Biological network analysis

Rui Benfeitas
NBIS - National Bioinformatics Infrastructure Sweden
Science for Life Laboratory, Stockholm
Stockholm University

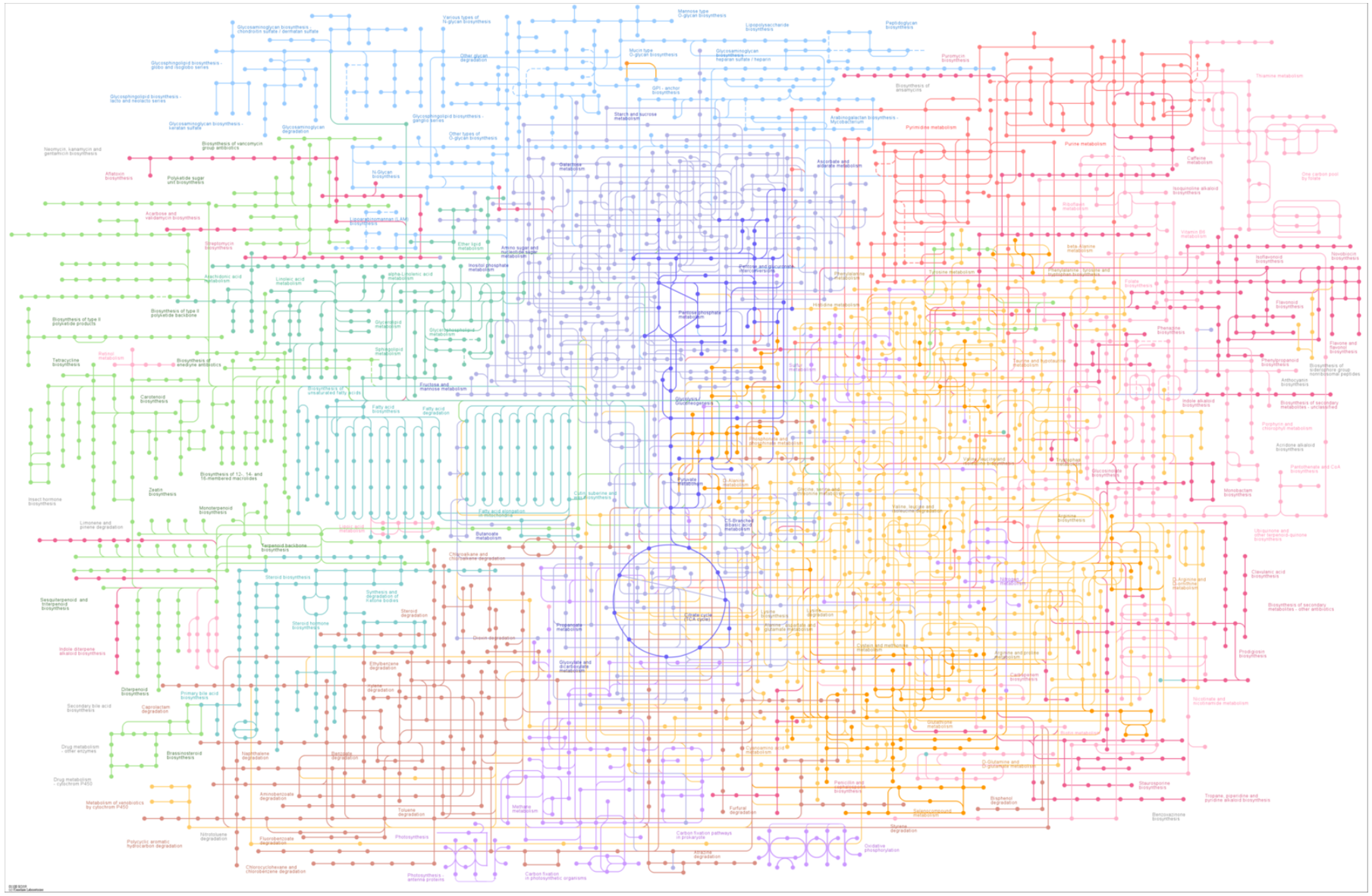
rui.benfeitas@scilifelab.se

Overview

1. Introduction to network analysis
2. Terminology
3. Network inference
4. Key network properties
5. Community analysis

Introduction

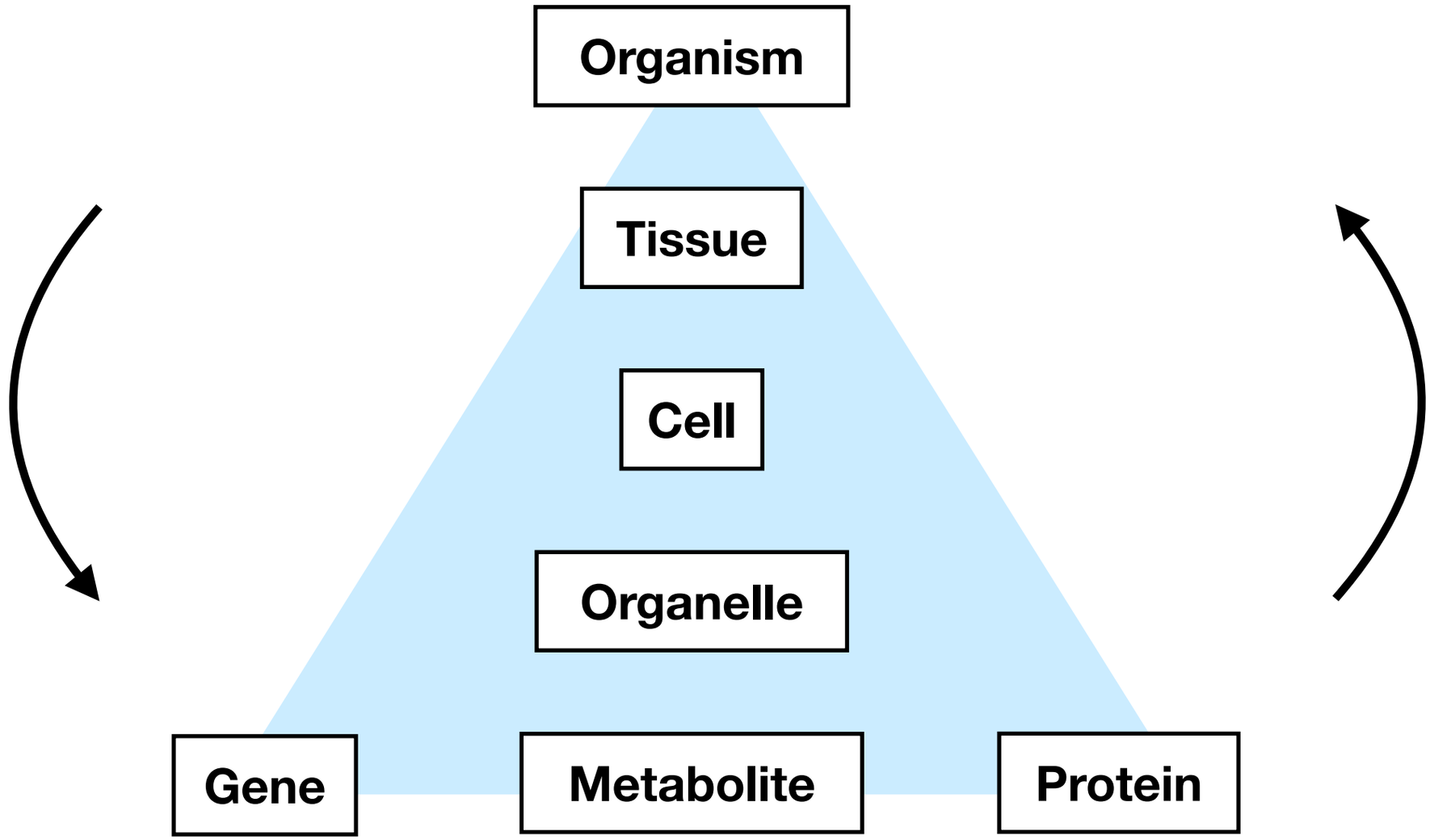
How to tackle biological complexity?



Lobo 2008
genome.jp/kegg/

How to tackle biological complexity?

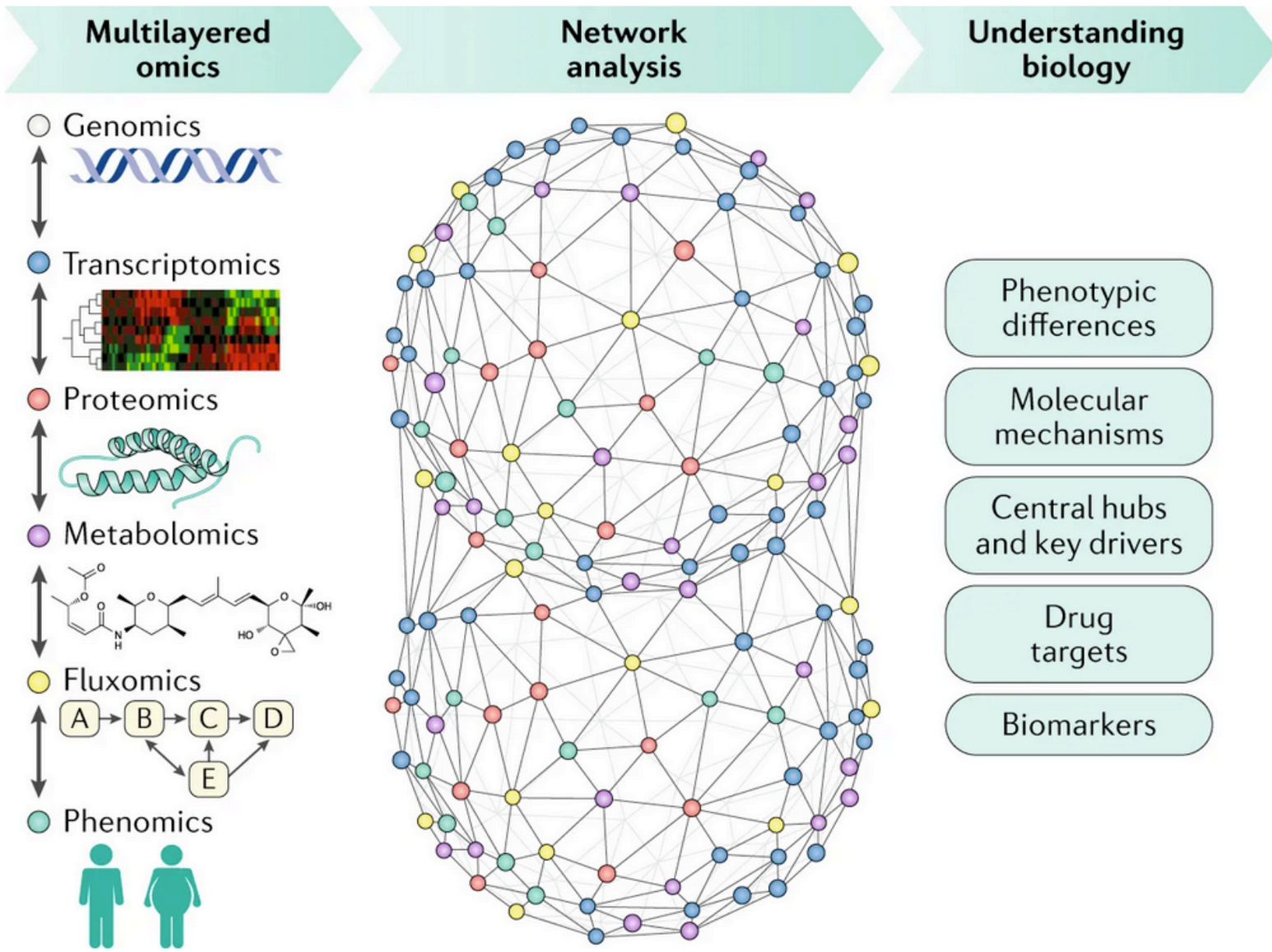
Moving from reductionist approaches towards global characterisations



How to tackle biological complexity?

Integrative approaches, and global patterns

- Feature association
- Modeling
- Network analysis



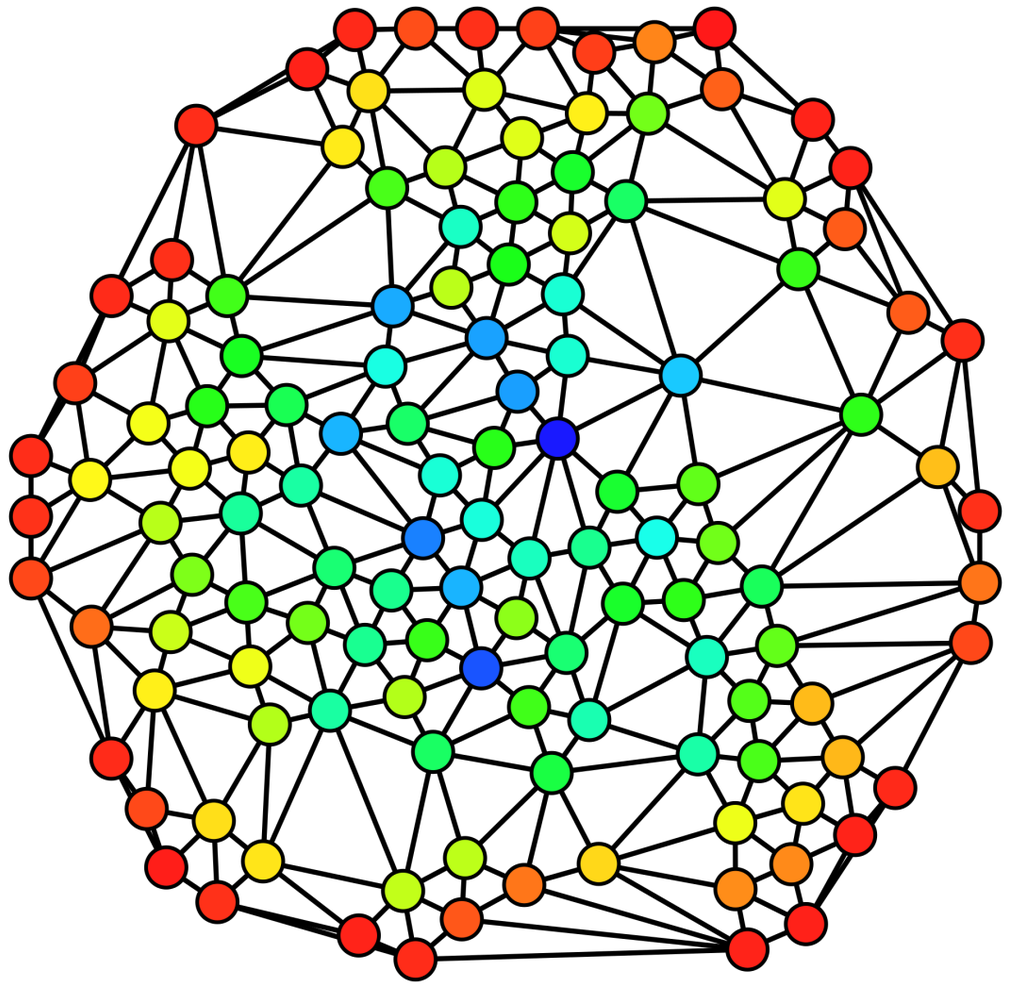
What are networks?

Networks are representations of complex systems

Permit defining and studying global properties of interacting components

Give us insight not easily achieved by other approaches:

- Comprehensive
- Coordinated



What are biological networks?

Protein - Protein interaction (PPI) networks

Transcription-factor regulatory networks

Gene - gene co-expression networks

Signal transduction networks

Drug-disease association networks

Aim
Functional characterisations

What are biological networks?

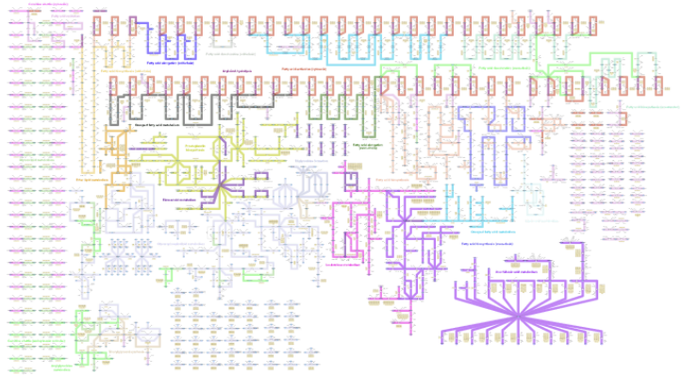
Metabolite - Enzyme - Signal - Genes (GEMs)

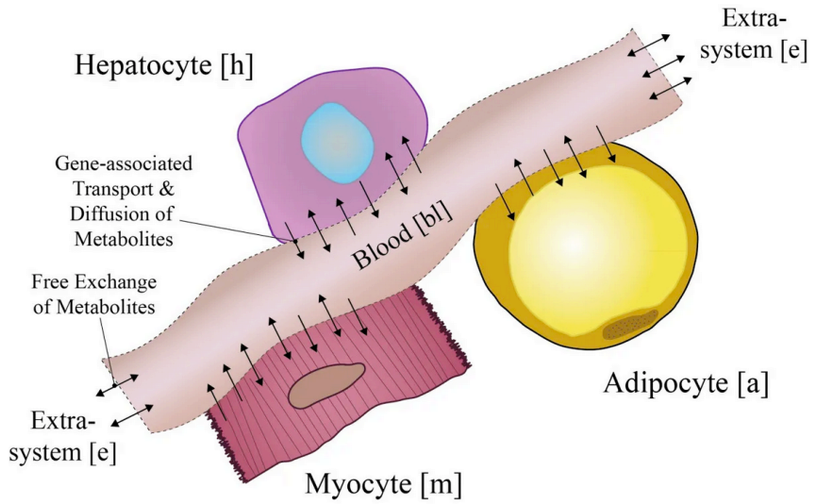
Multi-tissue networks

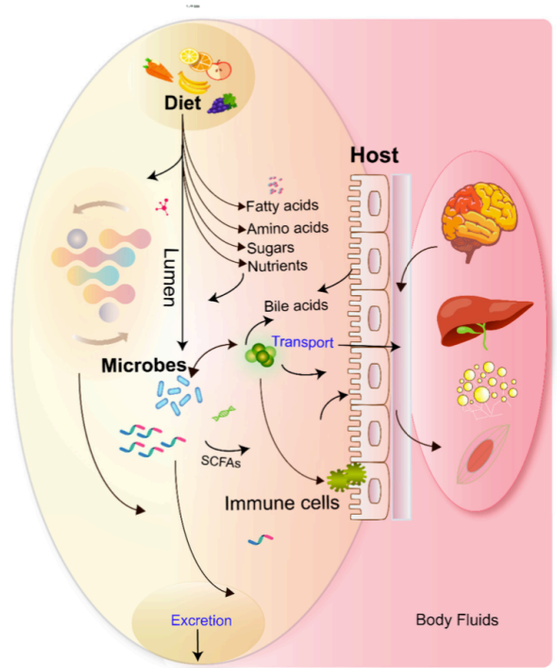
Multi-species networks

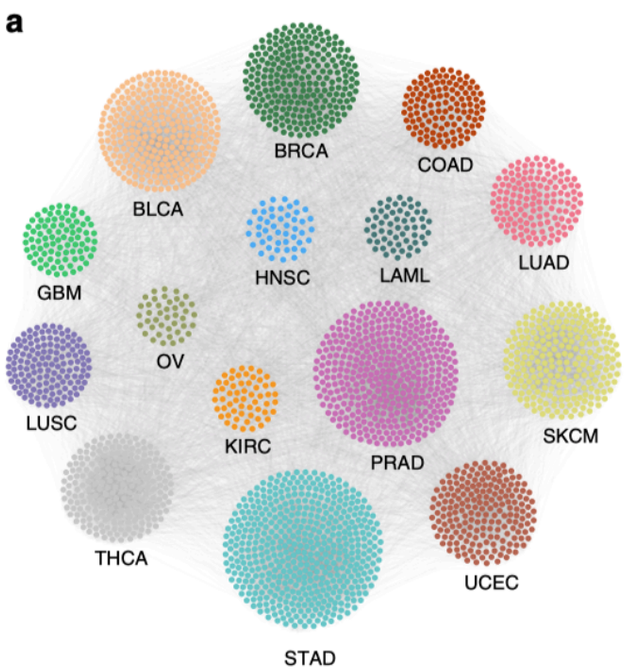
Disease networks

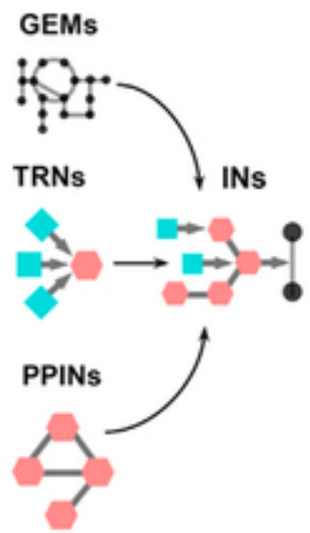
Integrated networks





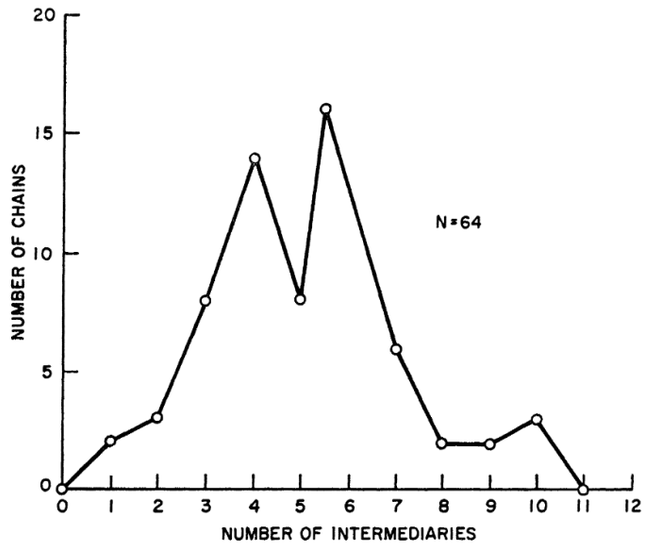




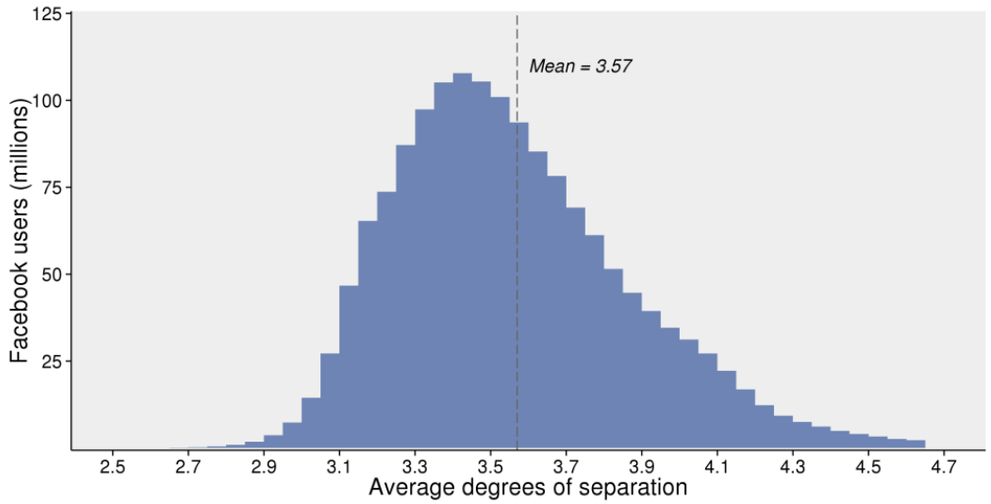


Small world

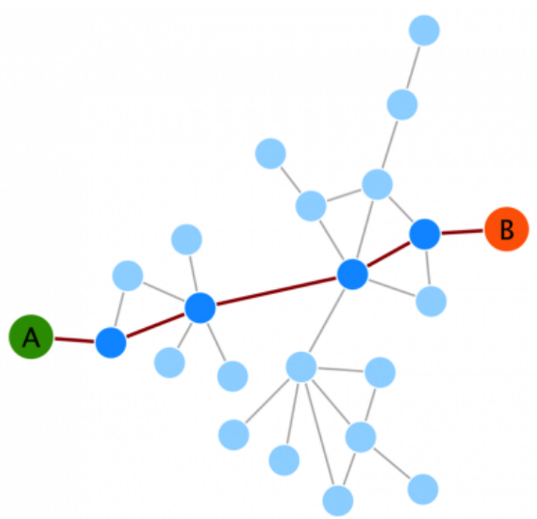
Stanley Milgram (1967) - 6 degrees



Backstrom et al. (2016) - 3.6 degrees



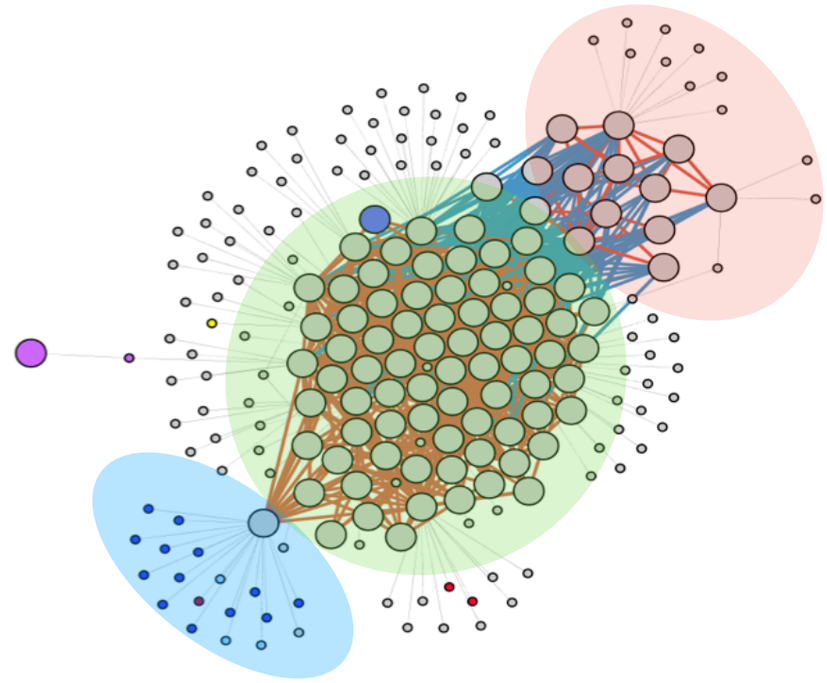
Biological Networks



Why look at network topology?

Use networked systems to:

- Identify global / local patterns
- Identify functional properties
- Make predictions



Examples:

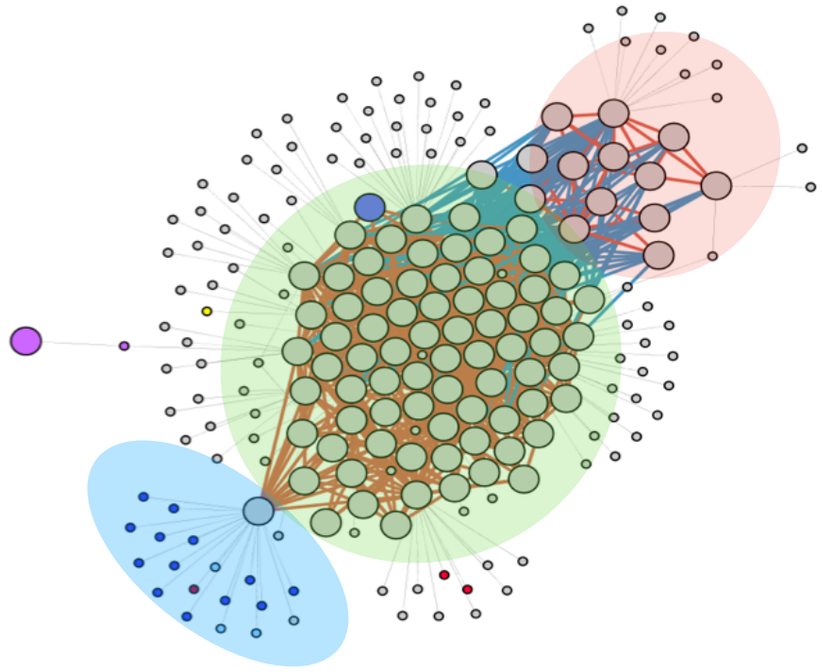
- How associated are the elements of my network?
- What are its first-hand associated elements?
- What are the “neighbourhoods”?
- What are their common functional relationships?
- What are the “key” elements in my network?
- What are the "weakest" links in the network?

What is my biological network?

Any adjacency matrix may be translated to a network format

Many standard analyses may be employed regardless of data type

…but care must be taken in generating the network

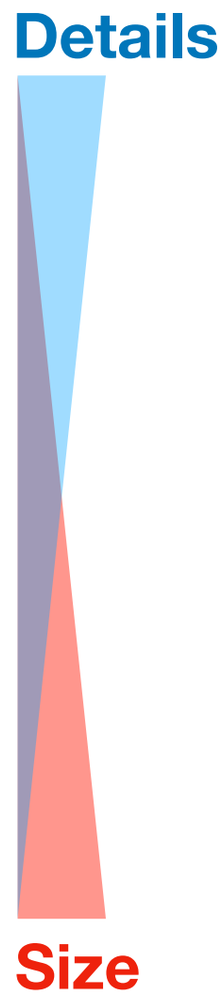


Some of the functional analyses depend on annotation

Motivation

What modeling formalism suits your data and biological question?

	Pros	Cons
Kinetic	Detailed Quantitative Dynamic / Steady state	Small Requires detailed parameterization
Stoichiometric	Comprehensive Semi-quantitative Steady state	Static
Topological	Comprehensive Only topological information	Hard to examine dynamically



Adapted from Hartmann et al.

Terminology and initial properties in graph analysis

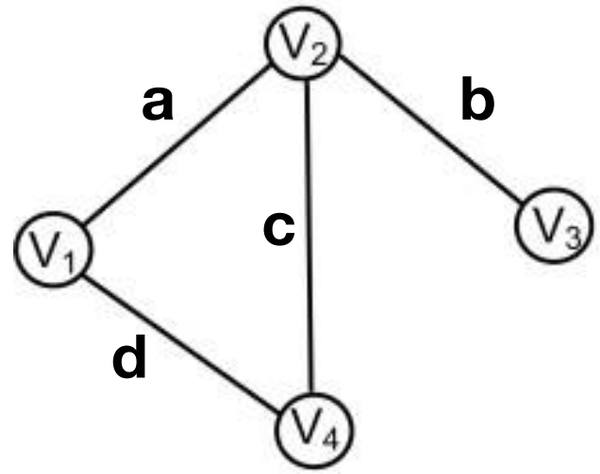
Graphs, nodes, edges

Graph *G* consists of a set of **nodes** (V) interconnected by **edges** (E)

$$G = (V, E)$$

$$V=\{v1,v2,v3,v4\}$$

$$E=\{a,b,c,d\}$$

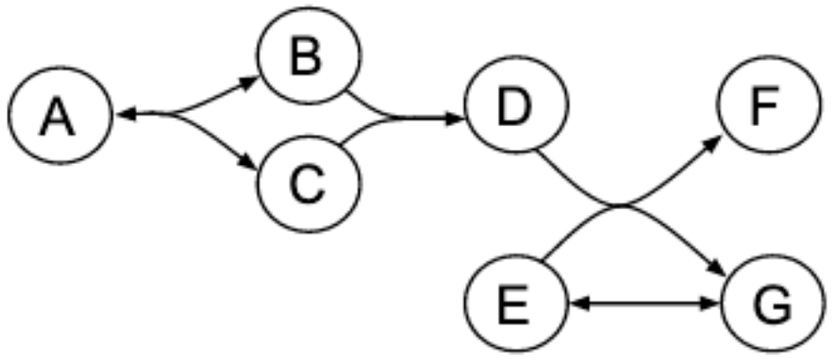


Nodes sometimes called **vertices**

Two connected nodes are called **neighbours, adjacent,** or **end-nodes**

Hypergraphs

Hypergraphs contain edges that connect any number of nodes



Reaction 1: $A \rightarrow B + C$
Reaction 2: $B + C \rightarrow D$
Reaction 3: $D + E \rightarrow F + G$
Reaction 4: $E \rightarrow G$
Reaction 5: $B + C \rightarrow A$
Reaction 6: $G \rightarrow E$

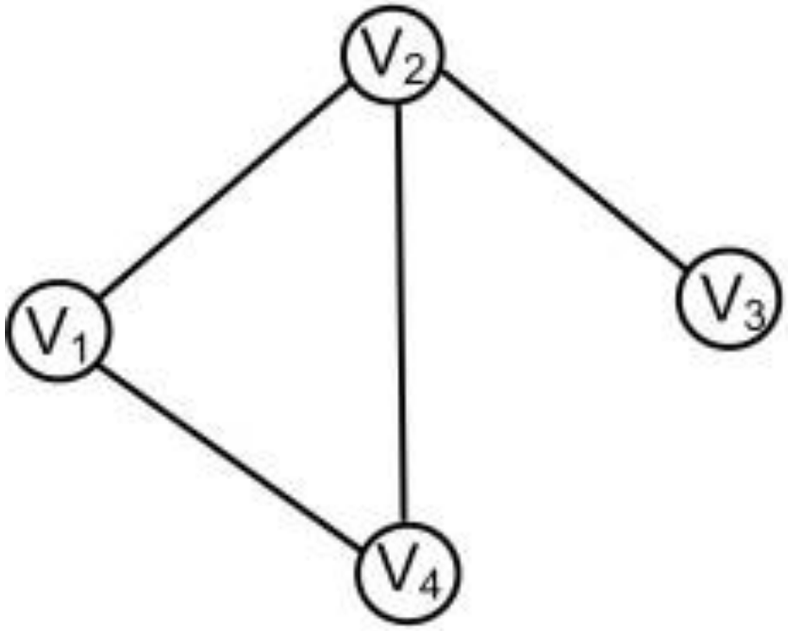
(a) Reaction network

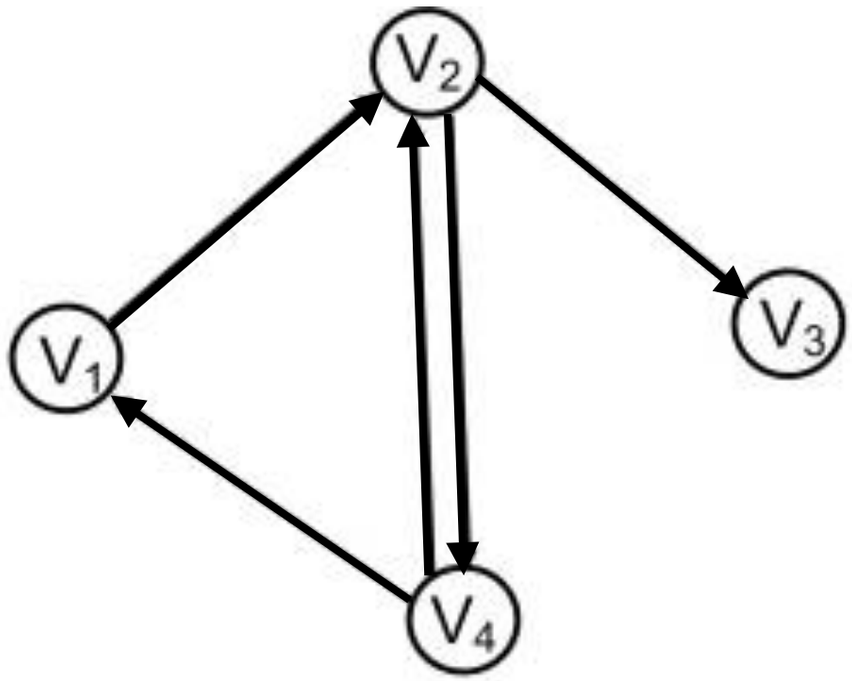
Directed vs undirected graphs

Examples:

- **Undirected graphs:** co-expression networks

- **Directed graphs:** metabolic networks



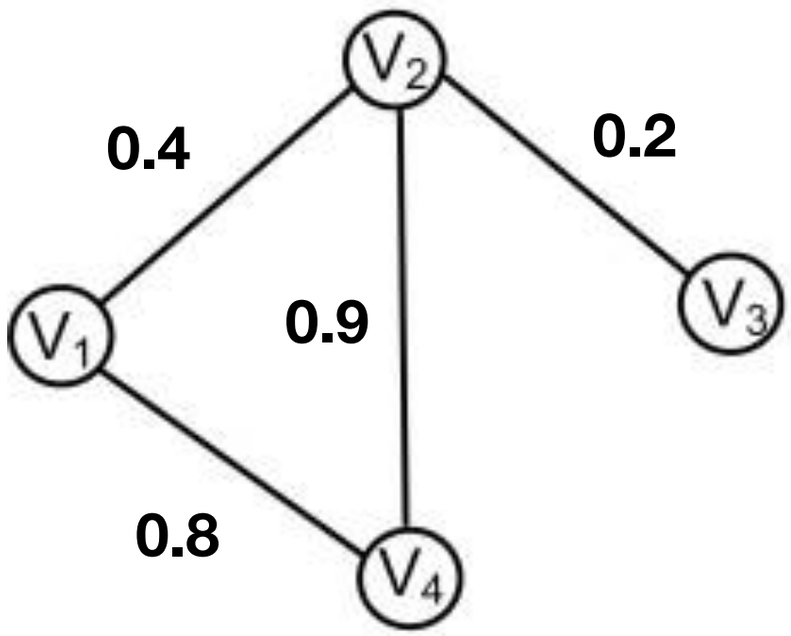


Weighted vs unweighted graphs

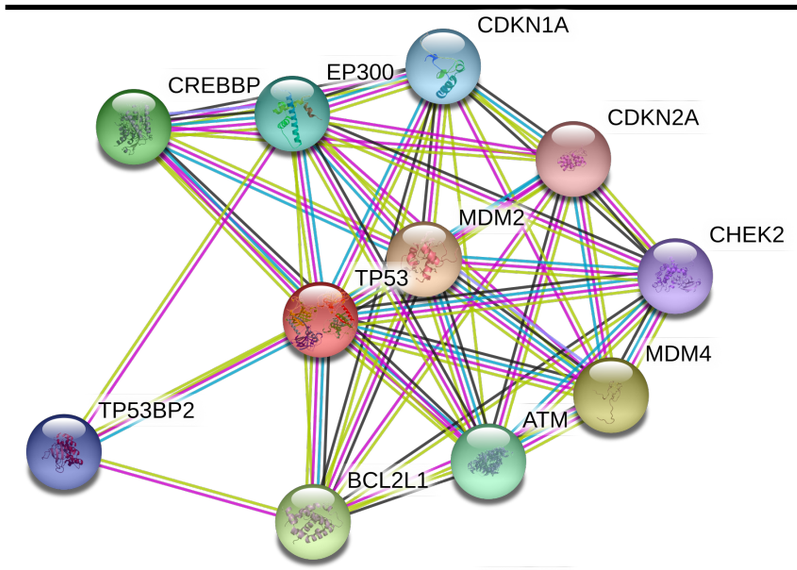
Weighted edges associate a value to an interaction between two nodes. Usually give the confidence in the interaction.

Negative weights?

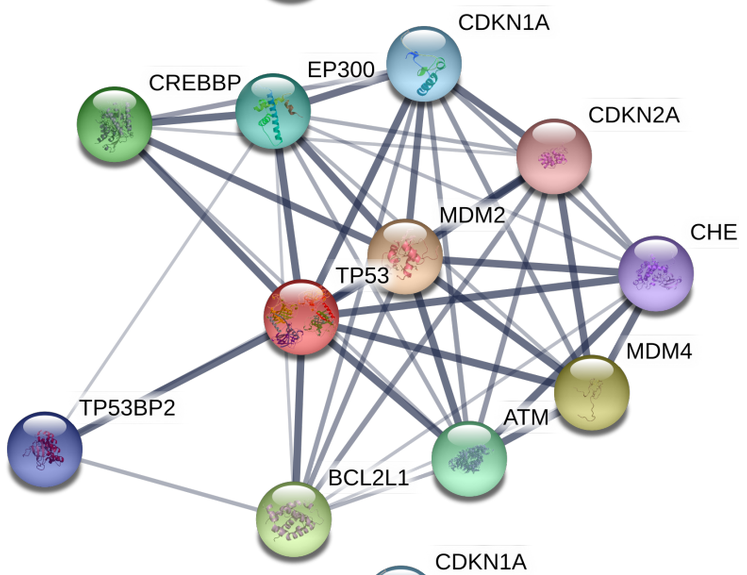
E.g. weighted co-expression networks



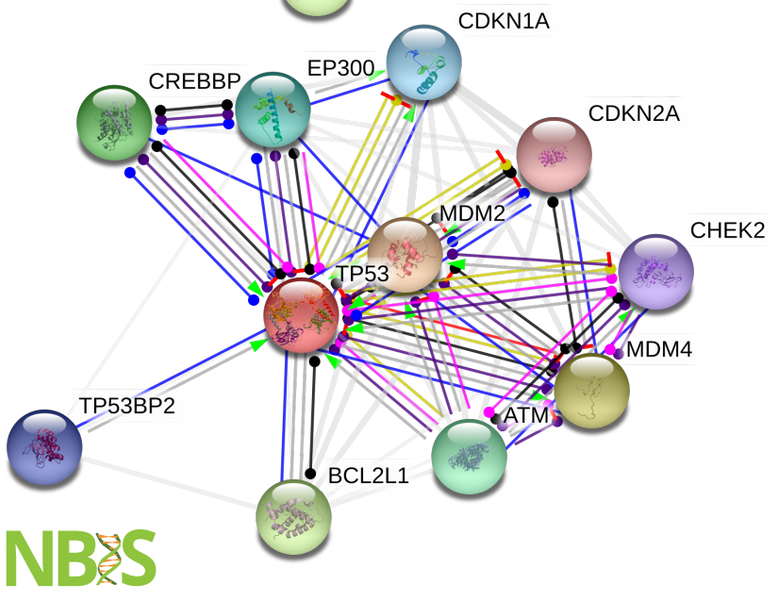
STRING-db.org: TP53



Multi-edged



Weighted multi-edged



Multi-edged directed

Bipartite graphs

A graph

$$G=(V,E)$$

may be partitioned into two sets of nodes (V_1, V_2) such that

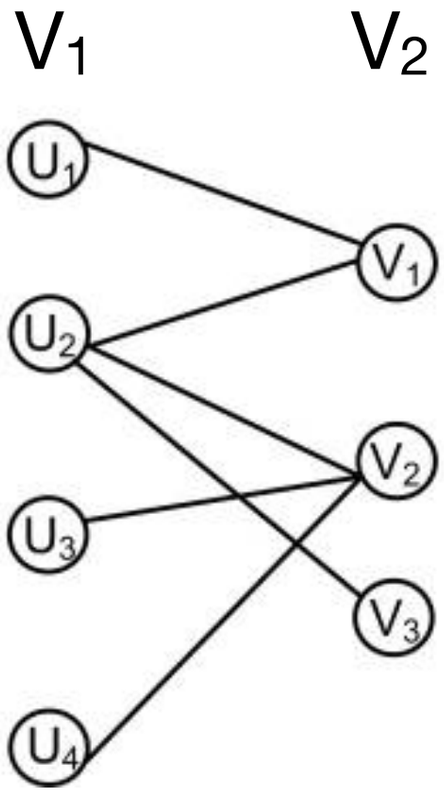
$$u \in V_1 \text{ and } v \in V_2$$

All e_i has end-nodes in V_1, V_2

A **subgraph** of G will thus be given by

$$G_1 = (V_1, E_1)$$

Examples?



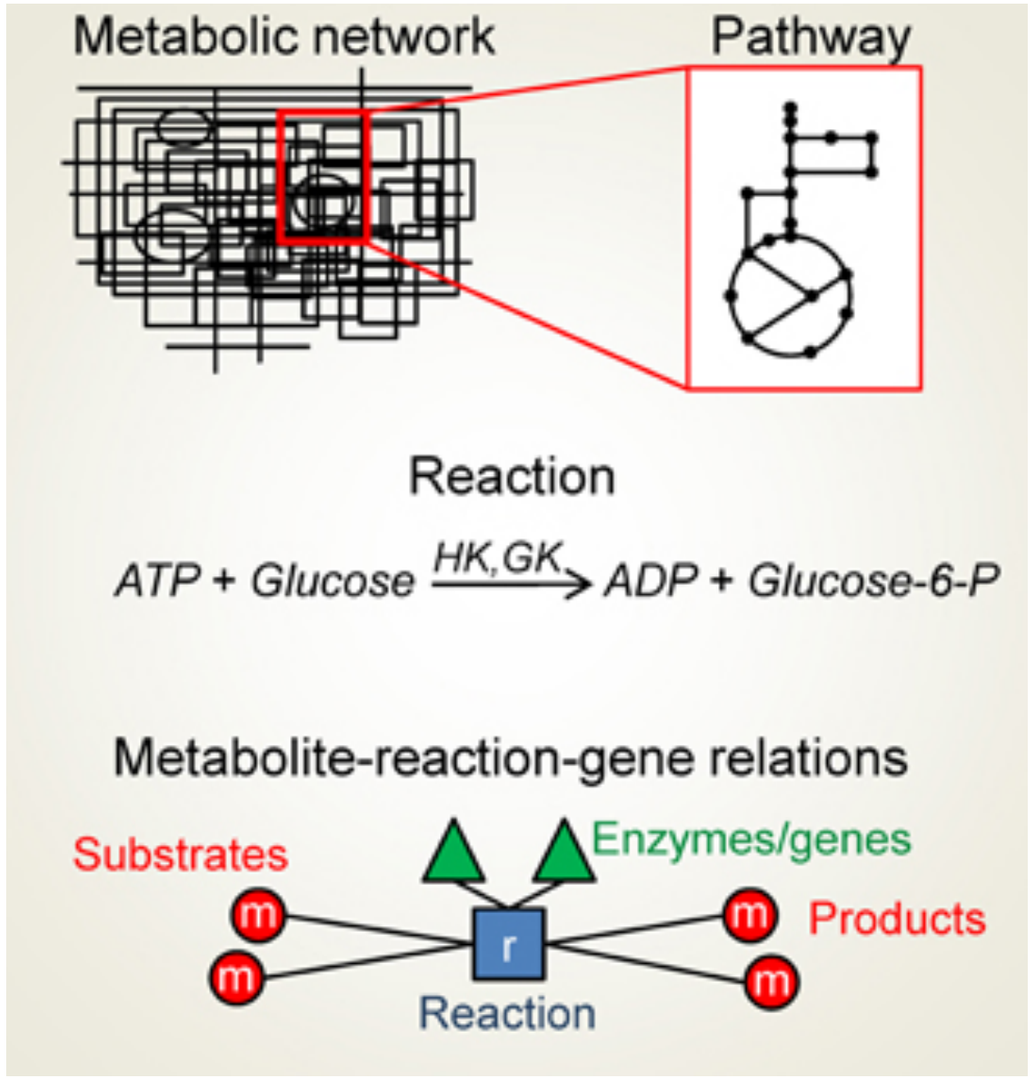
Bipartite and *k*-partite graphs

Example of bipartite graph:

Enzyme - Reaction

Metabolite - reaction - enzyme

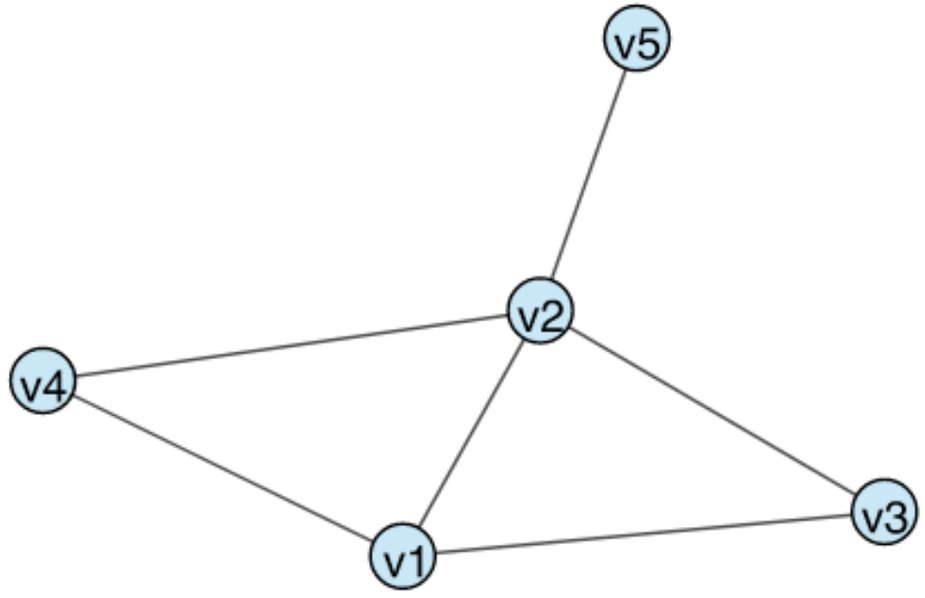
k-partite graphs display *k*-types of nodes



Adjacency matrix (undirected graphs)

Vertex association (undirected network)

n1	n2
v1	v2
v1	v4
v2	v4
v2	v3
v2	v5
v1	v3



Adjacency matrix is symmetric (undirected graphs)

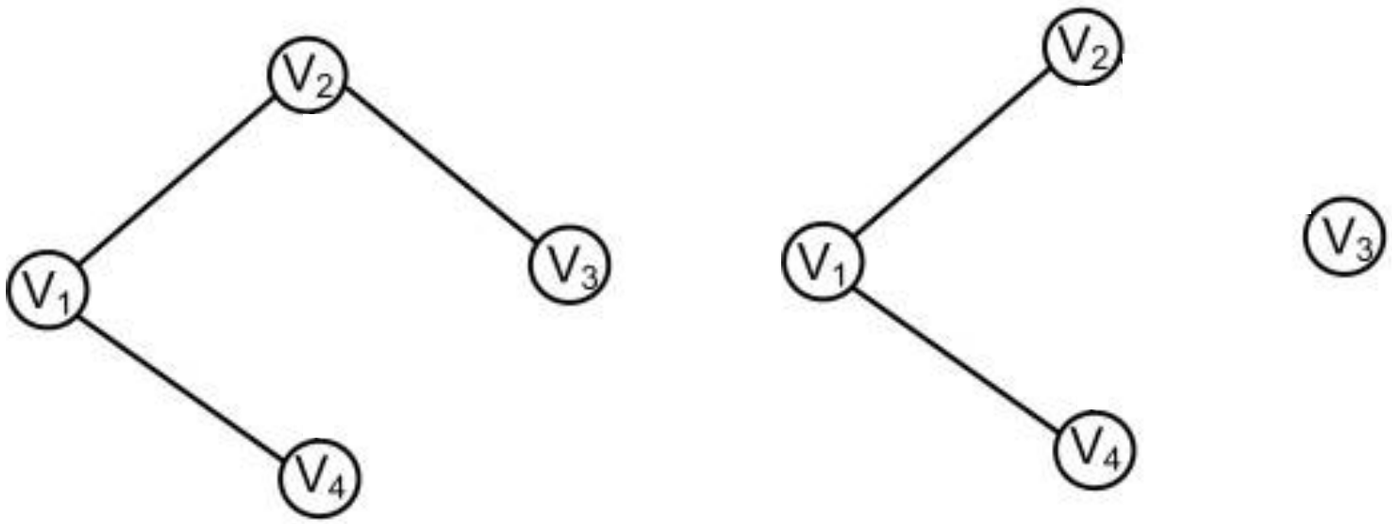
	v1	v2	v3	v4	v5
v1	0	1	1	1	0
v2	1	0	1	1	1
v3	1	1	0	0	0
v4	1	1	0	0	0
v5	0	1	0	0	0

Connected vs disconnected networks

Connected network: there is at least 1 path connecting all nodes in a network

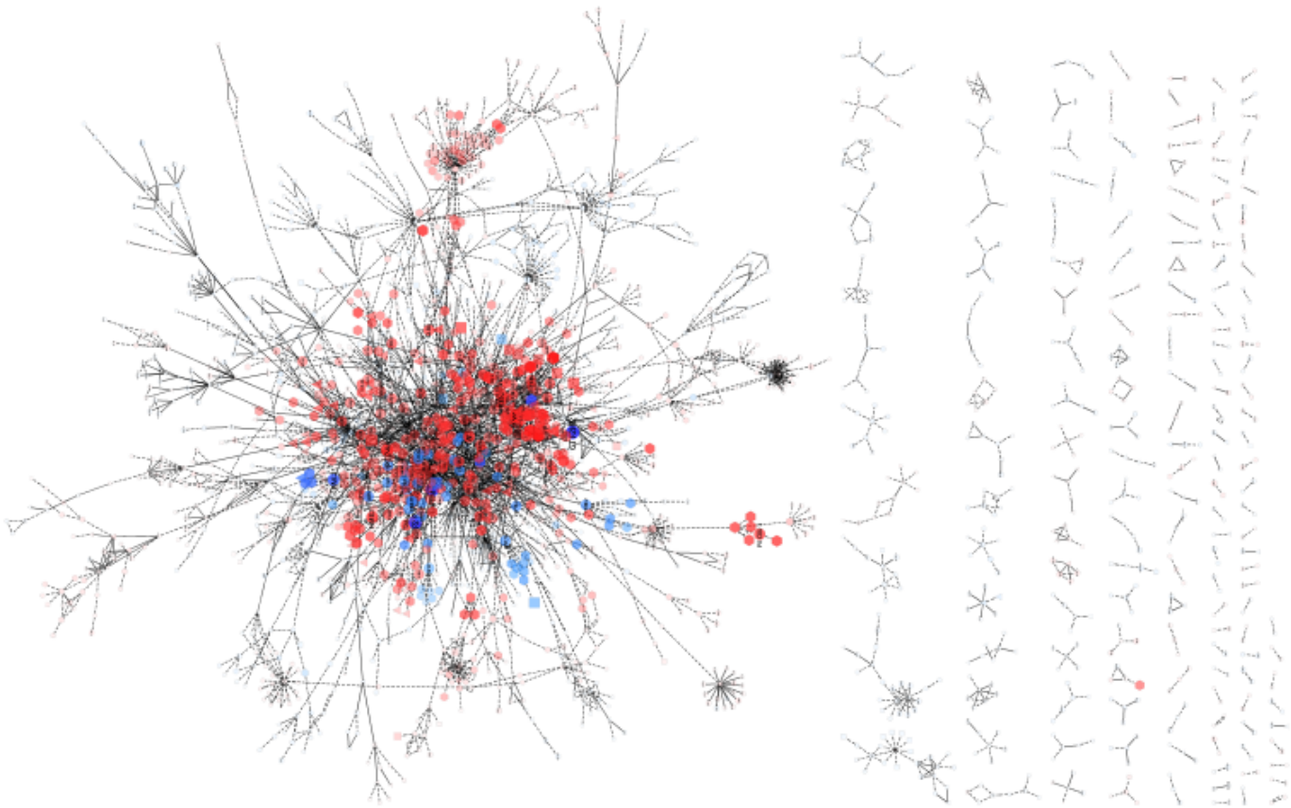
Disconnected network: some of the nodes are unreachable

Connected components



Connected components

In biological networks, often the most insightful properties come from the **largest connected component(s)**



CC>1? CC>5? CC>30?

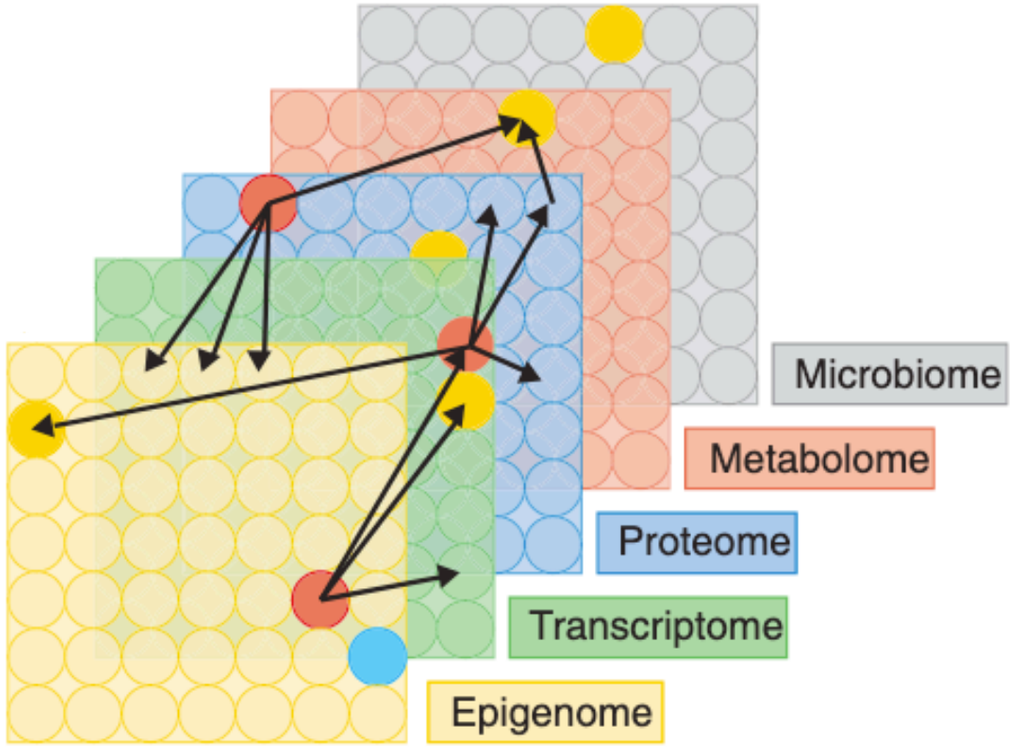
Biological network inference

Interomic vs Intraomic networks

Networks may be build for individual omics or for their integration

Should I even integrate different omics? What is my biological question?

- Do I want to analyse vertical relationships between features?
- Why integrate omics with different coverage such as transcriptomic and proteomic data
- Do I want to extract functional properties?
- Am I predicting biomarkers?

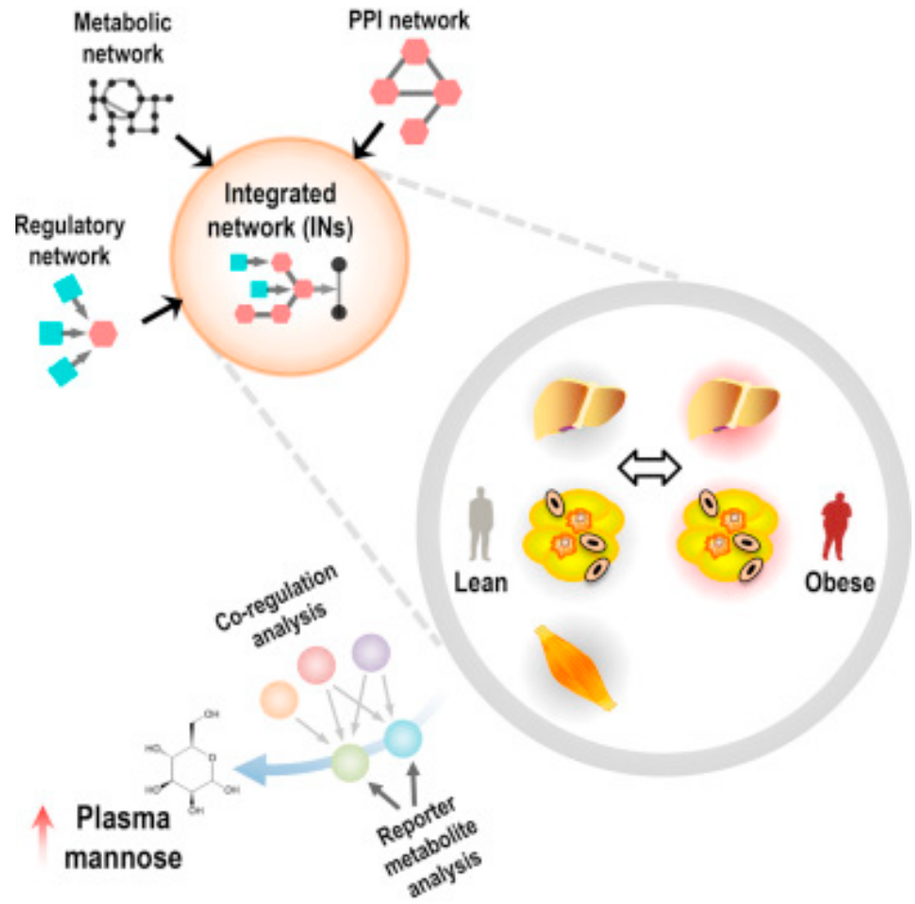


Interomic vs Intraomic networks

Multi-modal (*k*-partite) networks may be generated from different sources

- Transcription-factor - Gene (DNAseq)
- Gene-gene (Co-expression, PPI, GEMs)
- Gene-metabolite (GEM)
- Metabolite-metabolite (GEM)

Integrated Networks



Different approaches for network inference

1. Feature association

2. Genome-scale metabolic models

   Pathway-based

3. Network deconvolution

No prior graph structure

Based on available information

Filter indirect effects

Different approaches for network inference

1. Feature association
2. Genome-scale metabolic models
3. Network deconvolution

Pathway-based

- KEGG
- Interactome
- Reactome
- ...

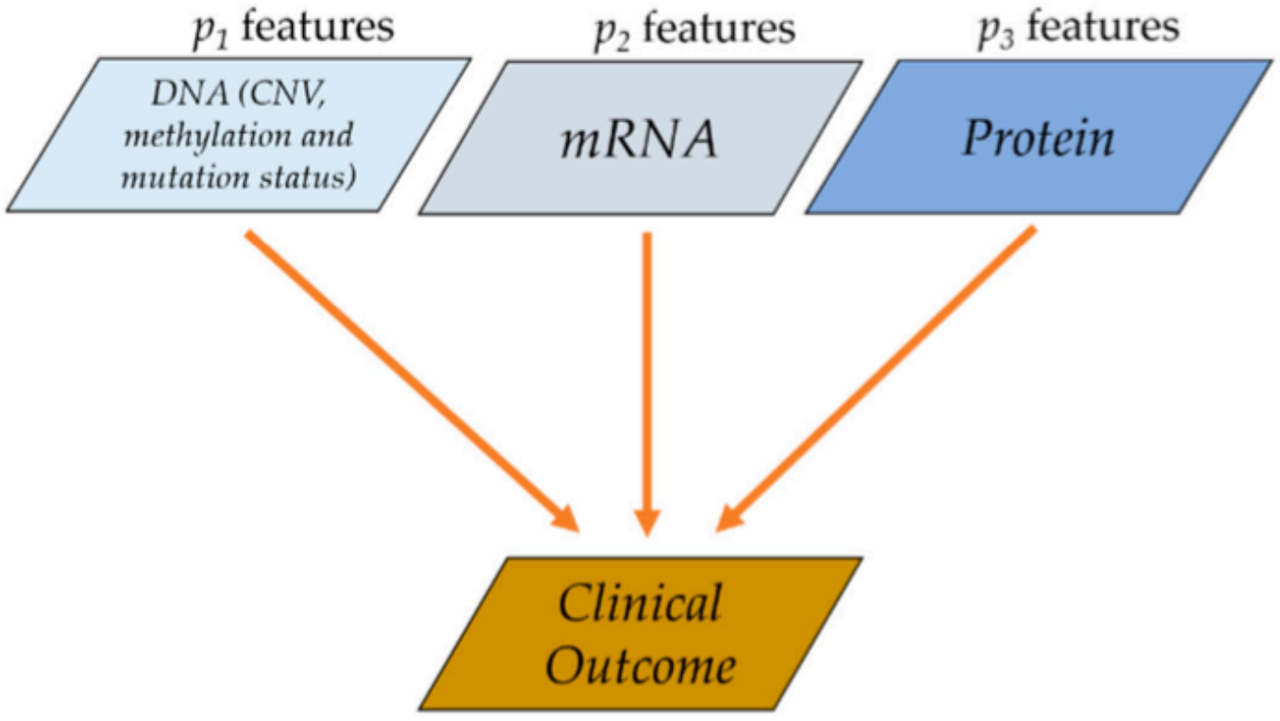
1. Feature association

Balanced dataset, standard pre-processing of each omics needed

Normalization may be needed to make omic datasets comparable (e.g. standardization)

Common approach: compute correlations between different features

Extend known associations



Wu 2019

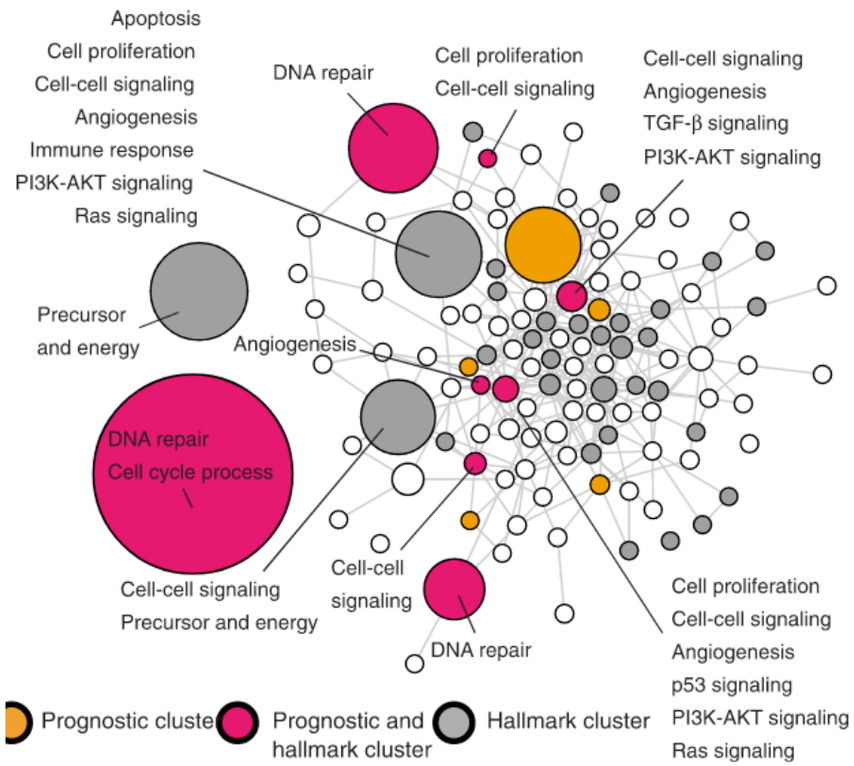
1. Association analysis

Easy to interpret

Unweighted vs weighted ($-1 \leq \rho \leq 1$)

Unbalanced and dense networks

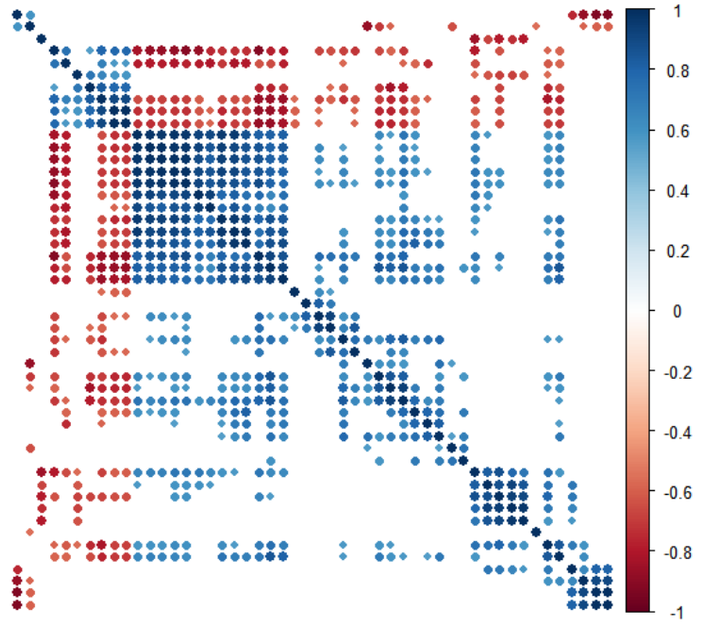
Prone to type I errors



Filtering

- FDR vs Bonferroni
- Correlation coefficient cutoff
- Number of edges (KNN)

Need adjustment to possible confounding factors



Wu 2019
Uhlen 2017

1. Association analysis

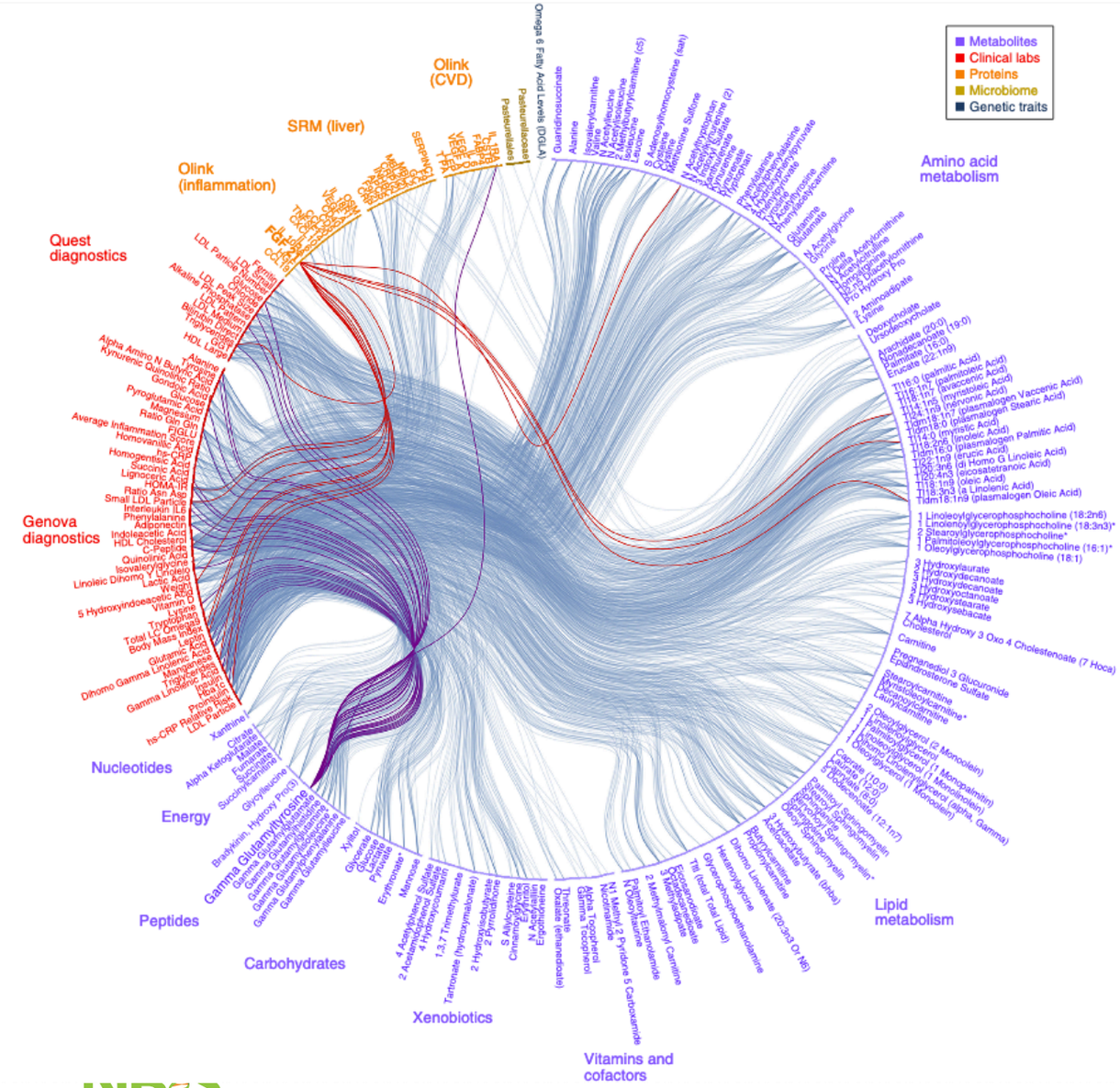
Adjusting for confounding factors: partial correlation analysis

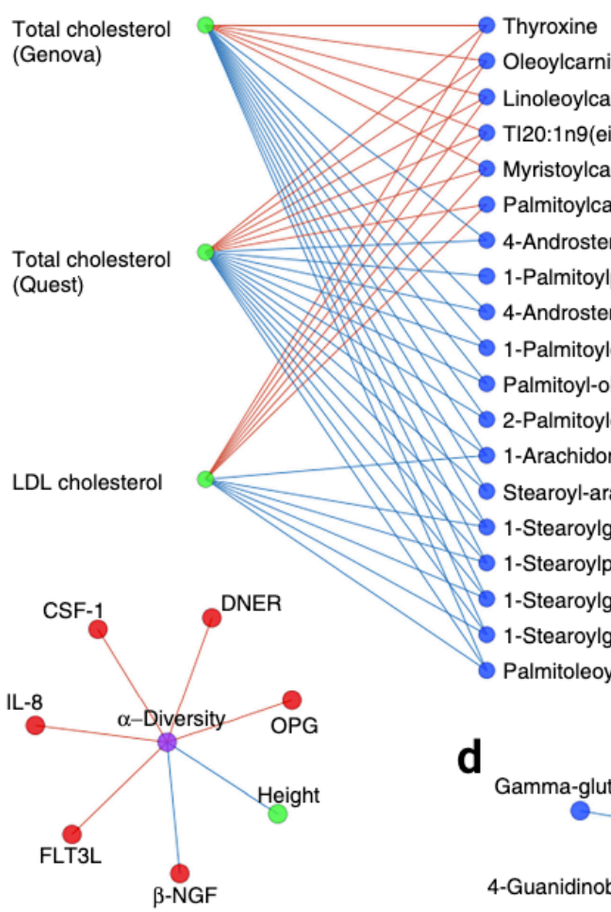
Still considers linear relationship between variables

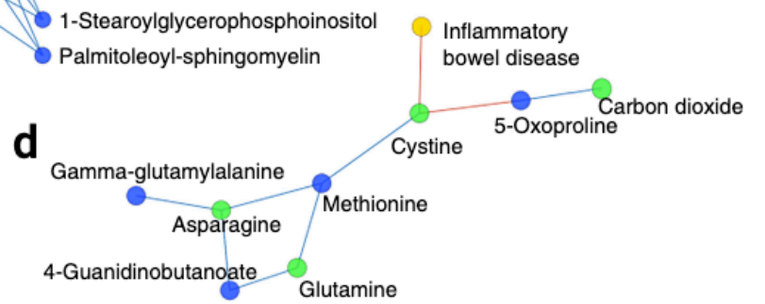
Below:

- gender and age are known confounding factors

- feature regression on confounding factors, followed by correlation on the residuals of each model







2. Using genome-scale metabolic models for graph creation

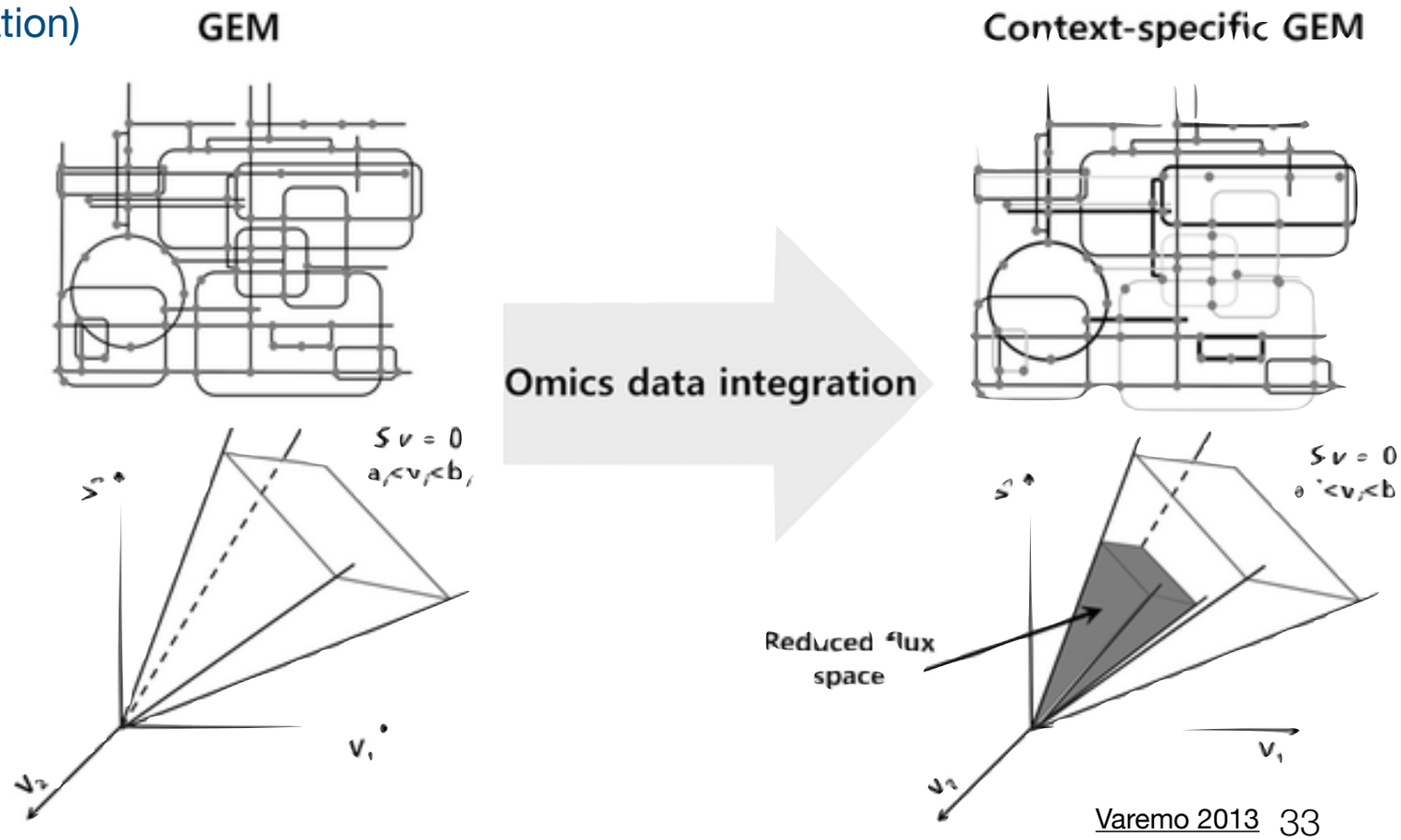
GEMs may be used to find such missing relationships, but there is a coverage issue

The overall strategy follows

1. Integrate proteomic, transcriptomic, metabolomic, fluxomic
2. Flux distribution
3. Compute metabolite-reaction-gene relationships
4. Extract relevant relationships (met-met, gene-gene, met-gene)

4b. Exclude unnecessary interactions (e.g. cofactors)

5. Downstream analysis (e.g. topology, stratification)

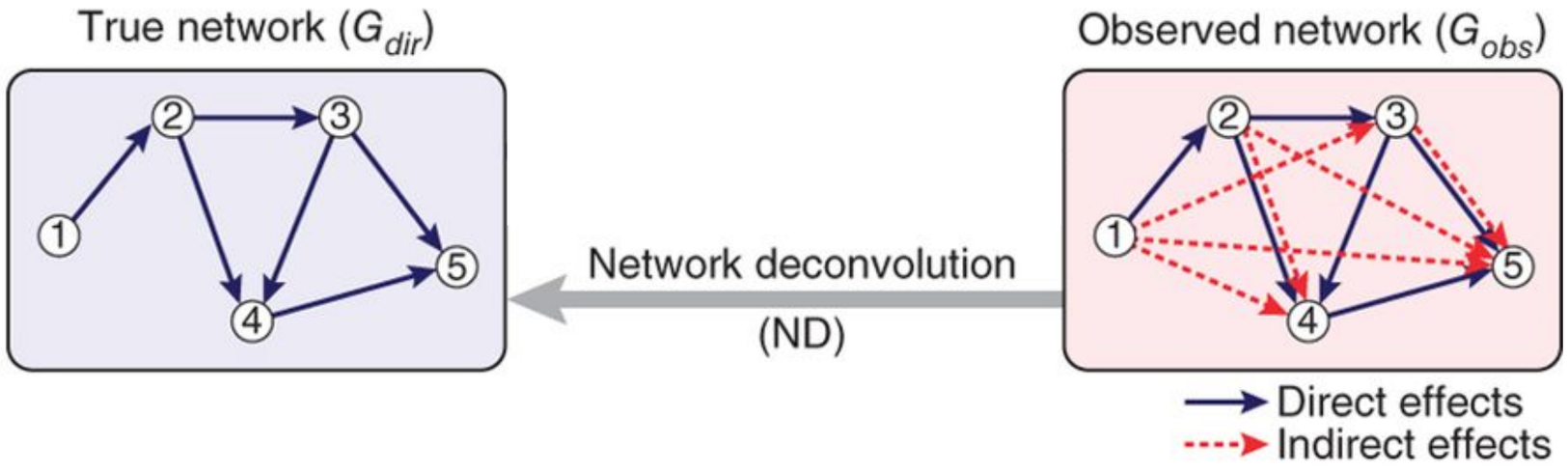


3. Network deconvolution

Biology is ***noisy***, which may result in edges that are not true

- 1 → 2
- 2 → 3
- 1 ⇢ 3

Direct and indirect effects

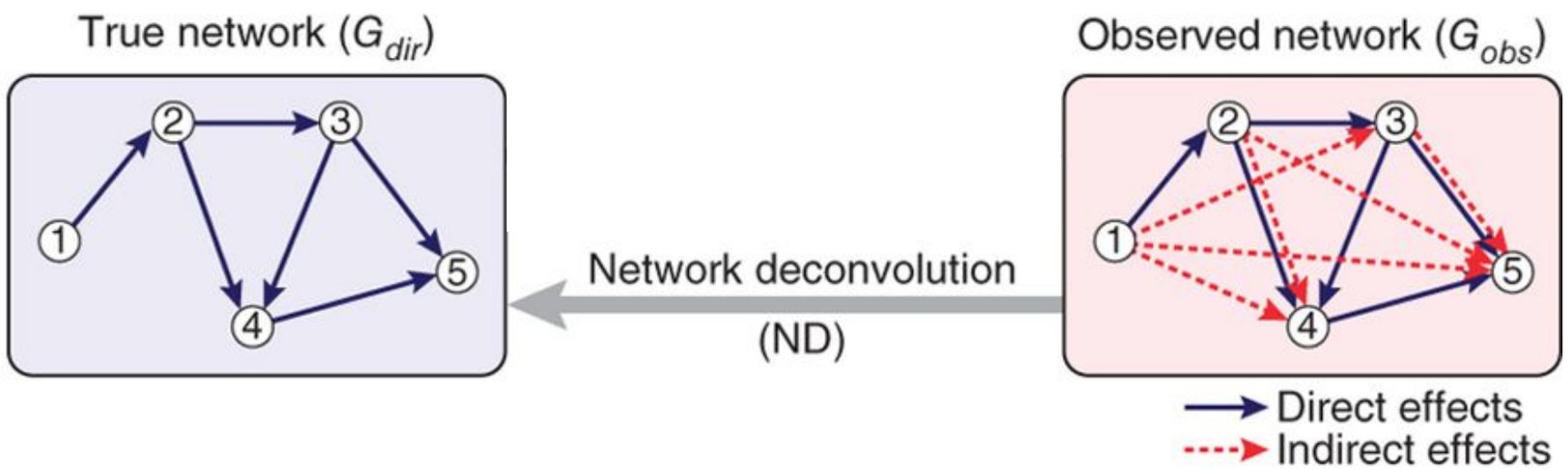


Indirect effects can be of arbitrary length:

- 2 → 3 → 5
- 1 → 2 → 3 → 5
- …

Decreasing effect with increasing path length

4. Network deconvolution



$$G_{obs} = G_{dir} + G_{indir}$$

$$G_{indir} = G_{dir}^2 + G_{dir}^3 + G_{dir}^4 + \ldots$$

3rd order effects (paths of length 3)

2nd order effects (paths of length 2)

$$G_{obs} = G_{dir} + \underbrace{G_{dir}^2 + G_{dir}^3 + G_{dir}^4 + \ldots}_{\text{Indirect effects}}$$

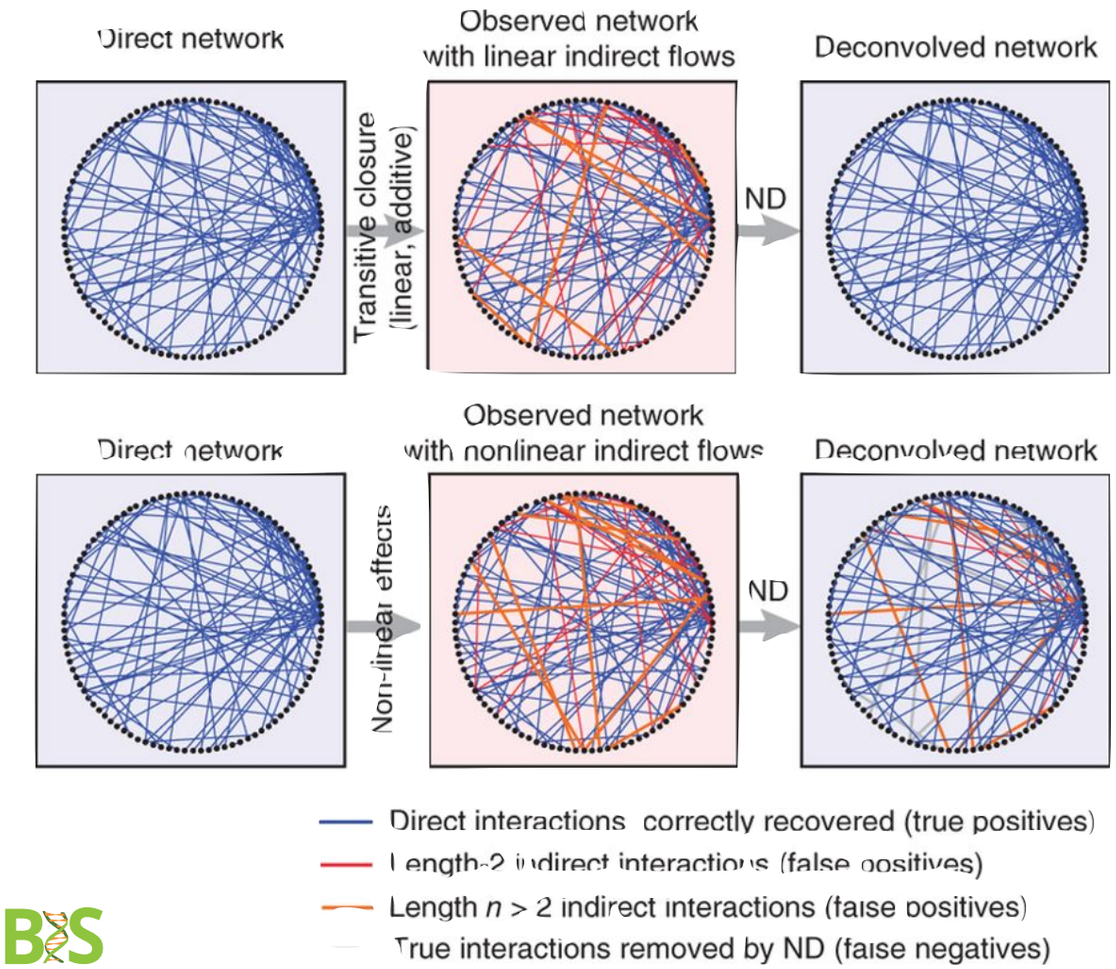
4. Network deconvolution

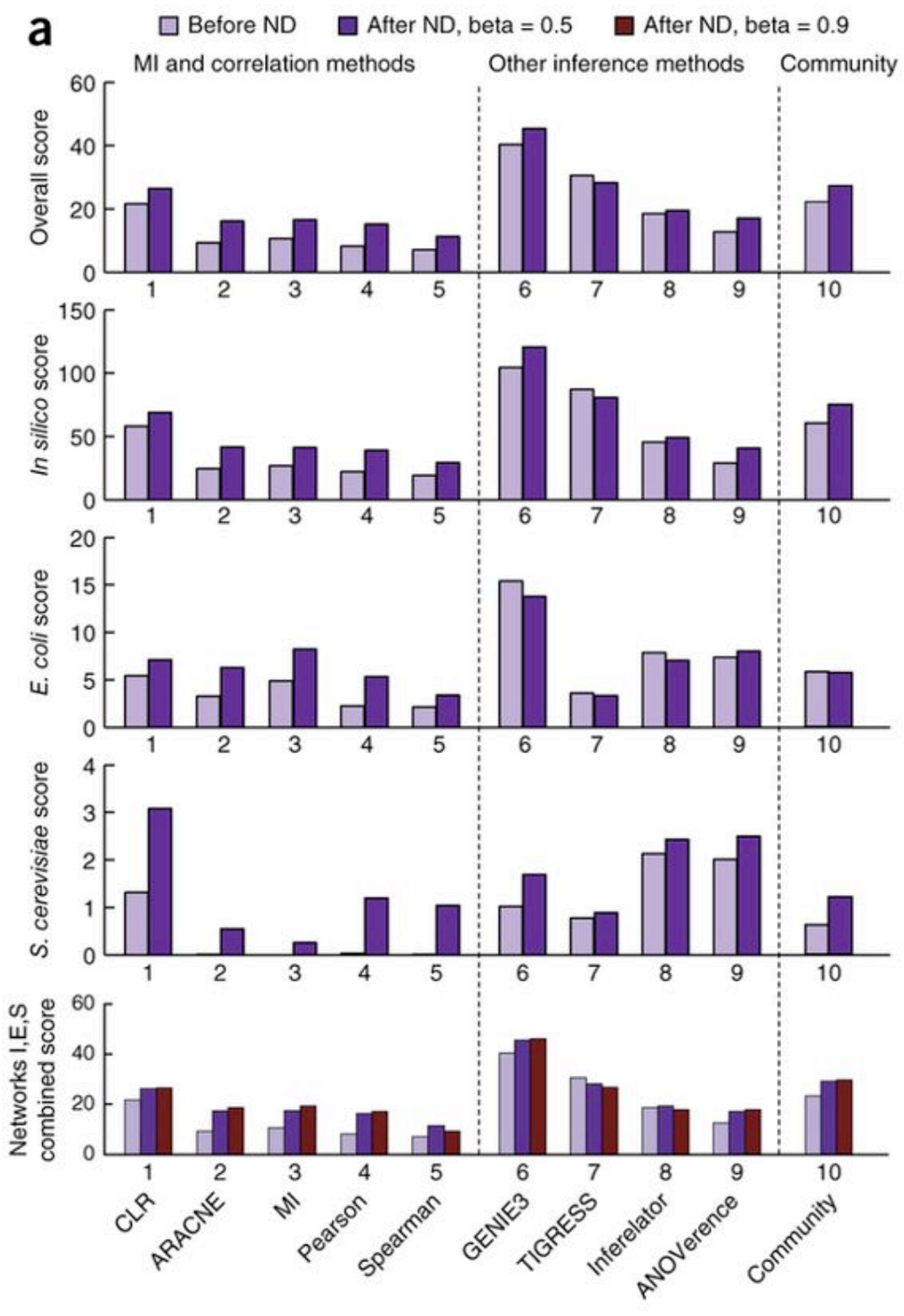
Nonlinear indirect effects are not always captured

May remove some direct interactions

Does not take into consideration edge weight

Still improve predictions (true edges: experimental)

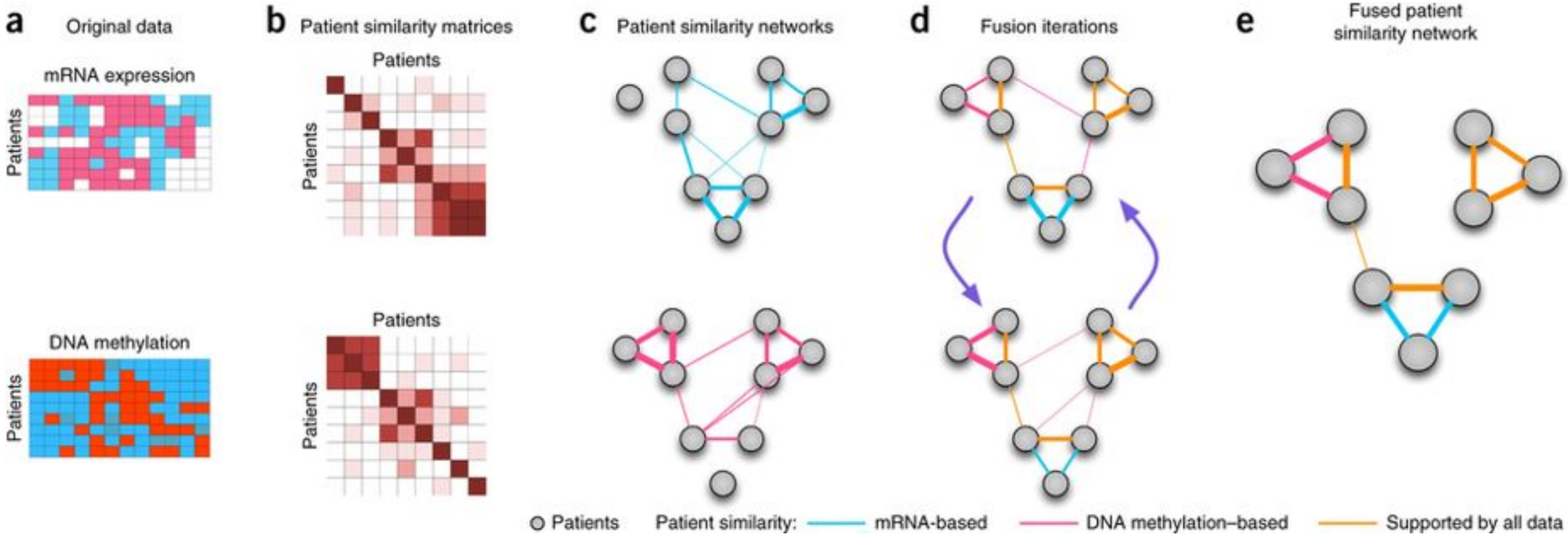




Similarity network fusion

Sample-sample clustering based on multi-omic data improves clustering

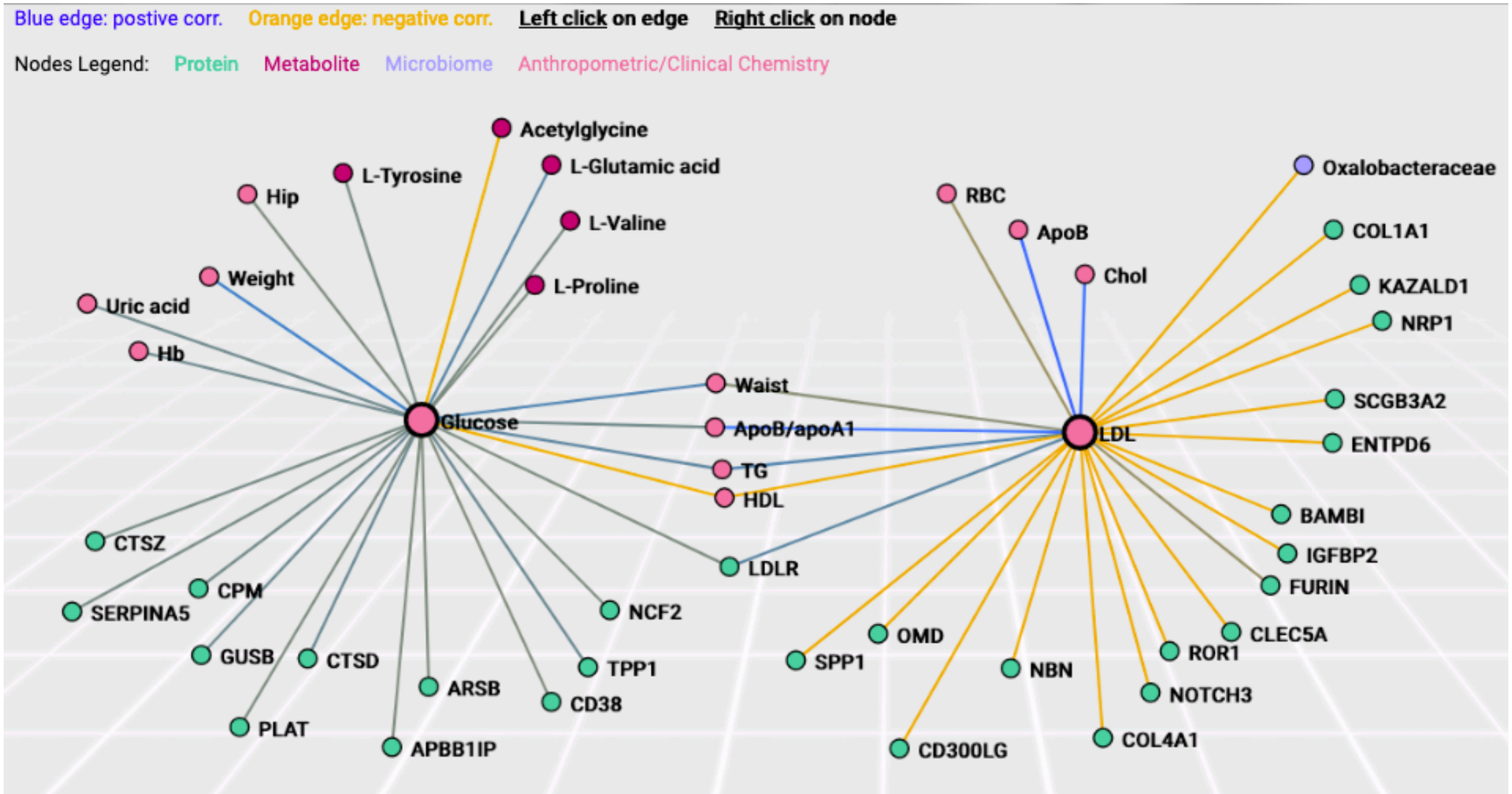
Enables further comparisons between clusters



Key network properties

Motivation

You have built an association network (e.g. PPI, multi-omic, GEM-derived). How to identify pivotal features, their organization, and biological characteristics?



Key network properties to discuss

1. Network density
2. Paths
3. Centrality
4. Degree distributions
5. Small world

1. Network density

A **dense graph** is a graph where the number of edges approximates the maximum possible number of edges for the given node number.

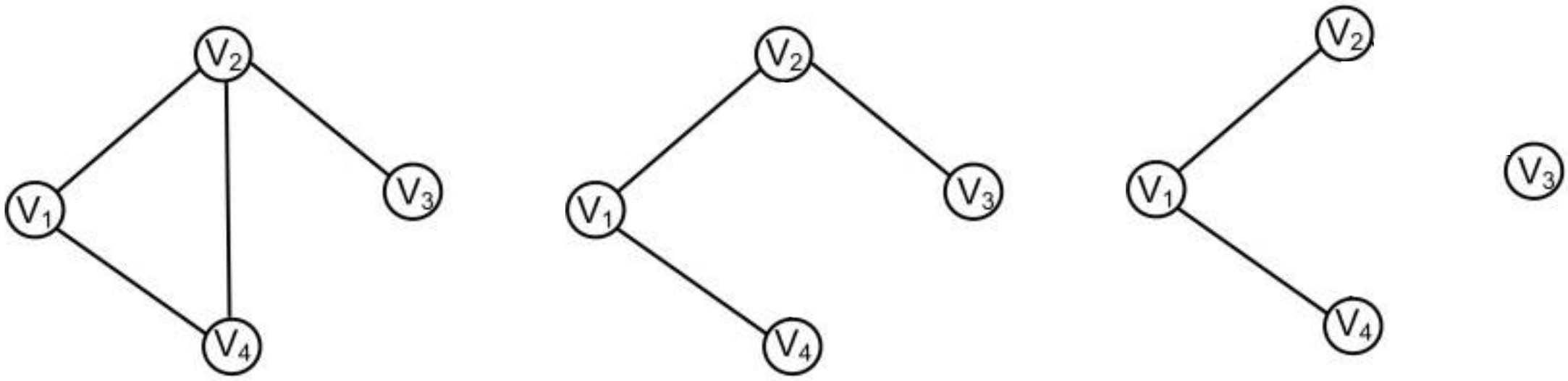
We can thus compute the network **density** (or **global connectivity**) as

$$D = \frac{E}{V \cdot (V - 1)}$$

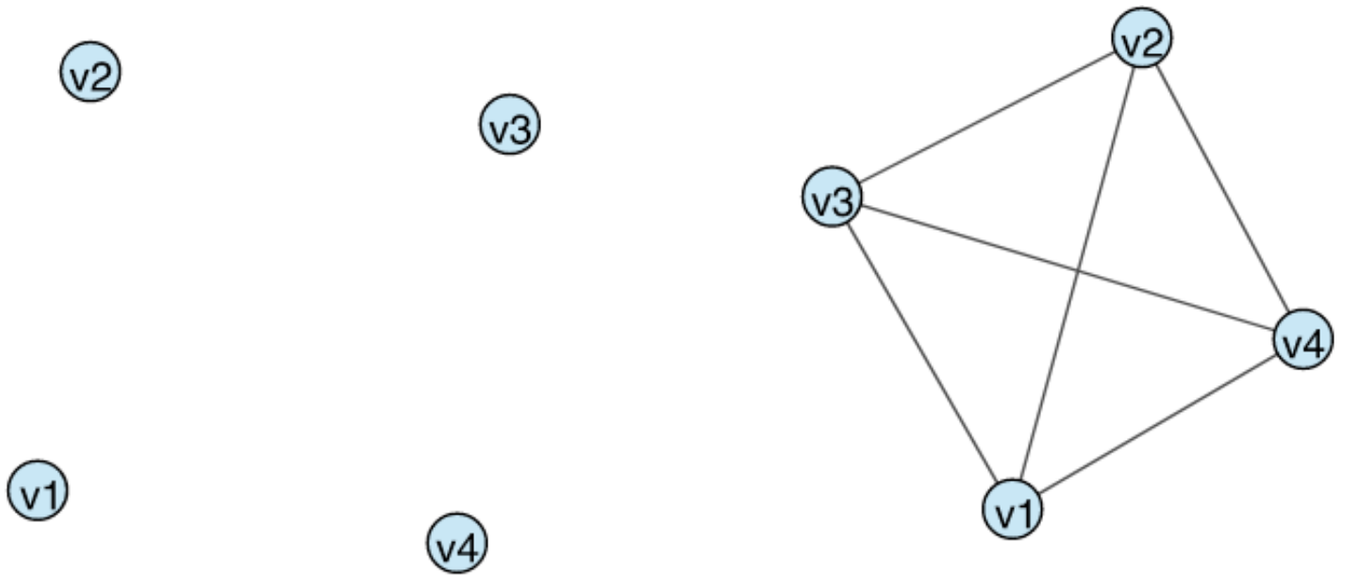
(≠ directed graphs)

E : number of edges

V : number of vertices

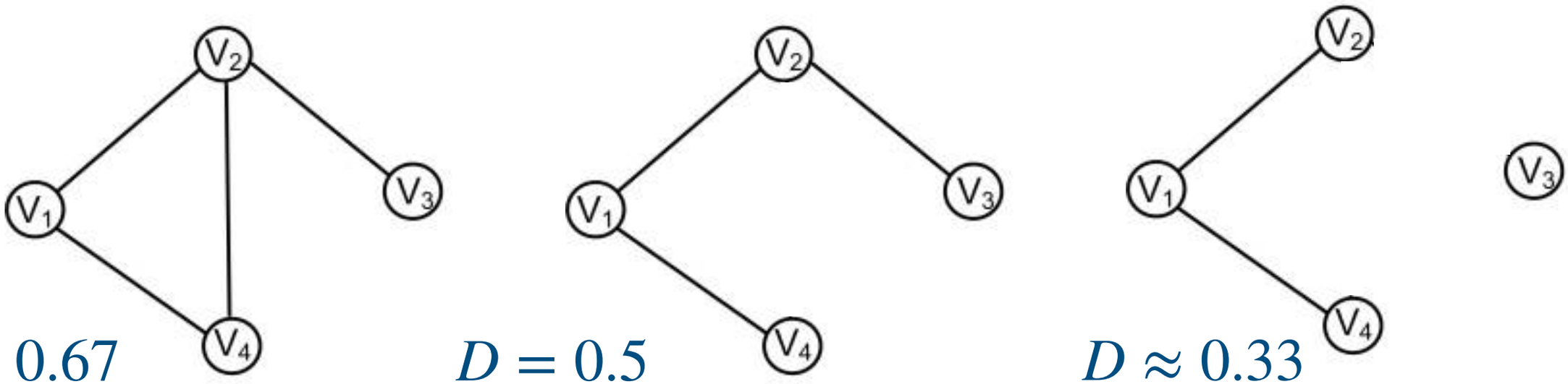


1. Network density



$$0 \leq D \leq 1$$

Higher density indicates higher associations in the network, which implies lower resilience to changes.



$D \approx 0.67$ $D = 0.5$ $D \approx 0.33$

1. Biological network density

Evolutionary analysis of biological networks indicates general sparsity

Network structure must balance robustness to mutation, stochasticity and environmental queues

Sparse networks show higher robustness when accounting for costs and benefits of complexity

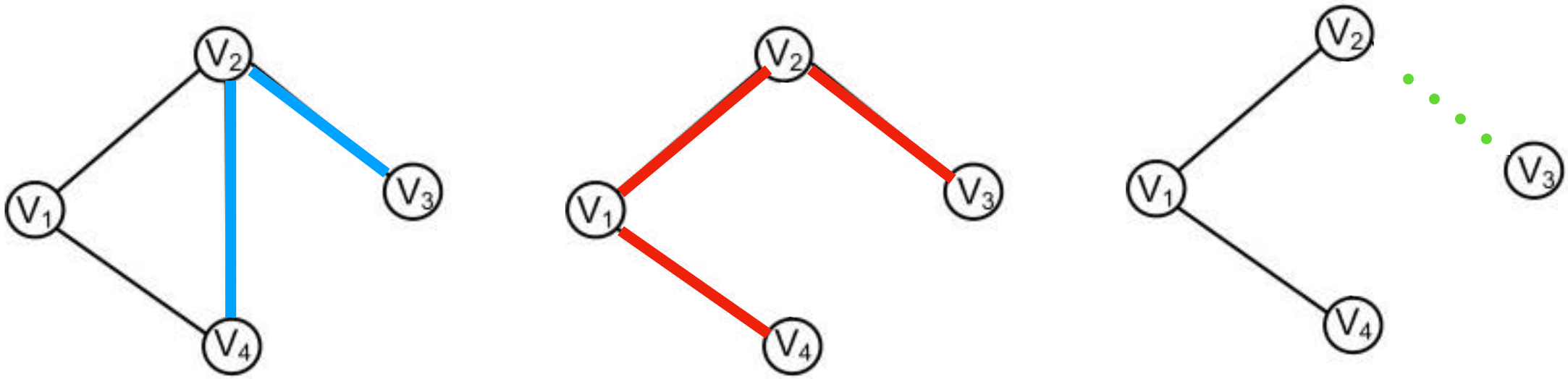
Table I Biological networks are sparsely connected

Organism	Interactions	Genes	D	K
Drosophila melanogaster	29	14	0.148	2.07
D. melanogaster	45	25	0.072	1.8
Sea urchin	82	44	0.0065	1.86
Saccharomyces cerevisiae	1052	678	0.0023	1.55
S. cerevisiae	3969	2341	0.0007	1.7
S. cerevisiae	106	56	*0.0338*	1.9
Escherichia coli[a]	578	423	0.0032	1.37
Arabidopsis thaliana[b]	18 625	6760	0.0004	2.75

2. Paths

Distance between nodes is measured in path length

Important in many centrality measures



d(v4,v3) = 2

d(v4,v3) = 3

d(v4,v3) = inf

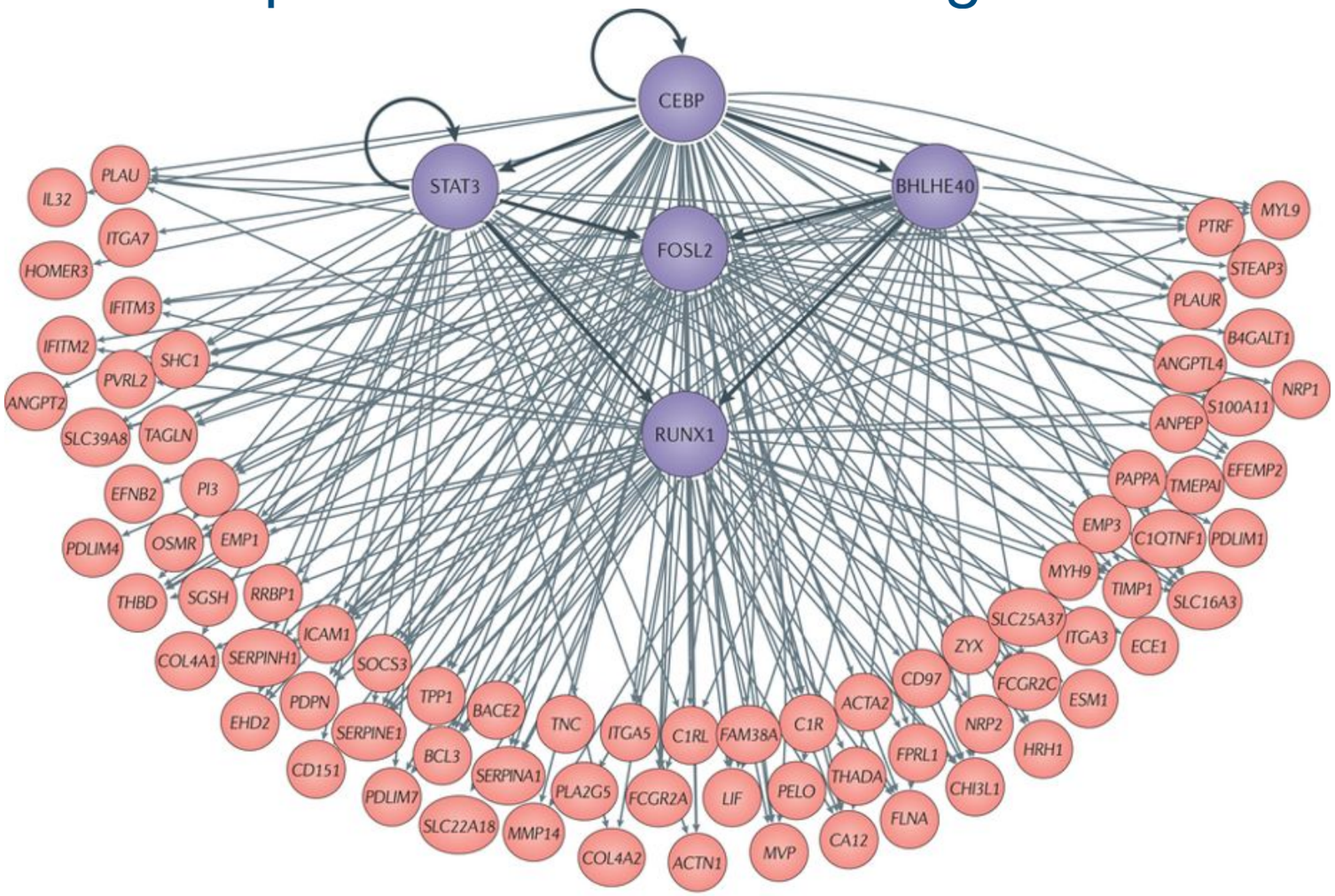
3. Centrality

Indicate the most central nodes in a network

Why look at the central nodes?

Hubs

Example: Transcription Factor Master Regulators



Nature Reviews | Cancer

3. Centrality

Indicate the most central nodes in a network

Central nodes **possibly** most important in the network

There are many different measures of centrality:

- **Degree**
- **Eccentricity**
- *Closeness*
- *Betweenness*
- *Eigenvector*
- Katz
- PageRank
- Percolation
- Cross-clique

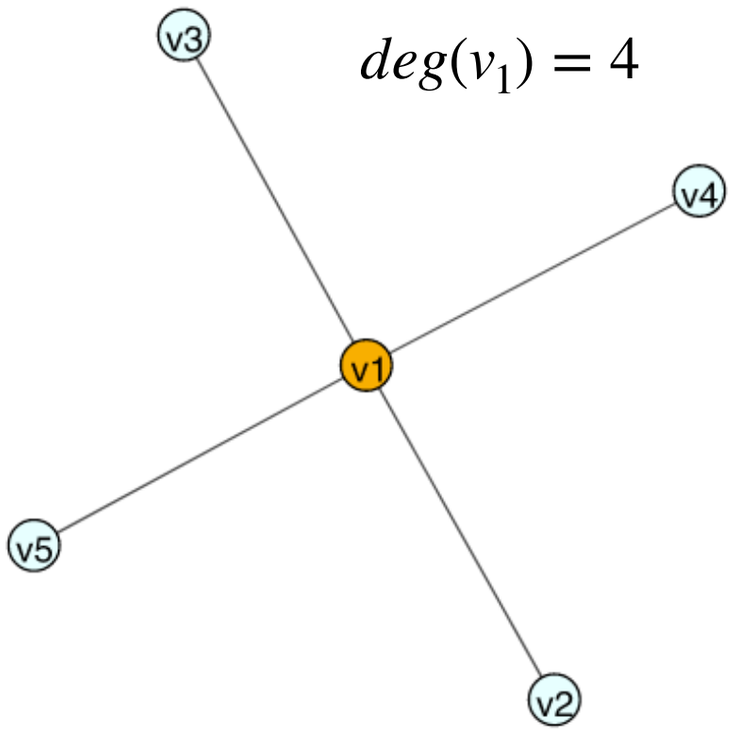
…

3. Centrality: degree centrality

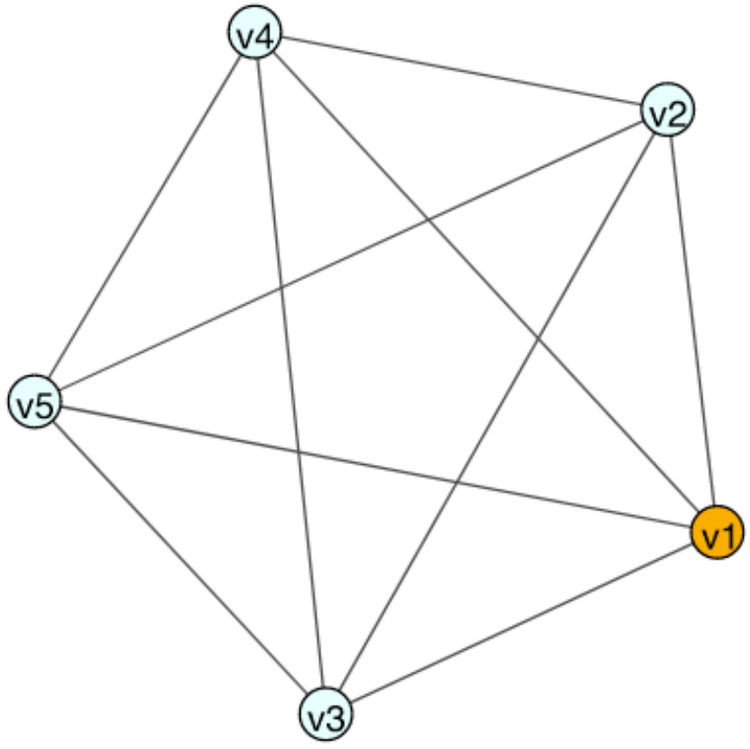
Degree indicates the number of connections with a node

$$d(v) = |N(i)|$$

where *N(i)* is the number of 1st neighbours of a node.



$deg(v_1) = 4$



$deg(v_1) = 4$

3. Centrality: degree centrality

Degree centrality

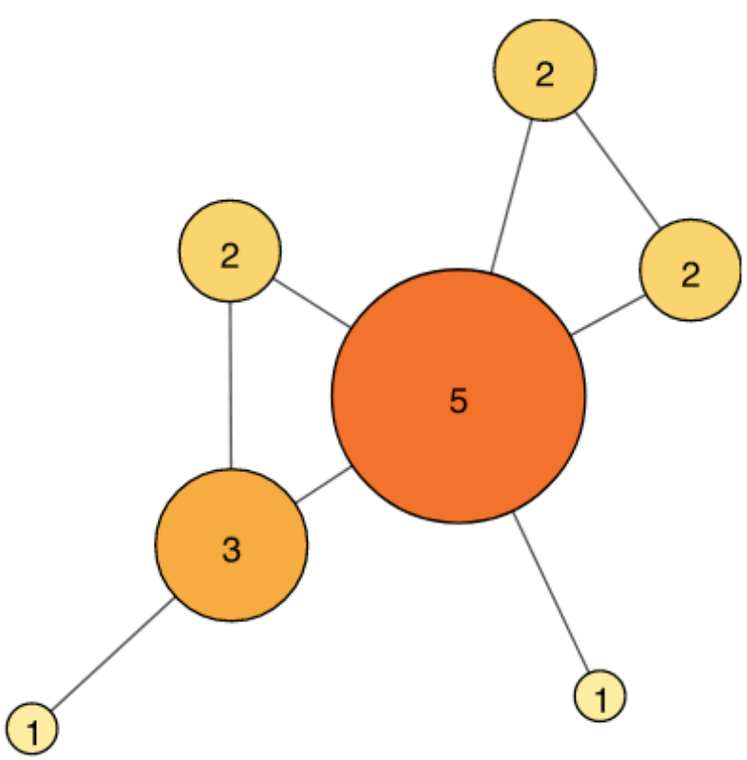
$$C_D(v_i) = \sum_{j=1}^{N} e_{ij}$$

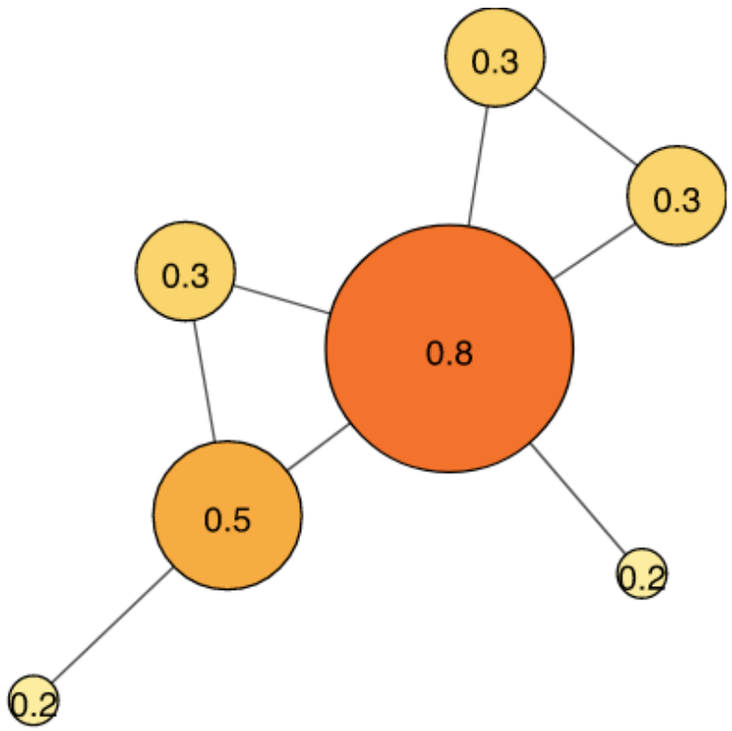
Normalized degree centrality

$$C_D(v_i) = \frac{\sum_{j=1}^{N} e_{ij}}{N-1}$$

Normalized degree centrality accounts for the total possible number of connections

Centrality normalization allows for comparison between networks of different sizes





Network inference and nodes with highest degree

3. Centrality: eccentricity centrality

Network Diameter is the maximum distance

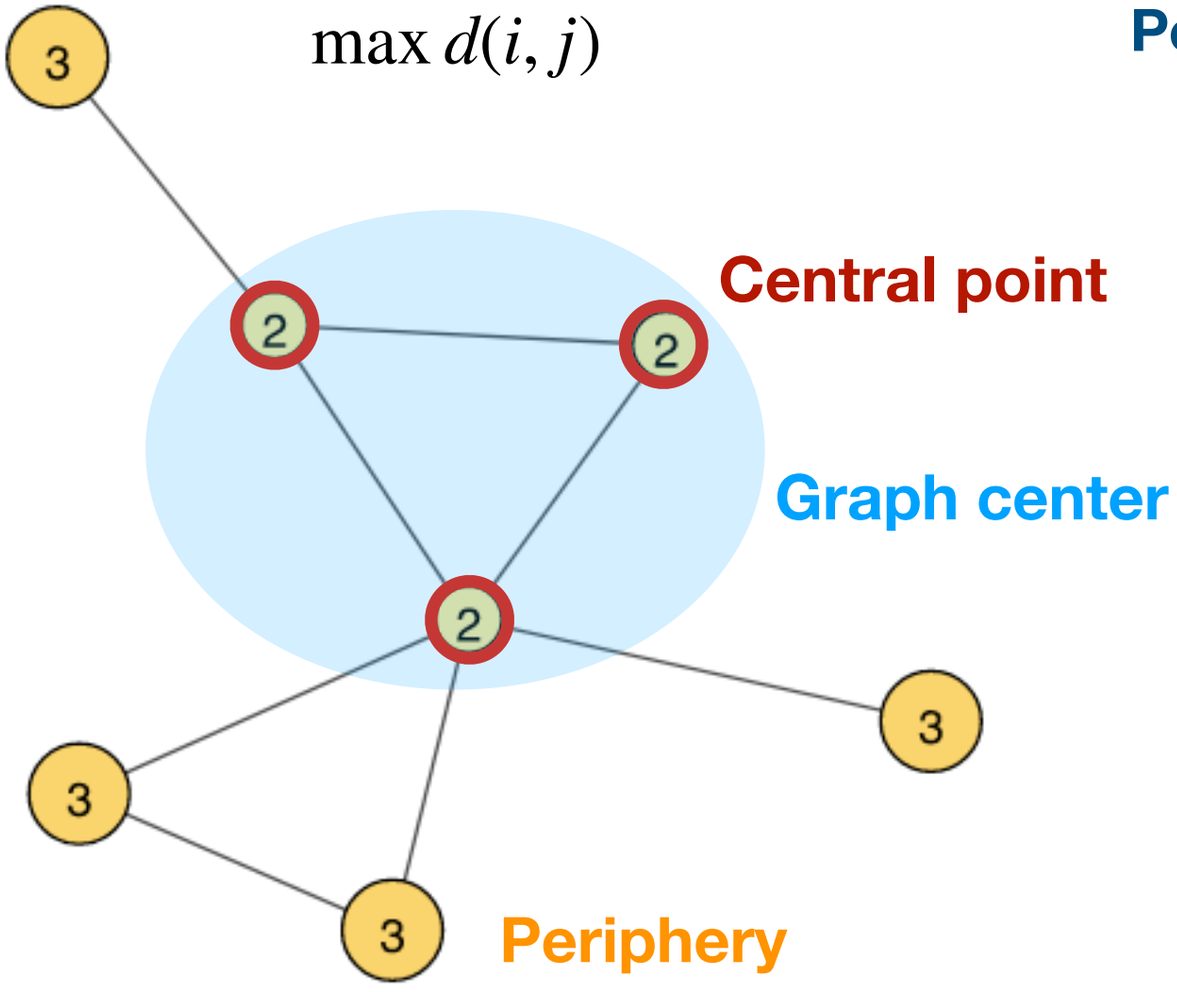
Radius is the minimum distance

Central point is that which has $d(i, j) =$ radius

Graph center: set of nodes with minimum ecc

Periphery: set of nodes with $d(i, j) =$ diameter

$$\max d(i, j)$$



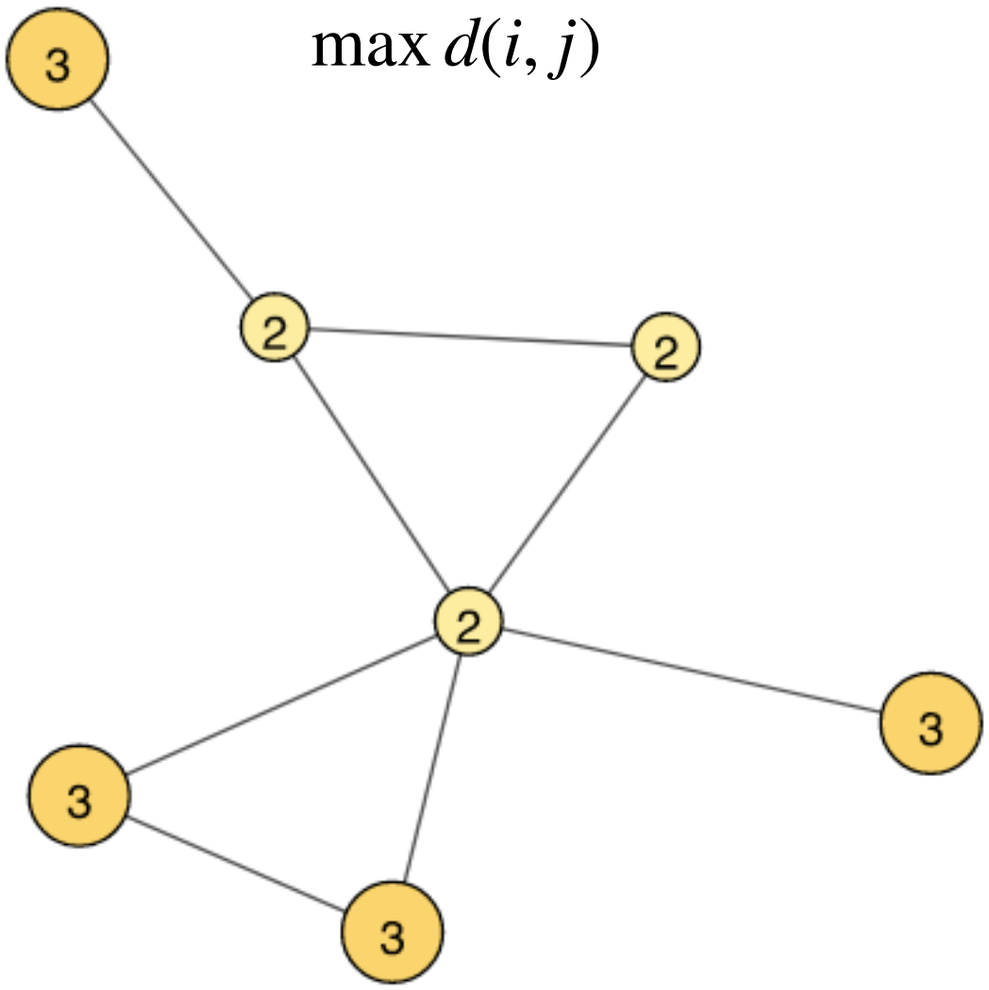
Diameter = 3

Radius = 2

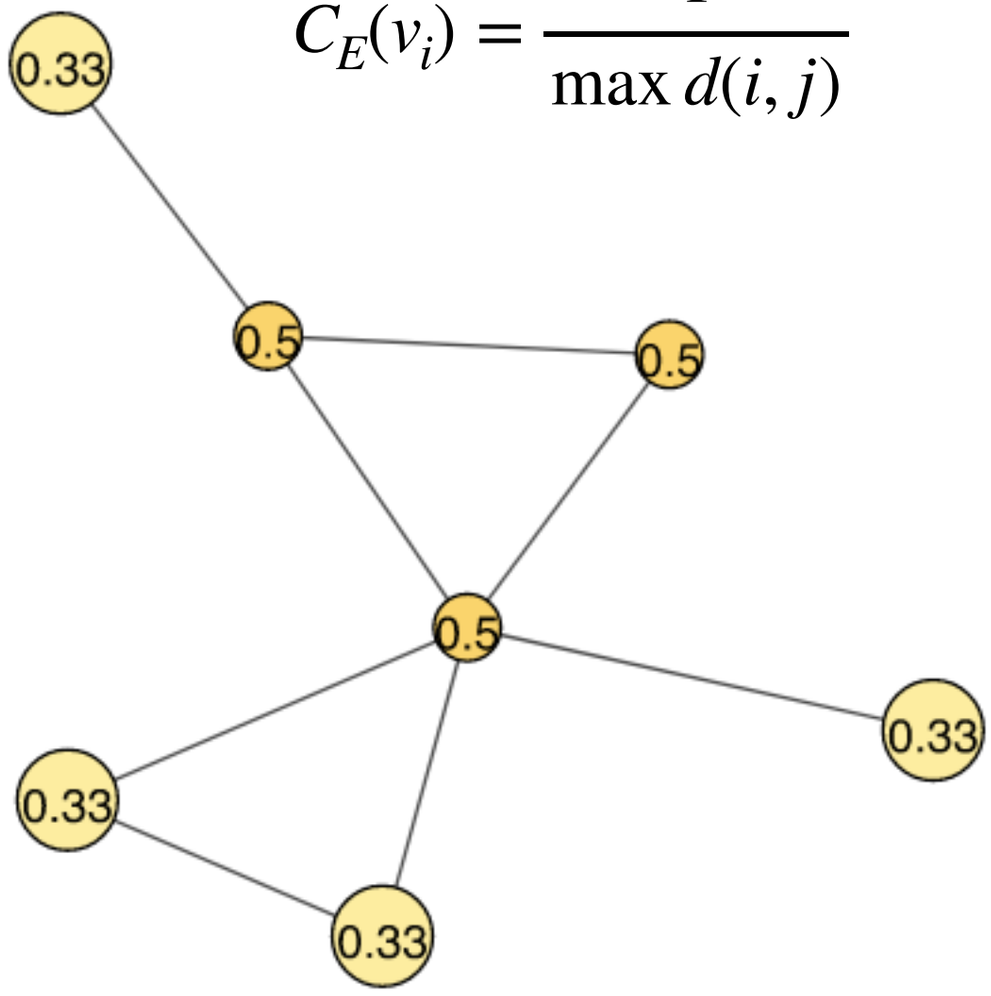
3. Centrality: eccentricity centrality

Eccentricity considers a node's max path to all other nodes

$$\max d(i, j)$$



$$C_E(v_i) = \frac{1}{\max d(i, j)}$$



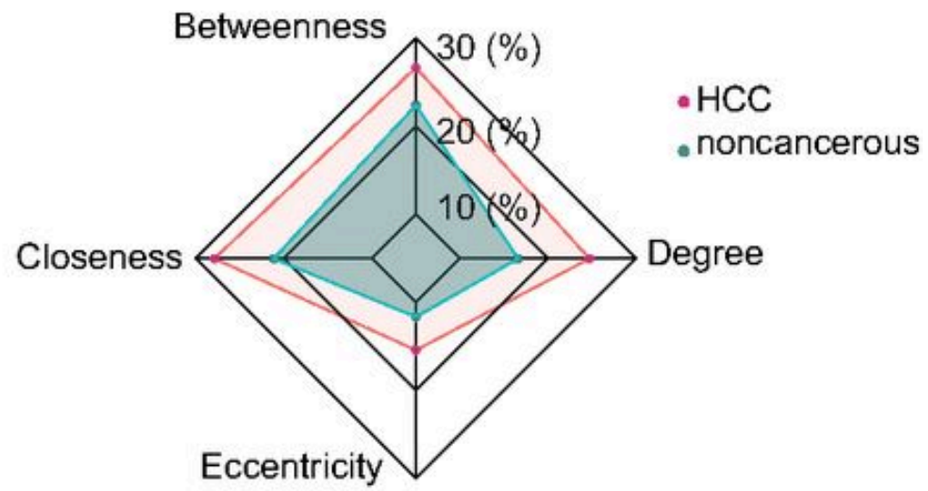
3. Centrality: limitations

Node centrality does not necessarily imply **importance**

How to tackle this?

1. Complement with experimental observations

2. Compute multiple metrics and summarise joint observations



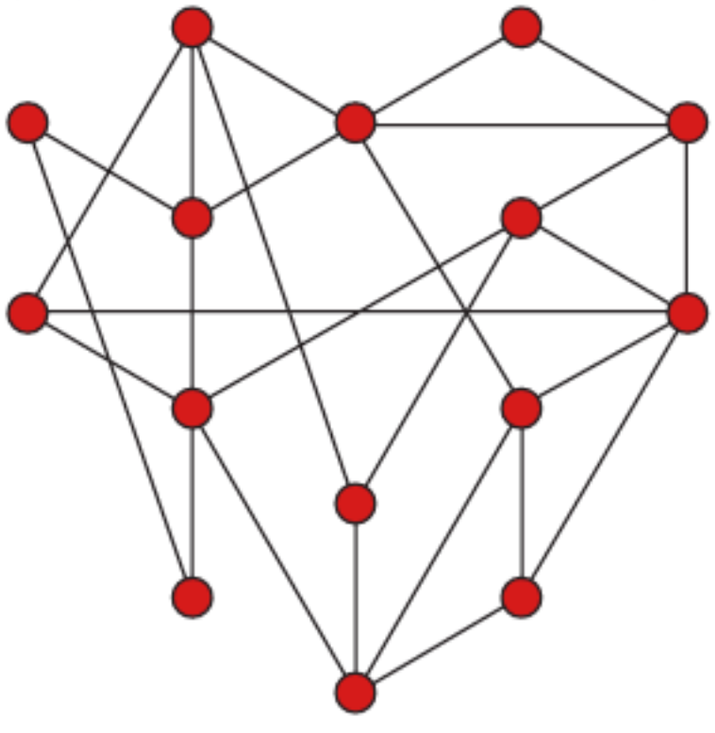
Bidkhori 2018

3. Degree distribution

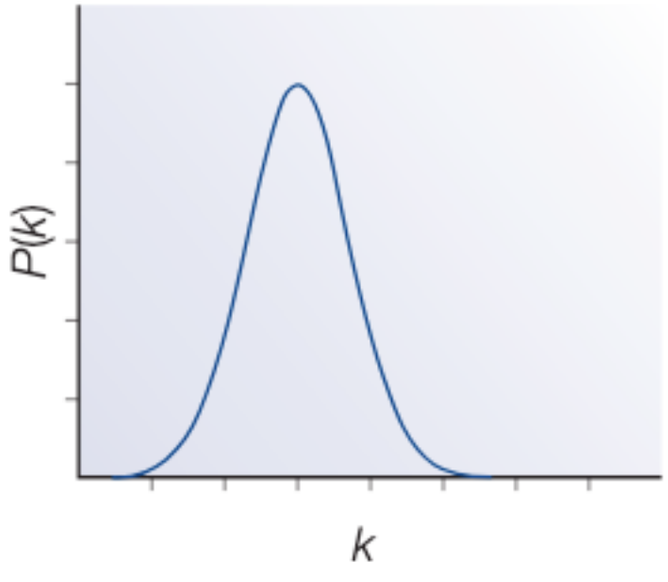
P(k) gives the probability that a selected node has exactly *k* edges

Allows distinguishing different kinds of networks

Random network
(e.g. Erdös-Rényi model)



Poisson degree distribution
shows no highly connected nodes

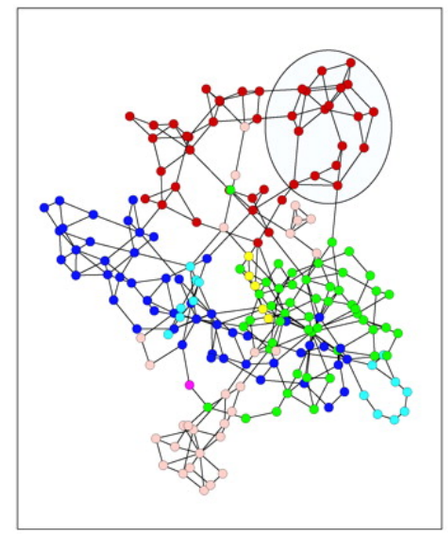


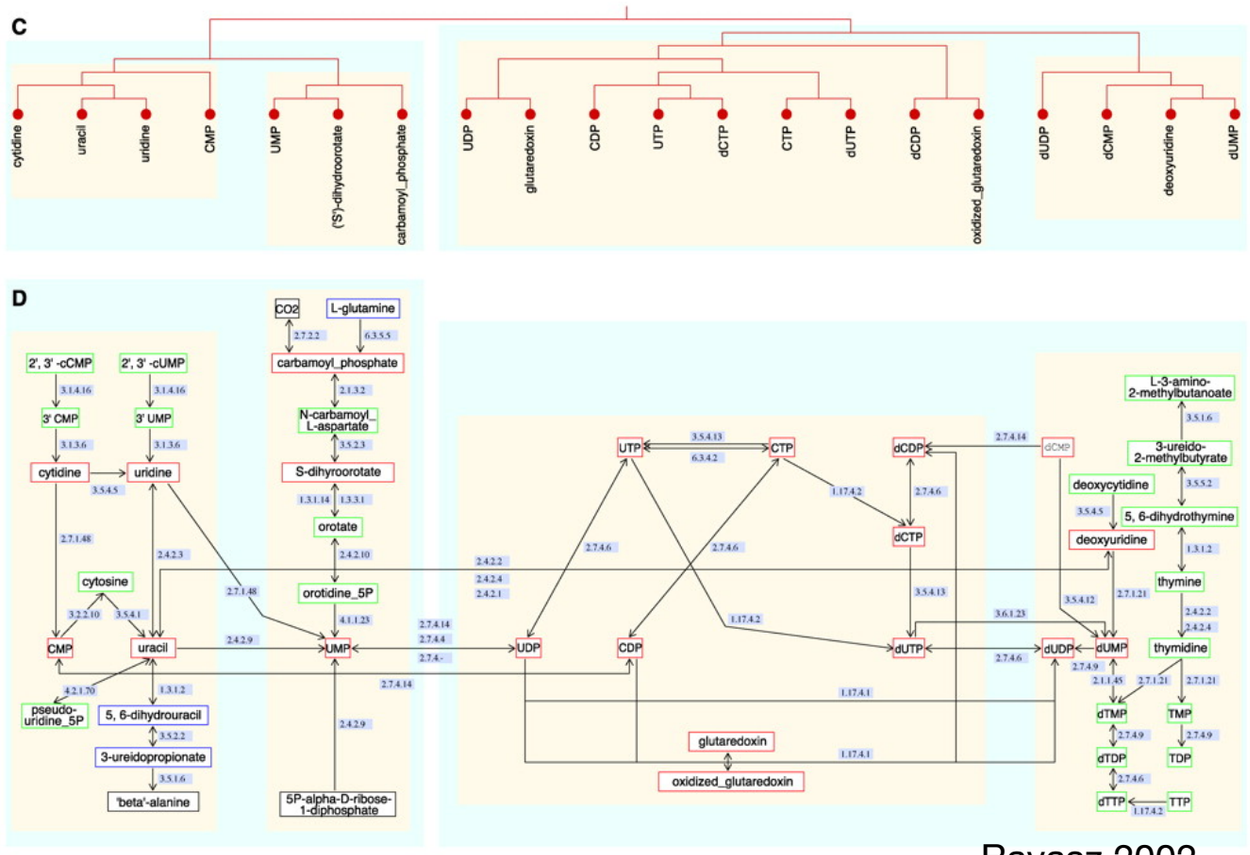
Most nodes have near <*k*>

Metabolic networks show hierarchical topology

Metabolic networks of 43 organisms are organised into **small, tightly connected modules**

Their combination shows a hierarchical structure





Ravasz 2002

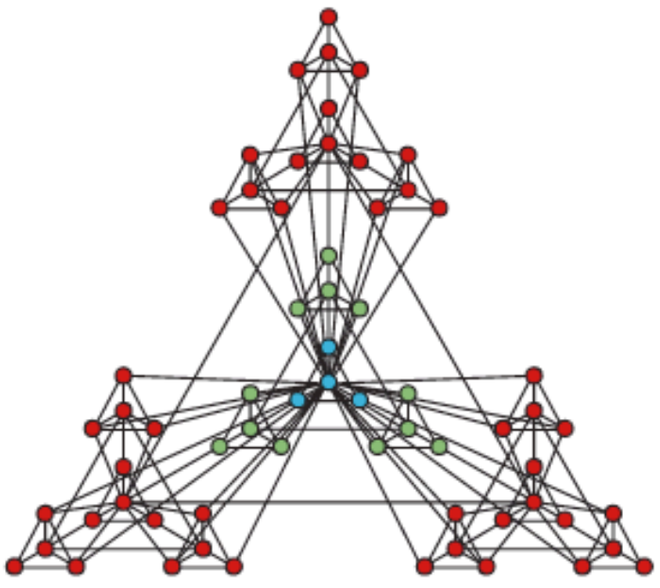
3. Degree distribution

Biological networks do not follow topology features of random networks.

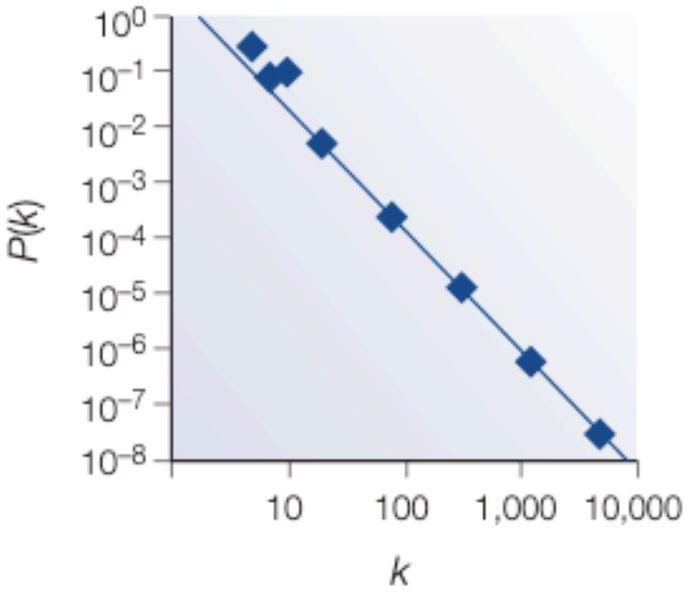
Degree distribution *follows* the power-law $P(k) \propto k^{-\gamma}$, and is termed **scale-free**

Degree exponent $2 < \gamma < 3$, where smaller γ indicates larger degree

Scale-free networks tend to display high robustness to node failure: removal of <80% nodes still retains paths between any two nodes



Degree distribution
shows many with low degrees
a few highly connected nodes



In practice:

$\gamma < 3$:
$P(k) \propto k^{-\gamma}$ and $P(k) \propto N$

$\gamma > 3$:
network behaves like random

Barabasi 2004
Jeong 2000

4. Small world

Any two nodes can be connected in a small number of steps.

This is a property seen in **random networks** where the mean path length

$$l(G) \approx logN$$ for a network of size *N*

Scale-free networks show **ultra-small world**:

$$l(G) \approx log(logN)$$

In practice, this indicates that perturbations may quickly spread throughout the network

Highly central hubs tend **not** to be connected in biological networks: they are **disassortative**

(social networks: **assortative**)

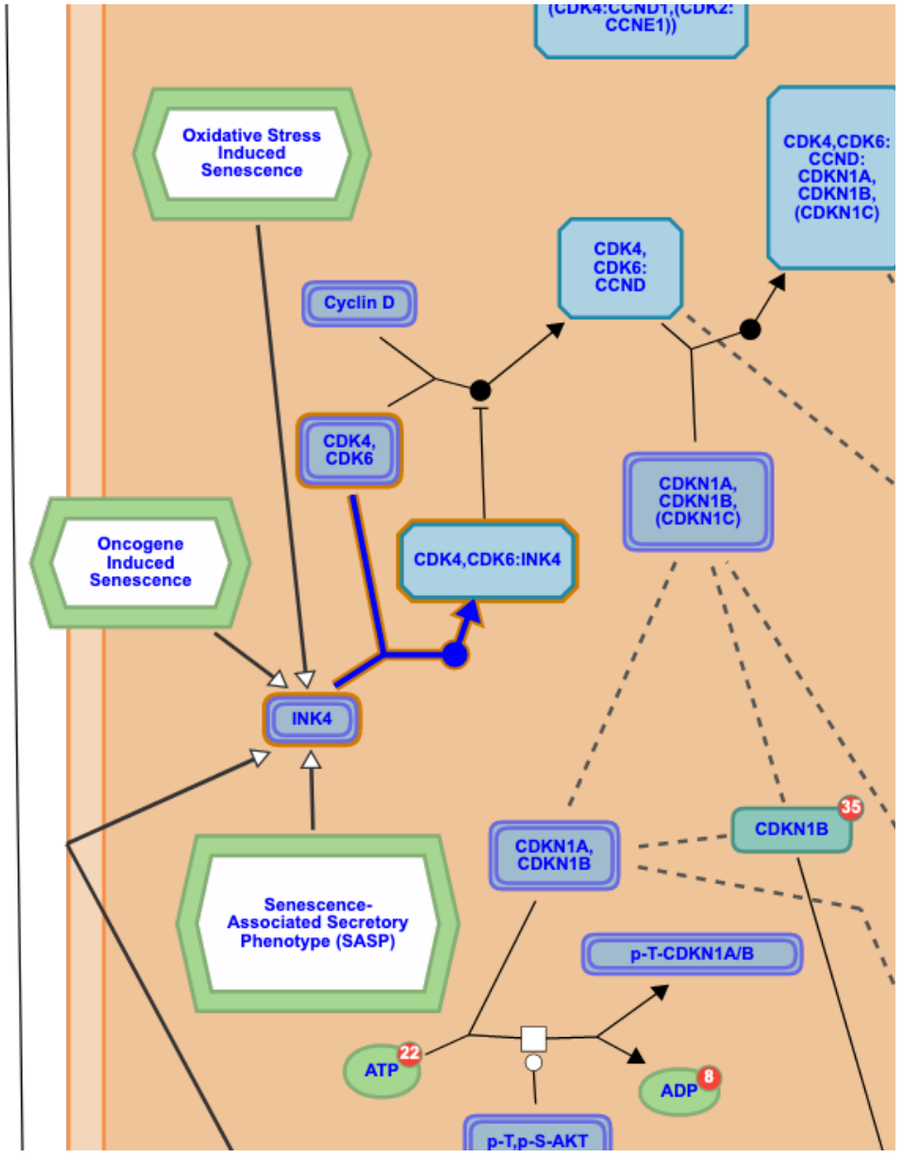
Barabasi 2004

Community and functional analysis

What are modules?

Modules are physically or functionally associated nodes that work together to achieve a distinct function

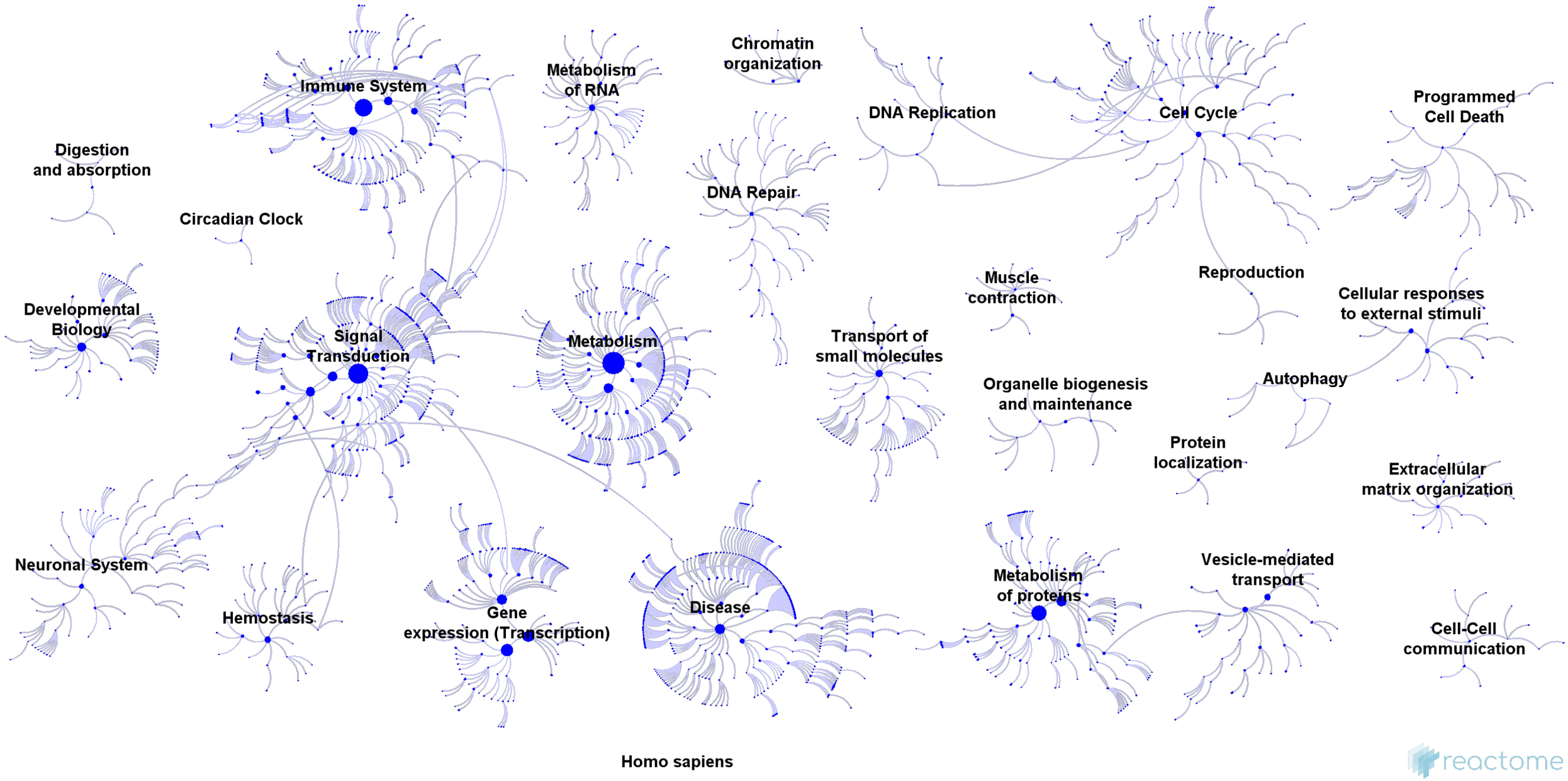
Protein complexes are physical modules



What are modules?

Pathway-associated proteins may represent functional modules

Gene Ontology



What are modules?

In addition to physical or functional modules, one may identify other types of modules

Topological: derived from their high within-module degree

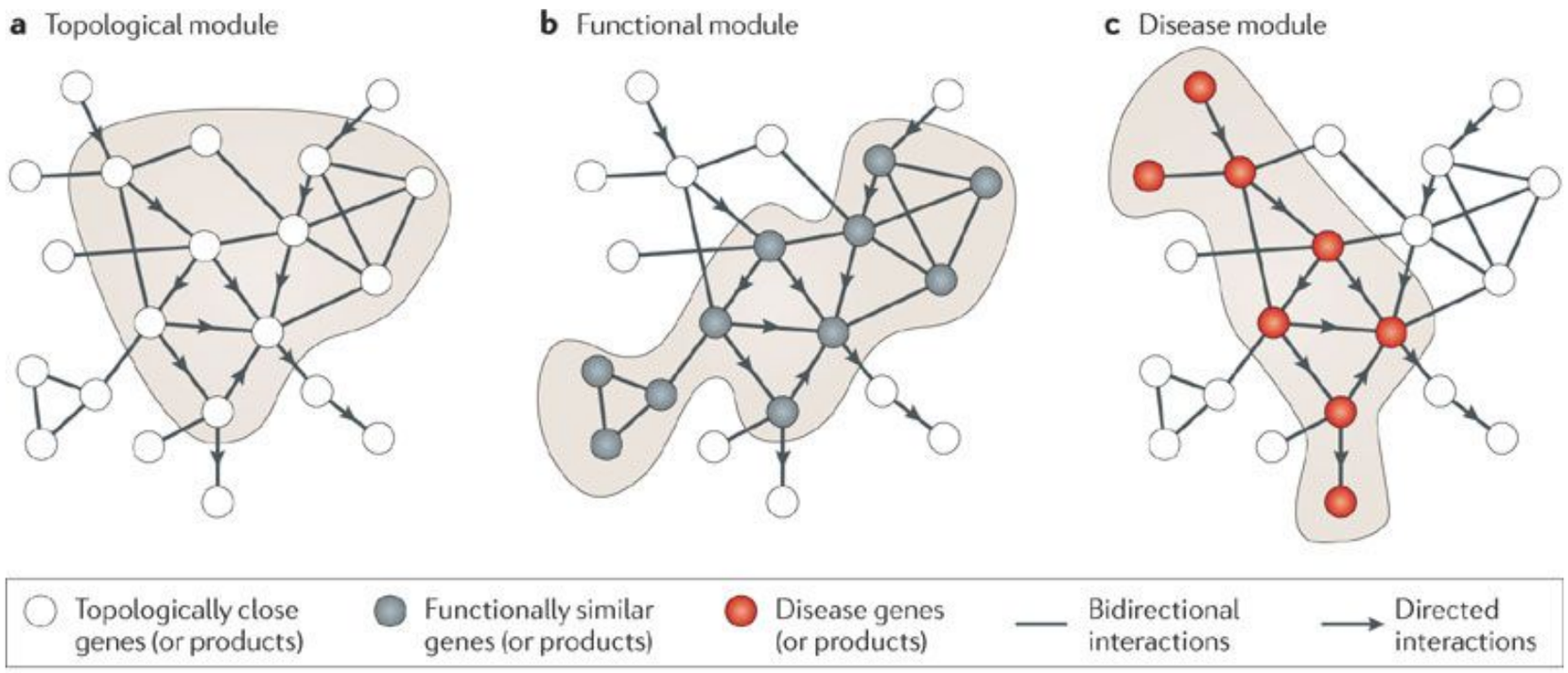
Disease: highly interconnected nodes associated with a disease response

Drug: highly interconnected nodes associated with a drug response

Subgroup: highly interconnected nodes associated with a sample subgroup (e.g. cancer subtype)

Tissue-, cell-type-specific: highly interconnected nodes associated with a specific tissue or cell type

Highly interlinked local regions of a network



Barabasi 2011

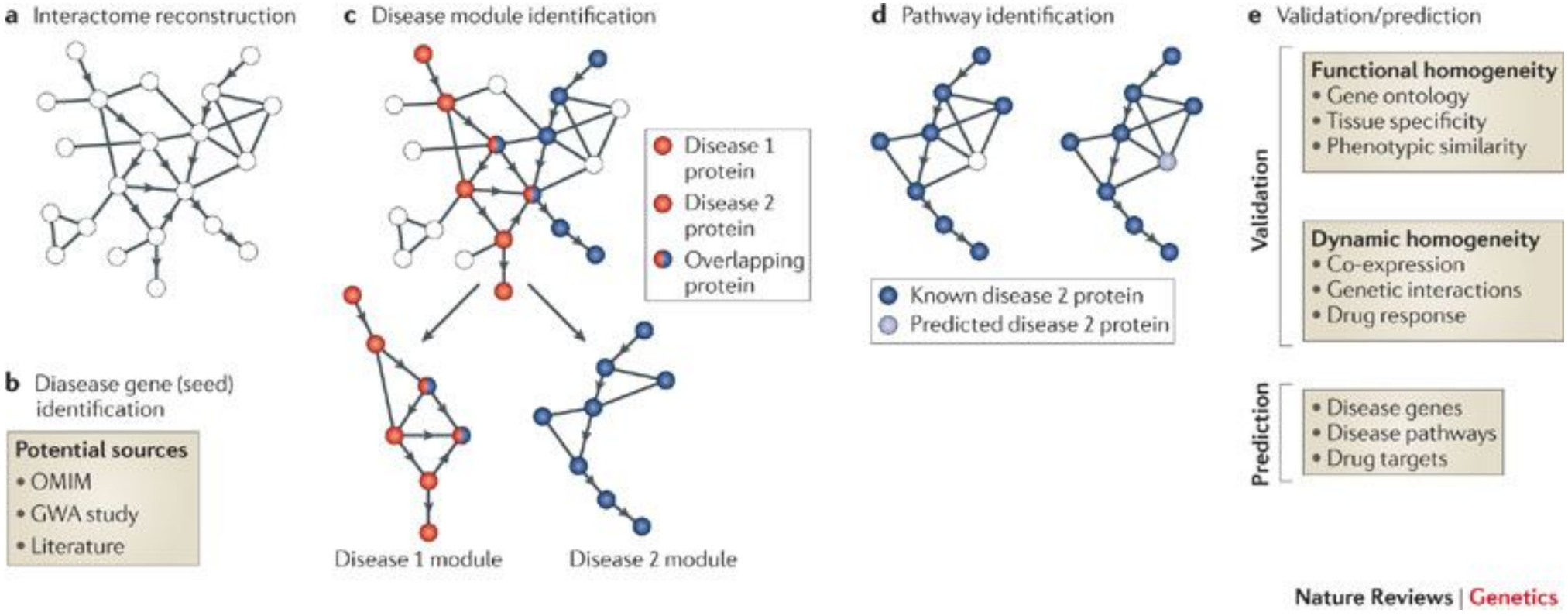
The challenge: identify and characterise modules

Moving from full network to modular characterisation

Different features (diseases, biological processes, etc.) may be associated with the same module

Prediction: *in silico*, relies on available knowledge

Validation: experimental responses (e.g. drug or essentiality screens)



Modularity

Modularity is a property of the network

Modularity (Q) measures the tendency of a graph to be organised into modules

Modules computed by comparing probability that an edge is in a module vs what would be expected in a random network

For a given partitioning of the network into individual groups s, compute

$$Q \propto \sum_{s \in S} [(e_s) - (e_{s\_expected})]$$

edges in group s

Random network with
same number of nodes, edges and
degree per node

Modularity

Number of expected edges e if network is random, given the degree for its nodes

$$-1 < Q < 1$$

Q = 1: much higher number of edges than expected by chance

Q = -1: lower number of edges than expected by chance

Q > 0.3 - 0.7 means significant community structure

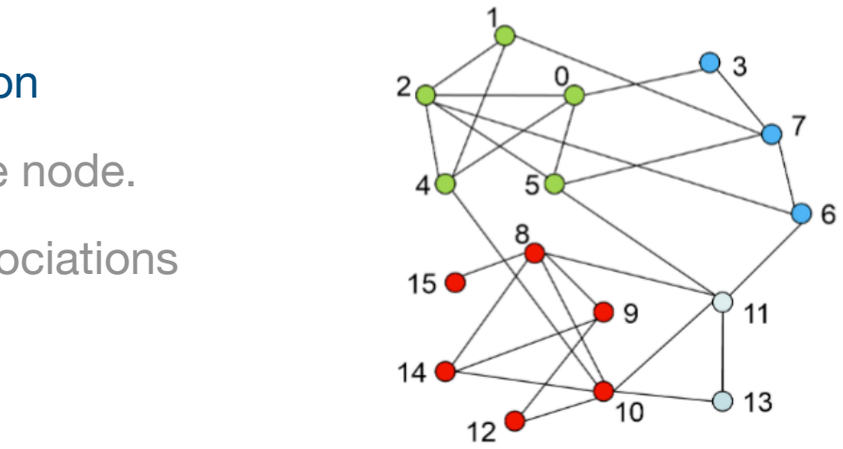
Module detection: Louvain algorithm

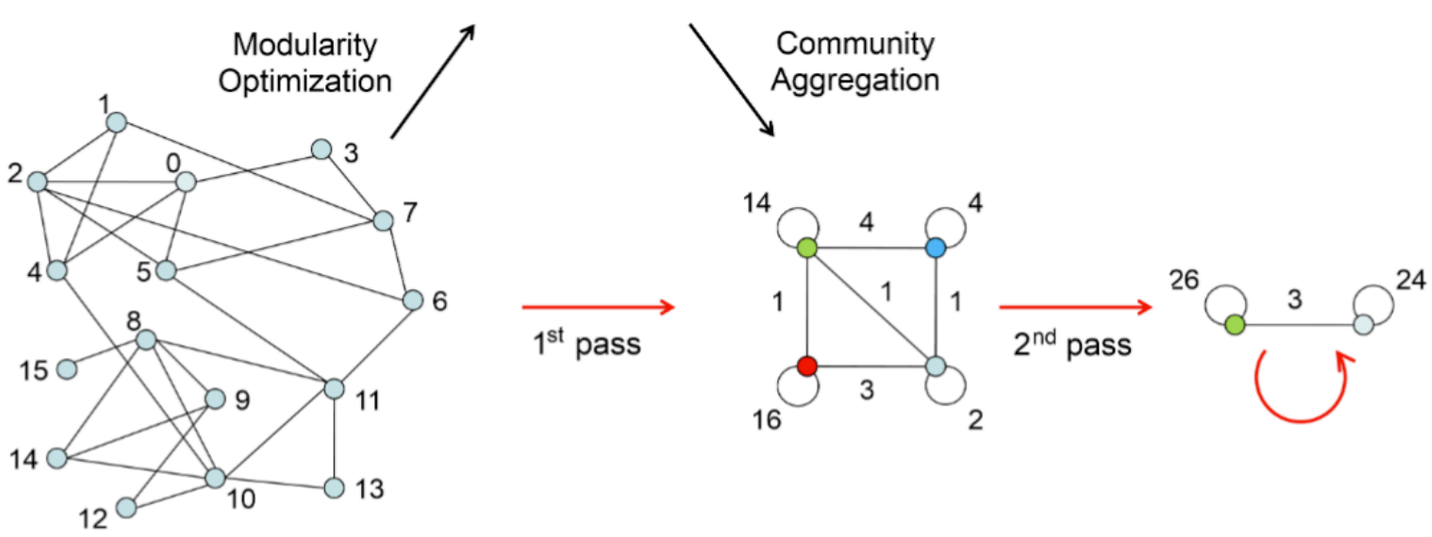
Phase 1: greedy modularity optimisation

1. Start with 1n/community
2. Compute Q by moving *i* to the community of *j*
3. If ΔQ>1, node is placed in community
4. Repeat 1-3 until no improvement is found. Ties solved arbitrarily

Phase 2: coarse grained community aggregation

5. Link nodes in a community into single node.
6. Self loops show intra-community associations
7. Inter-community weights kept
8. Repeat phase 1 on new network





[Leiden algorithm]

Campigotto 2014
Traag 2019

Community characterisation

Hypothesis: community-associated features show coordinated changes associated with common biological processes

GSEA calculates overrepresentation by comparison of gene-level statistics against those of the gene-set, considering sample and feature permutation

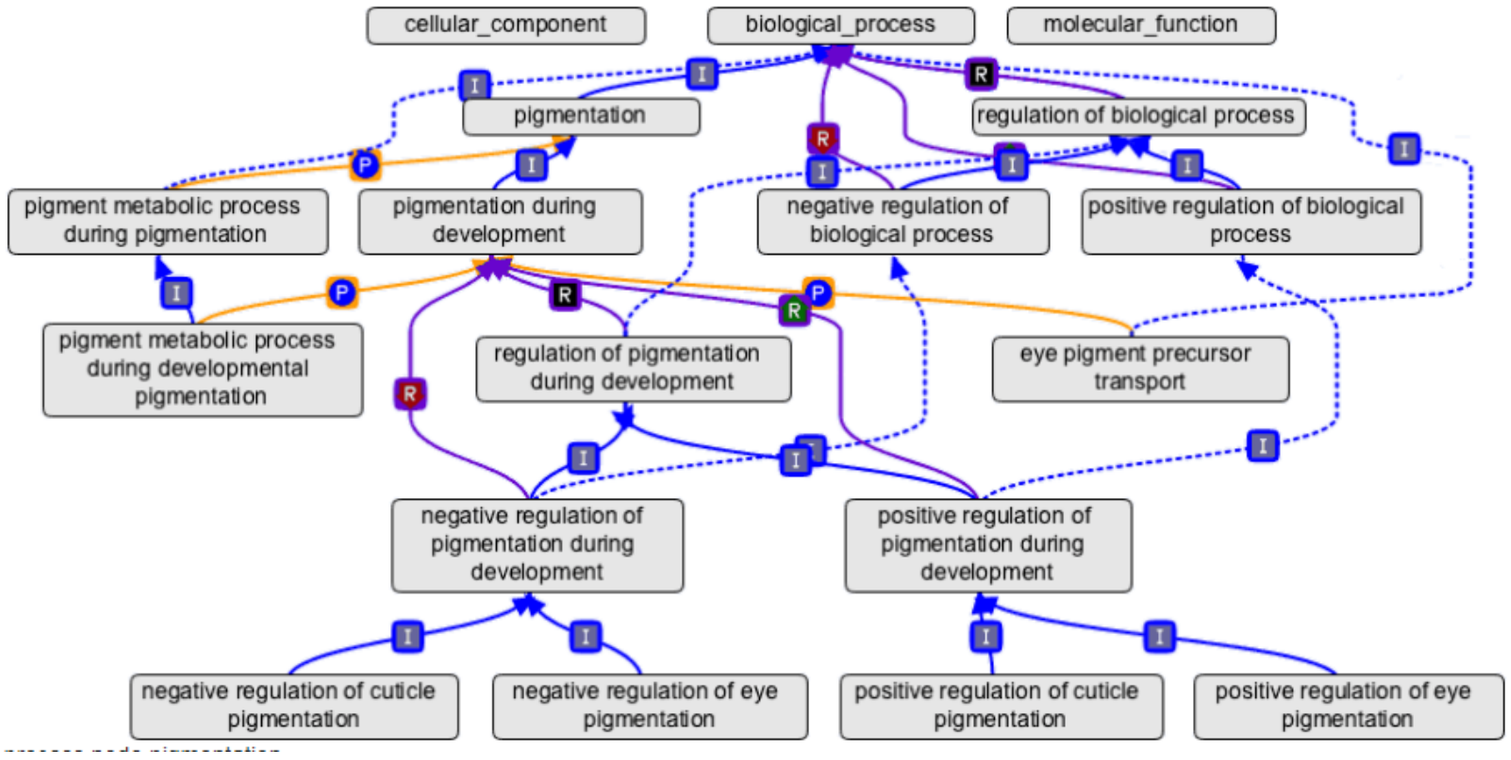
Enrichment analysis

GO-terms, pathways, subcellular location, TF-targets, disease, drug, other?

Tests for significant overlap between groups

Some biological processes may have no biological meaning in your analysis

Does <u>not</u> serve as validation



http://geneontology.org/

Enrichment analysis

MSigDB

GSEA
Gene Set Enrichment Analysis

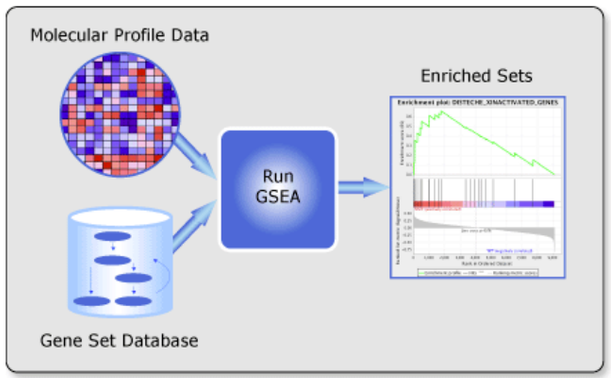
GSEA Home | Downloads | Molecular Signatures Database | Documentation | Contact

Overview

Gene Set Enrichment Analysis (GSEA) is a computational method that determines whether an a priori defined set of genes shows statistically significant, concordant differences between two biological states (e.g. phenotypes).

From this web site, you can:

- **Download** the GSEA software and additional resources to analyze, annotate and interpret enrichment results.
- **Explore the Molecular Signatures Database (MSigDB)**, a collection of annotated gene sets for use with GSEA software.
- **View documentation** describing GSEA and MSigDB.



What's New

20-Aug-2019: MSigDB 7.0 released. This is a major release that includes a complete overhaul of gene symbol annotations, Reactome and GO gene sets, and corrections to miscellaneous errors. See the release notes for more information.

20-Aug-2019: GSEA 4.0.0 released. This release includes support for MSigDB 7.0, plus major internal updates for Java 11 support and performance improvements. See the release notes for more information.

16-Jul-2018: MSigDB 6.2 released. This is a minor release that includes updates to gene set annotations, corrections to miscellaneous errors, and a handful of new gene sets. See the release notes for more information.

Follow @GSEA_MSigDB

License Terms

GSEA and MSigDB are available for use under these license terms.

Please register to download the GSEA software, access our web tools, and view the MSigDB gene sets. After registering, you can log in at any time using your email address. Registration is free. Its only purpose is to help us track usage for reports to our funding agencies.

Contributors

GSEA and MSigDB are maintained by the GSEA team. Our thanks to our many contributors. Funded by: National Cancer Institute, National Institutes of Health, National Institute of General Medical Sciences.

Citing GSEA

To cite your use of the GSEA software, please reference Subramanian, Tamayo, et al. (2005, PNAS 102, 15545-15550) and Mootha, Lindgren, et al. (2003, Nat Genet 34, 267-273).

Enrichr

Enrichr

Login | Register

21,153,478 lists analyzed
307,486 terms
154 libraries

Analyze | What's New? | Libraries | Find a Gene | About | Help

Gene-set Library	Terms	Gene Coverage	Genes per Term
Genes_Associated_with_NIH_Grants	32876	15886	9.0
Cancer_Cell_Line_Encyclopedia	967	15797	176.0
Achilles_fitness_decrease	216	4271	128.0
Achilles_fitness_increase	216	4320	129.0
Aging_Perturbations_from_GEO_down	286	16129	292.0
Aging_Perturbations_from_GEO_up	286	15309	308.0
Allen_Brain_Atlas_down	2192	13877	304.0
Allen_Brain_Atlas_up	2192	13121	305.0
ARCHS4_Cell-lines	125	23601	2395.0
ARCHS4_IDG_Coexp	352	20883	299.0
ARCHS4_Kinases_Coexp	498	19612	299.0
ARCHS4_TFs_Coexp	1724	25983	299.0
ARCHS4_Tissues	108	21809	2316.0
BioCarta_2013	249	1295	18.0
BioCarta_2015	239	1678	21.0
BioCarta_2016	237	1348	19.0
BioPlex_2017	3915	10271	22.0
ChEA_2013	353	47172	1370.0
ChEA_2015	395	48230	1429.0
ChEA_2016	645	49238	1550.0
Chromosome_Location	386	32740	85.0
Chromosome_Location_hg19	36	27360	802.0
CORUM	1658	2741	5.0
Data_Acquisition_Method_Most_Popular_Genes	12	1073	100.0
dbGaP	345	5613	36.0
DepMap_WG_CRISPR_Screens_Broad_CellLines_2019	558	7744	363.0
DepMap_WG_CRISPR_Screens_Sanger_CellLines_2019	325	6204	387.0
Disease_Perturbations_from_GEO_down	839	23939	293.0
Disease_Perturbations_from_GEO_up	839	23561	307.0
Disease_Signatures_from_GEO_down_2014	142	15406	300.0

Enrichment analysis

Important databases with gene-sets:

- MSigDB (gene)
- Enrichr (gene)
- KEGG (metabolite, gene)
- DIANA (miRNA)
- MetaboAnalyst (metabolite)
- DAVID (web)
- Reactome (web)

Creating custom sets and joint sets

Tools for Enrichment analysis

STRING

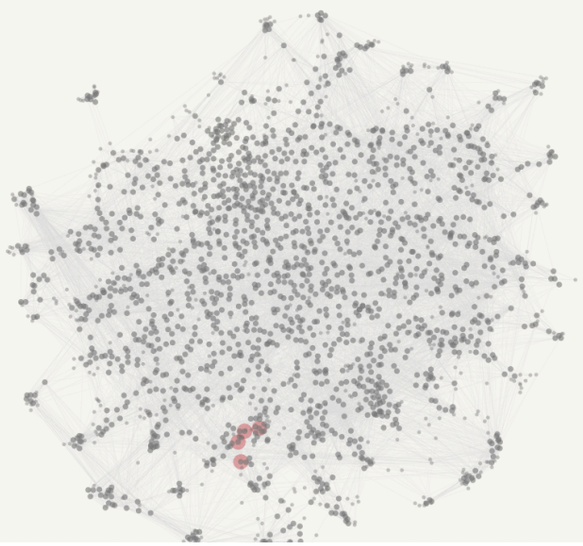
Search Download Help My Data

Your input data

GO:0000083 — 5 genes in your input

1: PKP1	Plakophilin-1	-8.326649
2: CDSN	Corneodesmosin	-8.130157
3: SERPINB5	Serpin B5	-8.065760
4: DSC1	Desmocollin-1	-7.917077
5: DSG1	Desmoglein-1	-7.838328
6: CALML5	Calmodulin-like protein 5	-7.706114
7: ZNF750	Zinc finger protein 750	-7.527767
8: SERPINB7	Serpin B7	-7.497837
9: LCE2B	Late cornified envelope protein 2E	-7.467221
10: CHP2	Calcineurin B homologous protein	-7.423878
11: GJB6	Gap junction beta-6 protein	-7.301189
12: COL17A1	Collagen alpha-1(XVII) chain	-7.263660
13: C19orf33	Immortalization up-regulated prot	-7.195207
14: SBSN	Suprabasin	-7.140458
15: LY6D	Lymphocyte antigen 6D	-7.056120
16: TRIM29	Tripartite motif-containing protein	-7.034785
17: FLG	Filaggrin	-7.031575
18: CRCT1	Cysteine rich C-terminal 1	-7.022690
19: KRT15	Keratin, type I cytoskeletal 15	-6.867025

Full proteome network



Your detected functional enrichments

Biological Process (GO)

GO-term	description	enrichment score	direction	pathway size	false discovery rate
GO:0000083	regulation of transcription involved in G1/S transition of mitotic cell cycle	5.93242	bottom of input	29	0.0044 (afc)
GO:0007094	mitotic spindle assembly checkpoint	5.84439	bottom of input	21	0.0055 (afc)
GO:0030071	regulation of mitotic metaphase/anaphase transition	5.83409	bottom of input	49	0.00045 (afc)
GO:0051983	regulation of chromosome segregation	5.76051	bottom of input	97	0.00018 (afc)
GO:0051784	negative regulation of nuclear division	5.68192	bottom of input	47	0.0023 (afc)

(more ...)

Molecular Function (GO)

GO-term	description	enrichment score	direction	pathway size	false discovery rate
GO:0030280	structural constituent of epidermis	6.53875	top of input	14	0.00043 (afc)
GO:0005198	structural molecule activity	2.84694	top of input	679	0.00043 (ks)
GO:0032559	adenyl ribonucleotide binding	2.1506	bottom of input	1514	0.00046 (ks)
GO:0030554	adenyl nucleotide binding	2.04171	bottom of input	1524	0.00071 (ks)
GO:0005524	ATP binding	2.03433	bottom of input	1462	0.0018 (ks)

(more ...)

Cellular Component (GO)

GO-term	description	enrichment score	direction	pathway size	false discovery rate
GO:0001533	cornified envelope	6.04702	top of input	64	4.13e-12 (ks)
GO:0009925	basal plasma membrane	5.97917	top of input	29	0.0080 (afc)
GO:0097209	epidermal lamellar body	5.83055	top of input	4	0.0093 (afc)
GO:0000794	condensed nuclear chromosome	5.66137	bottom of input	96	0.00092 (afc)
GO:0030057	desmosome	5.59763	top of input	25	1.63e-06 (afc)

(more ...)

Reference publications

publication	(year) title	enrichment score	direction	pathway size	false discovery rate
PMID:27426474	(2017) Clinical, microscopic and microbial characterization of exfoliative superficial pyoderma-associated epid...	9.43445	top of input	4	0.00018 (afc)
PMID:25496350	(2015) Expression patterns of superficial epidermal adhesion molecules in an experimental dog model of acut...	9.43445	top of input	4	0.00018 (afc)
PMID:24324345	(2013) Epidemiology of 'fragile skin': results from a survey of different skin types.	9.43445	top of input	3	0.00018 (afc)
PMID:23810772	(2013) Aberrant distribution patterns of corneodesmosomal components of tape-stripped corneocytes in atopi...	9.43445	top of input	3	0.00018 (afc)
PMID:23378711	(2012) Surgical Therapy by Sandwich Transplantation using a Dermal Collagen-Elastin Matrix and Full Thickne...	9.43445	top of input	4	0.00018 (afc)

(more ...)

local STRING network cluster

Additional reading

- Network Science - A textbook on graph theory and network analysis.
- Communication dynamics in complex brain networks - Whether and how network topology may be applied to study the brain networks.
- A Systematic Evaluation of Methods for Tailoring Genome-Scale Metabolic Models - General review and discussion on methods to use in genome-scale metabolic models.
- Analysis of Biological Networks - Introduction into biological networks, network notation, and analysis, including graph theory.
- Multi-omics approaches to disease - Introduction to how integrative approaches may be applied in disease
- Analysis of Biological Networks - Introduction into biological networks, network notation, and analysis, including graph theory.
- Using graph theory to analyze biological networks - overview of the usage of graph theory in biological network analysis
- Survival of the sparsest: robust gene networks are parsimonious - analysis of network complexity and robustness.
- Network biology: understanding the cell's functional organization - Overview of key concepts in biological network structure
- Graph Theory and Networks in Biology - extended perspective on how graph analysis is applied in biology
- Scale free networks are rare
- Modularity and community structure in networks

Additional references displayed as hyperlinks in each slide.