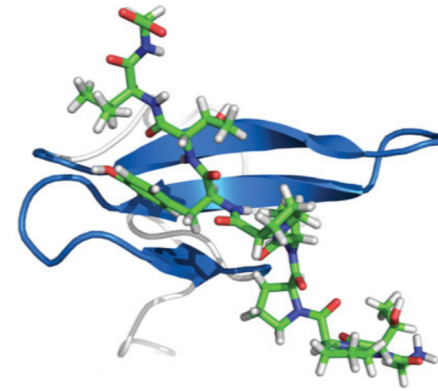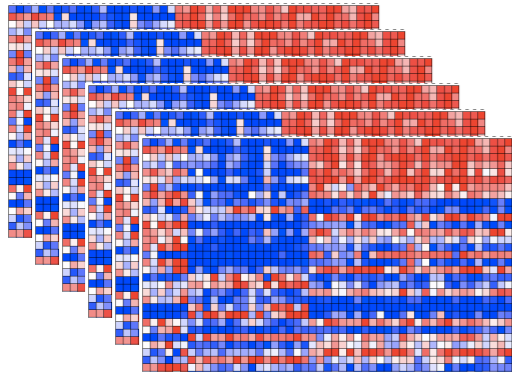# Large-scale integration of heterogeneous biological networks

Jian Peng
Department of Computer Science
University of Illinois at Urbana-Champaign

Heterogeneous data

Metabolic network

From wiki
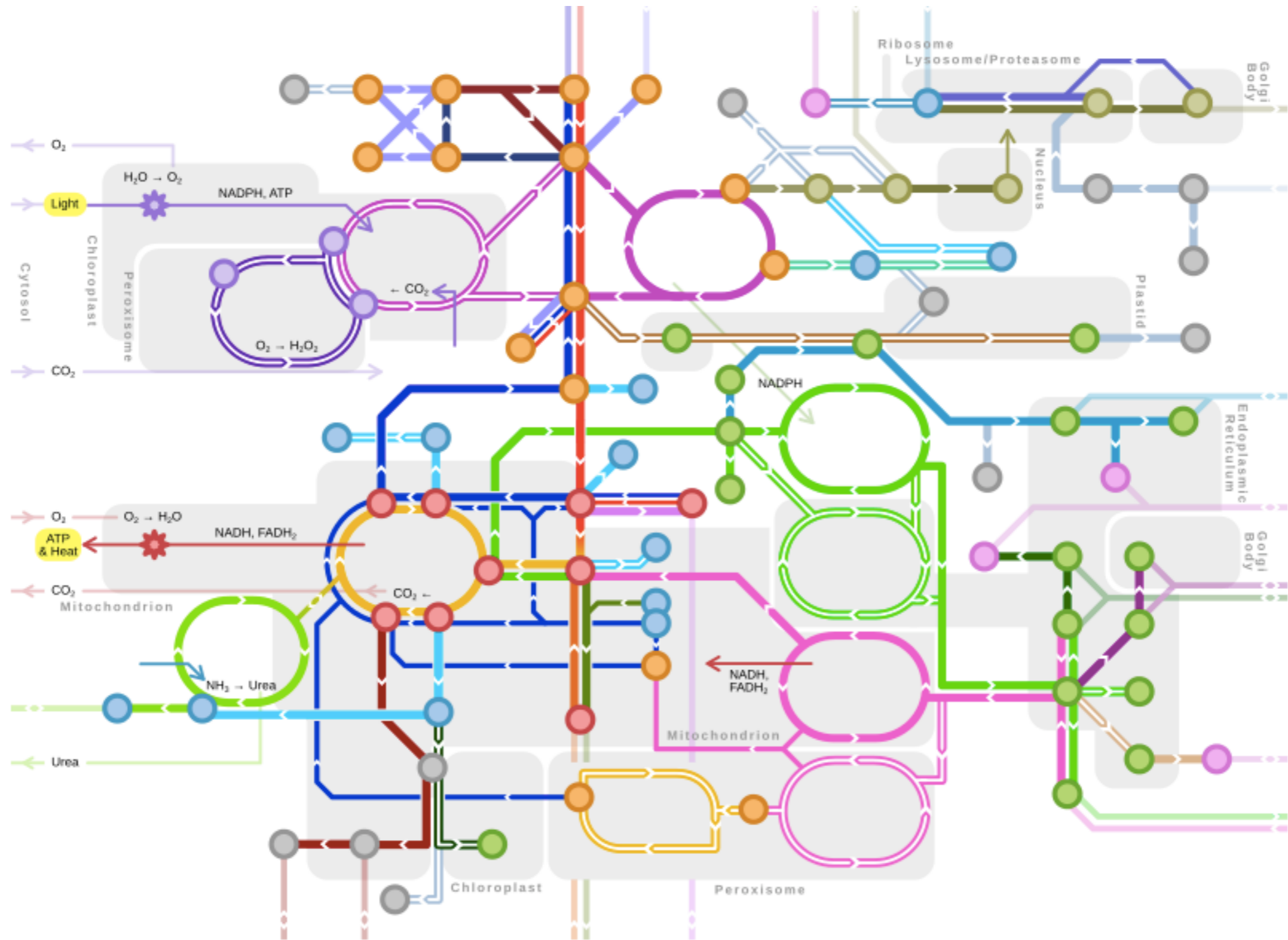
Protein-protein interaction networks
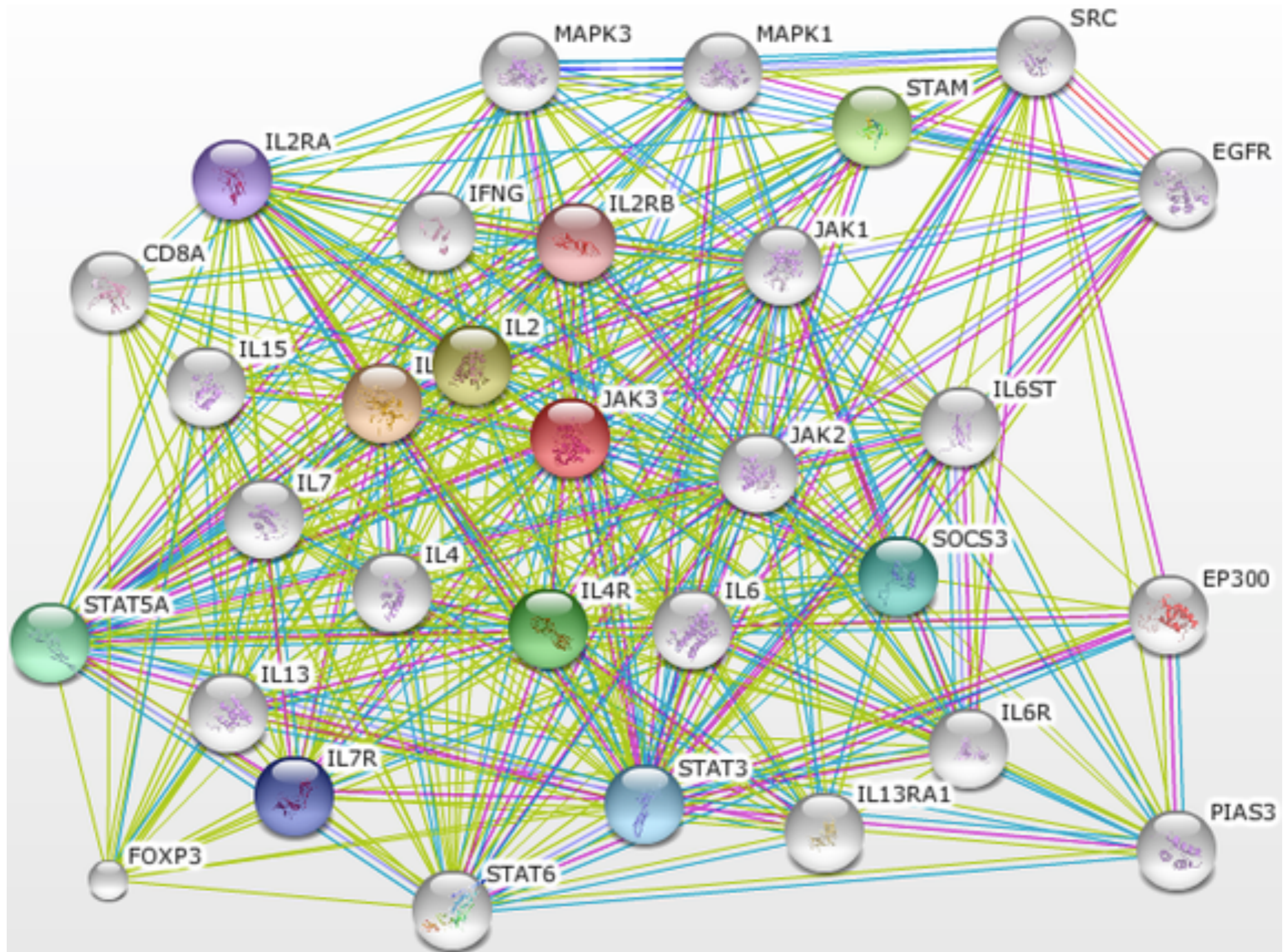
generated by String database

Heterogeneous association network
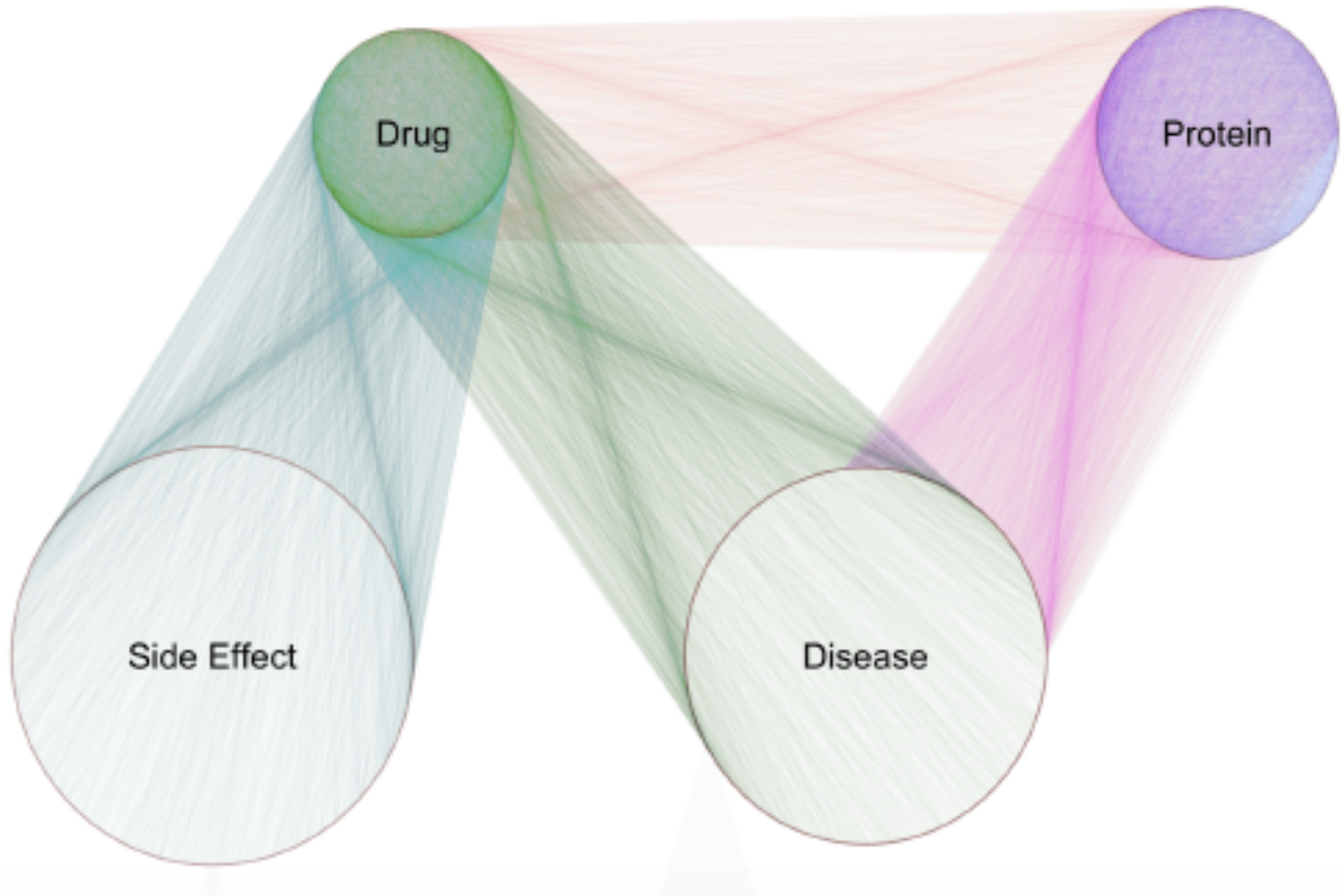
How to make sense of these networks?
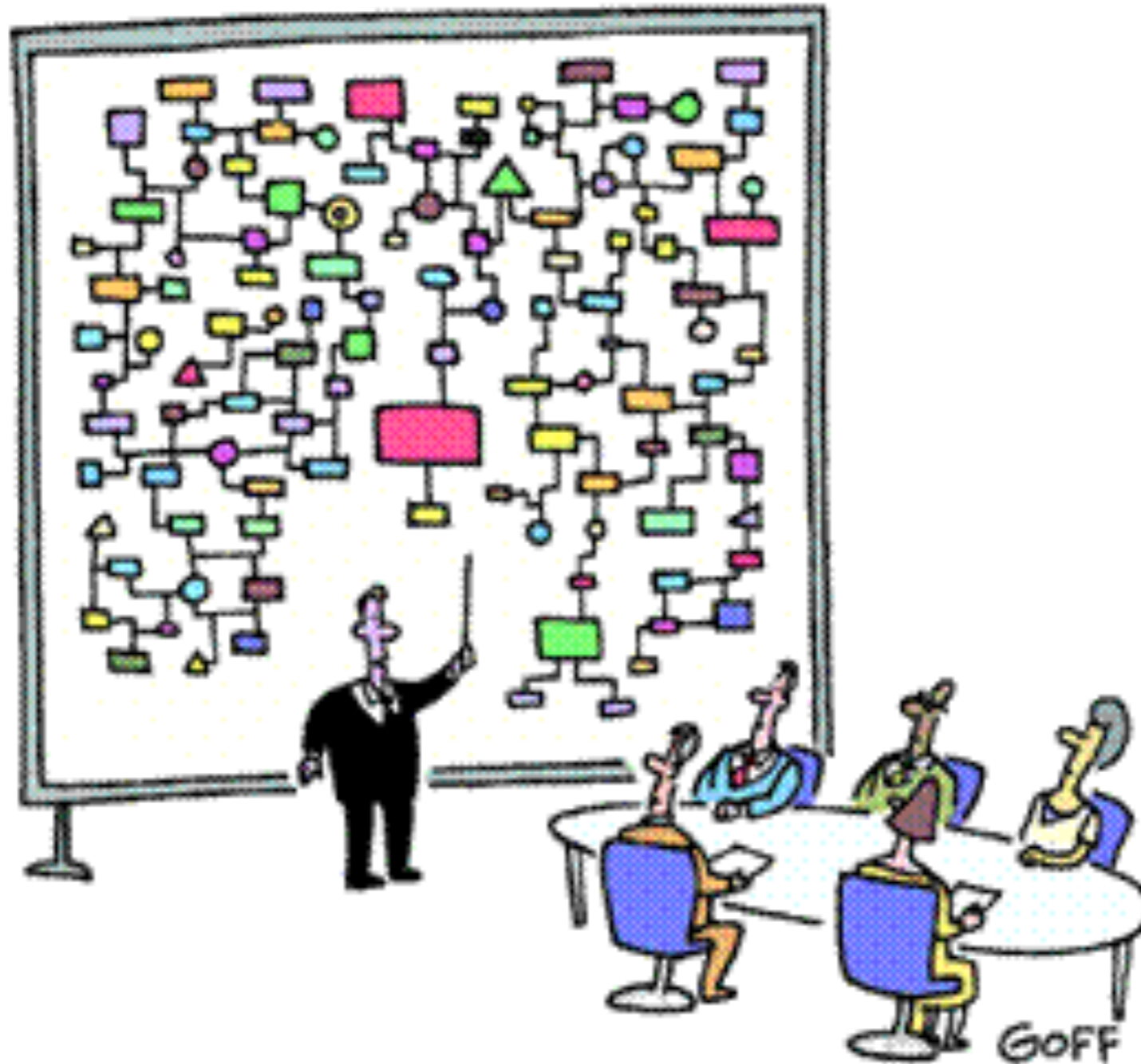
"And that's why we need a computer."

Dimensionality reduction

High-dimensional space

Node vector space

Node 3
Node 5
Node 2
Node 4
Node 1

Learning

Prediction

Low-dimensional space

Diffusion Component Analysis: Unraveling Functional Topology in Biological Networks
Hyunghoon Cho, Bonnie Berger and Jian Peng
RECOMB, 2015
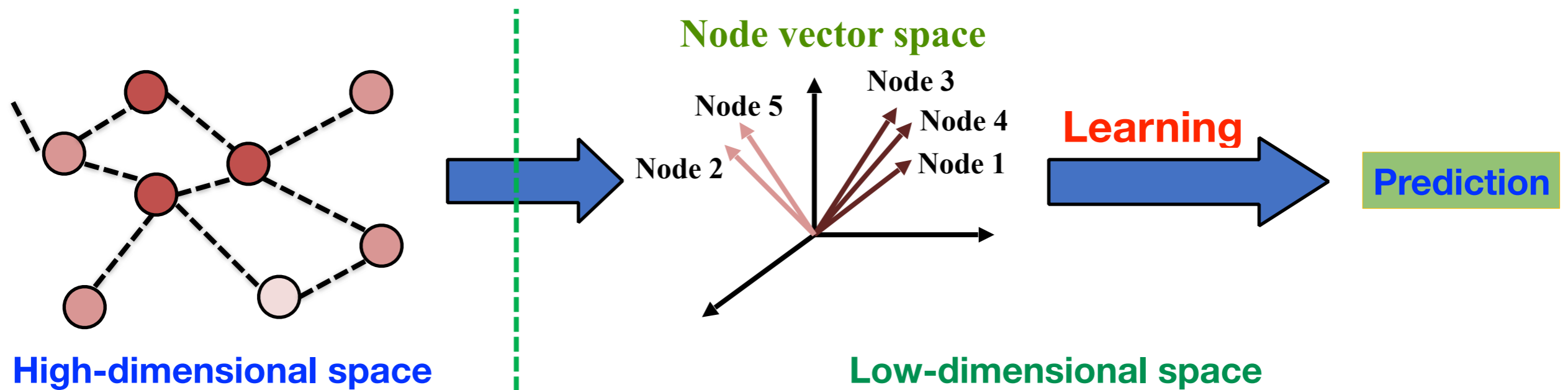
Mashup: network integration

Example: Gene function prediction

Human, GO BP, 31-100
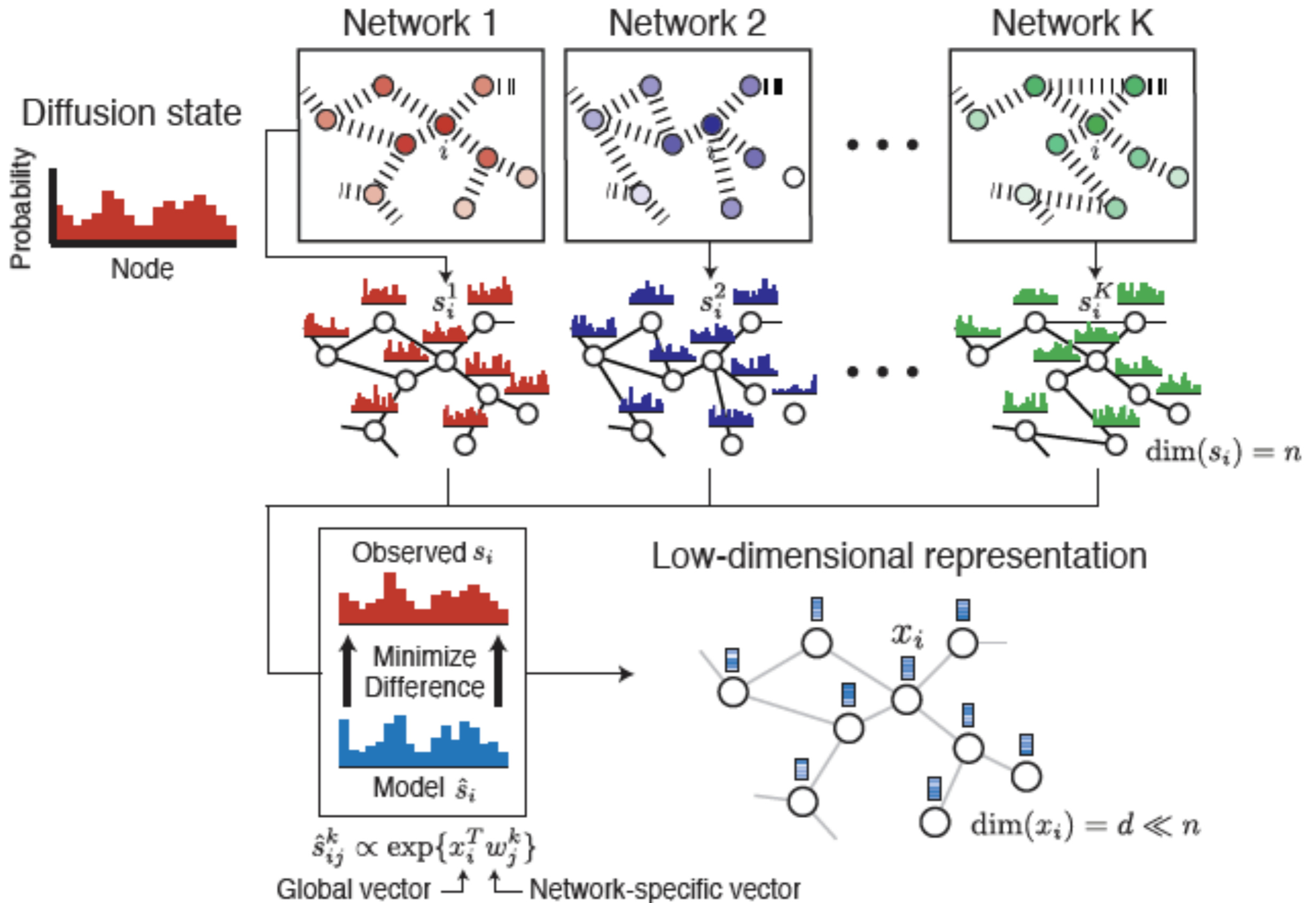
Example: Gene function prediction

A incomplete list of successful applications

Drug target prediction

Disease gene identification

Module detection

Gene essentiality prediction

Interaction prediction

Drug efficacy prediction

Cancer patient stratification

......

Drug discovery through model organisms
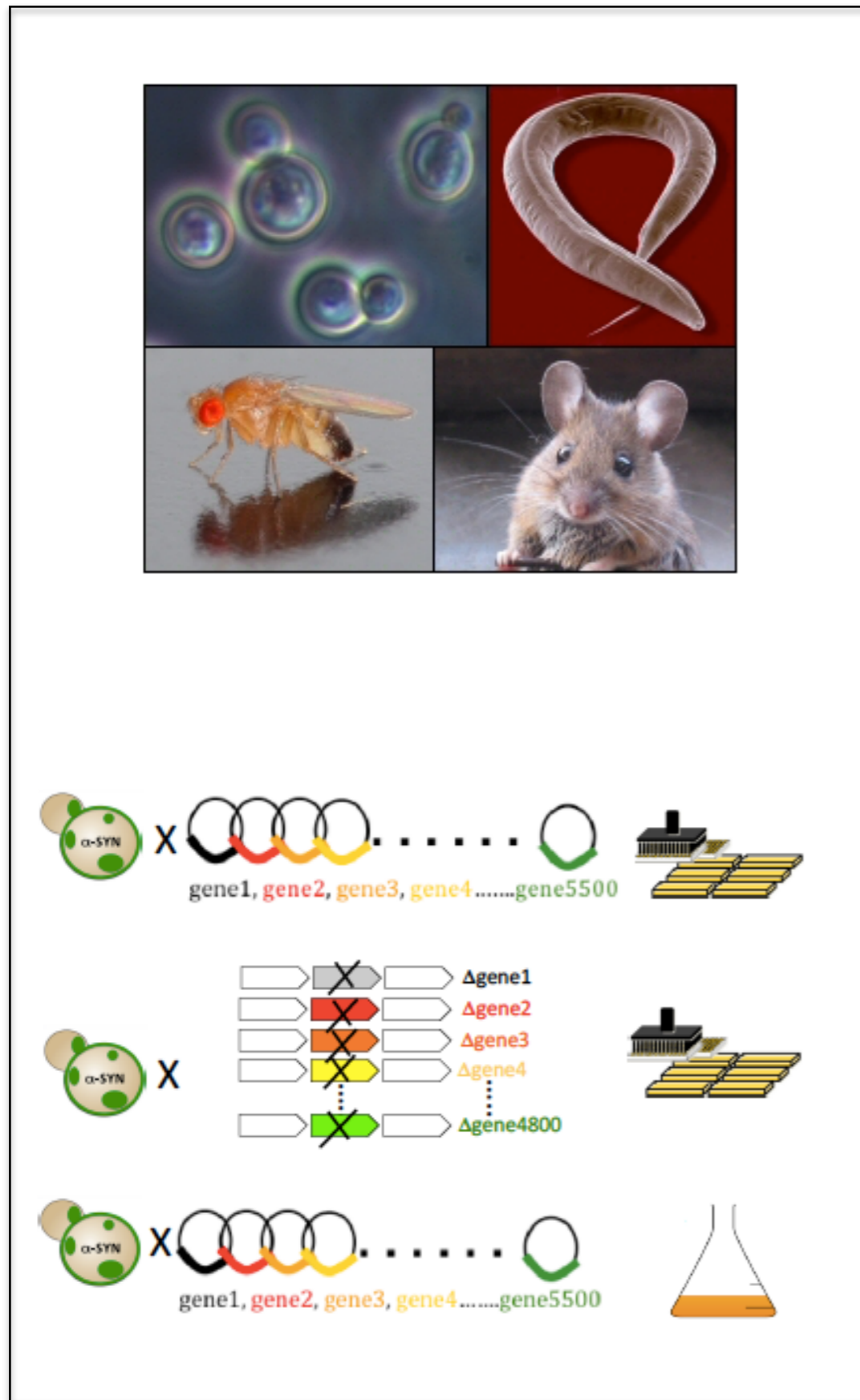
TransposeNet: from yeast to human

Yeast-to-Human Homology (DCA mapping)

- Sequence (BLAST)
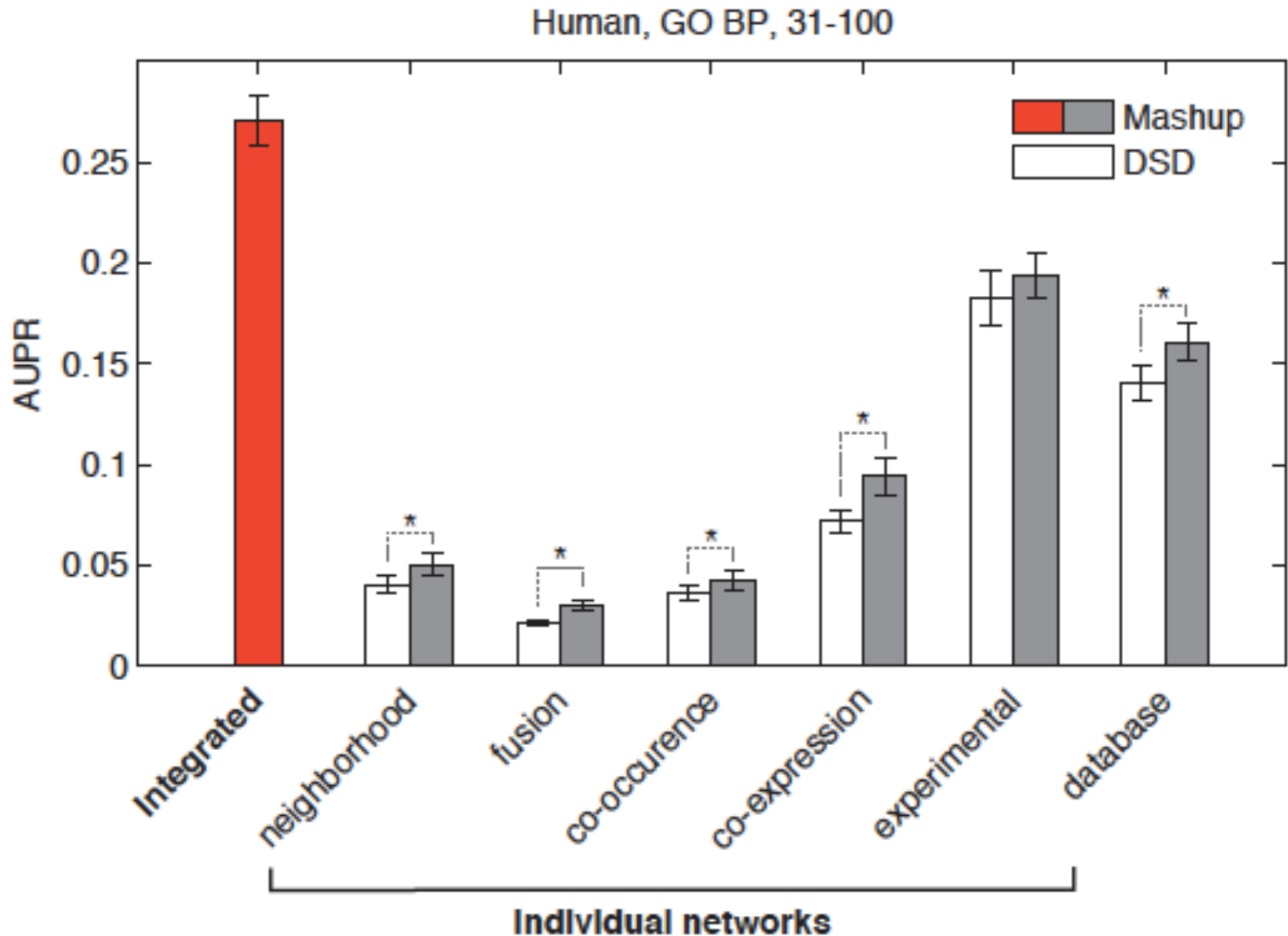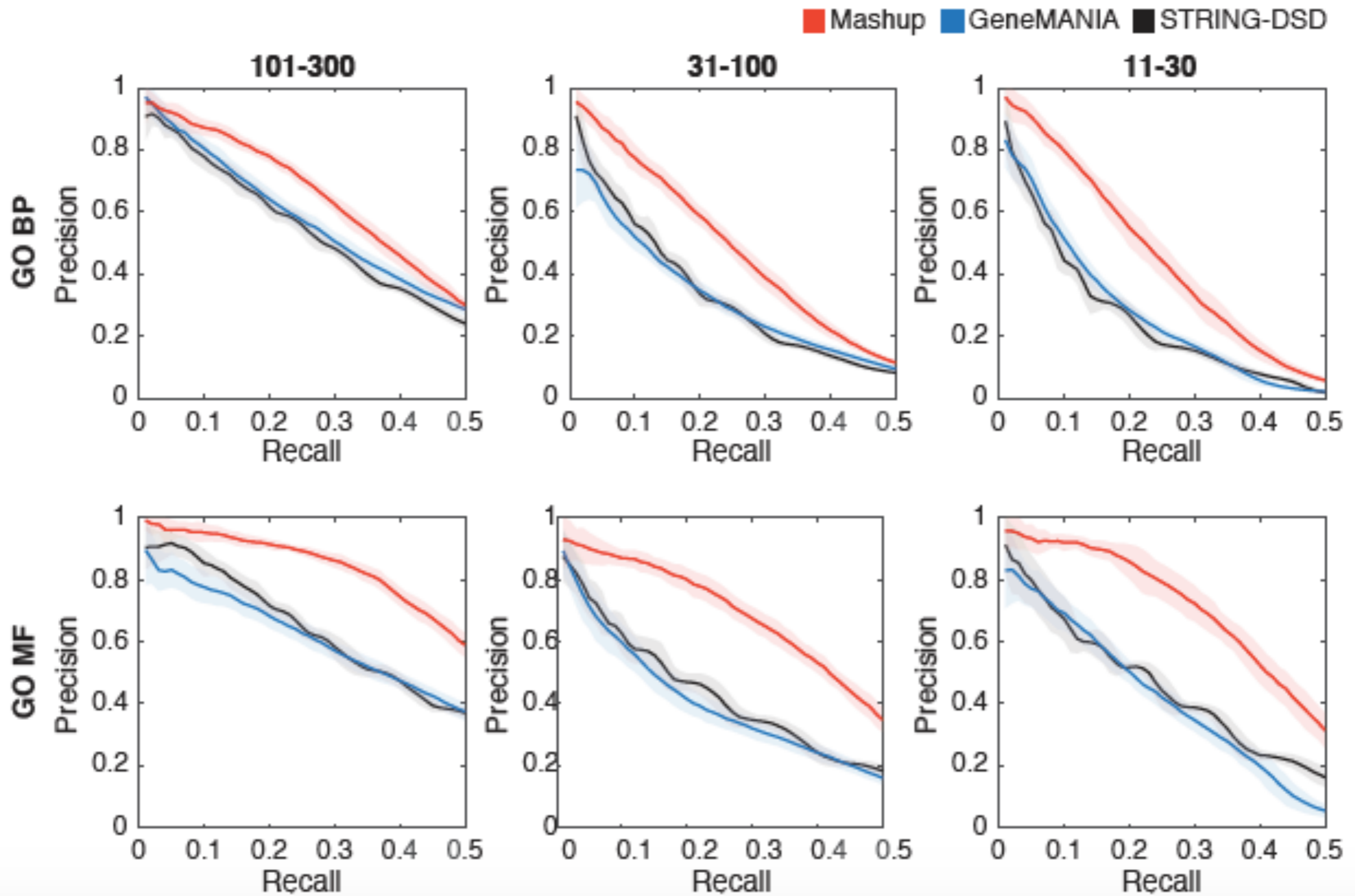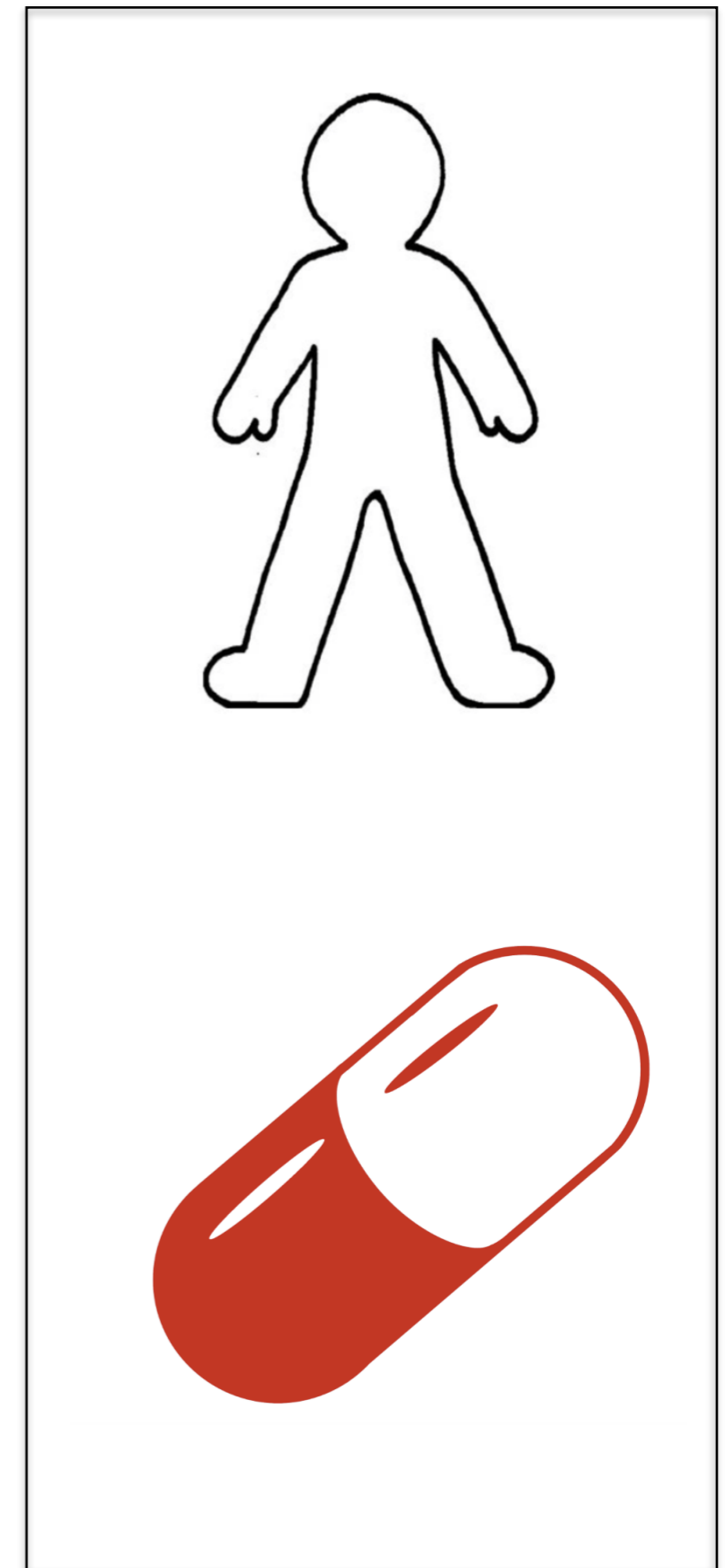- Structural prediction (HHPRED)
- Network topology

77 genetic Modifiers (yeast over-expression screen)

GTEX Filter: → "BRAIN"

Human homologs

Steinerforest Ensemble in which the human interactome is augmented with transposed yeast interactions

"Humanized" α-Syn network (over-expression screen)

AP1B1
PARK1
SNCA
STUB1
NSF
LRRK2
PARK8
ATP13A2
PARK9
SYVN1
RAB1A
PARK18
EIF4G1
PPP2R1A

YEAST HIT — HUMAN HOMOLOG OF YEAST HIT

A PREDICTED HUMAN GENE

"PARK" GENE

GENETIC/PHYSICAL/CURATED DATABASE INTERACTION

Parkinson's disease network
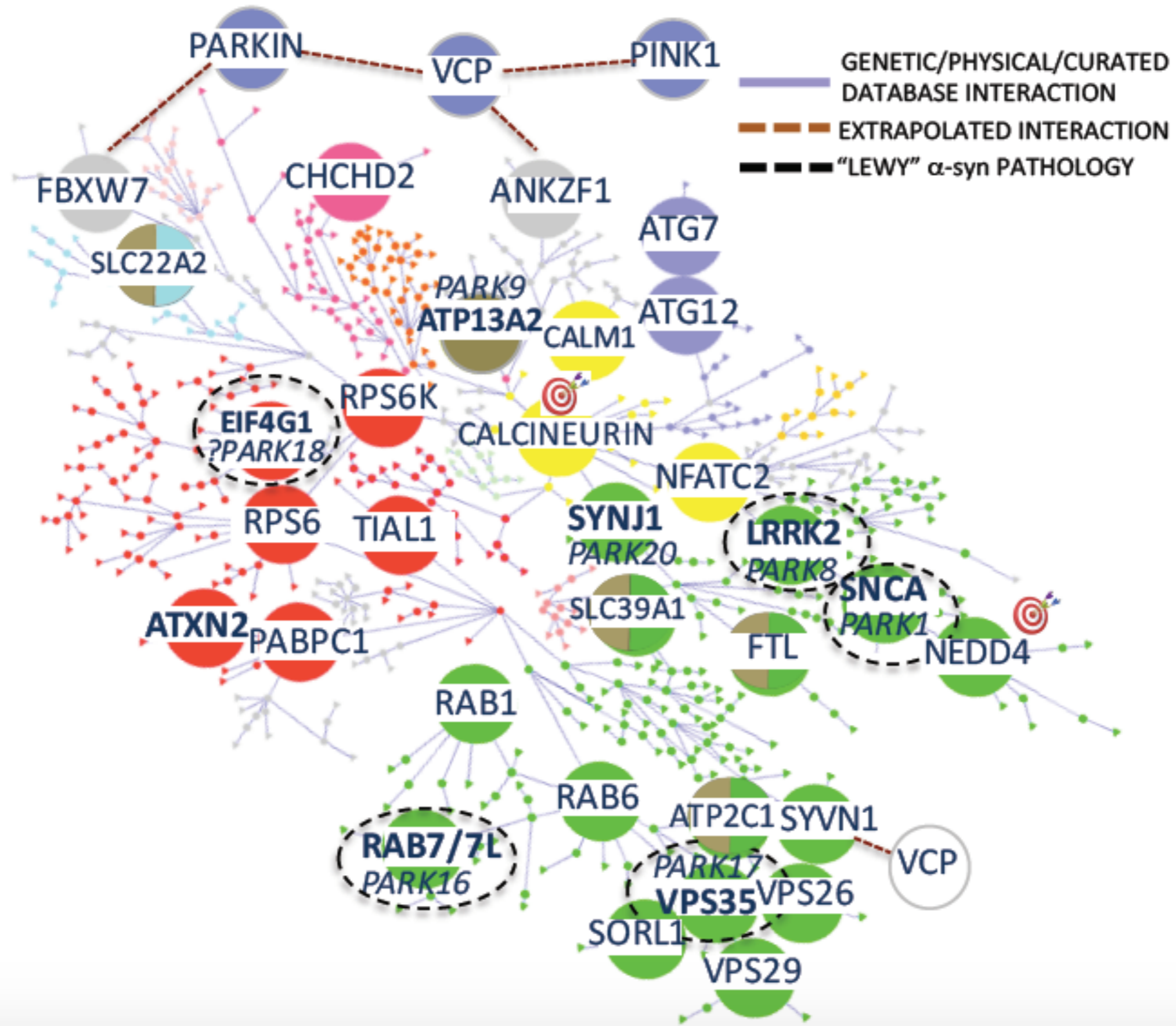
References

Mashup: Compact Integration of Multi-Network Topology for Functional Analysis of Genes
Hyunghoon Cho, Bonnie Berger, Jian Peng
Cell Systems, to appear

Genome-scale molecular networks link diverse neurodegenerative disease genes and processes to alpha-synuclein
Khurana V*, Peng J*, Chung CY*, Auluck PK, Tardiff DF, Fanning S, Bartels T, Koeva M, Benyamini H, Lou Y, Nutter-Upham A, Tuncbag N, Baru V, Freyzon Y, Costanzo M, San-Luis B, Schöndorf DC, Barrasa MI, Caraveo G, Ehsani S, Sanjana N, Zhong Q, Gasser T, Vidal M, Deleidi M, Boone C, Fraenkel E, Berger B, Lindquist S
Cell Systems, to appear

In situ proteome approaches connect alpha-synuclein directly to endocytic trafficking and mRNA metabolism.
Chung CY*, Khurana V*, Loh K, Yi S, Sahni N, Hill D, Peng J, Vidal M, Ting A, Lindquist S
Cell Systems, to appear

ProSNet: integrating homology with molecular networks for protein function prediction
Sheng Wang, Meng Qu, Jian Peng
PSB, 2017, to appear

Exploiting Ontology Graph for Predicting Sparsely Annotated Gene Function
Sheng Wang*, Hyunghoon Cho*, ChengXiang Zhai, Bonnie Berger and Jian Peng
ISMB/ECCB, 2015
Bioinformatics, 2015

Diffusion Component Analysis: Unraveling Functional Topology in Biological Networks
Hyunghoon Cho, Bonnie Berger and Jian Peng
RECOMB, 2015

Acknowledgements

knoweng BIG DATA TO KNOWLEDGE CENTER OF EXCELLENCE

ALFRED P. SLOAN FOUNDATION