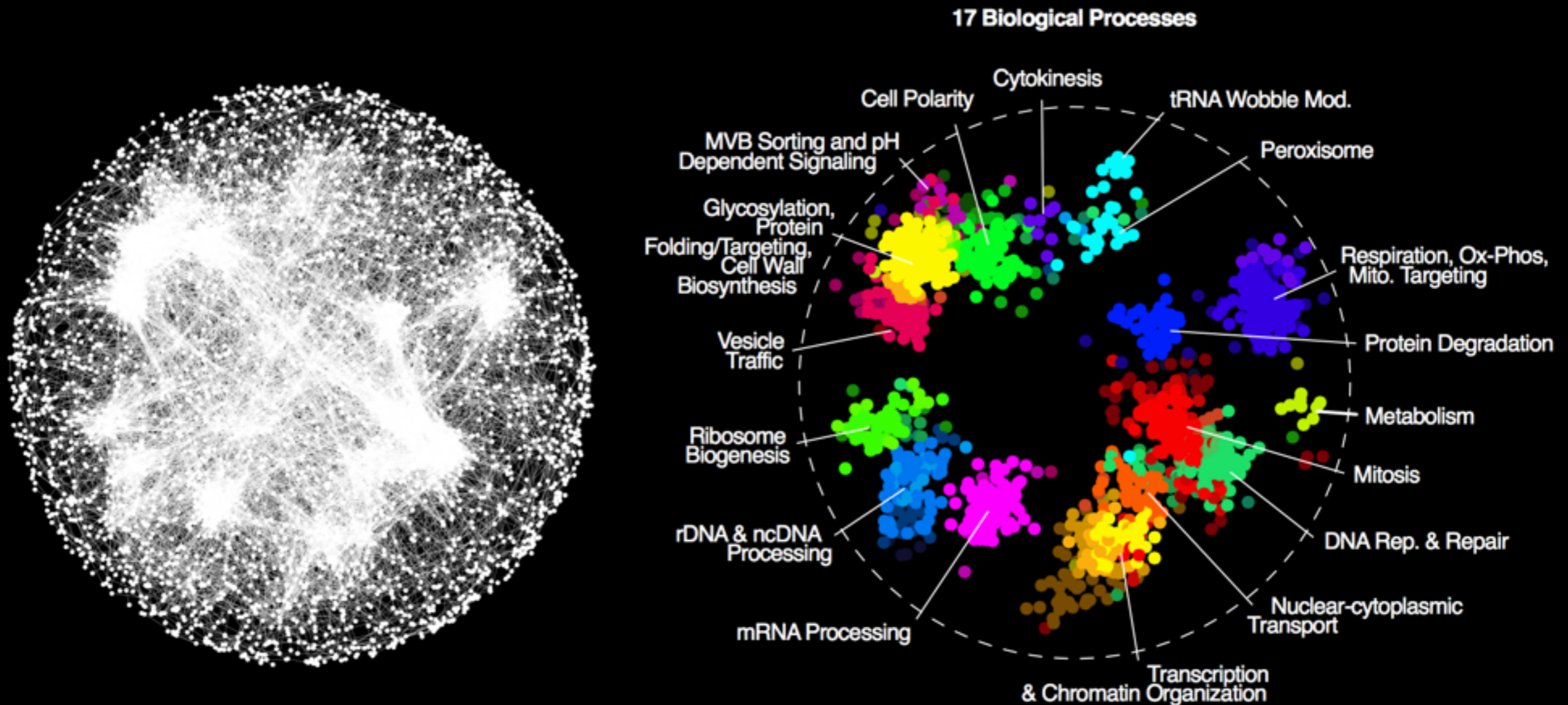


Non-essential Network



# 2016 Yeast Genetic Interaction Network

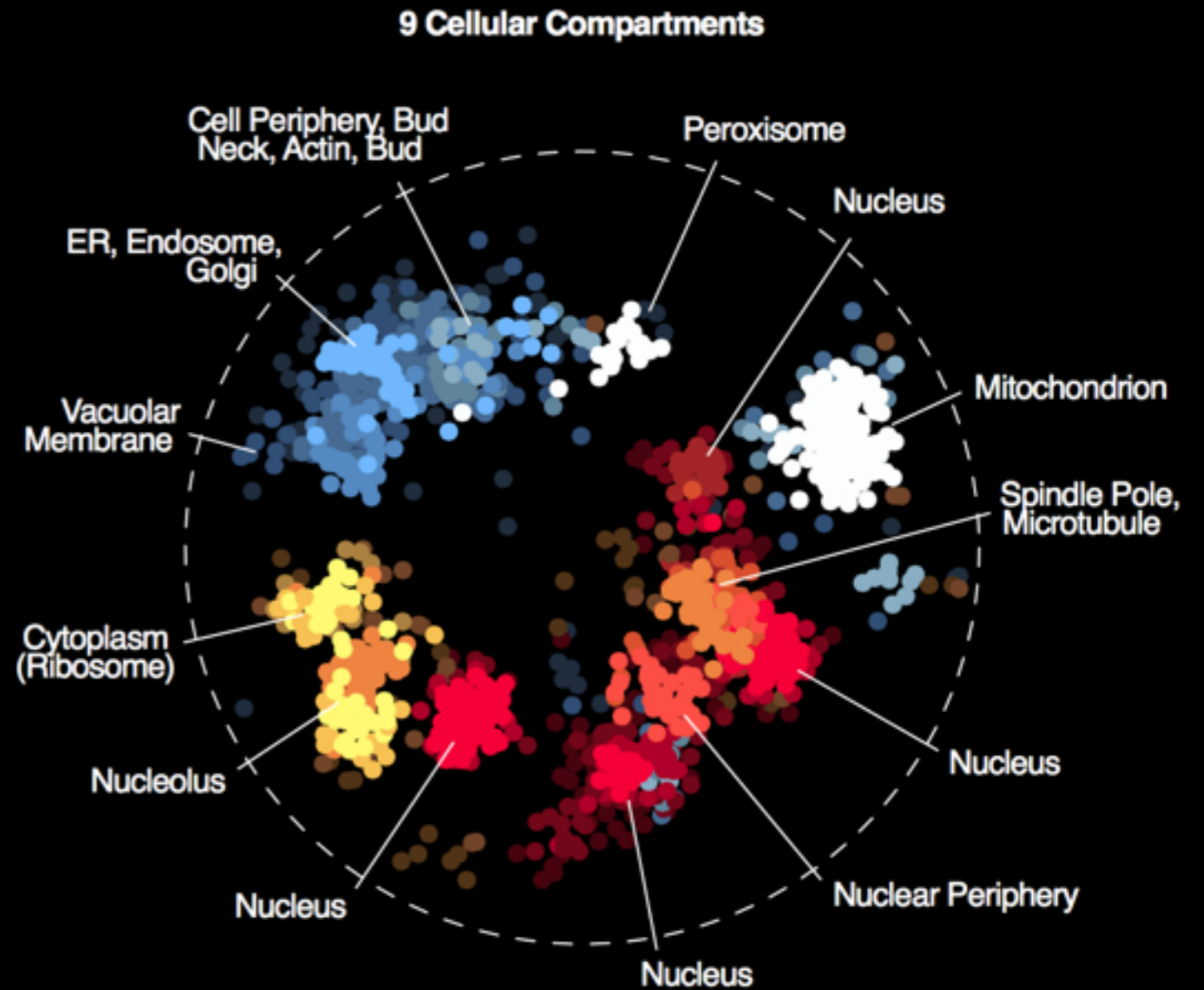
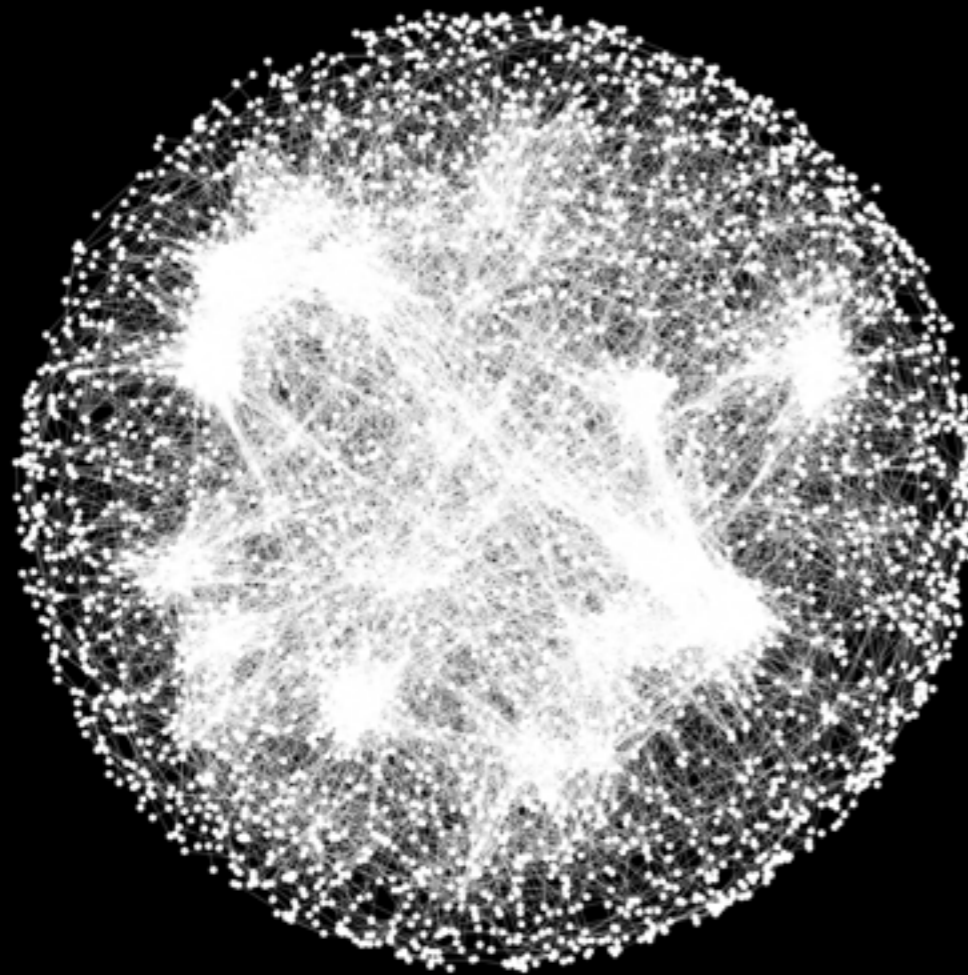
*Composite Network: with respect to GO biological process, similar to the 2010 network.*





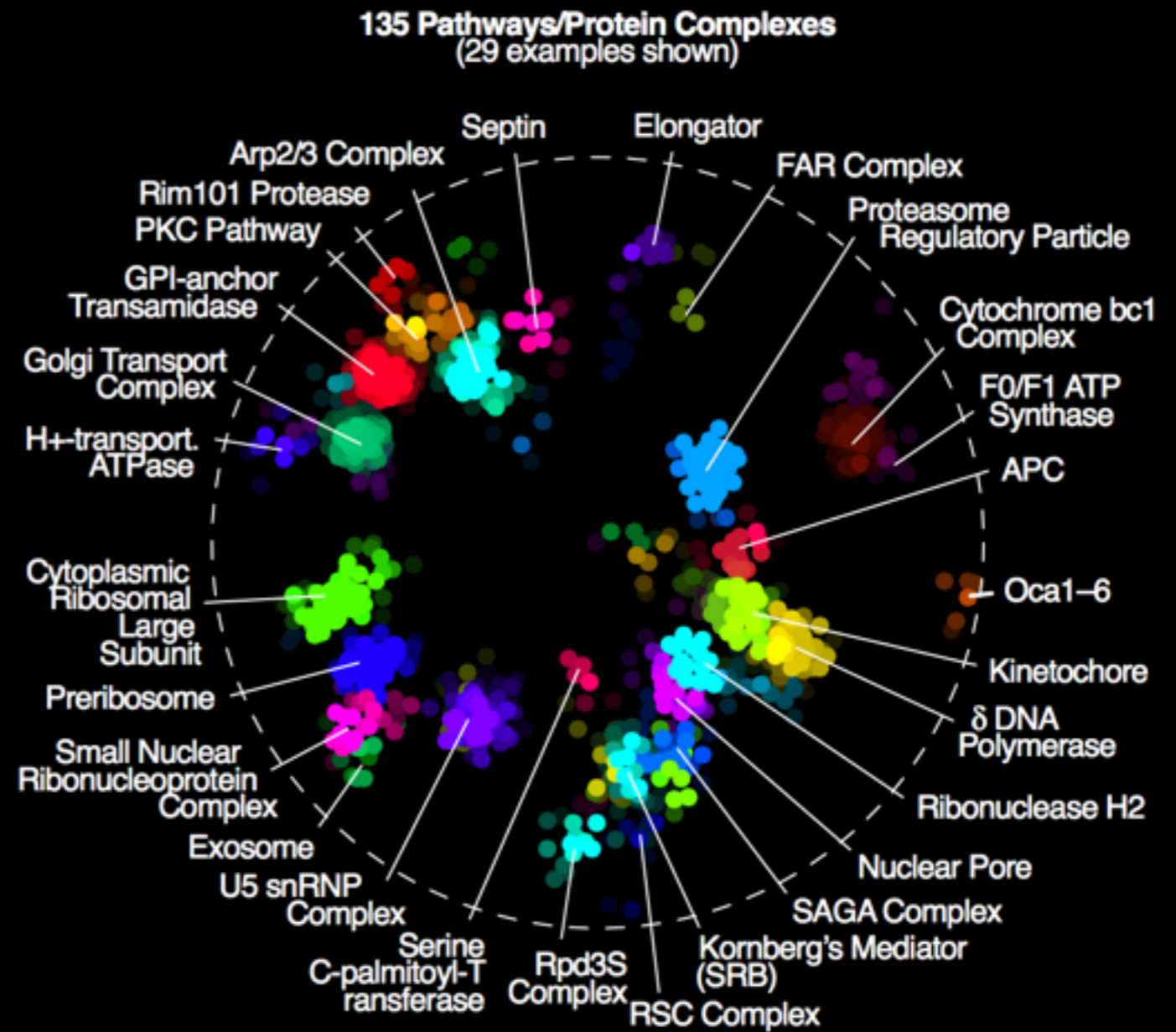
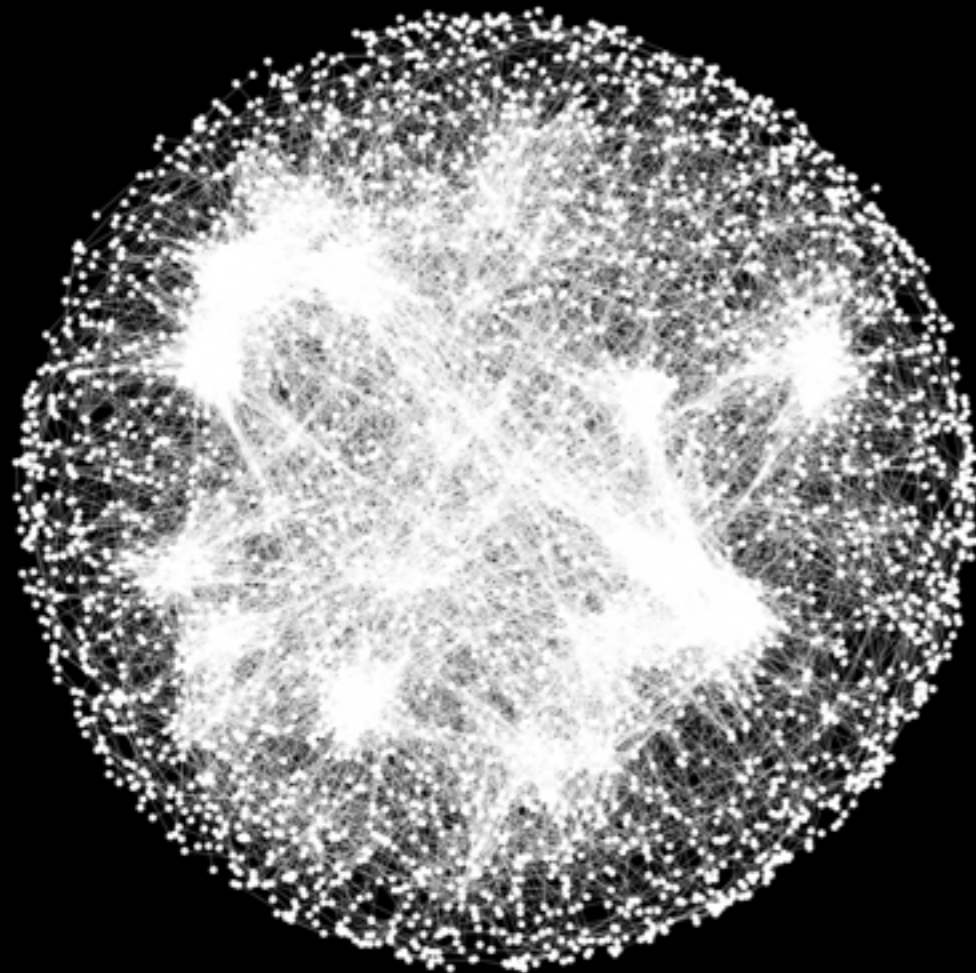
# 2016 Yeast Genetic Interaction Network

Huh et al., 2003 — Cellular localization: processes appear to organize into larger modules associated with cellular compartments.



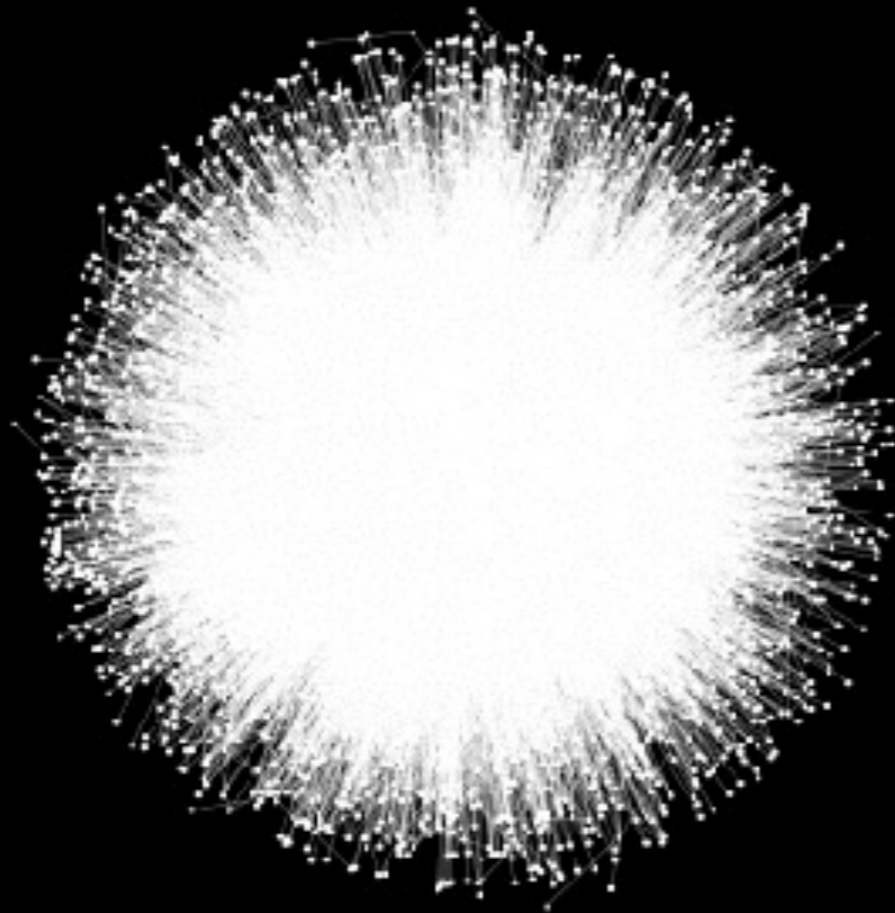
# 2016 Yeast Genetic Interaction Network

*Protein complexes: processes appear to subdivide into smaller modules associated with sets of protein complexes and/or pathways.*

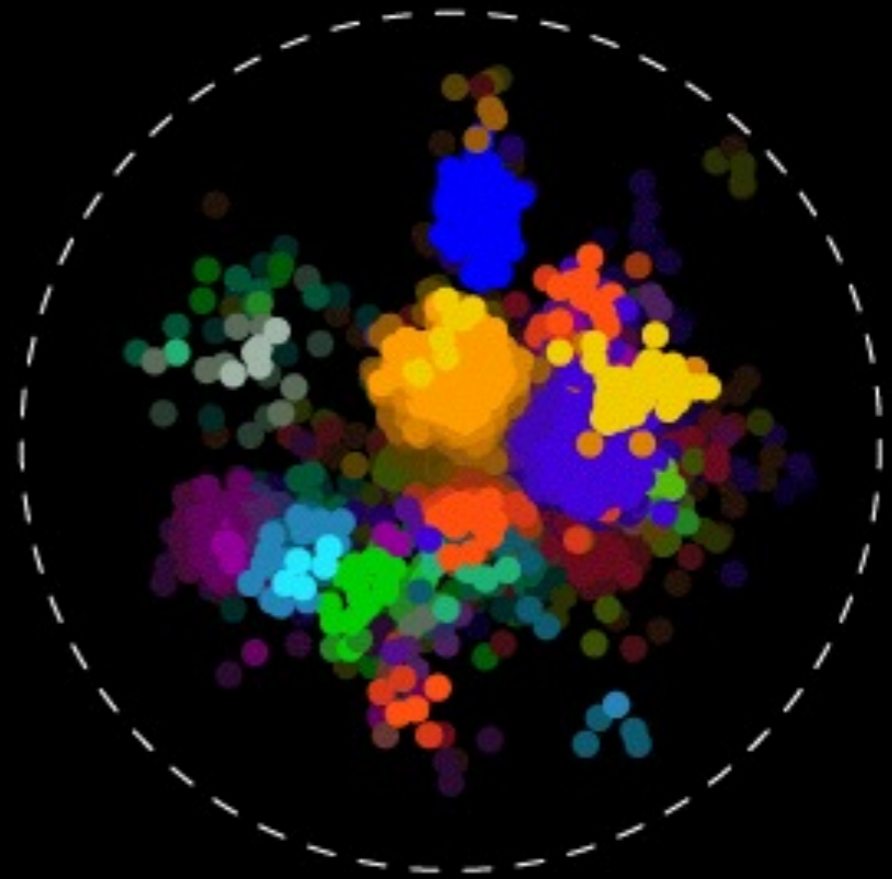


# SAFE starts revealing the functional structure of protein-protein interactions

5,699 nodes  
78,406 edges



21 functional domains  
(GO biological process)



# Acknowledgements

## Princeton University

Dmitriy Gorenshcheyn

### Undergrads

Brianna Richardson

Rachel Xu

### Treeview team

Chris Keil

Lance Parsons

Robert Leach

### YeastPhenome team

Kara Dolinski

Sven Heinicke

Rose Oughtred

Christie Chang

Jennifer Rust

Mark Schroeder

Fan Kang

## University of Toronto

Michael Costanzo

Charlie Boone

Brenda Andrews

Boone lab

Andrews lab

## University of Minnesota

Chad Myers

Benjamin VanderSluis

Elizabeth Koch

Carles Pons

Myers lab

SAFE manuscript on Biorxiv (& in press in Cell Systems):

<http://biorxiv.org/content/early/2015/11/03/030551>

SAFE code (Matlab) + Mac OS X app on Bitbucket:

<https://bitbucket.org/abarysh/safe>

SAFE app for Cytoscape:

Jason Montojo (in progress)

[www.baryshnikova-lab.org](http://www.baryshnikova-lab.org)

[www.yeastphenome.org](http://www.yeastphenome.org)