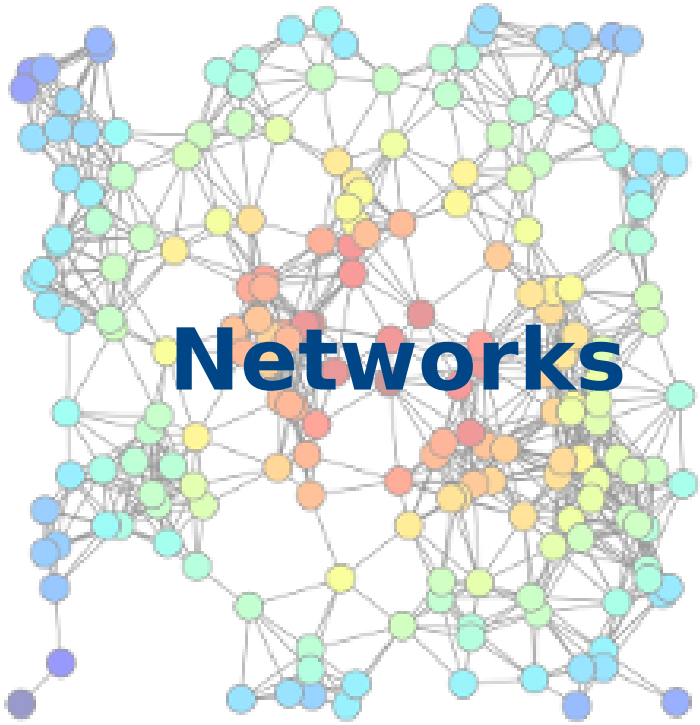
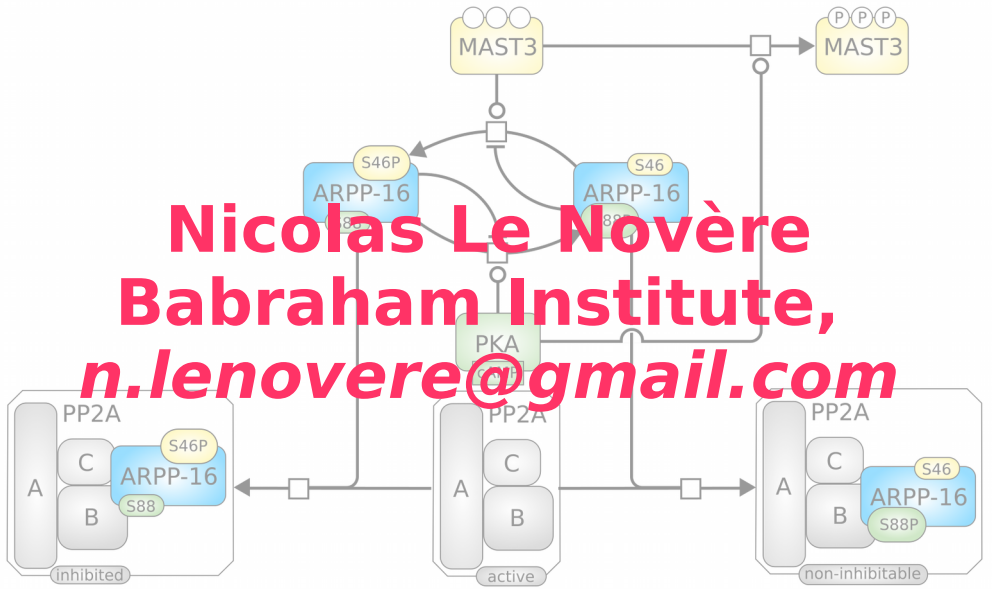


Networks



Nicolas Le Novère
Babraham Institute,
n.lenovere@gmail.com

Babraham Institute

Who am I? - Education

Scientific baccalaureate at the *Prytanée National Militaire*

"I was in one of the most famous schools of Europe"
René Descartes – Discours de la méthode

Preparatory classes for the French *Grandes Écoles*

Mathematics	Physics Chemistry	Natural Sciences

ENS ÉCOLE NORMALE SUPÉRIEURE

Bachelor in Physiology and Cellular Biology
Master of Biology-Biochemistry

UNIVERSITÉ PARIS SUD Evolution	Molecular Biophysics	UPMC SORBONNE UNIVERSITÉS Neuroendocrinology

Babraham Institute

Who am I? - Research

Paris

1992 Institut Pasteur 1999

PhD
Nicotinic receptors:
Bioinformatics,
Neuroanatomy
Behaviour

2001 Institut Pasteur 2003

CNRS Researcher
Nicotinic receptors:
Bioinformatics

"A tale of two cities"

2014-2015
20% institutCurie

Chief data officer

1999 2001

EMBO post-doc
Bacterial chemotaxis:
Bioinformatics,
Mathematical Modelling

2003 EMBL-EBI 2012

Group leader
Systems Biology
Synaptic signalling:
Modelling, knowledge
Representation, databases

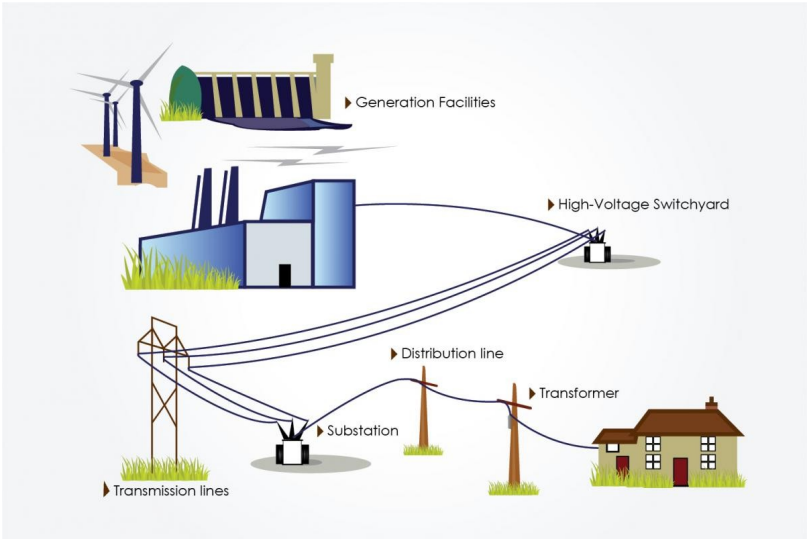
2012- Babraham Institute

Senior group leader
Signalling 50%
Epigenetic 30%
Lymphocytes 15%
Nuclear Dynamics 5%

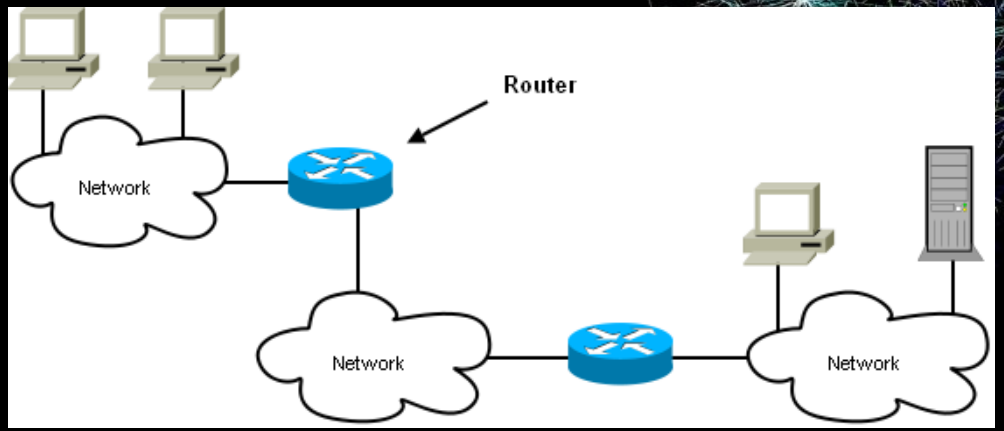
Cambridge

Networks are everywhere

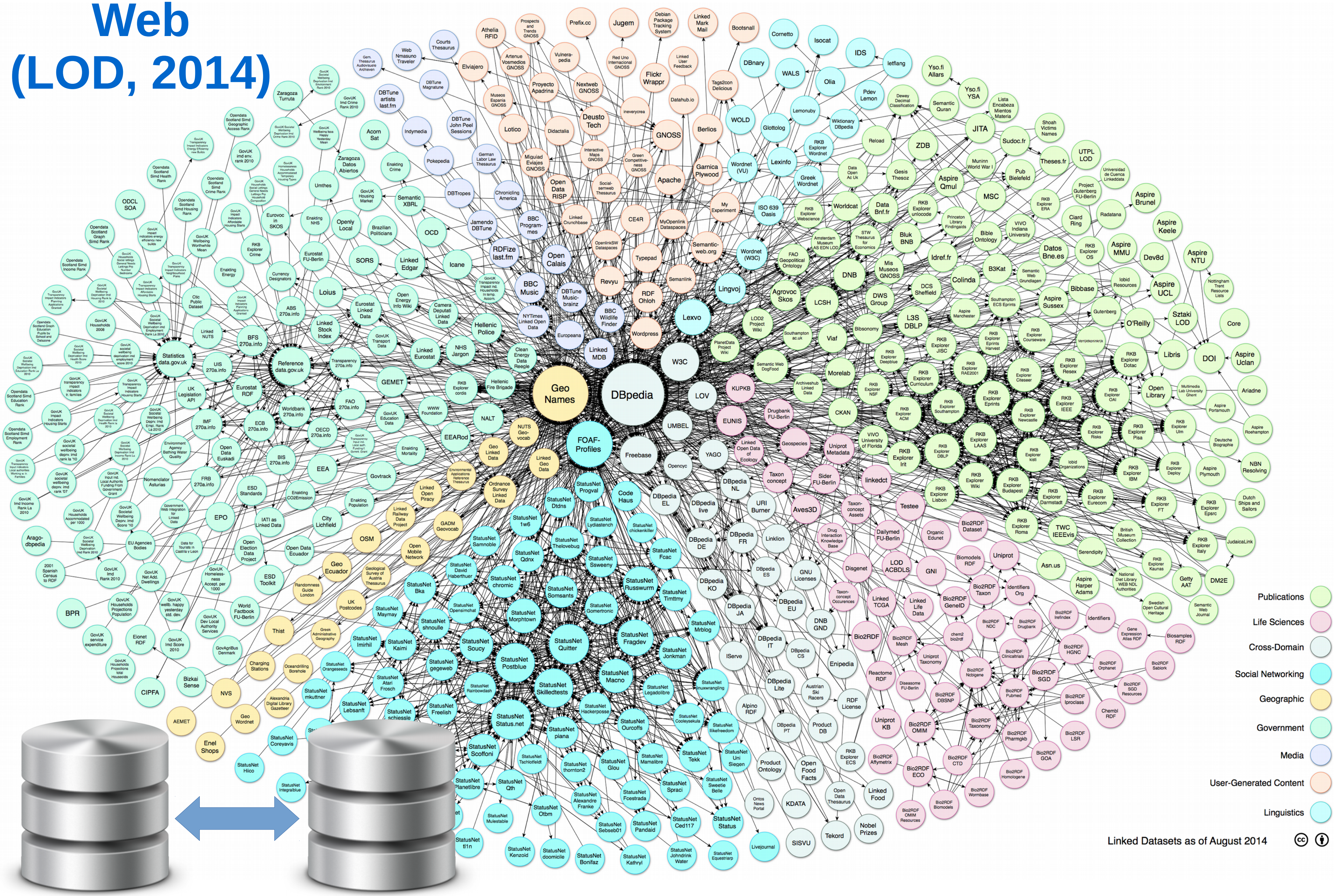
Power grid (detail, US)

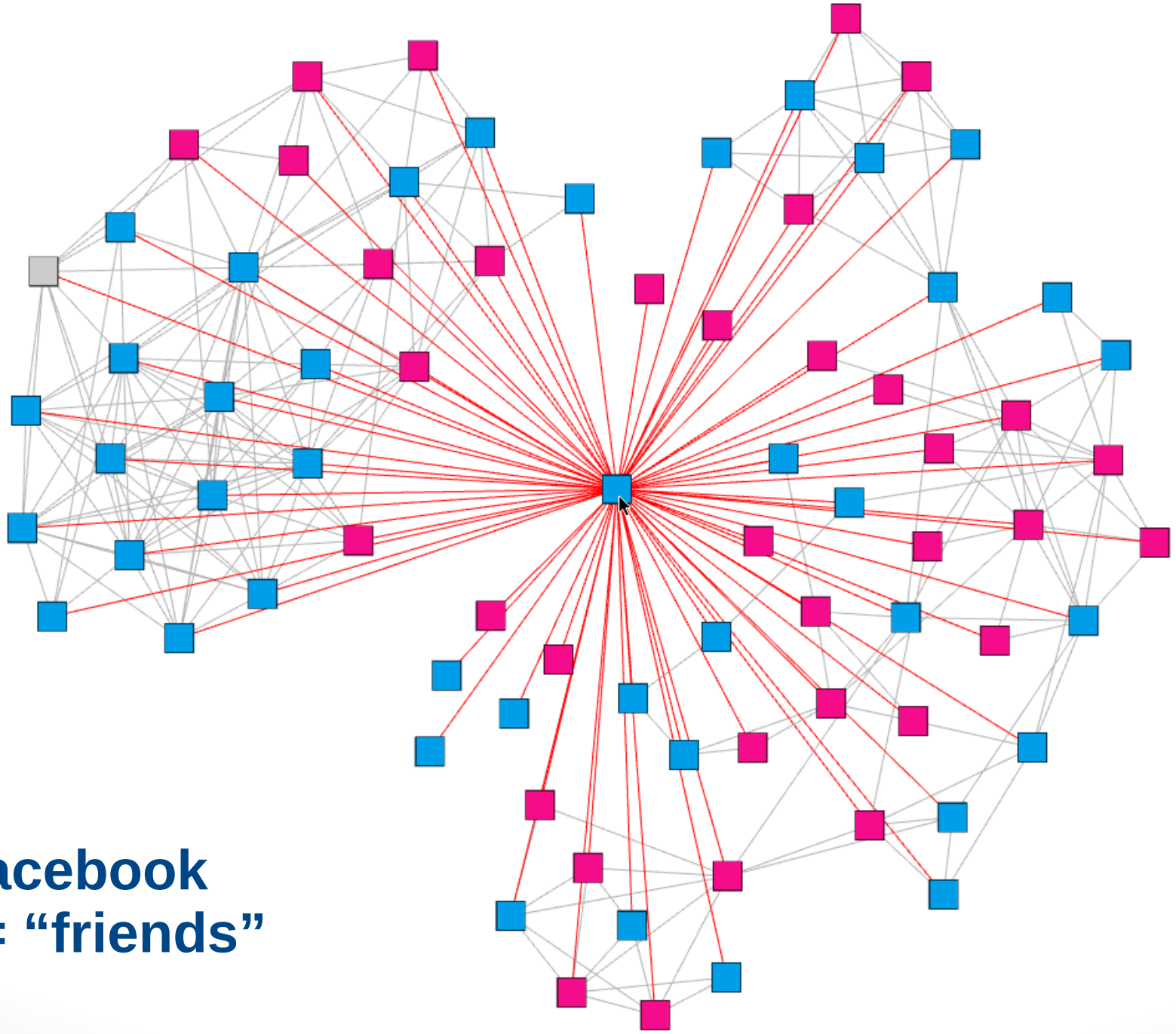


Internet (portion, 2005)



Web (LOD, 2014)

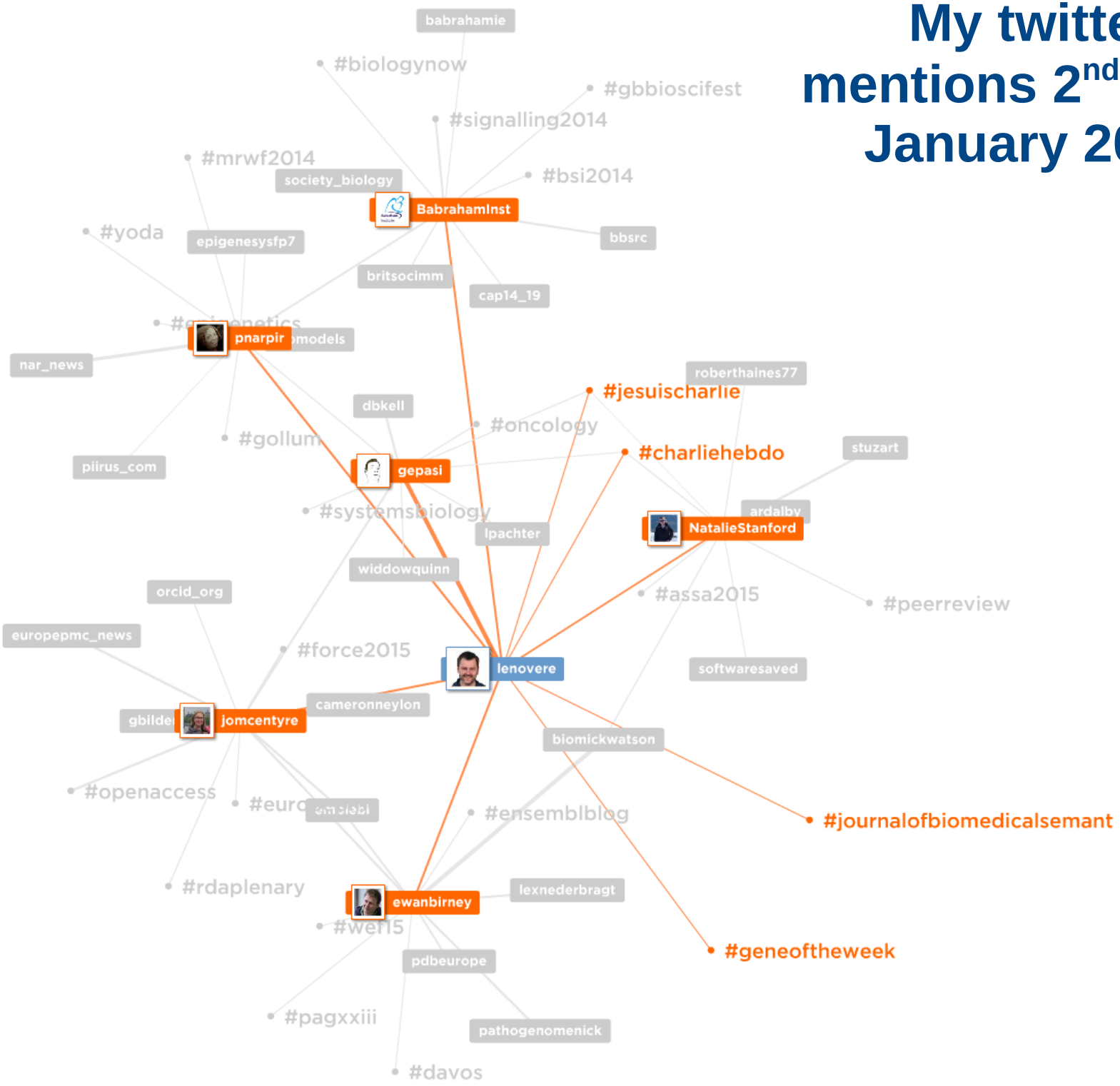

Geo Names
DBpedia
FOAF-Profiles
W3C
Lexvo
LOV
UMBEL
Freebase
Reference data.gov.uk
Statistics data.gov.uk
Publications
Life Sciences
Cross-Domain
Social Networking
Geographic
Government
Media
User-Generated Content
Linguistics
Linked Datasets as of August 2014



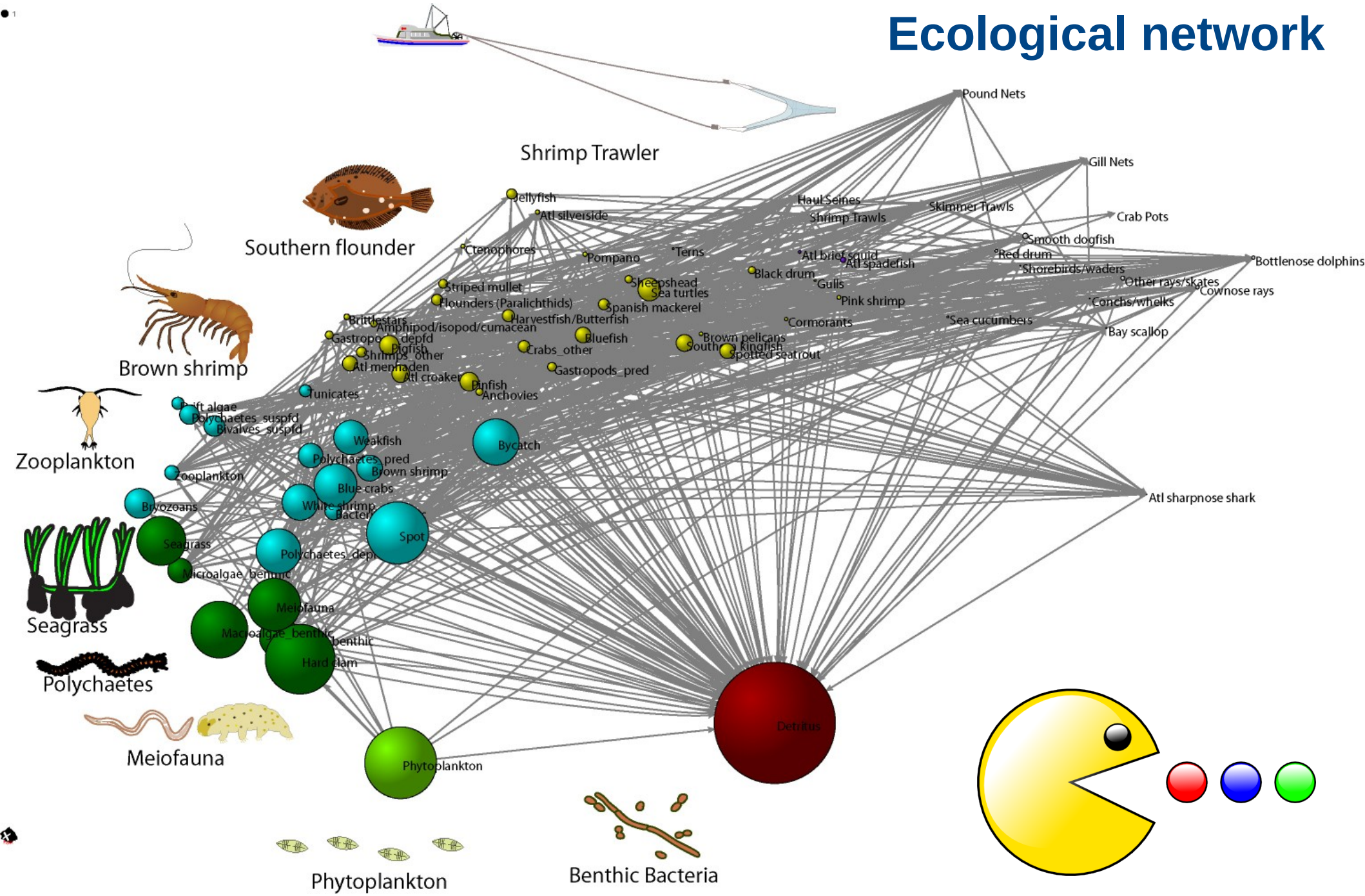
My Facebook nodes = “friends”

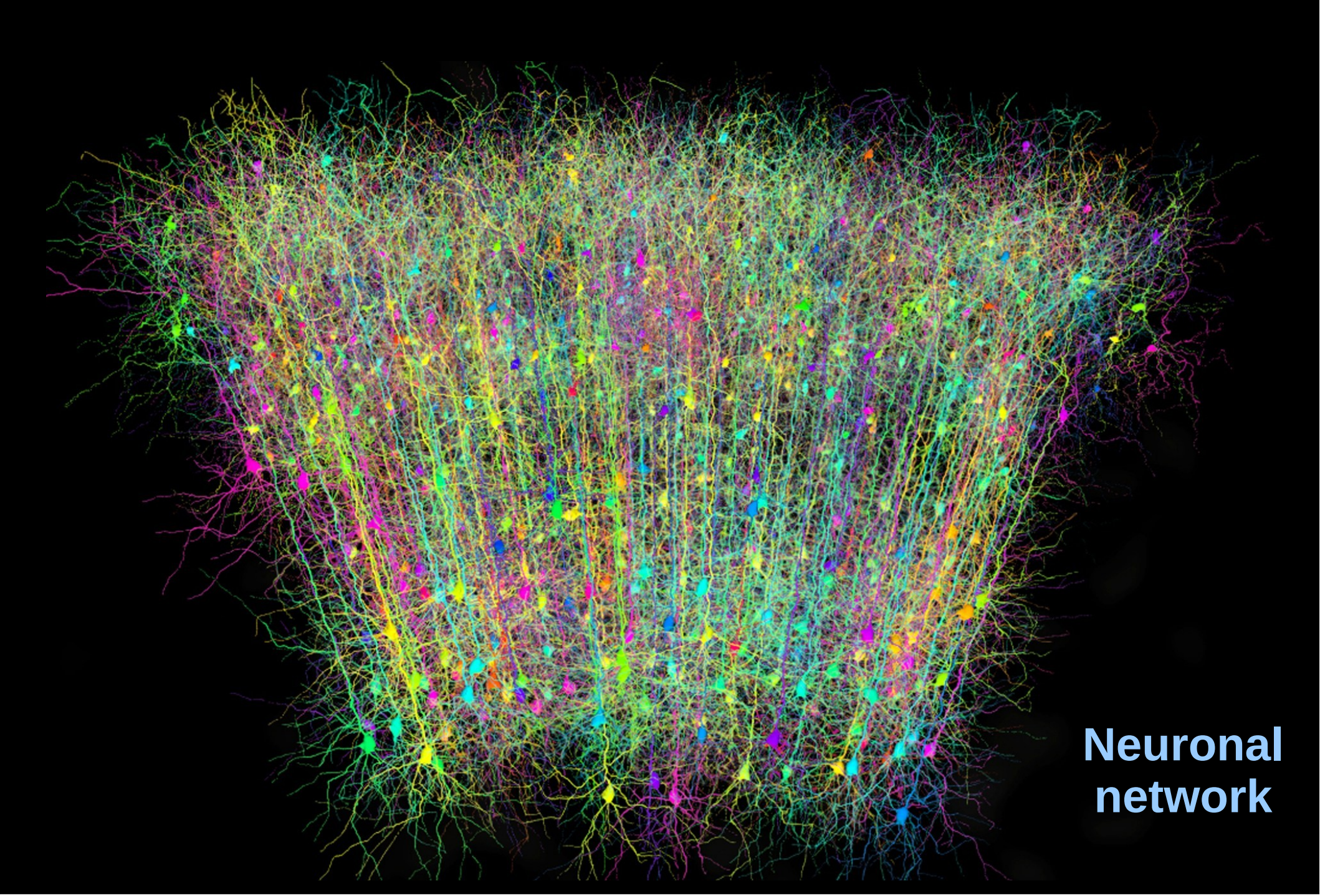

HERO CORP

My twitter mentions 2nd week January 2015



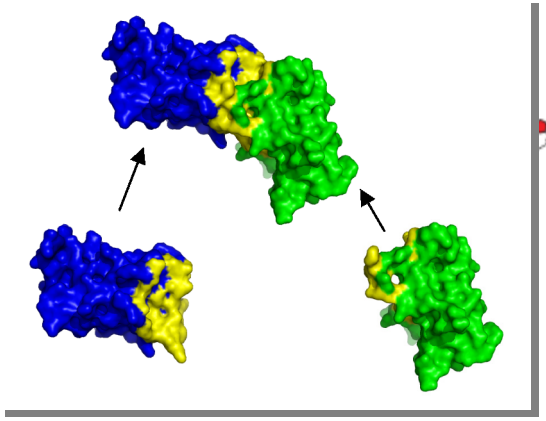
Ecological network



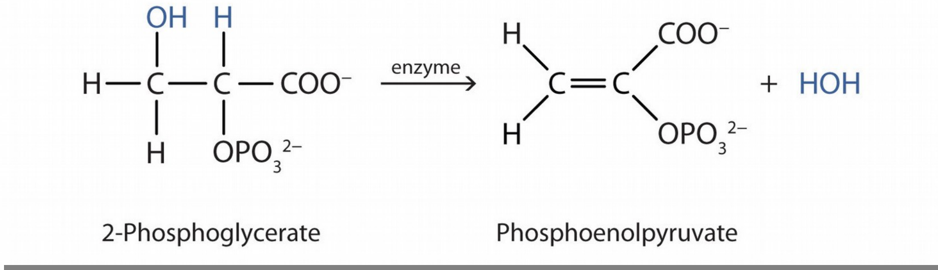


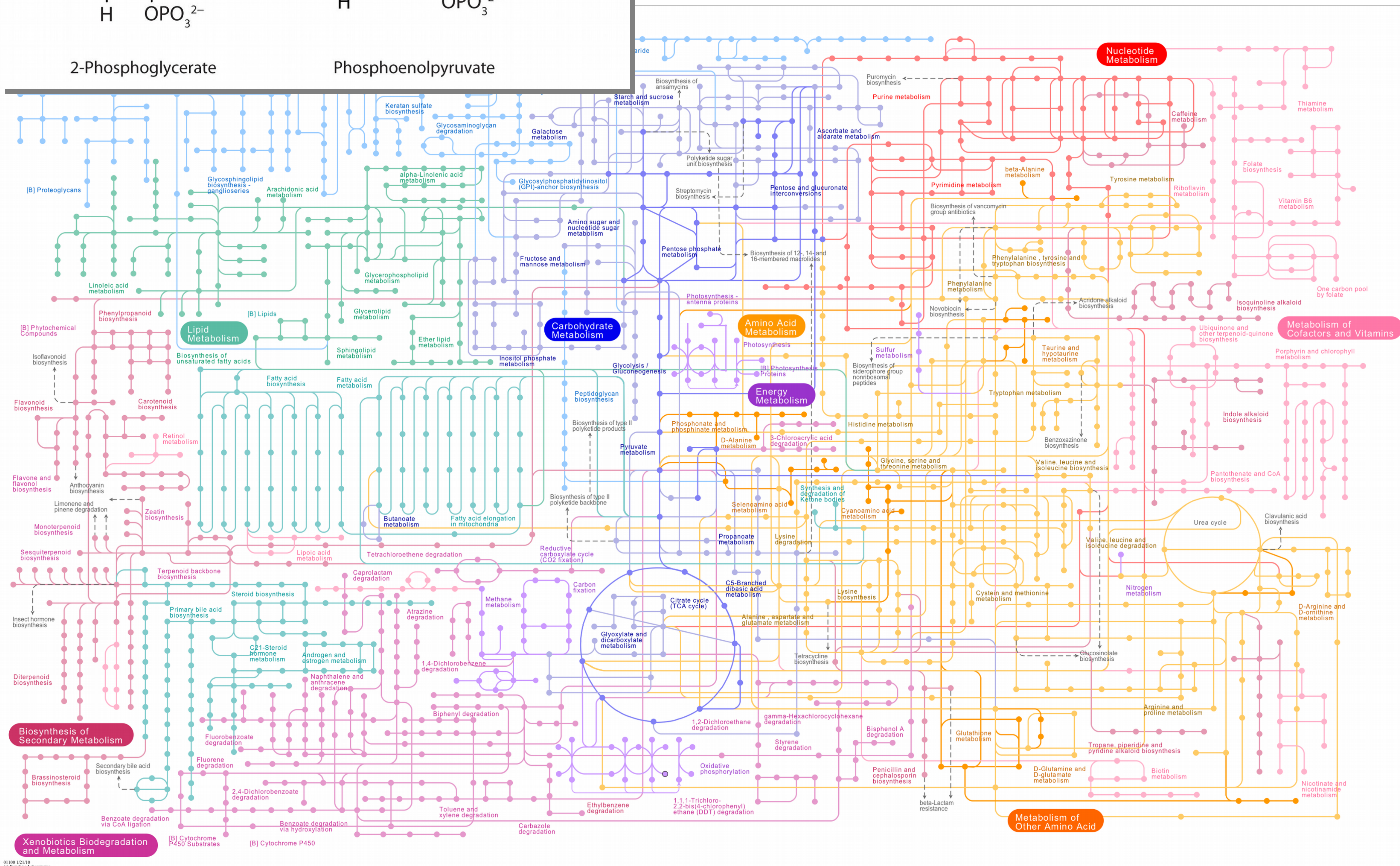
Neuronal network

Protein interactome (yeast)

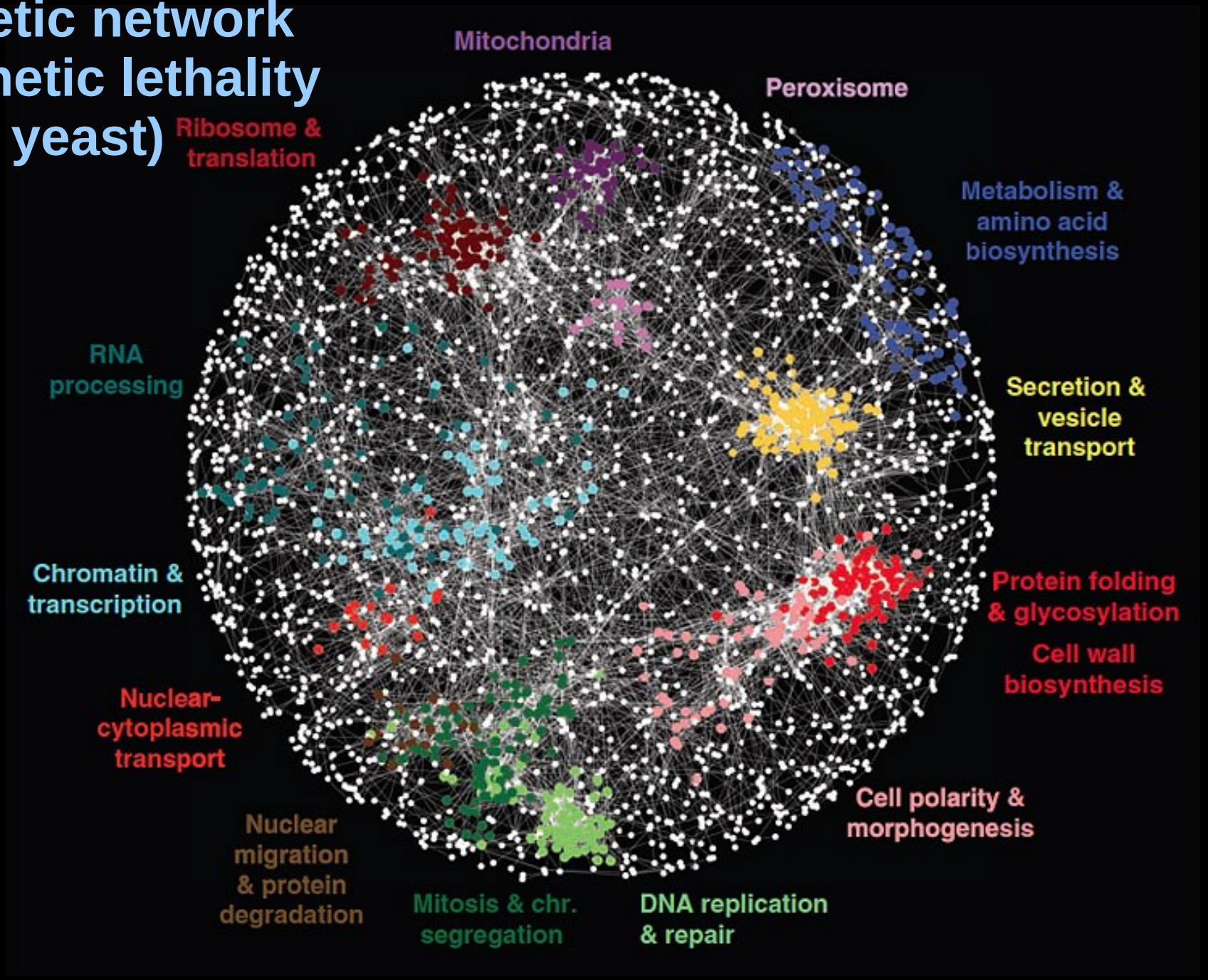


Metabolic network



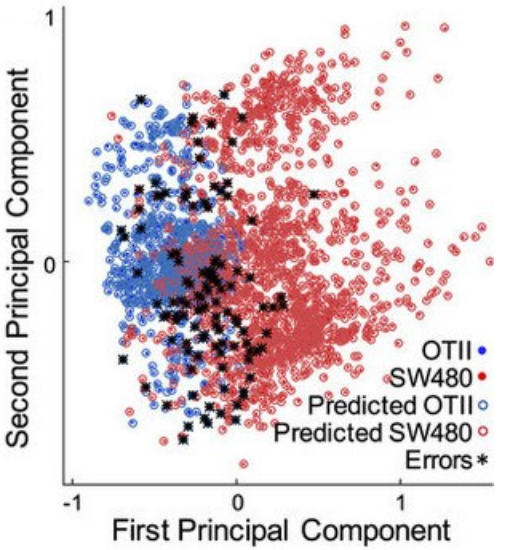


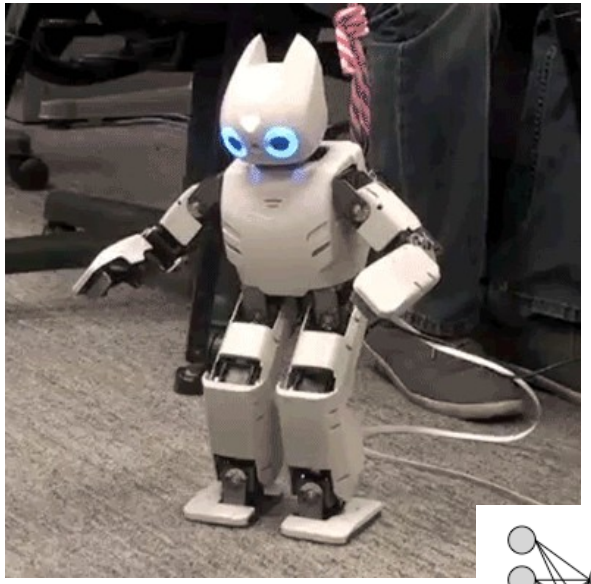
Genetic network (synthetic lethality yeast)

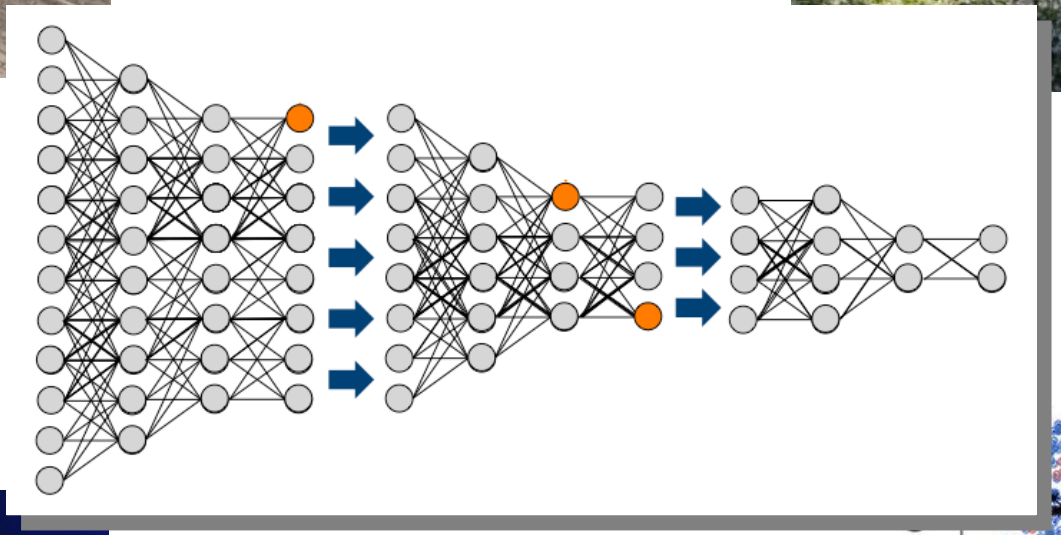


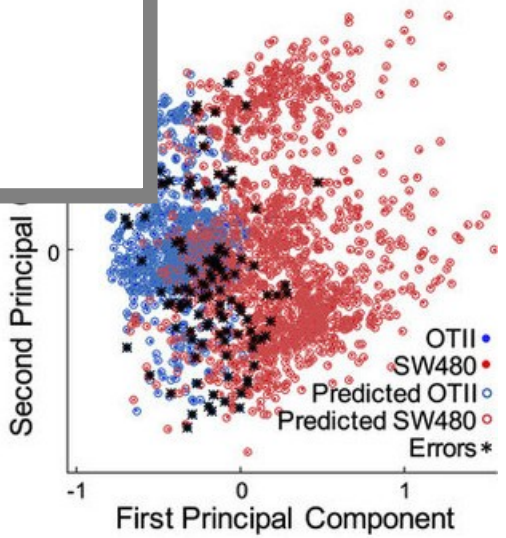

Google


AlphaGo
Lee Sedol


Second Principal Component
First Principal Component
1
0
-1
0
1
OTII
SW480
Predicted OTII
Predicted SW480
Errors







Babraham Institute

Reconstructing networks

Pre-existing knowledge

http://www.ebi.ac.uk/IntAct

Search: Q01097 Search Clear Show Advanced Fields » MIQL syntax reference

- Free text search will look by default for interactor identifier, species, interaction id, detection method, interaction type, publication identifier or author
- For a more specific search, use MIQL syntax or advanced search
- Search based on exact word matches eg. BRCA2 will not match BRCA2B
- Search for isoforms of 'P12345' by using 'P12345*'

- Home
- Advanced Search
- Tools
- Data Submission
- Downloads
- Documentation
- Acknowledgements
- Contact Us

FOLLOW US ON twitter

about 8 days ago IntAct release 158/159! 302465 binary interactions from 5787 manually curated publications ftp://ftp.ebi.ac.uk/pub/databases/intact/current

about 13 days ago IntAct is now available for download in PSI MITAB2.7 in addition to PSI-MI XML, MITAB2.5 and MITAB2.6.

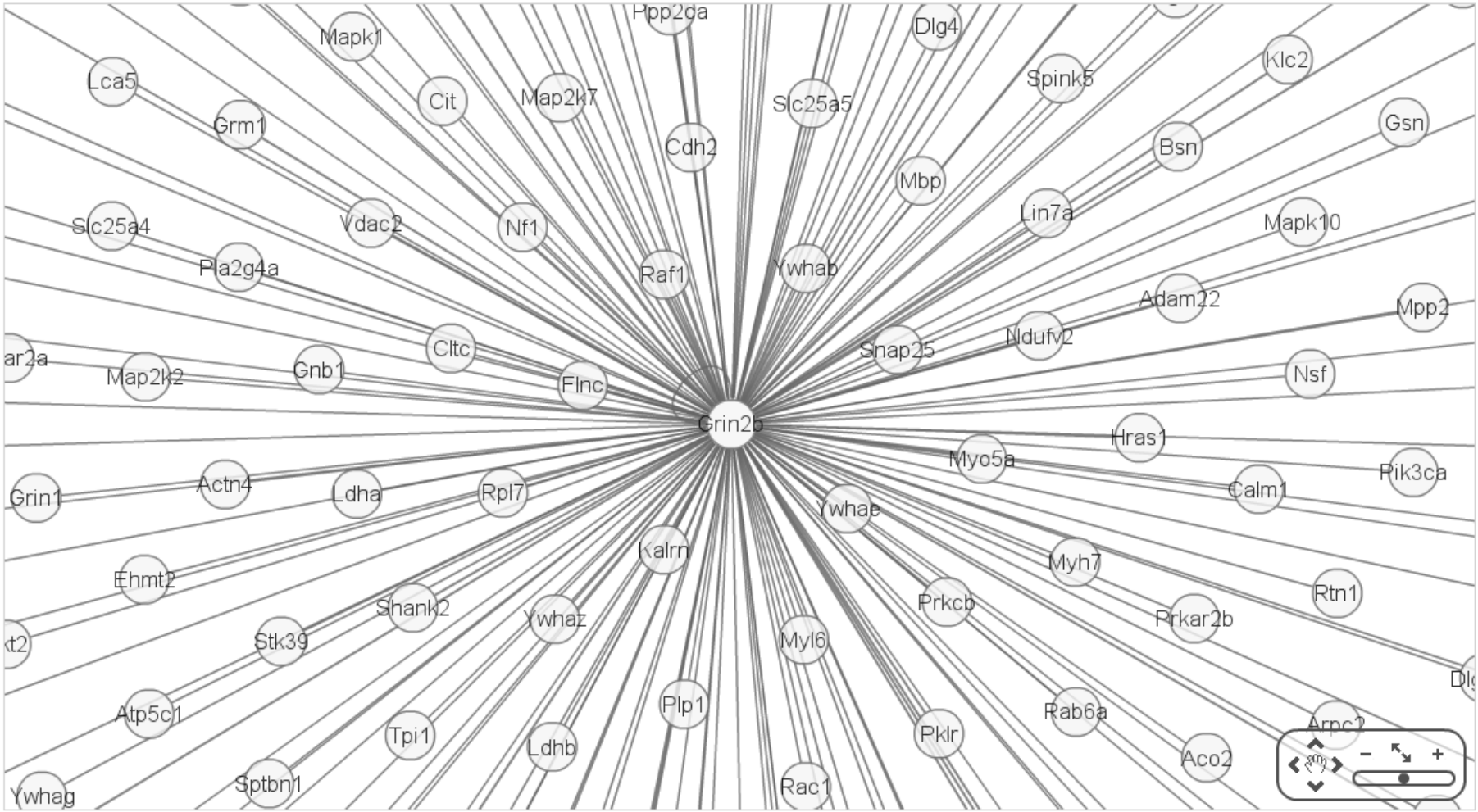
about 13 days ago Tarassov et al (PMID:18467557) low confidence interactors withdrawn in consultation with the authors

about 21 days ago Dataset of the month (November). Blandin et al. A Human Skeletal Muscle Interactome Centered on Proteins involved in Muscular Dystrophies.

about 34 days ago Now IntAct can

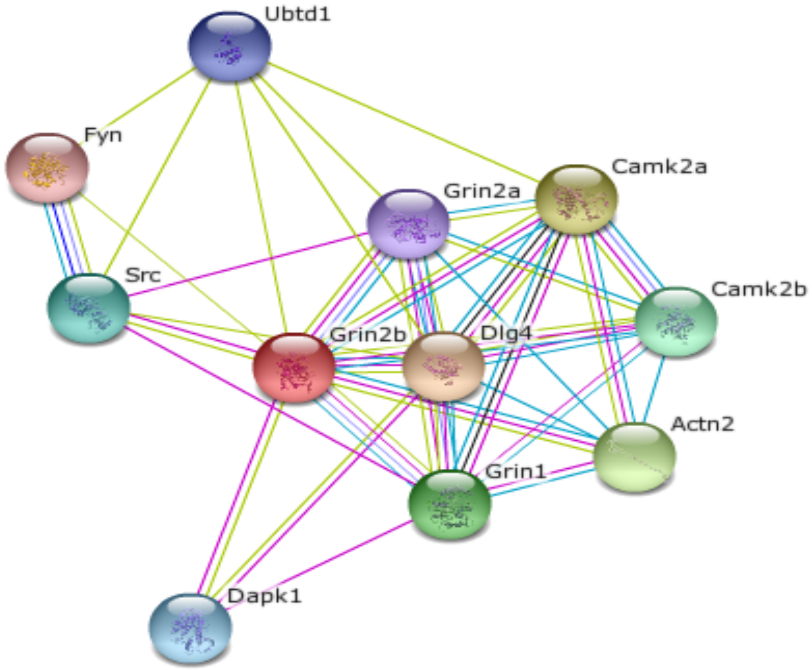
Home | Search | Interactions (254) | Browse | Lists | Interaction Details | Molecule View | Graph

Network visualisation



http://string-db.org/

Home · Download · Help/Info STRING 9.0



This is the **evidence view**. Different line colors represent the types of evidence for the association.

confidence evidence actions interactive advanced more less save

(requires Flash player 10 or better)

Your Input:

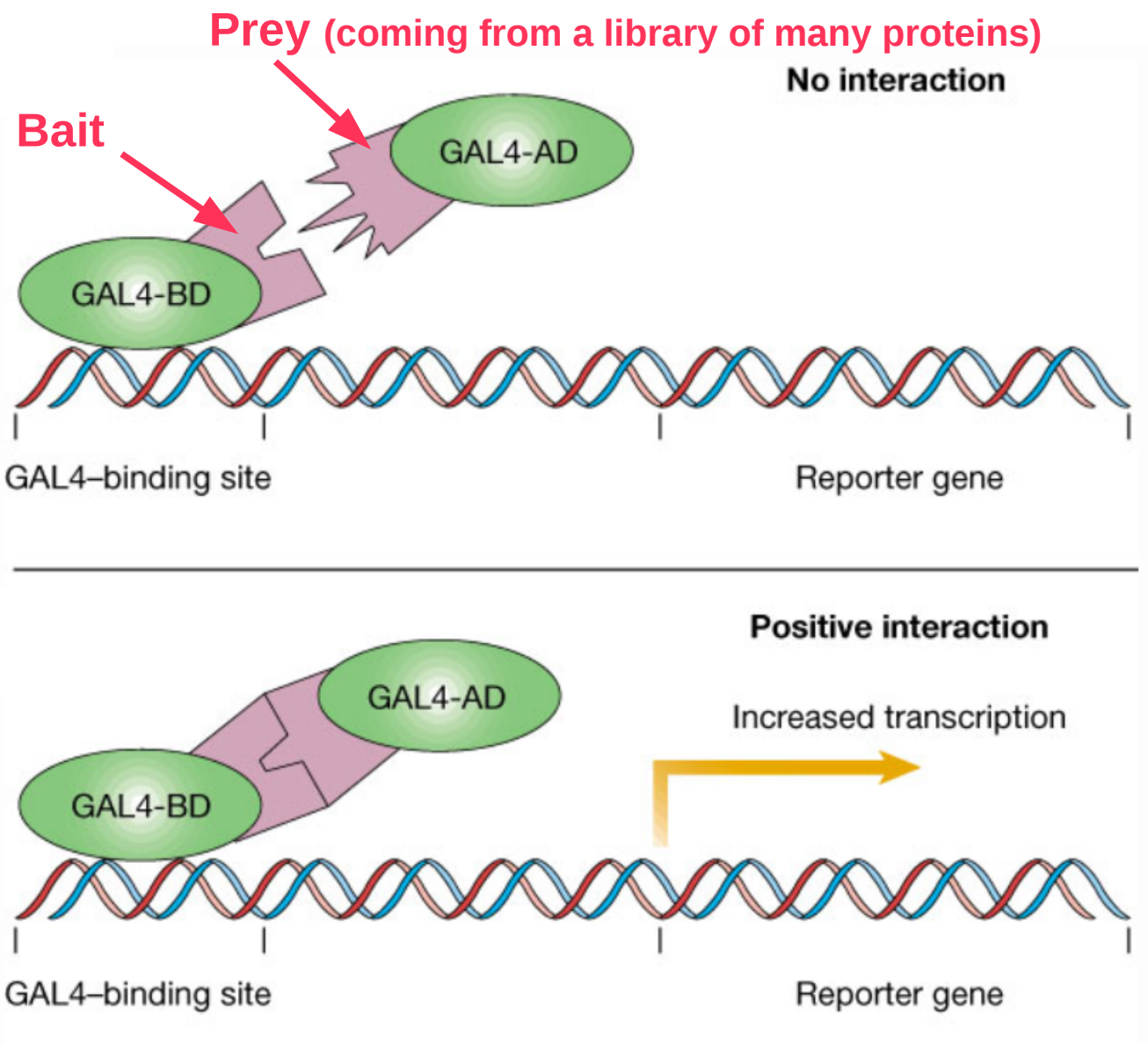
Grin2b glutamate receptor, ionotropic, NMDA2B (epsilon 2) Gene; NMDA receptor subtype of glutamate-gated ion channels with high calcium permeability and voltage-dependent sensitivity to magnesium. Mediated by glycine (1482 aa)
(Mus musculus)

Predicted Functional Partners:

Neighborhood Gene Fusion Cooccurrence Coexpression Experiments Databases Textmining [Homology] Score

Babraham Institute

observations: yeast two hybrid system



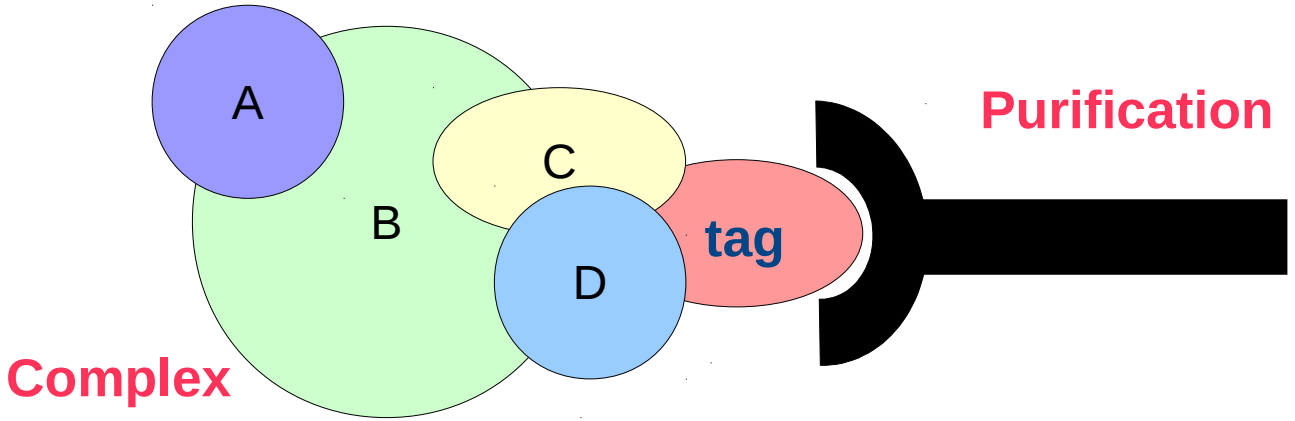
✓

Direct physical interactions

✗

All <u>possible</u> interactions (may never actually occur)

observations: protein complex precipitation

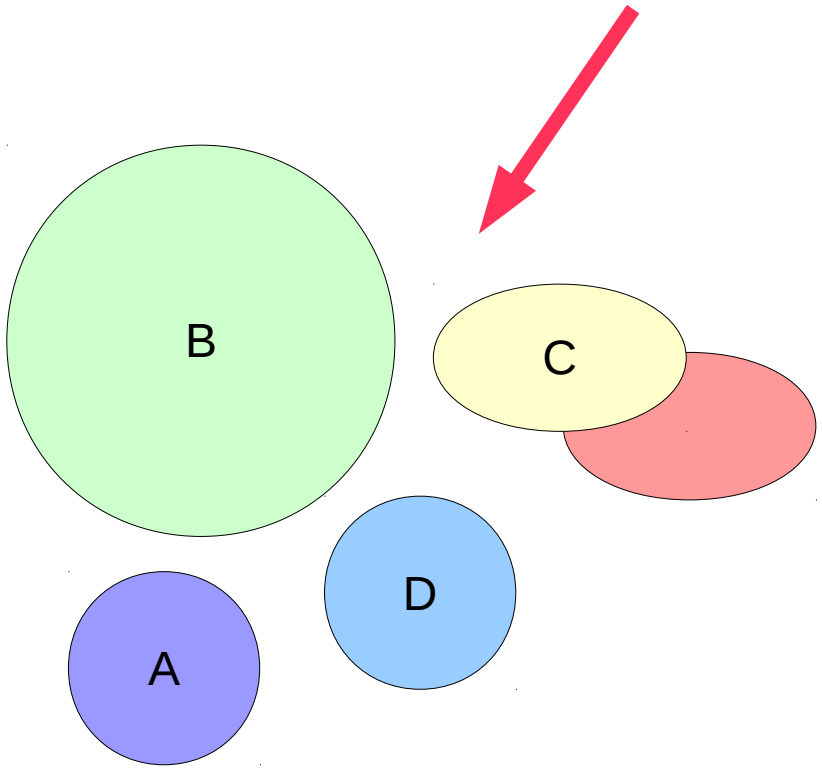


Purification

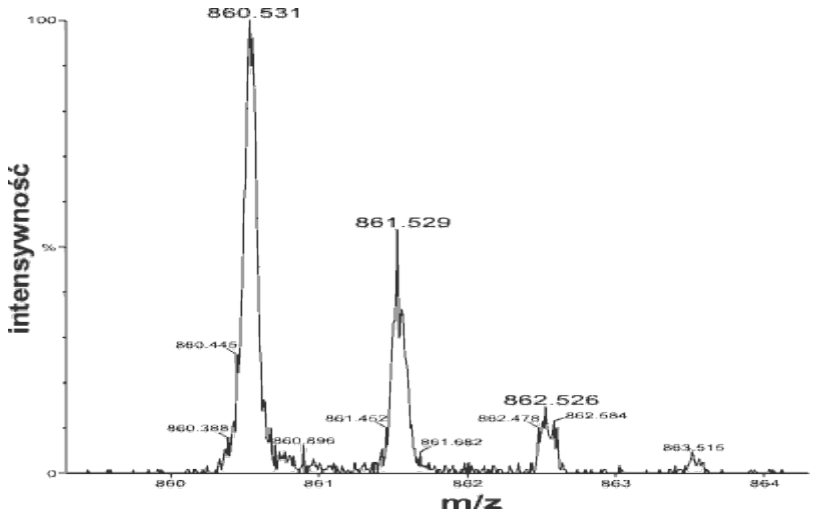
Complex

✓ actual complex

✗ not direct interactions (A-C and A-D wrongly predicted)

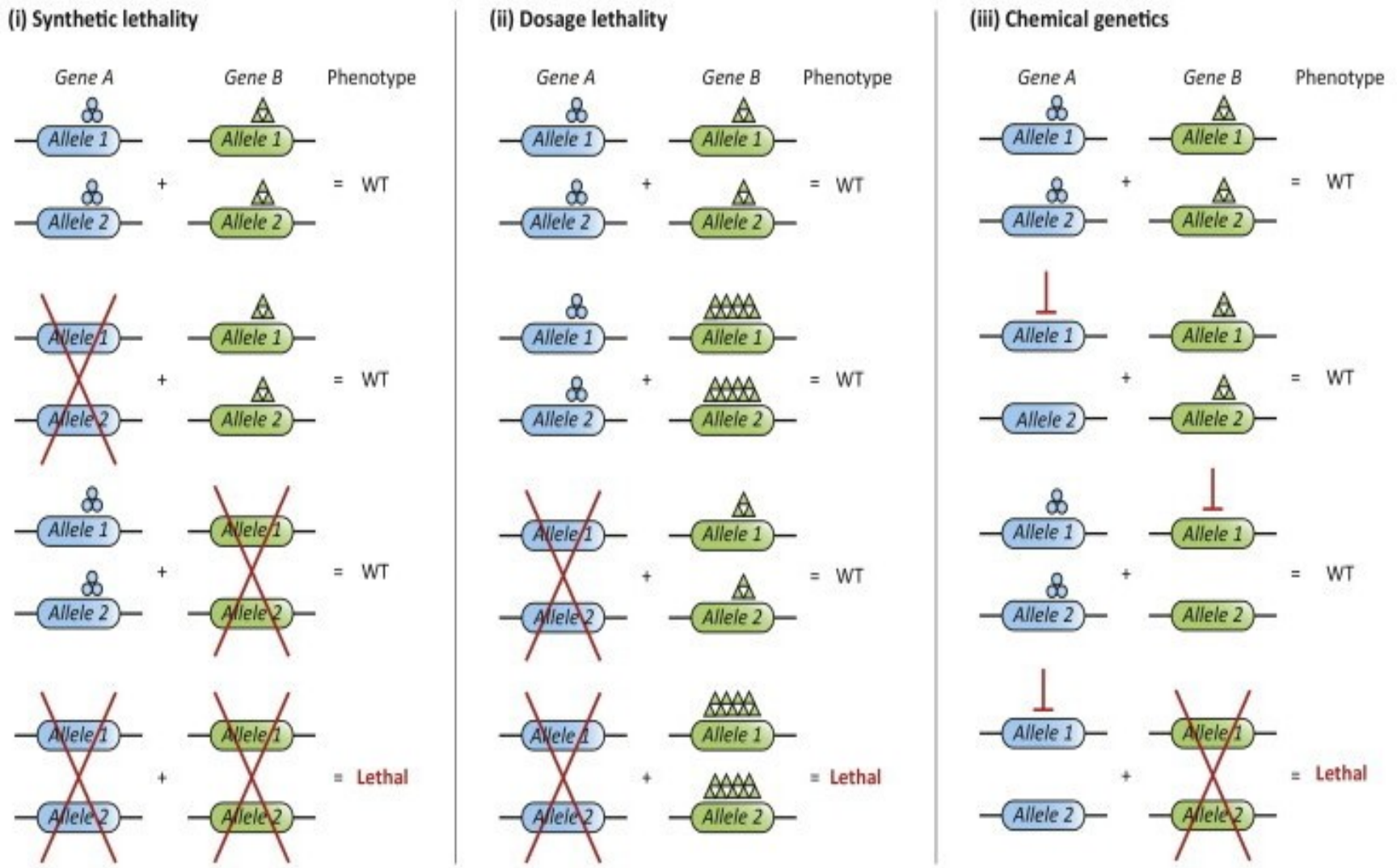


Identification

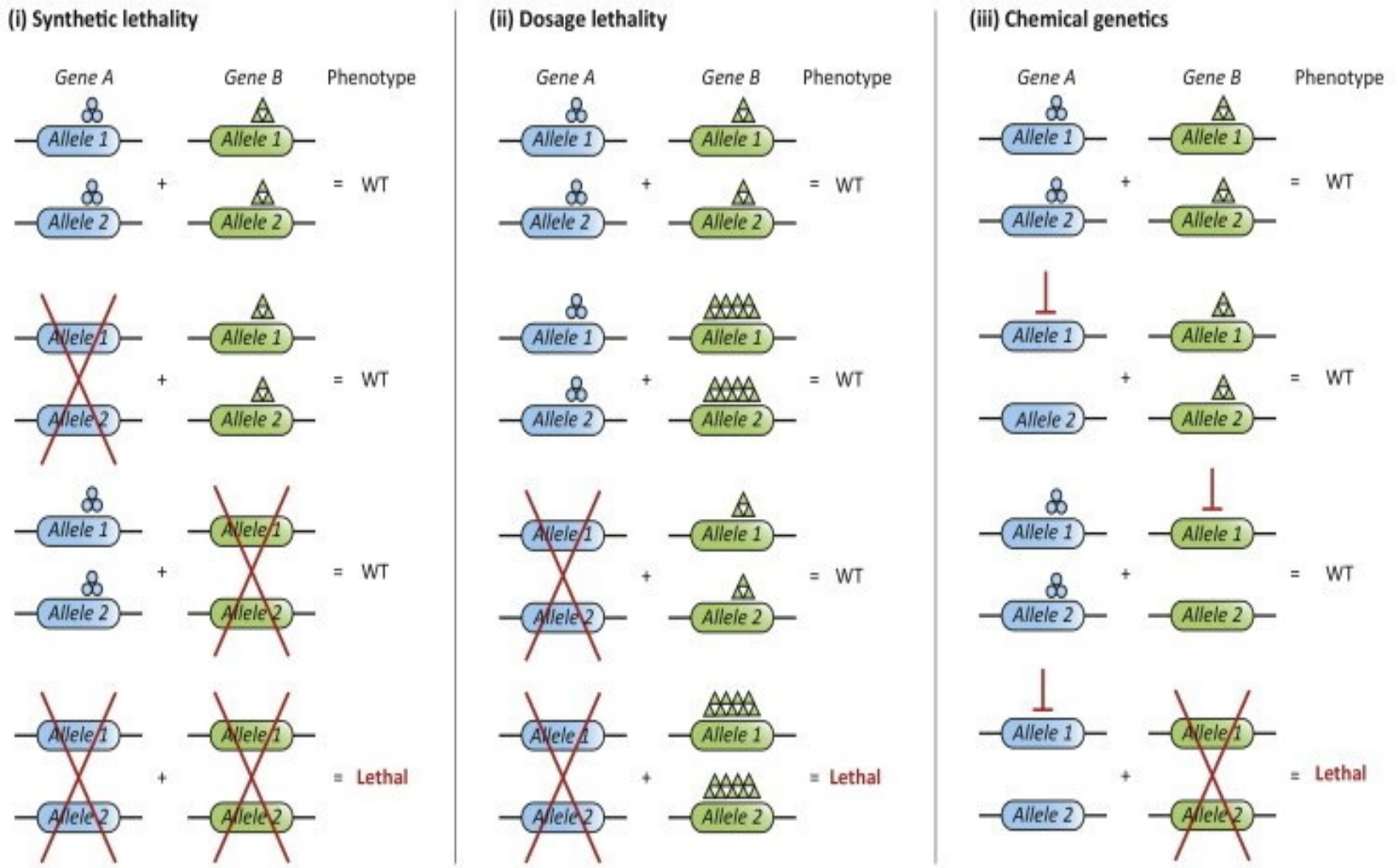


Babraham Institute

Synthetic lethality

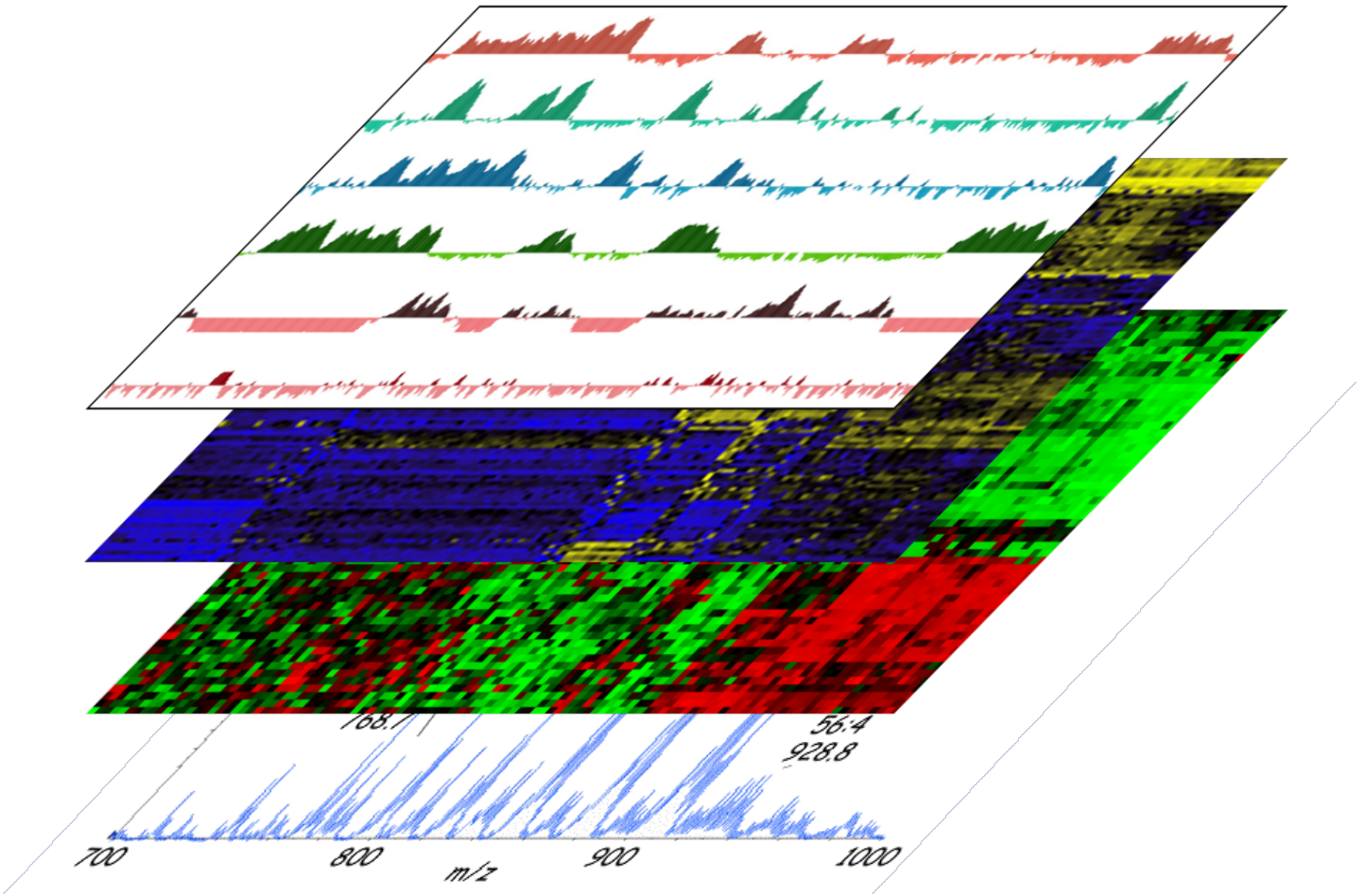


Synthetic lethality

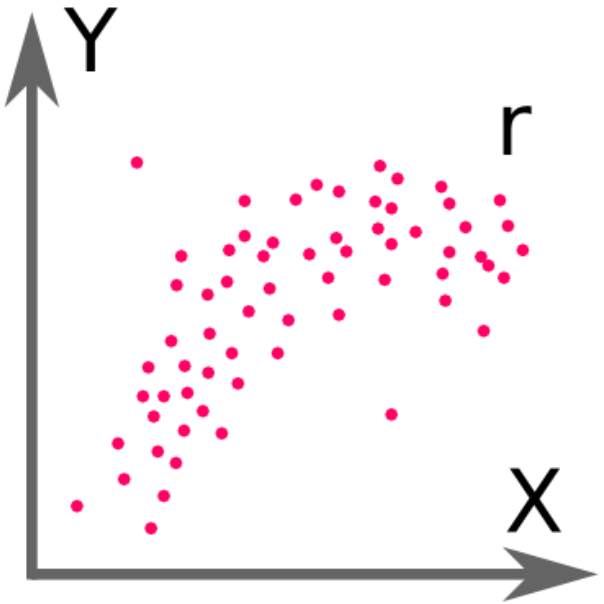


Can you think of a practical use of synthetic lethality in medicine?

network inference from data

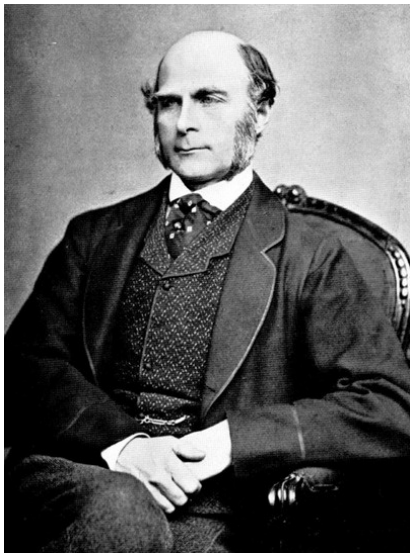


Network inference: Correlation



Values of X are somewhat related to values of Y

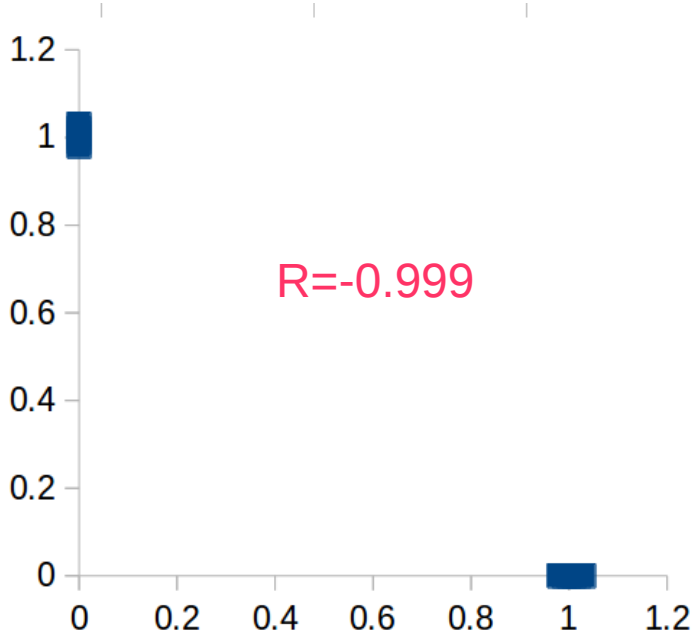
$$r = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2}\sqrt{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}}$$

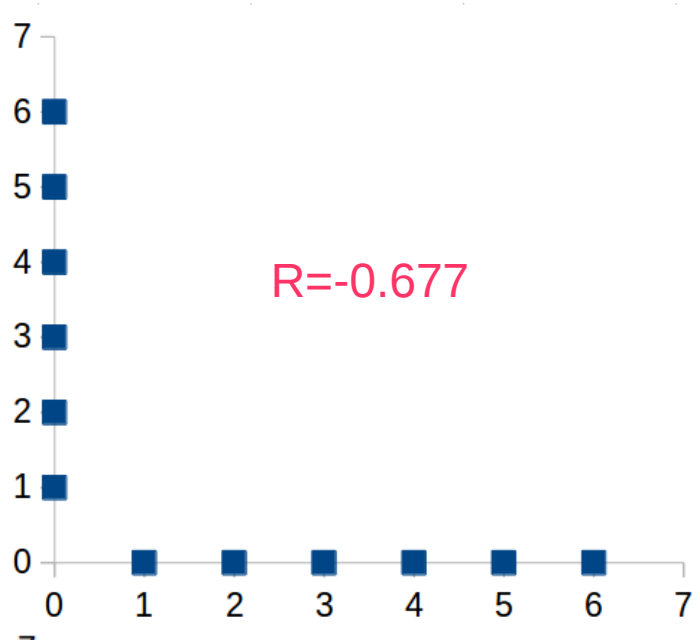


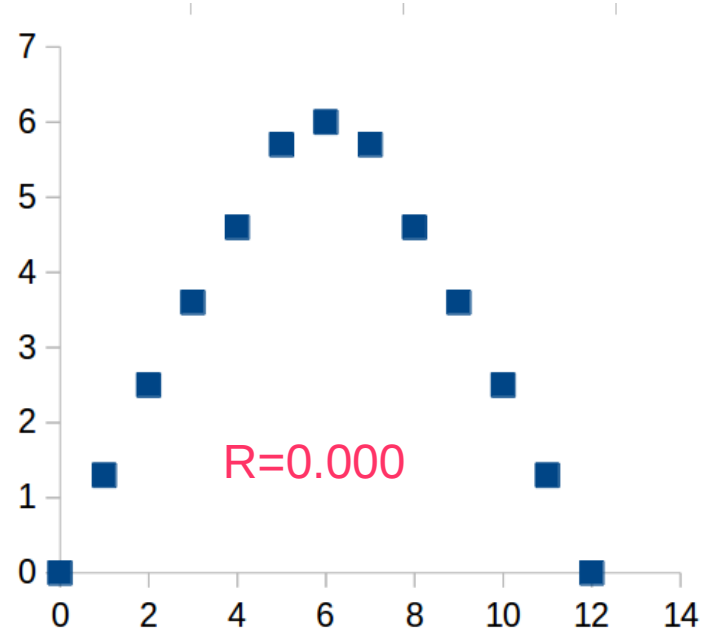
Francis Galton (studied at Trinity College, Camb)

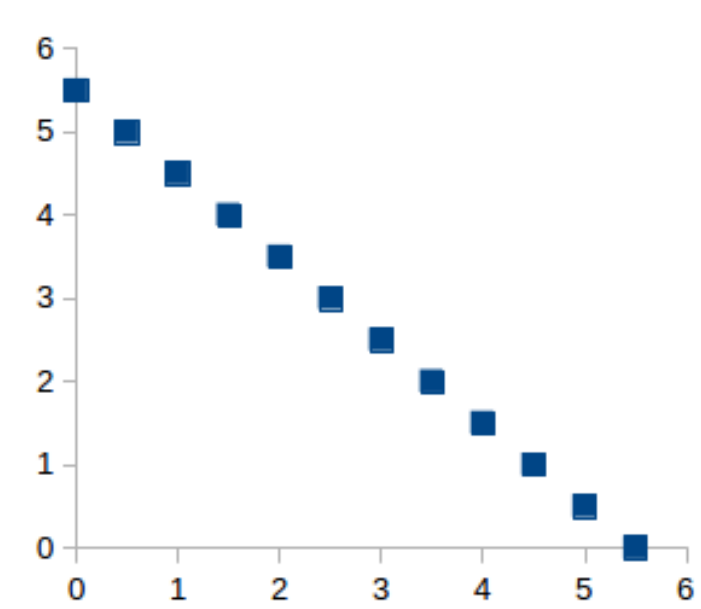
Karl Pearson (studied at King's College, Camb)

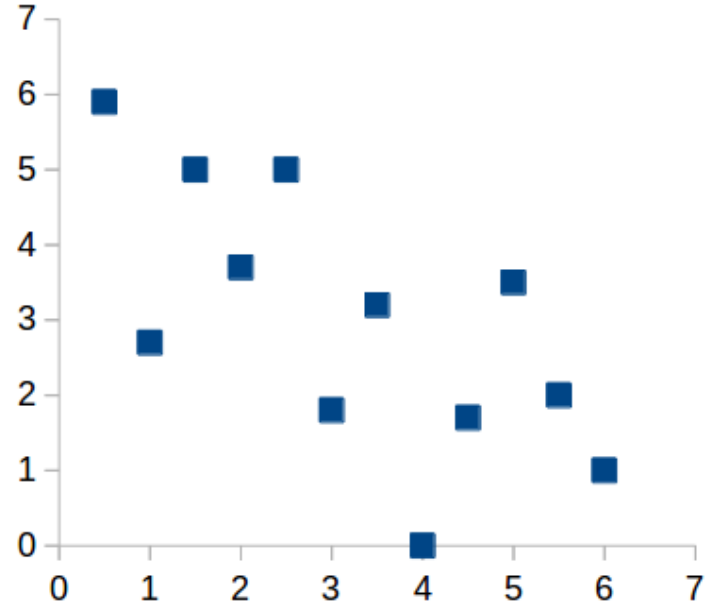
Correlations coefficients might be misleading

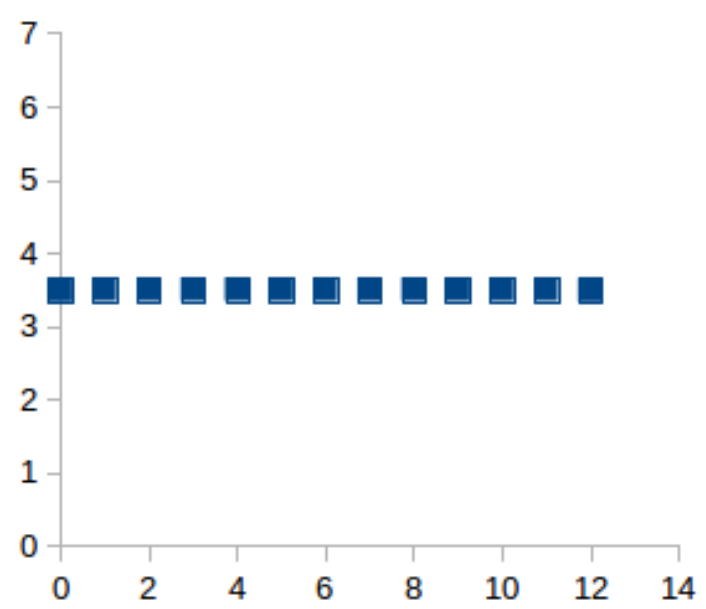






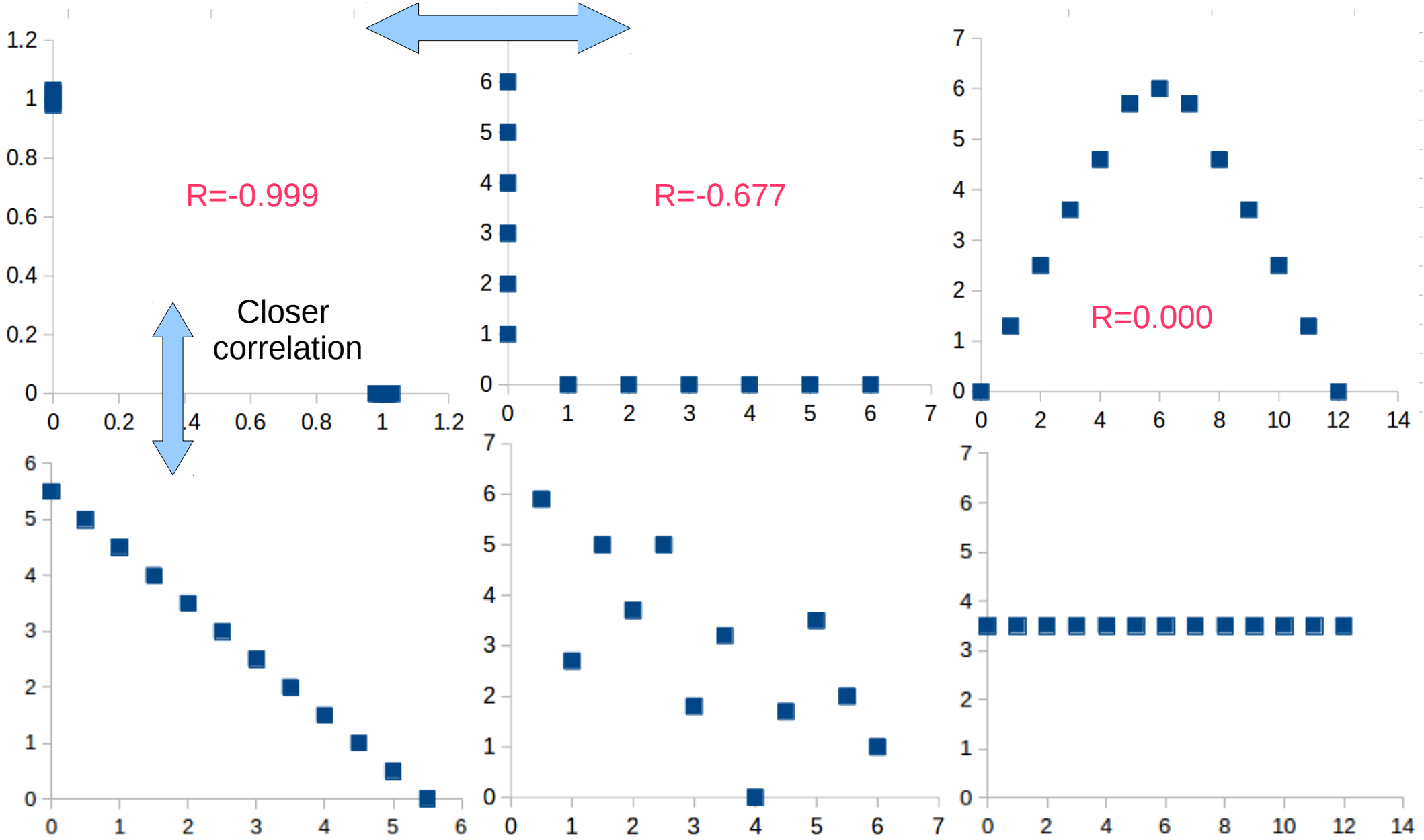




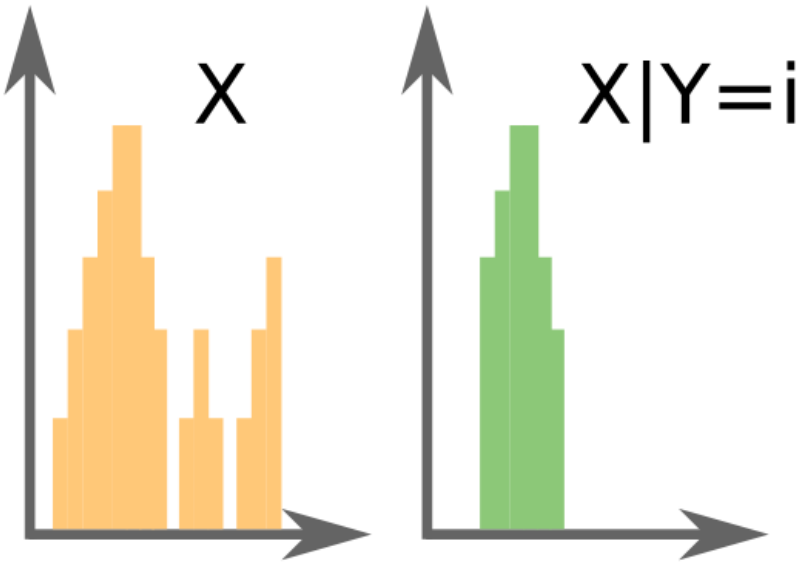


Correlations coefficients might be misleading

Closer biological meaning

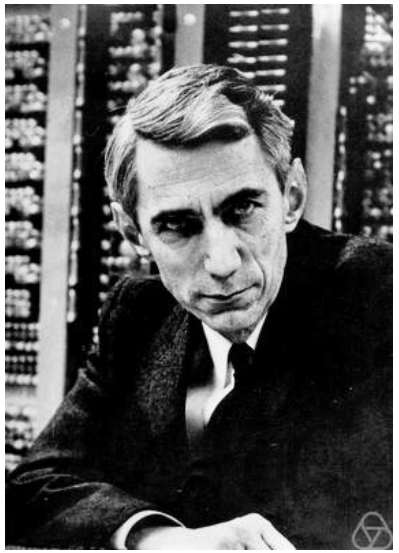


Network inference: Information Theory



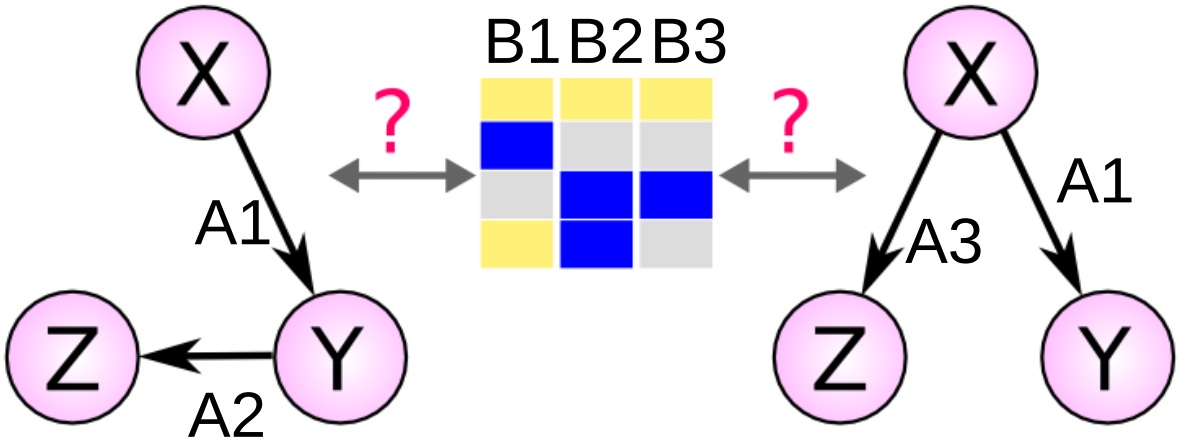
$$I(X,Y) = H(X) - H(X|Y)$$

E.g. mutual information:
knowledge of the value of Y reduces the uncertainty on the values of X



Claude Shannon

Network inference: Bayesian inference



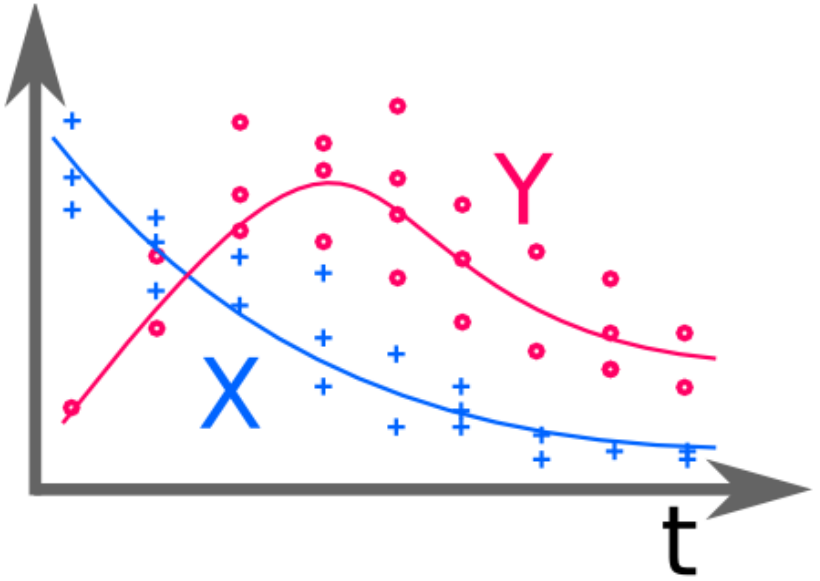
Which network most probably generated this dataset?

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Thomas Bayes

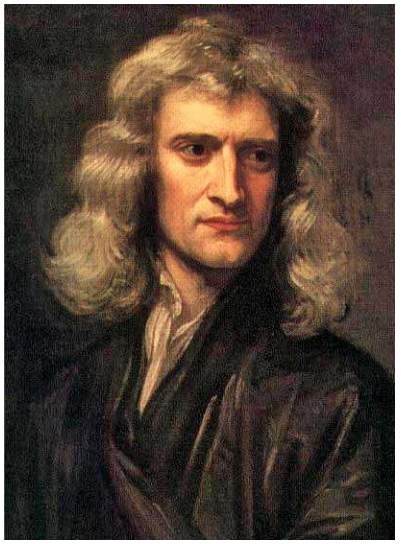
Pierre-Simon De Laplace

Network inference: Ordinary differential equations



The rate of change of X depends on the value of Y

$$\frac{dx_i}{dt} = \sum_{j=1}^{n} a_{i,j} x_j$$

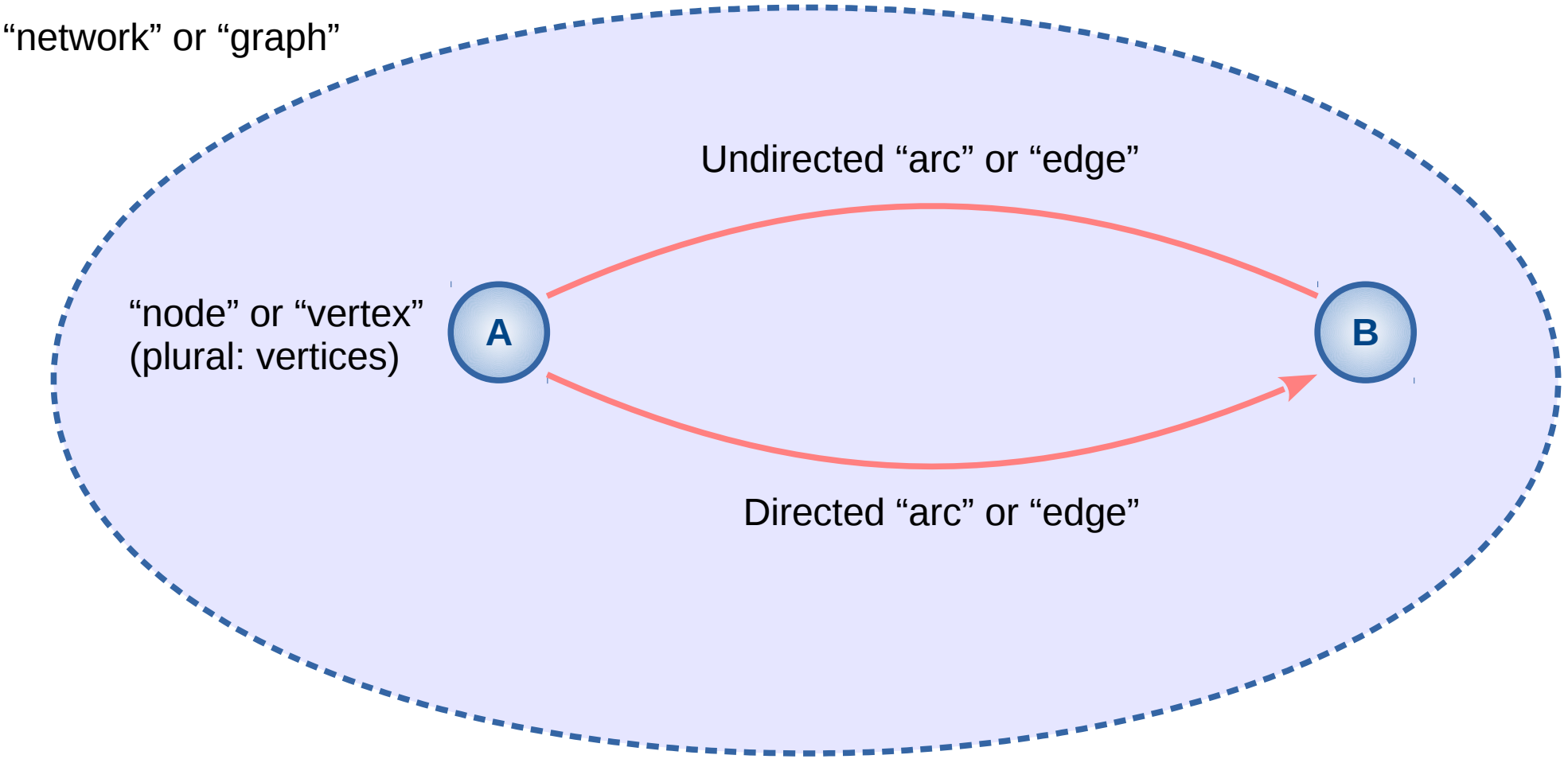


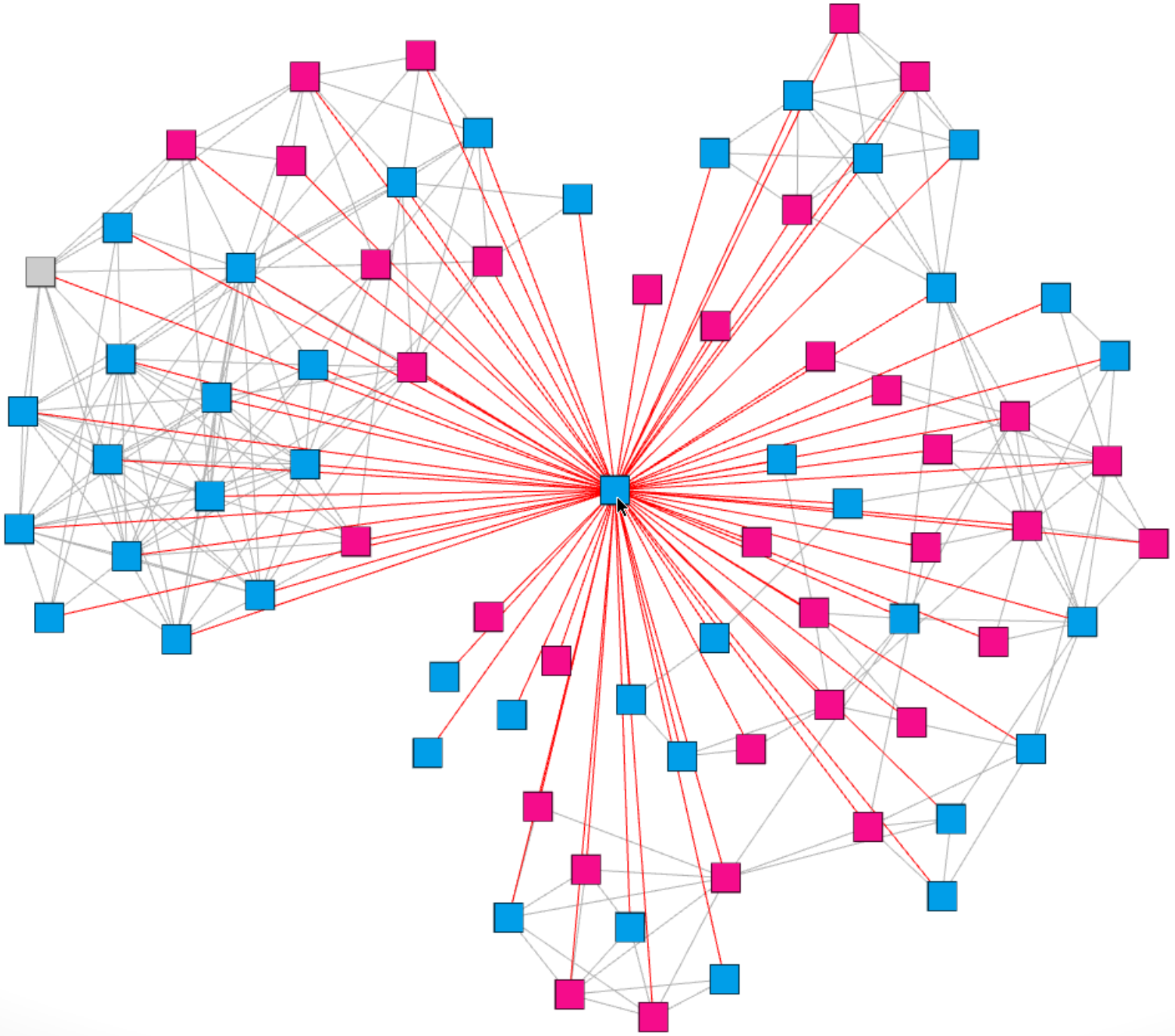
Isaac Newton (Trinity College, Camb)

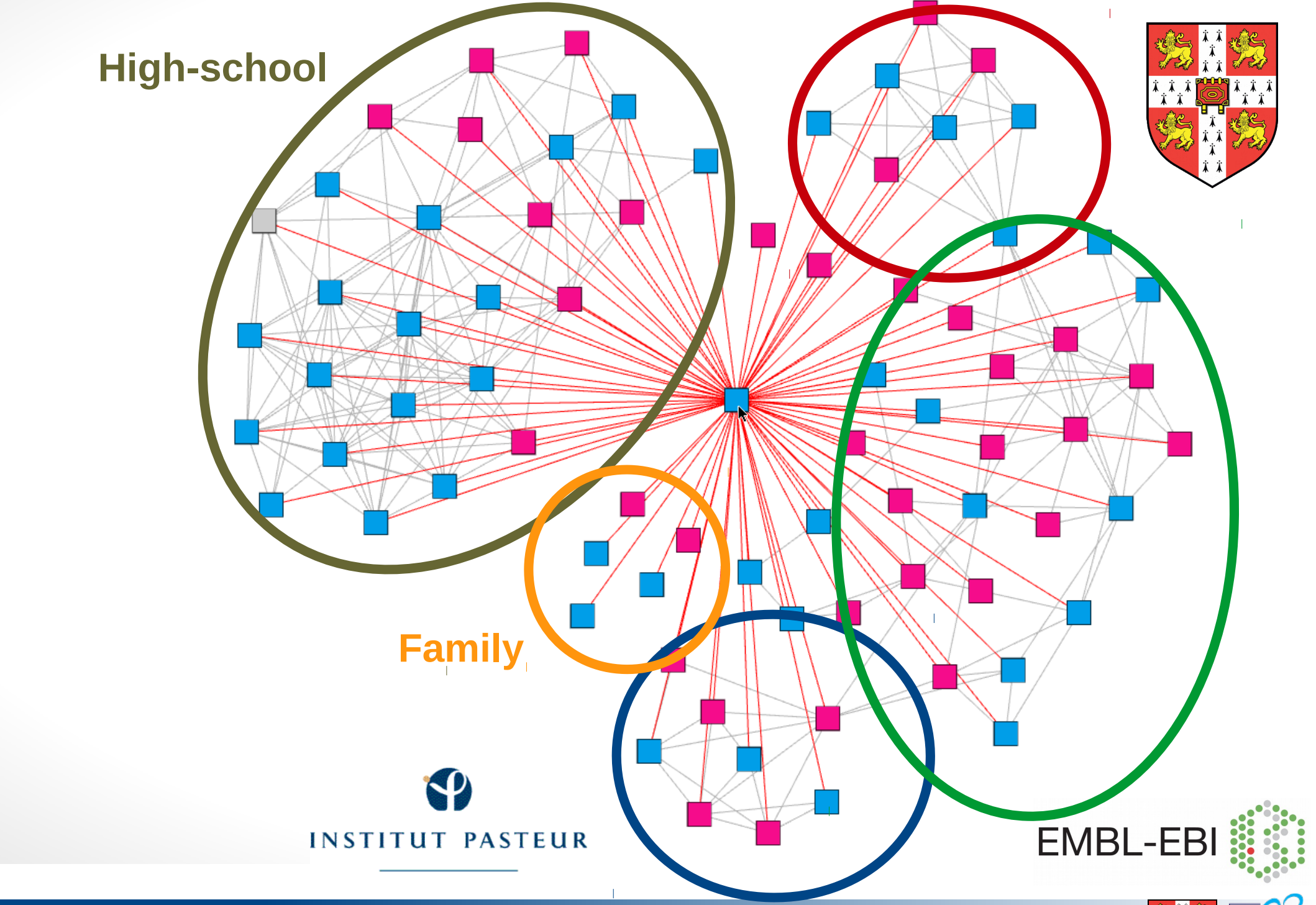
Gottfried Wilhelm Leibniz

Analysing the structure of networks

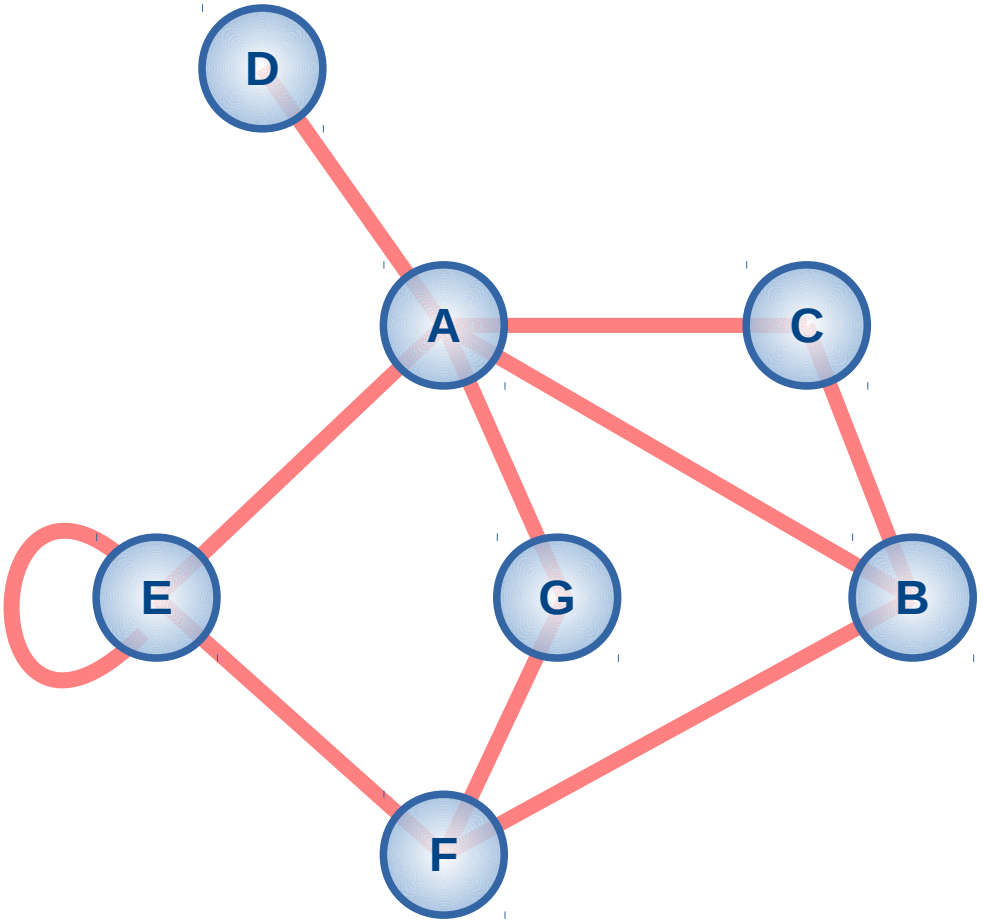
Definitions



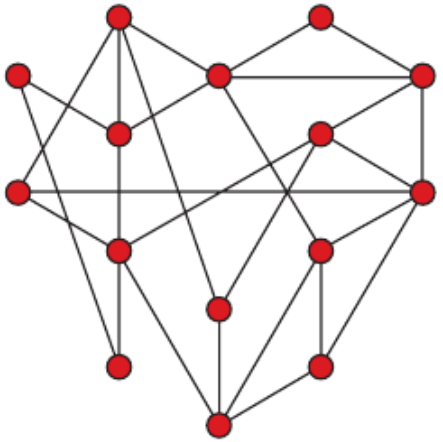


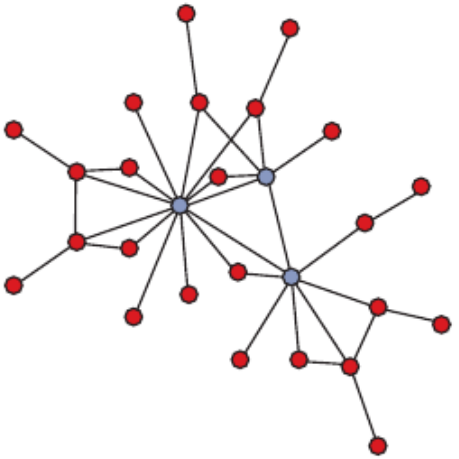

High-school
Family
INSTITUT PASTEUR
EMBL-EBI

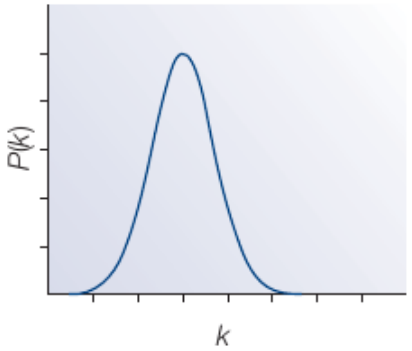
Node degree (k): number of ending edges

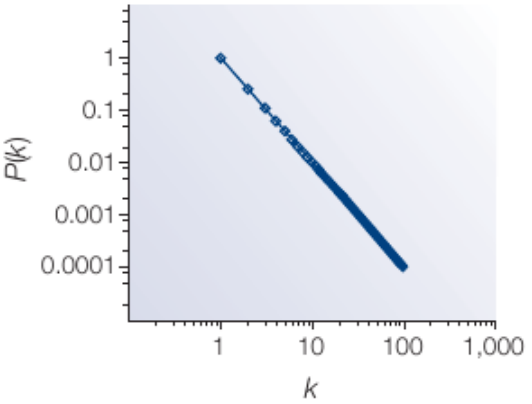


node	k
A	5
B	3
C	2
D	1
E	4
F	3
G	2





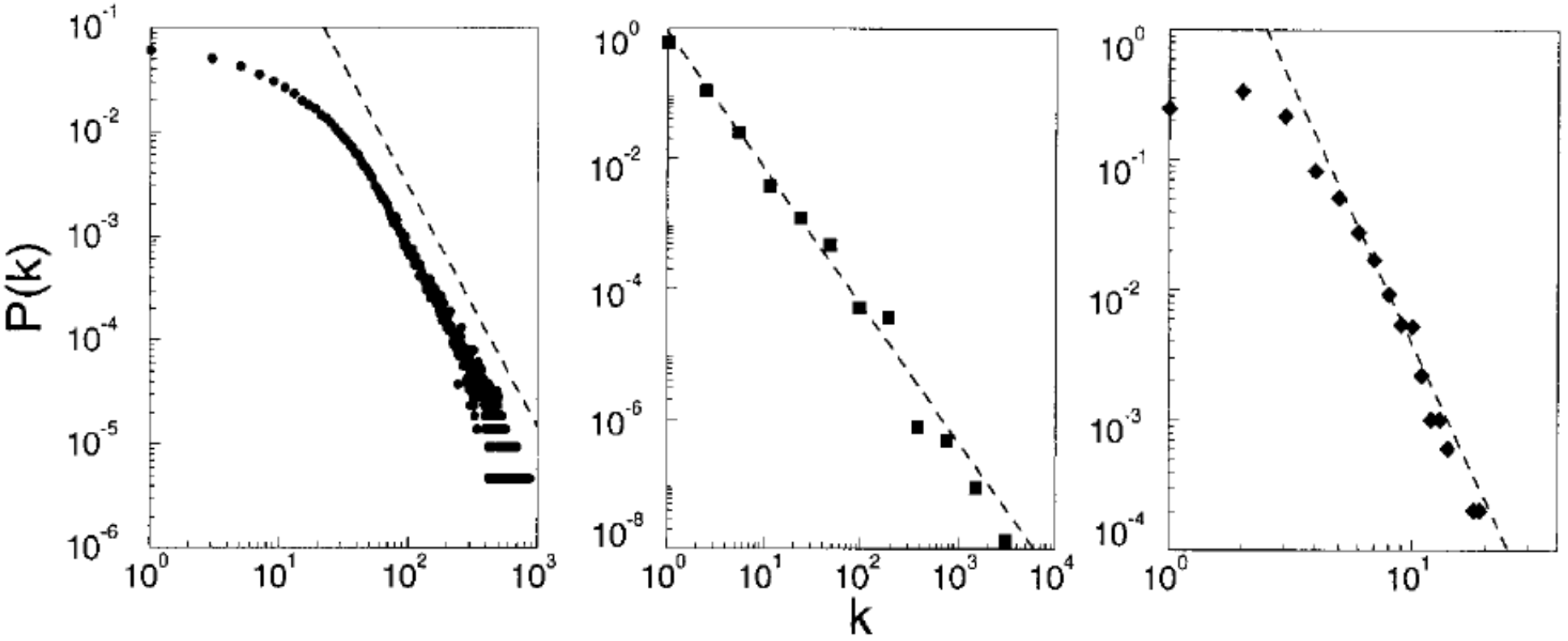




random networks (Erdős & Rényi)

Power law networks (Barabási & Albert)

Examples of power-law networks

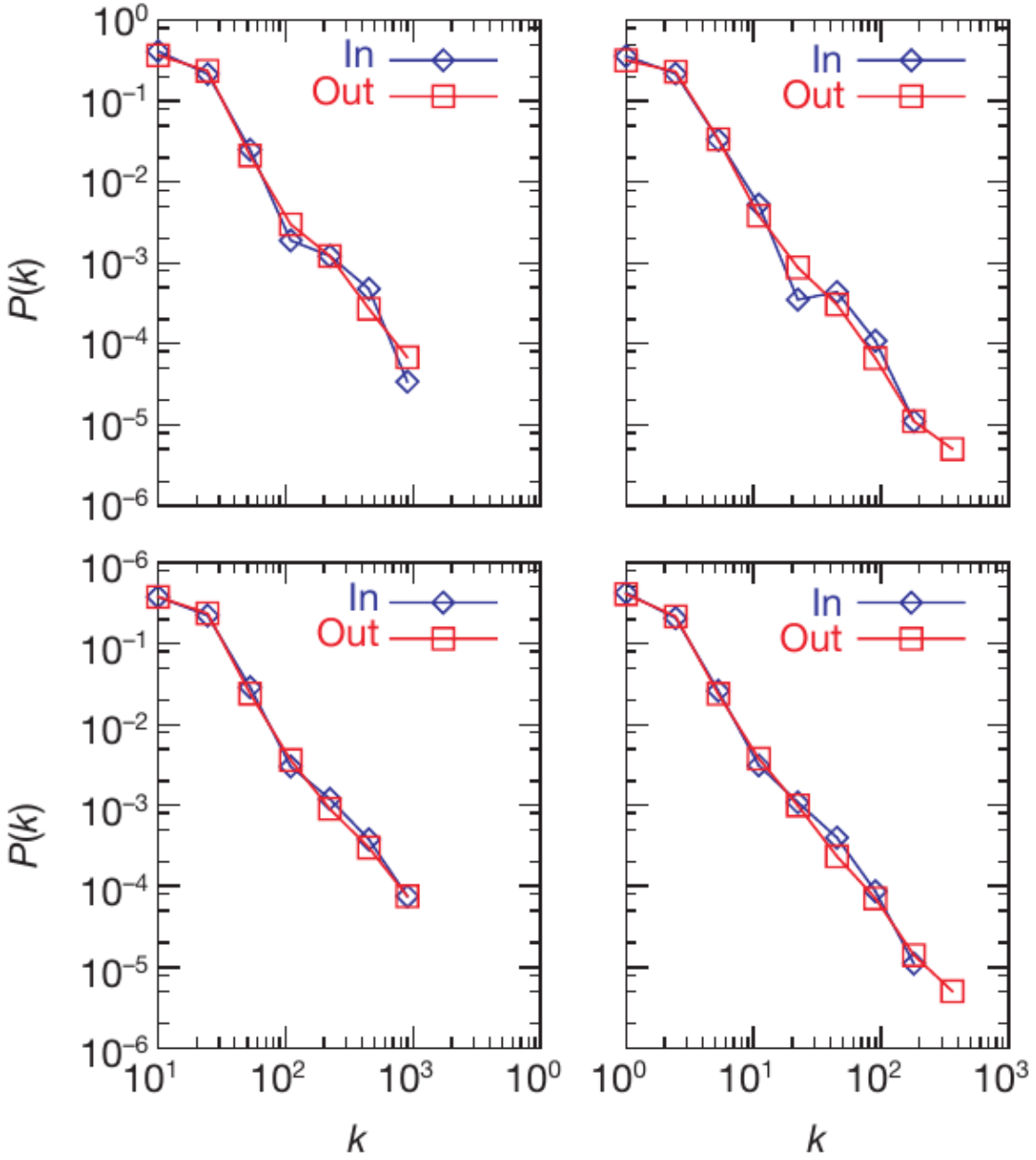


Actor collaboration

WorldWide Web

Power grid

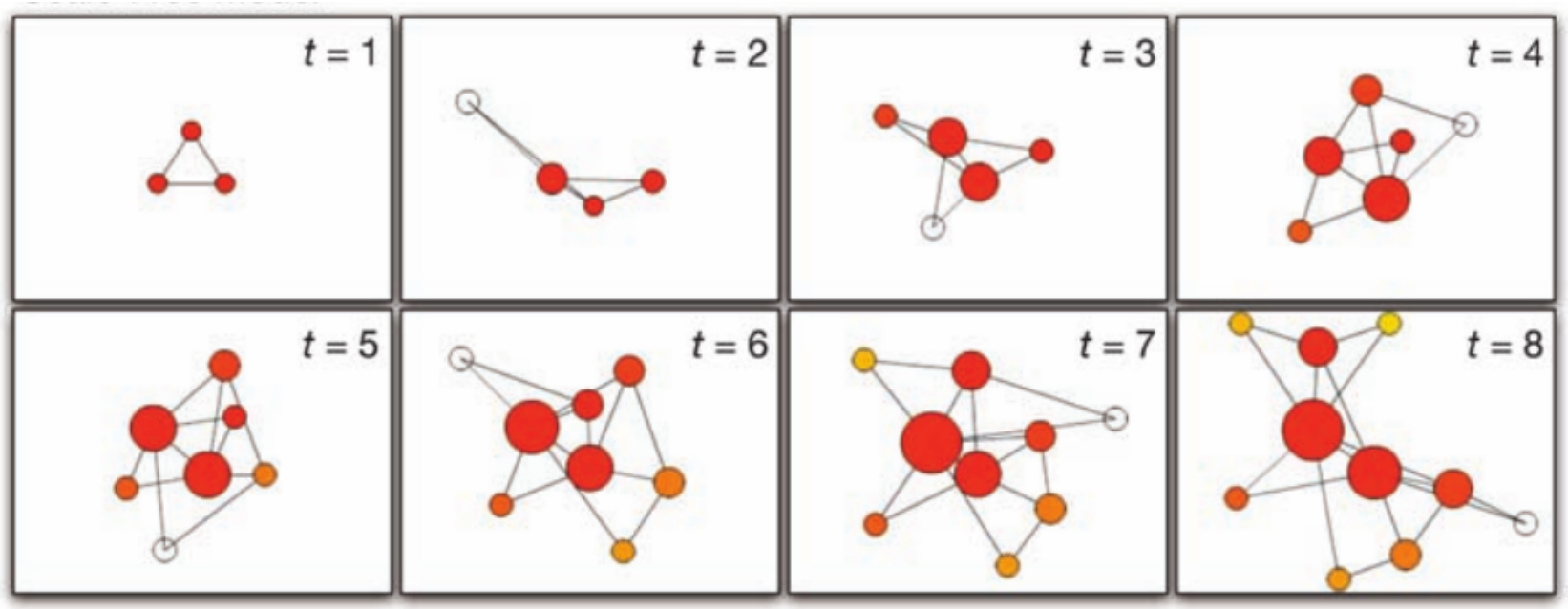
Metabolic networks are power-law



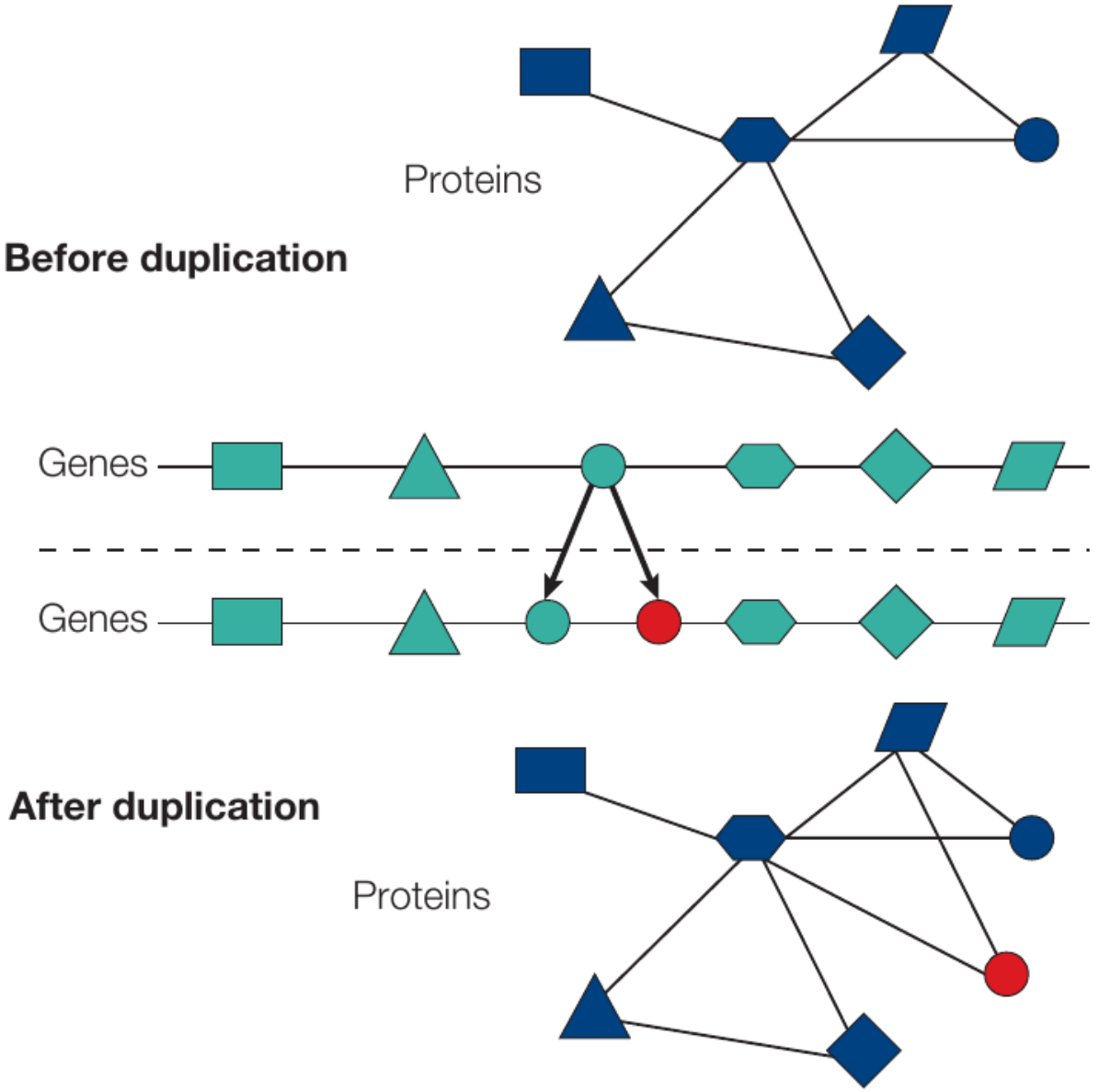
Metabolism of different organisms

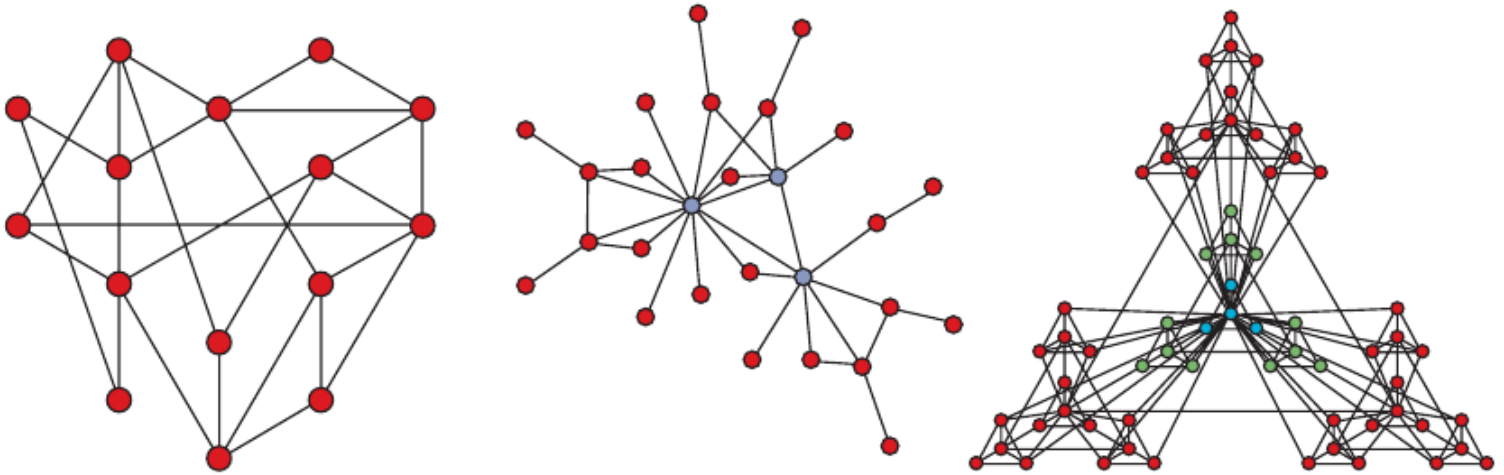
Creation = random, Evolution = power-law

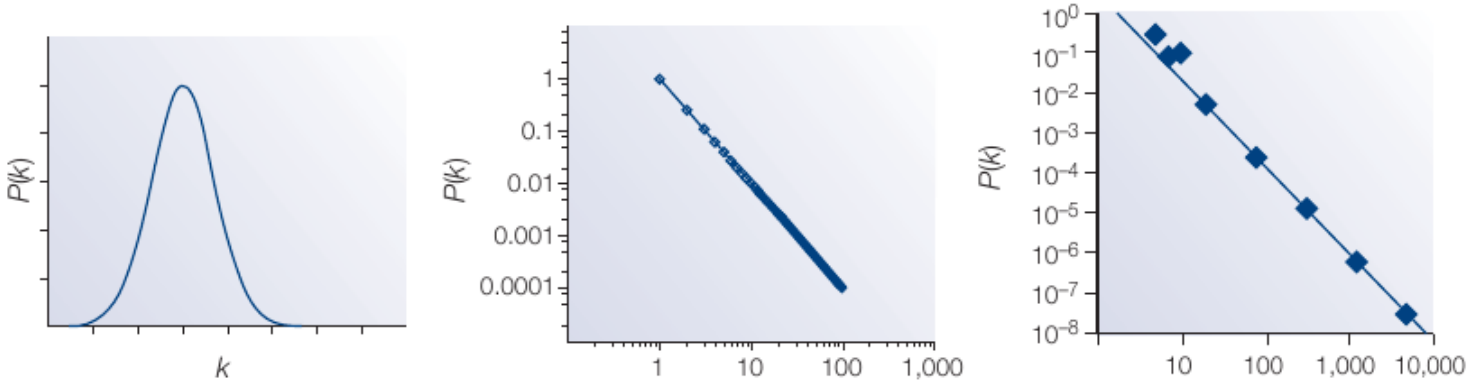
- Random creation at once:
  Do v times: pick a node among n, pick another among the same
  → random network
- Evolution:
  Repeat until n nodes: add a new node, link it to existing nodes with a probability depending on their existing degree
  → power-law network



Gene duplications and power-law interactomes





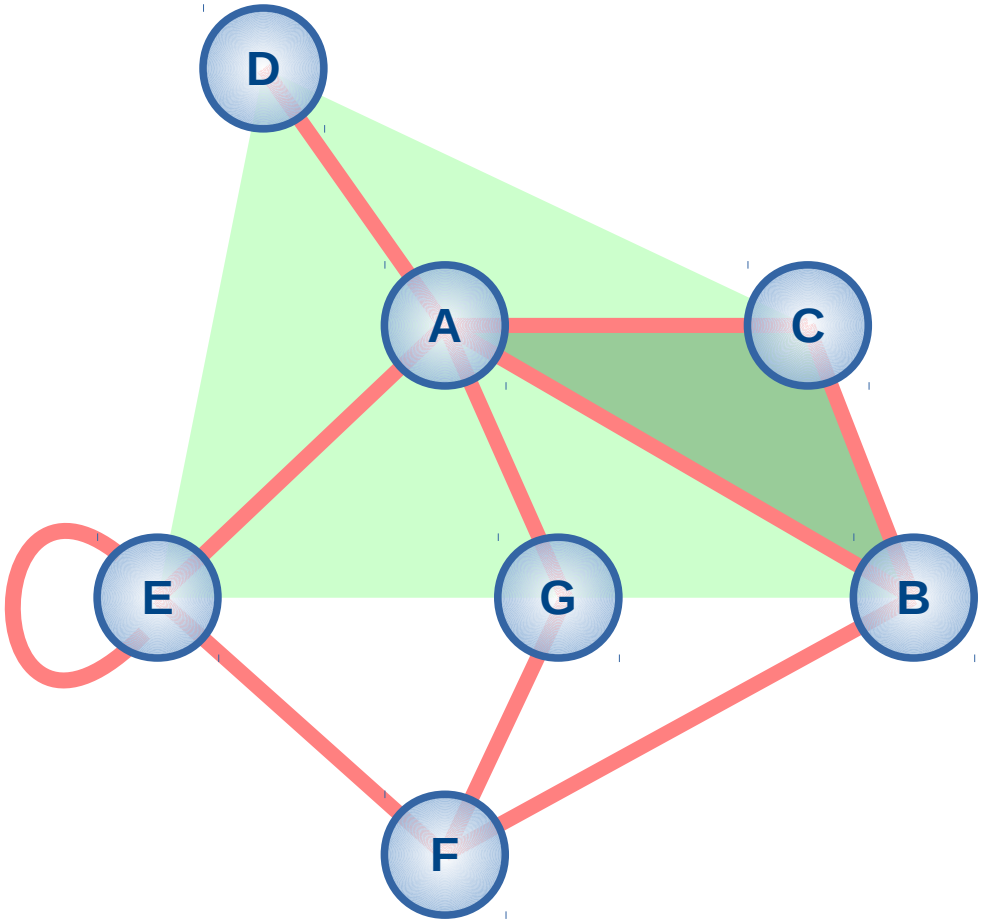


Random networks

Power law networks

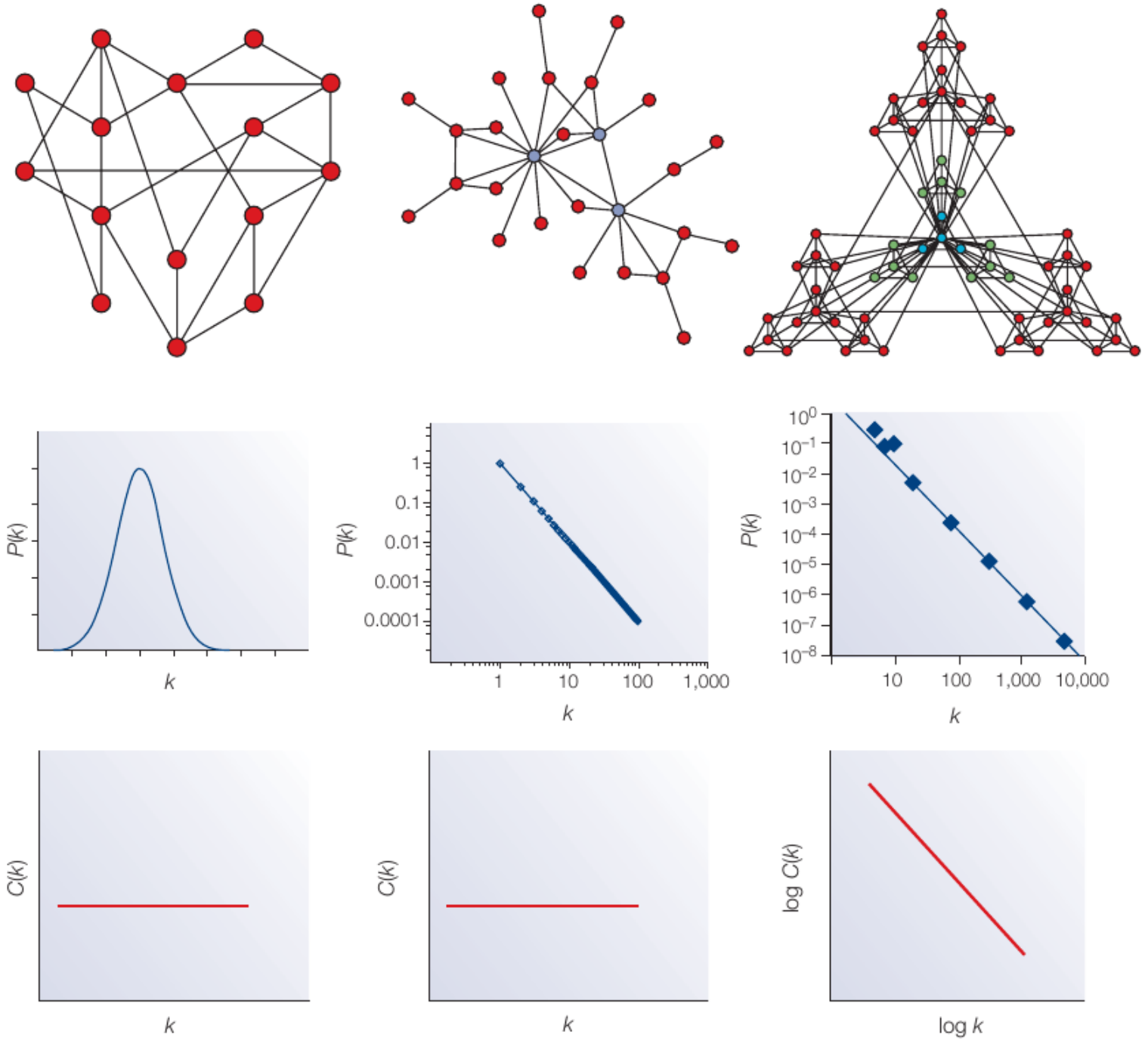
Hierarchical networks

Clustering (C): Proportion of possible cliques

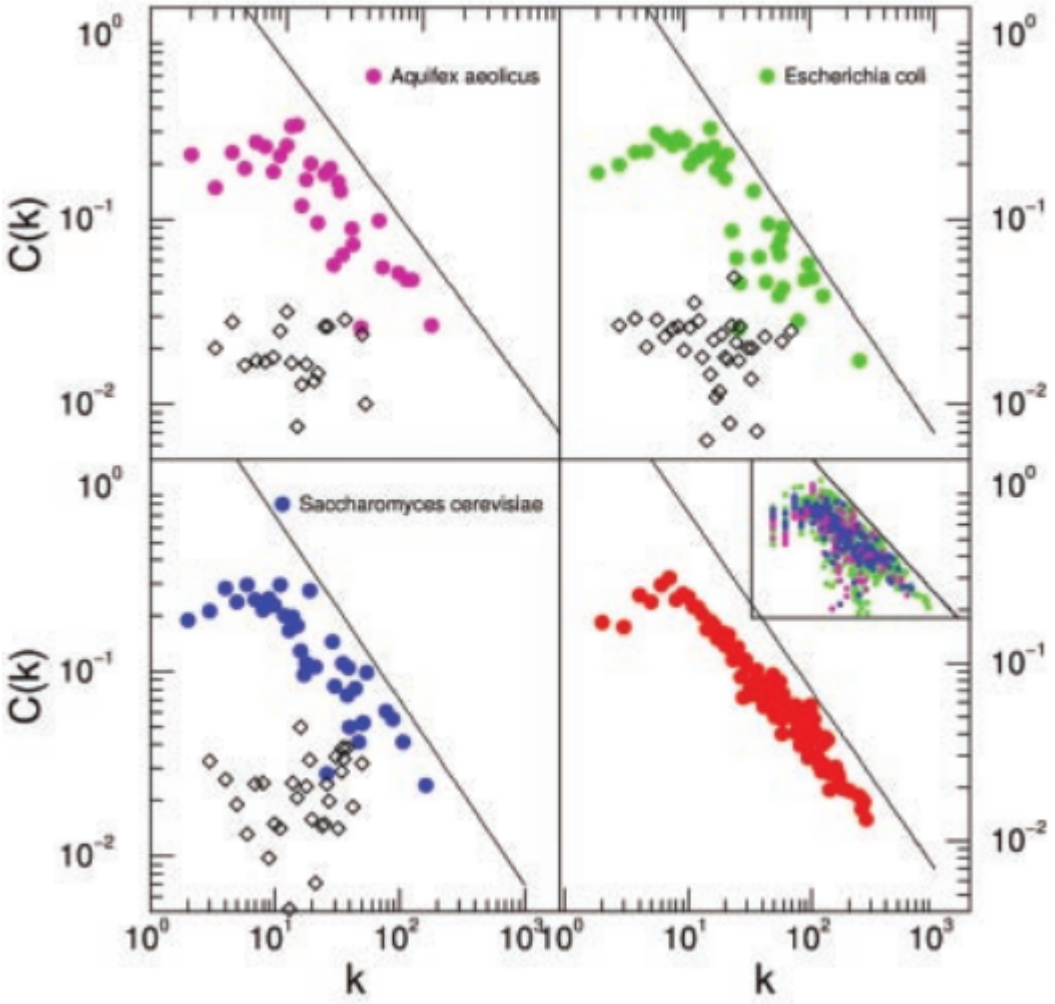


node	k	C
A	5	0.1
B	3	0.33
C	2	1
D	1	0
E	4	0
F	3	0
G	2	0

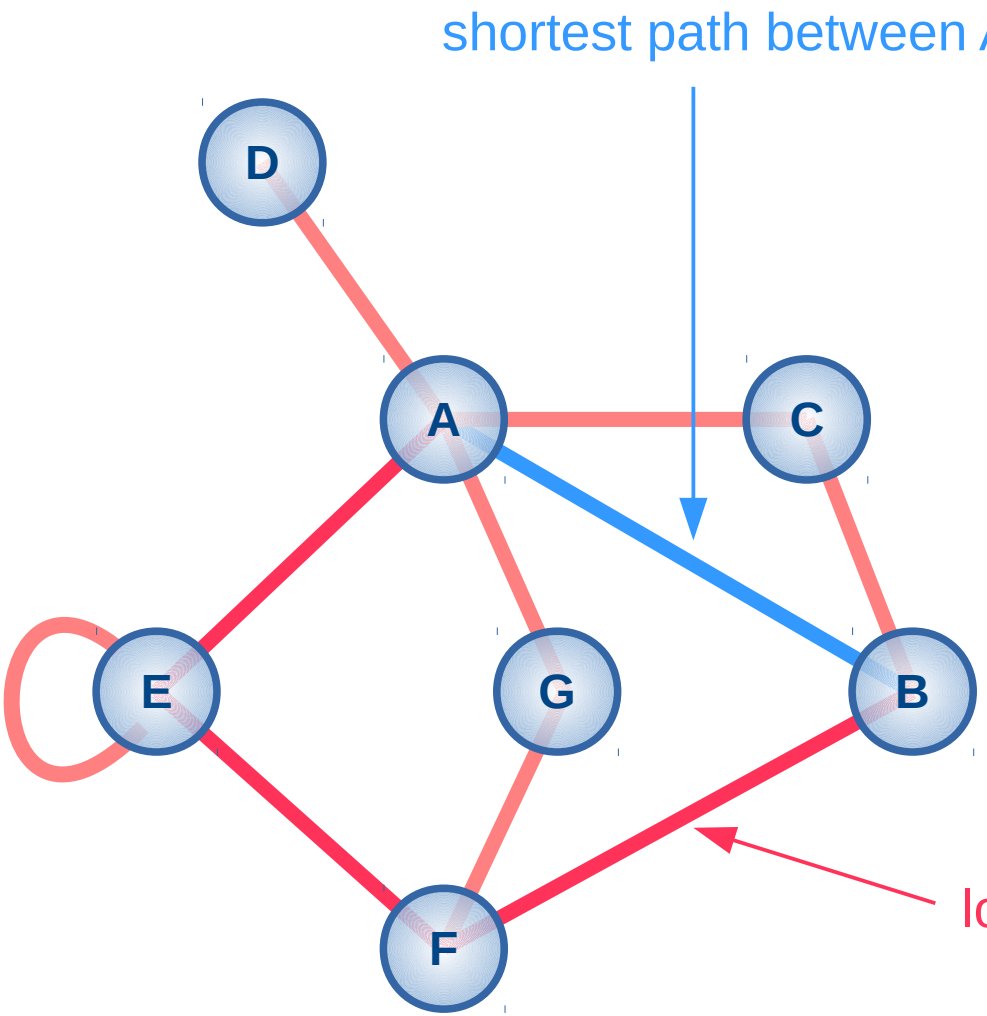
NB: nb of possible cliques = $\frac{k!}{2!(k-2)!}$


P(k)
k
P(k)
1
0.1
0.01
0.001
0.0001
1
10
100
1,000
k
P(k)
10^0
10^−1
10^−2
10^−3
10^−4
10^−5
10^−6
10^−7
10^−8
10
100
1,000
10,000
k
C(k)
k
C(k)
k
log C(k)
log k

Metabolic networks are hierarchical

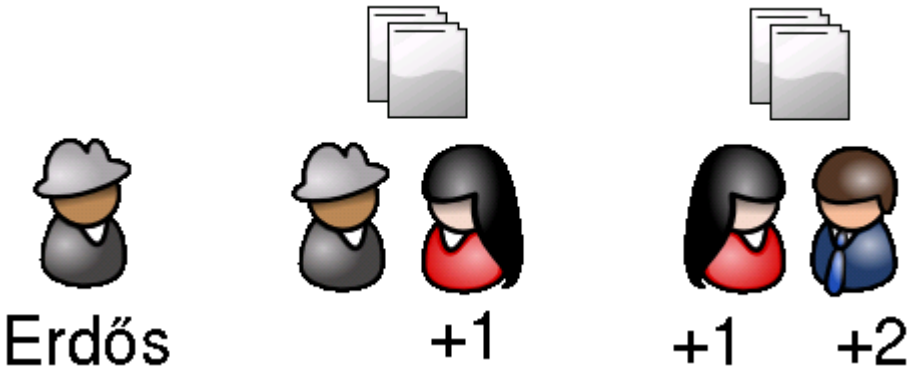


Shortest path (<SP>)



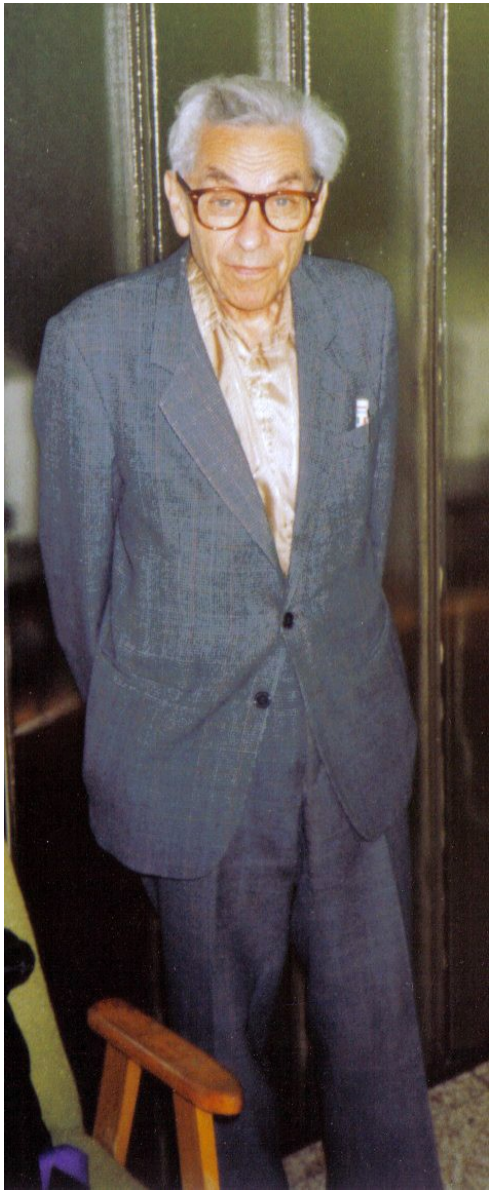
node	k	C	<SP>
A	5	0.1	1.167
B	3	0.33	1.5
C	2	1	1.67
D	1	0	2
E	4	0	1.67
F	3	0	1.67
G	2	0	1.67

Example of shortest path: Erdős number

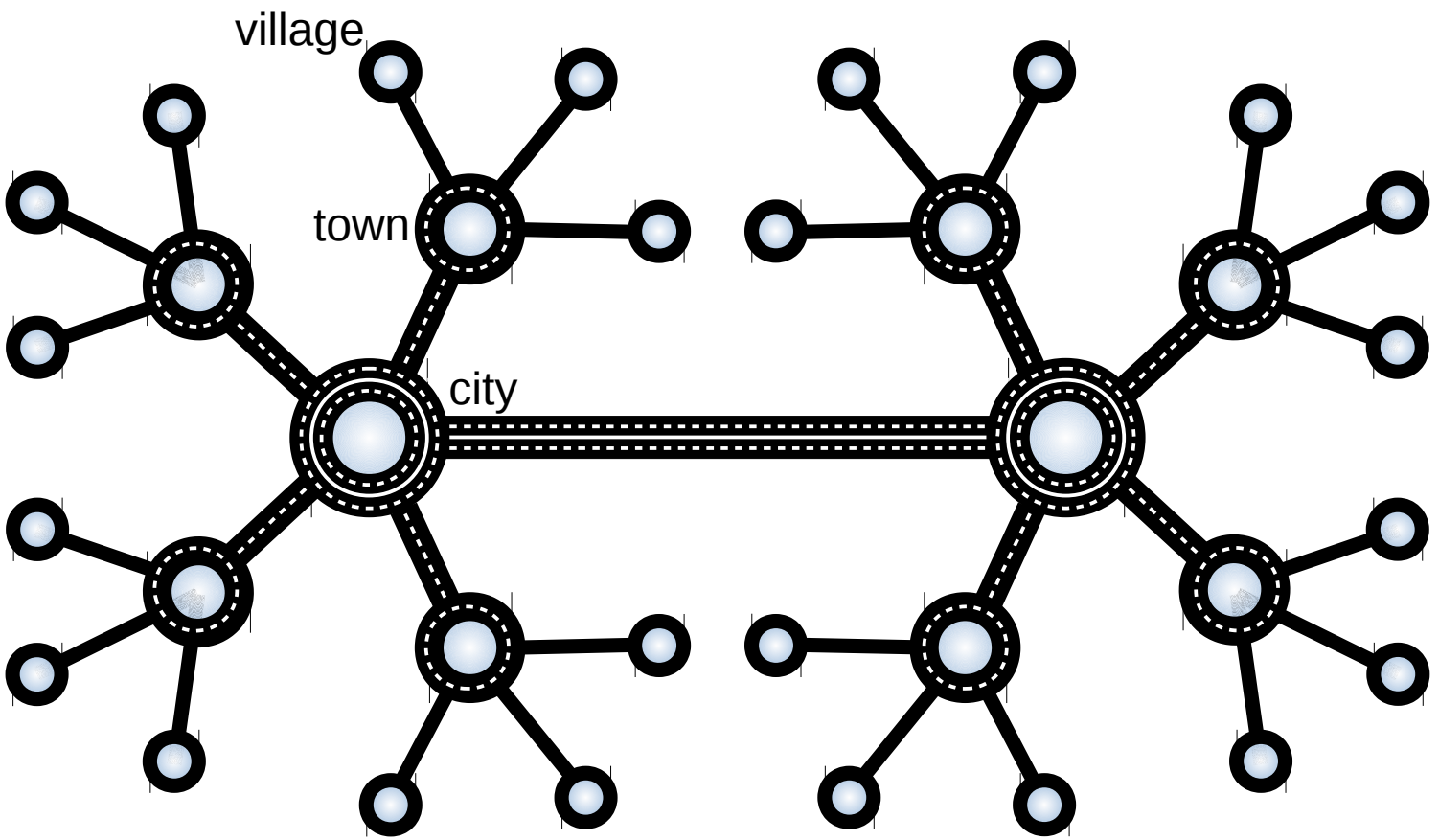


- 13 Field medals with En=2
- 2 Nobel Physics with En=2, 14 with En=3
- 4 Nobel Chemistry with En=3
- 3 Nobel Medecine with En=3
- 1 Nobel Economy with En=2. 8 with En=3
- My Erdős number is 4

  (Le Novère – Doyle – Boyd – Diaconis – Erdős
  Le Novère – Sauro – Nadim – Salamon - Erdős)

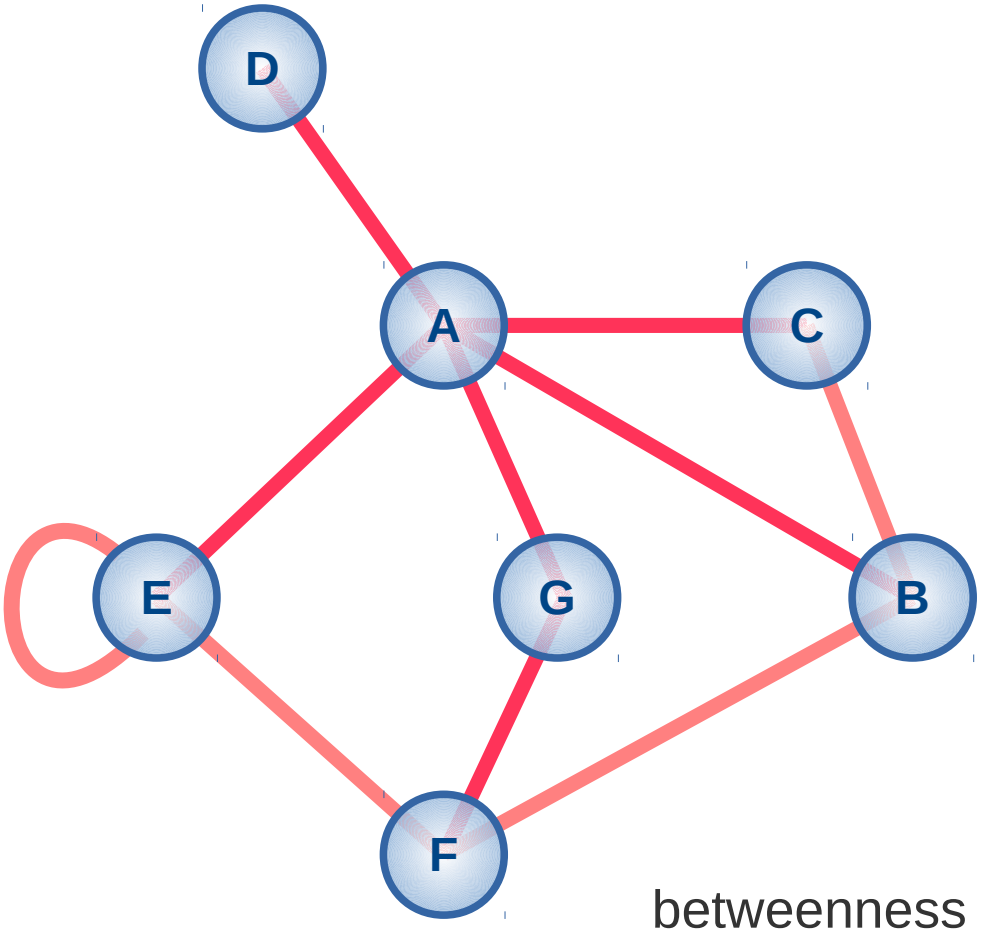


Power law often = small-world



The perfect road system=shortest average distance between two villages

Betweenness centrality (Cb)

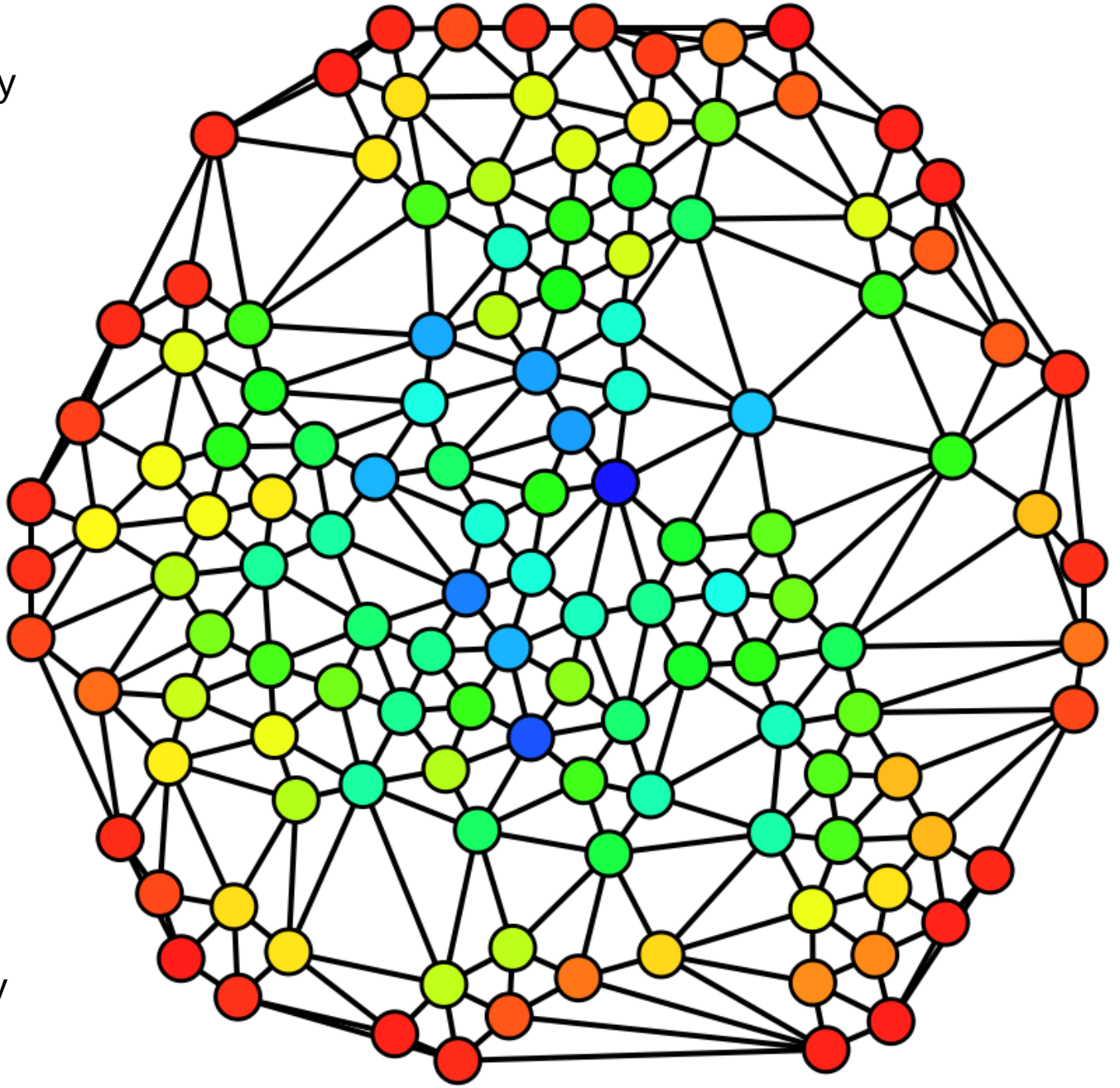


node	k	C	<SP>	Cb
A	5	0.1	1.167	0.567
B	3	0.33	1.5	0.111
C	2	1	1.67	0
D	1	0	2	0
E	4	0	1.67	0.044
F	3	0	1.67	0.1
G	2	0	1.67	0.044

betweenness = fraction of encountered shortest paths

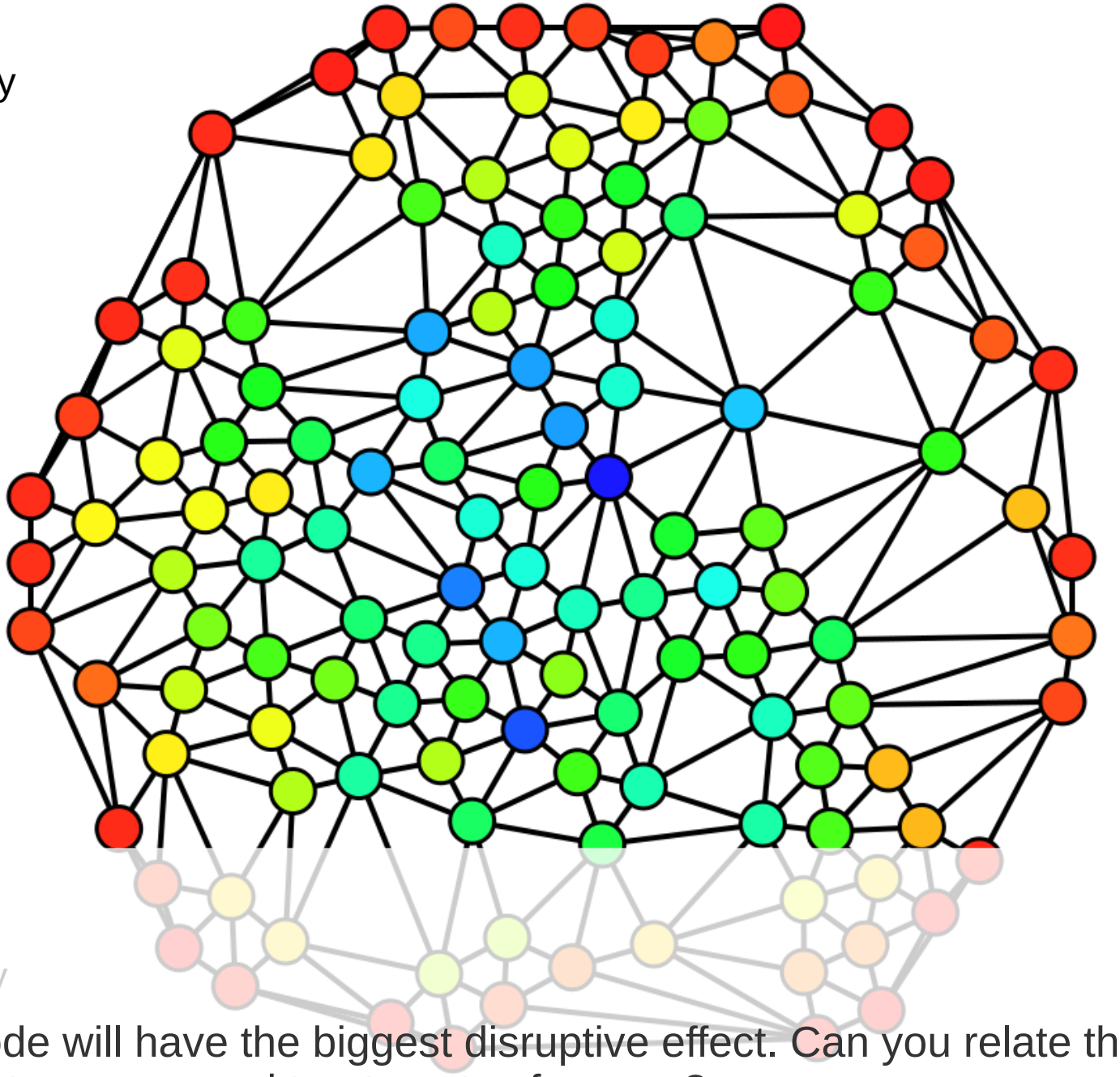
High betweenness centrality

Low betweenness centrality



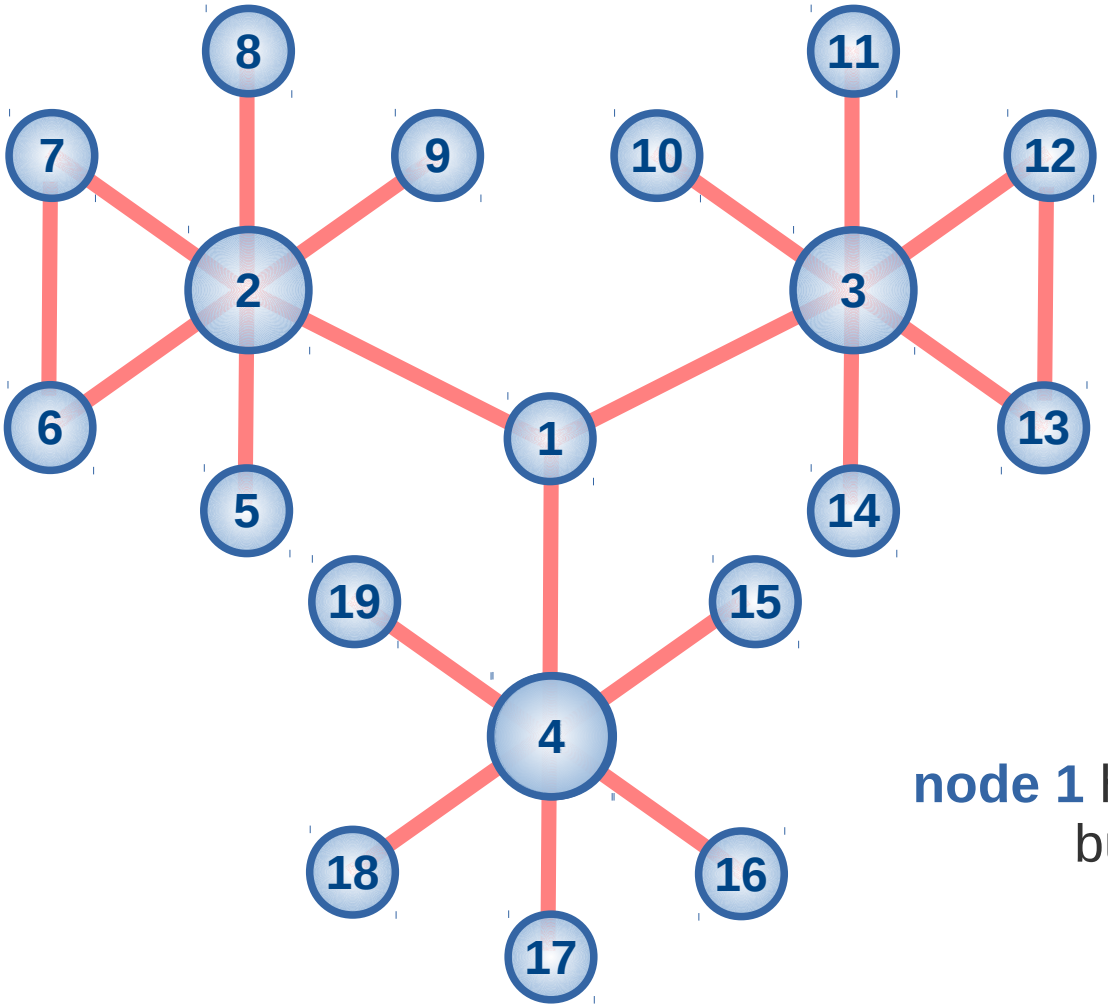
High betweenness centrality

Low betweenness centrality



Hitting the dark blue node will have the biggest disruptive effect. Can you relate that to causes and treatments of cancer?

No need to be highly connected to be important

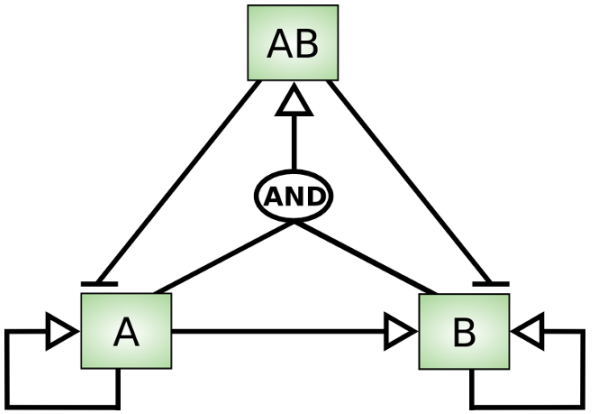


node 1 has low degree, low clustering but largest betweenness

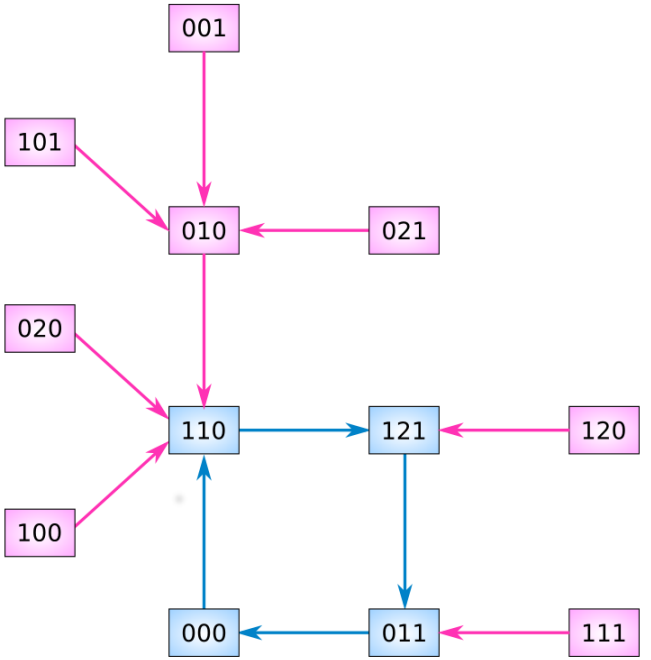
What's next?

Logic models

- Variables can take a discrete number of values, at least 2
- Transitions of output are expressed as logical combinations of input values
- Simulations can be:
  synchronous: all the nodes are updated at once
  asynchronous: nodes are updated one after the other
- One can add delays, inputs etc.



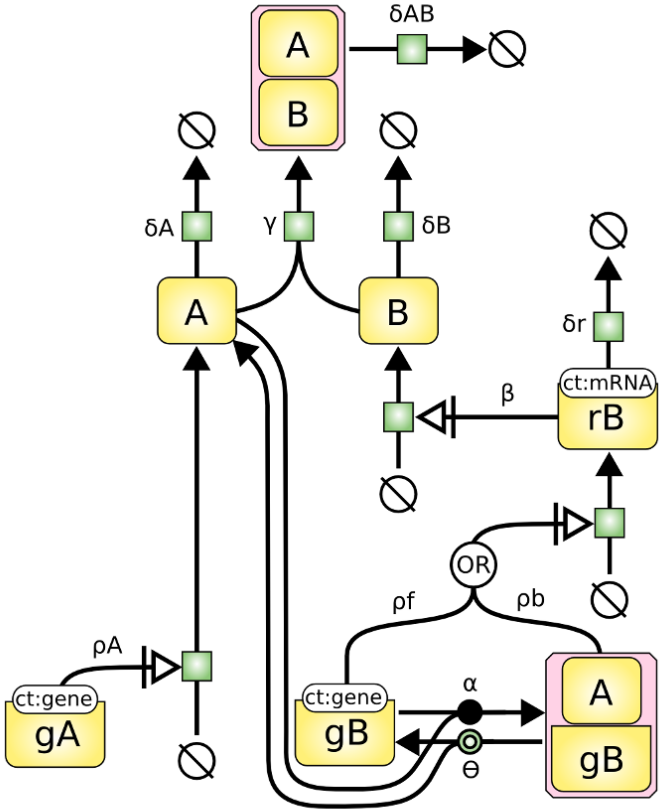
Influence diagram

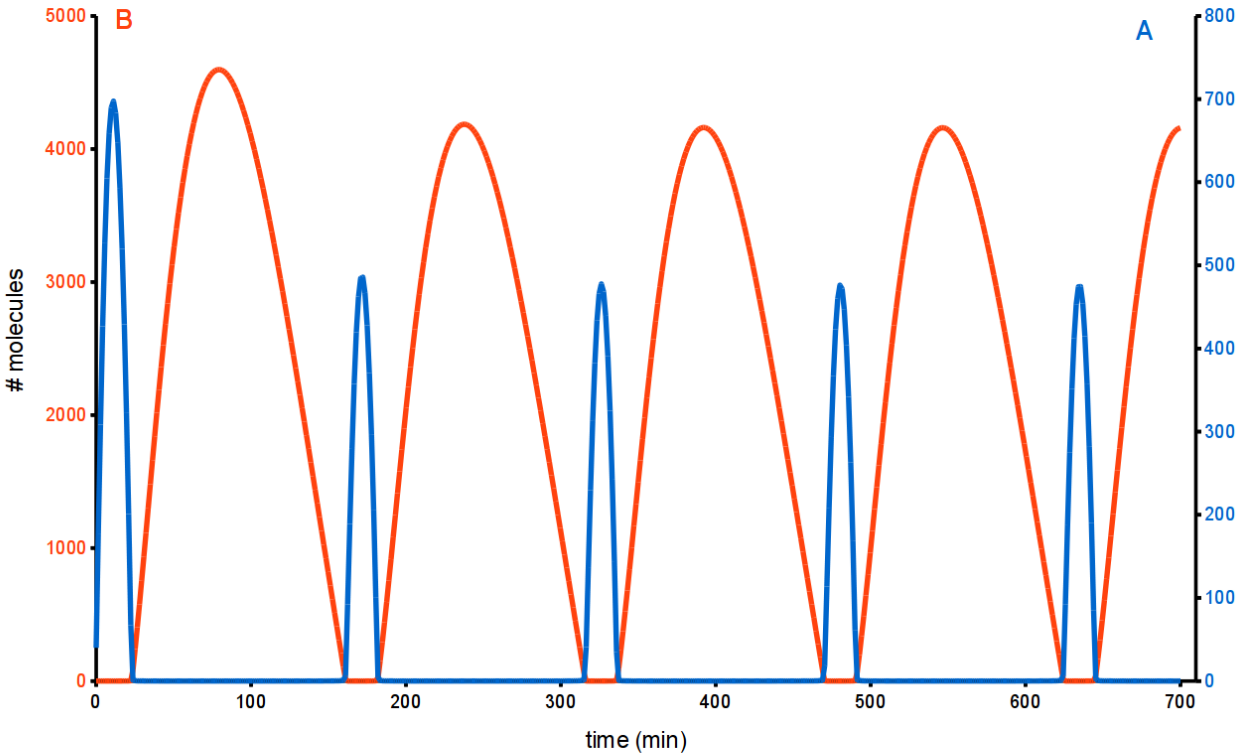


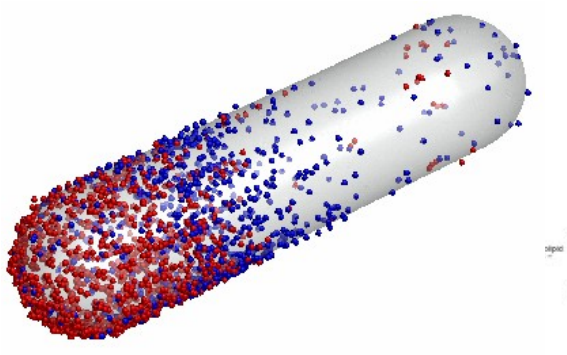
state diagram

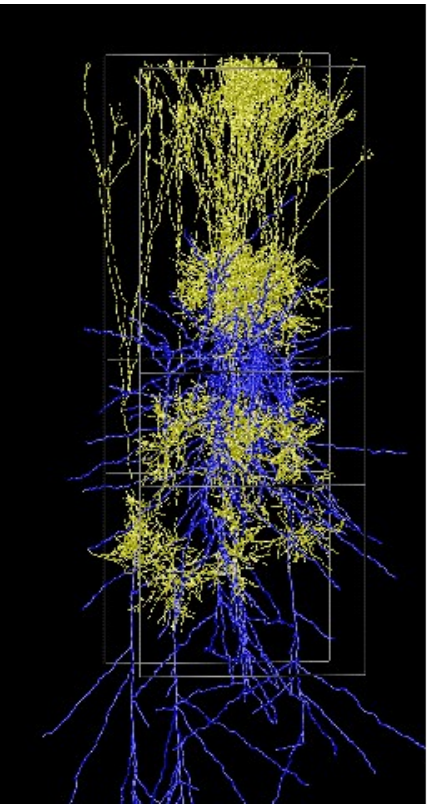
Quantitative models

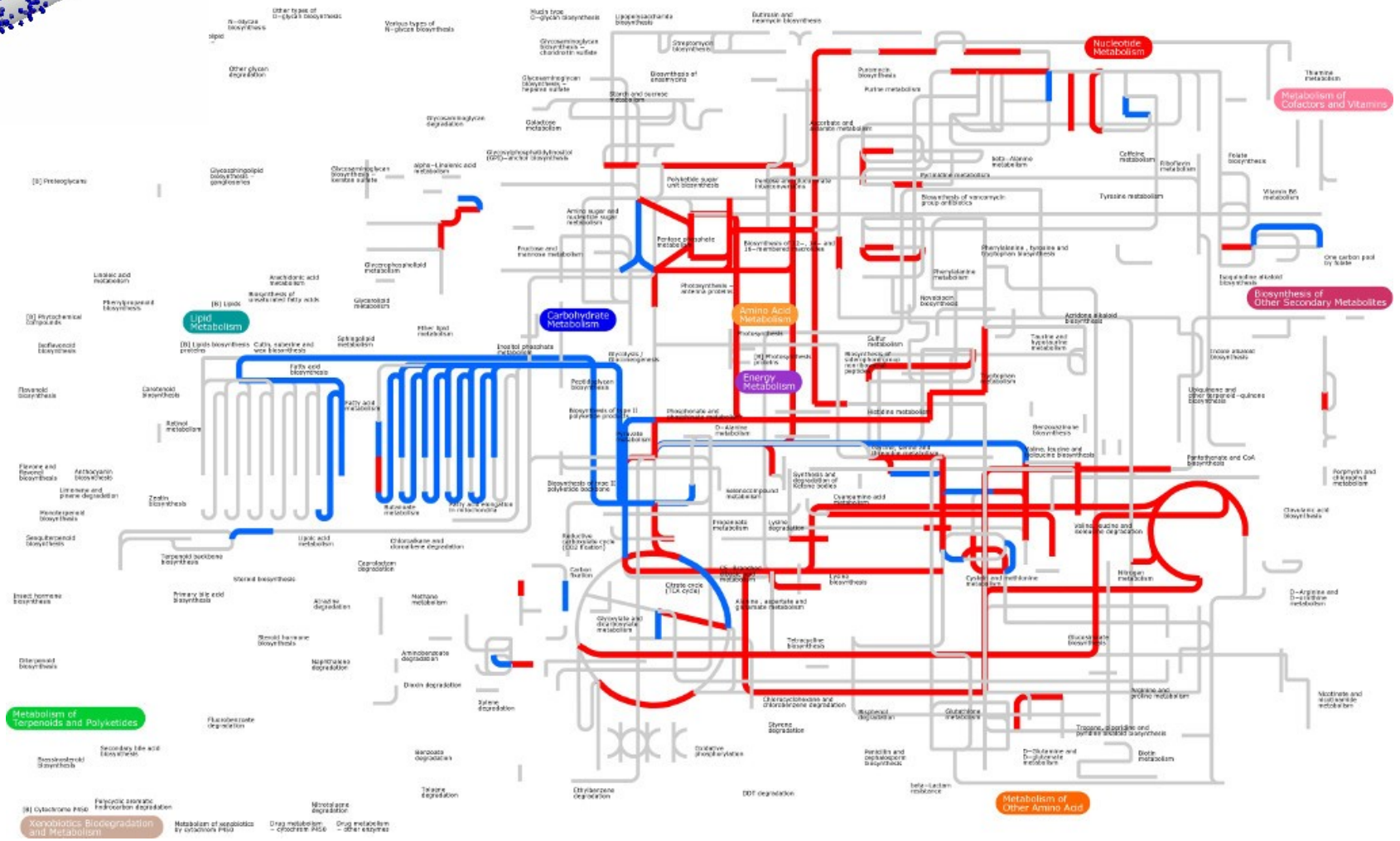
- Variables are concentrations or amounts
- Reactions rates are calculated based on quantitative values
- Simulations deterministic or stochastic











Babraham Institute

Any question: n.lenovere@gmail.com

Further information:

Barabási & Zoltán. Network biology: understanding the cell's functional organization.
Nature Reviews Genetics (2004) **5**: 101-113

Le Novère. Quantitative and logic modelling of molecular and gene networks.
Nature Reviews Genetics (2015) **16**: 146–158