@KorcsmarosLab

Overview lecture:
Computational analysis of protein-protein interactions in cell function and disease

EMBO Practical course
Bangalore, India

TAMAS KORCSMAROS
Group leader

Quadram Institute

Earlham Institute
Decoding Living Systems

The Korcsmaros group

The Table of Contents

- Protein-protein interactions
  from a computational and systems biology point of view

- Signalling pathways and networks
  - Integration of interaction data
  - Signal flow regulation on the PPI-level

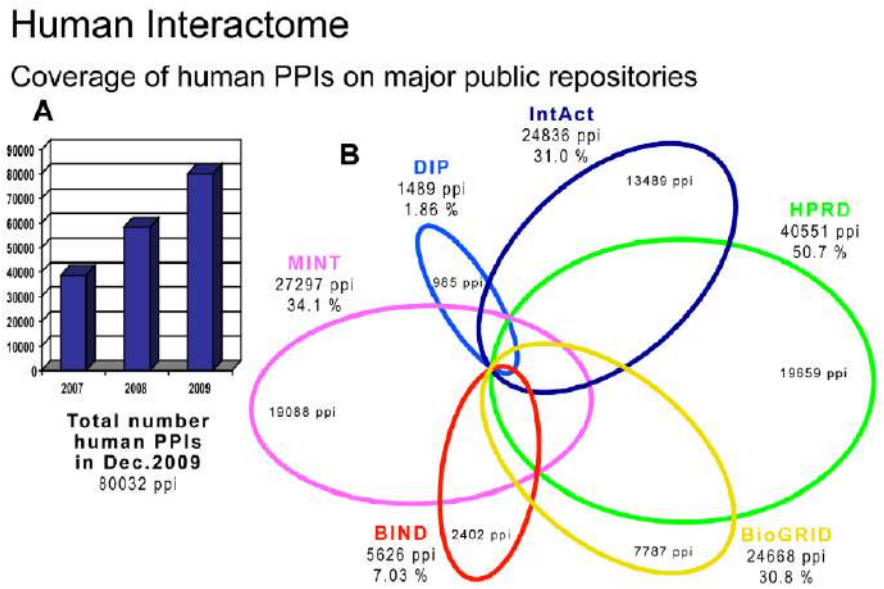
- Applications
  - Network medicine
  - Network pharmacology

Protein-protein interactions

Protein-protein interactions

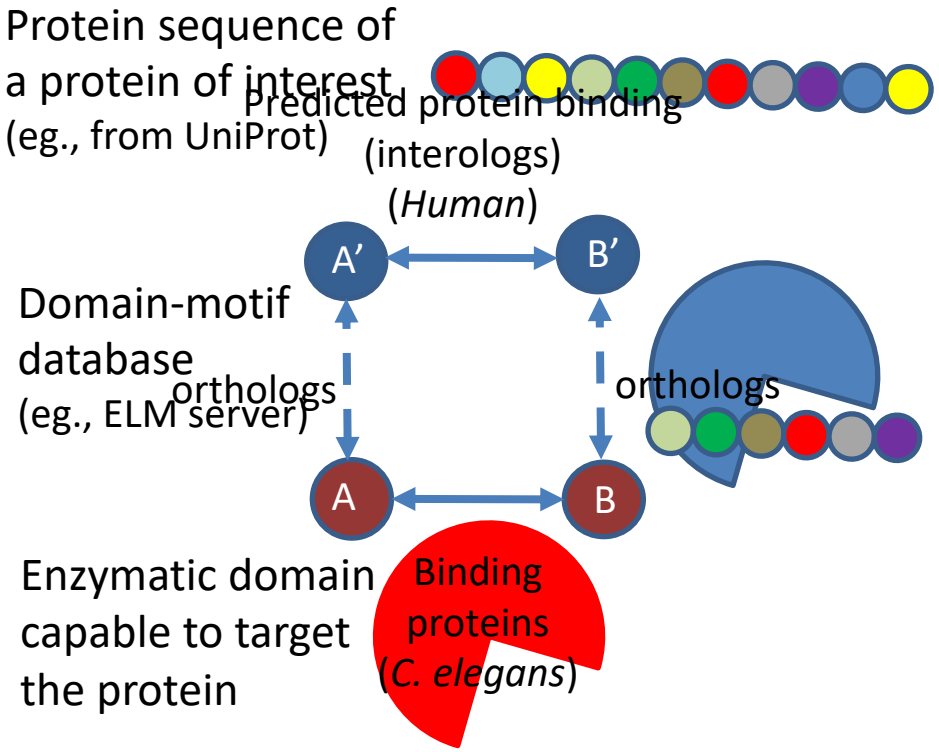
Experimentally found interactions

- Small-scale (precise, can be directed, should be collected)
- Large-scale (limited structural details, generally undirected, found in data repositories)
- HuRI (http://interactome.baderlab.org)
- BioGRID (http://thebiogrid.org)
- STRING (https://string-db.org)
- IntAct (http://www.ebi.ac.uk/intact)
- DroID, WI8, etc.



Predicted interactions

- Homology/orthology-based (interologs)
- Domain-motifs based (directed)
- Domain-domain based (interaction & direction)

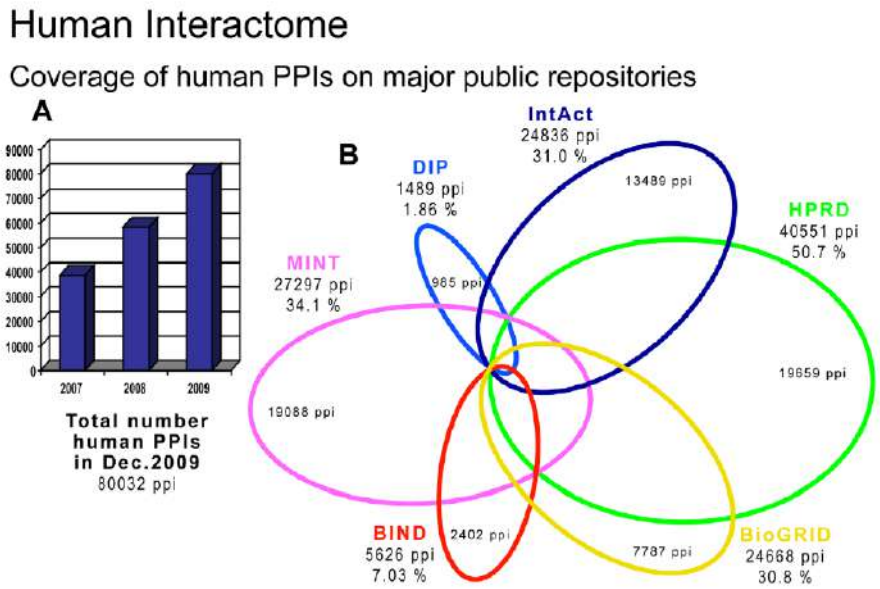


Predicted PPI based on domain-motif interaction

Protein-protein interactions

Experimentally found interactions

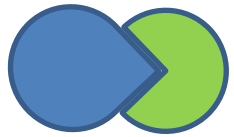
- Small-scale (precise, can be directed, should be collected)
- Large-scale (limited structural details, generally undirected, found in data repositories)
- HuRI (http://interactome.baderlab.org)
- BioGRID (http://thebiogrid.org)
- STRING (https://string-db.org)
- IntAct (http://www.ebi.ac.uk/intact)
- DroID, WI8, etc.



Predicted interactions

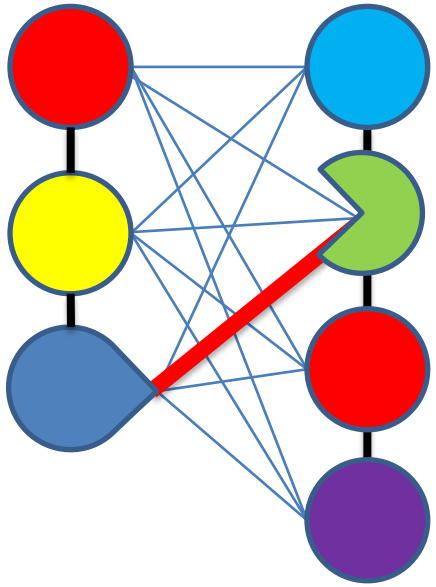
- Homology/orthology-based *(interologs)*
- Domain-motifs based (directed)
- **Domain-domain based** (**interaction** & direction)

Domain-domain interaction data (eg., DOMINE)



Protein-domain composition data (eg., PFAM)

Possible domain pairs

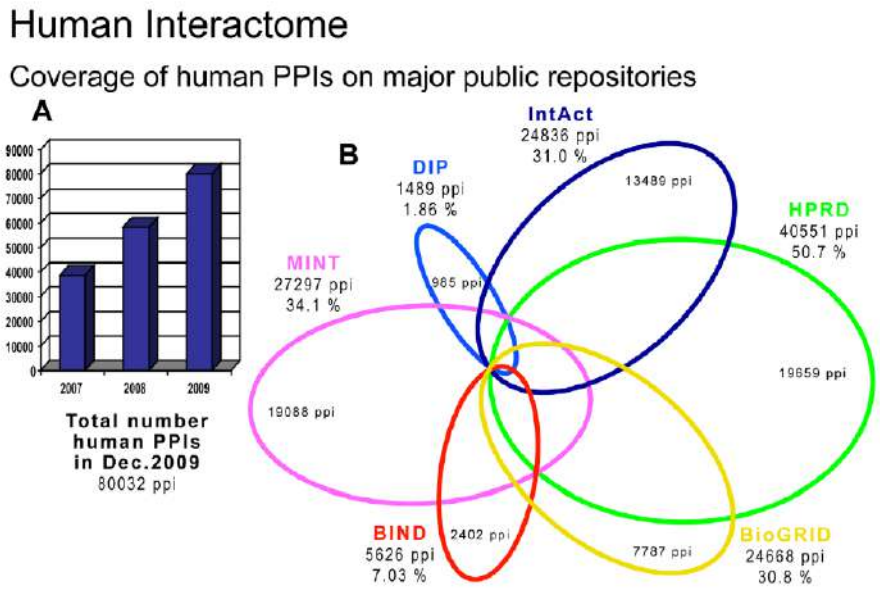


Predicted PPI based on domain-domain interaction

Protein-protein interactions

Experimentally found interactions

- Small-scale (precise, can be directed, should be collected)
- Large-scale (limited structural details, generally undirected, found in data repositories)
- HuRI (http://interactome.baderlab.org)
- BioGRID (http://thebiogrid.org)
- STRING (https://string-db.org)
- IntAct (http://www.ebi.ac.uk/intact)
- DroID, WI8, etc.

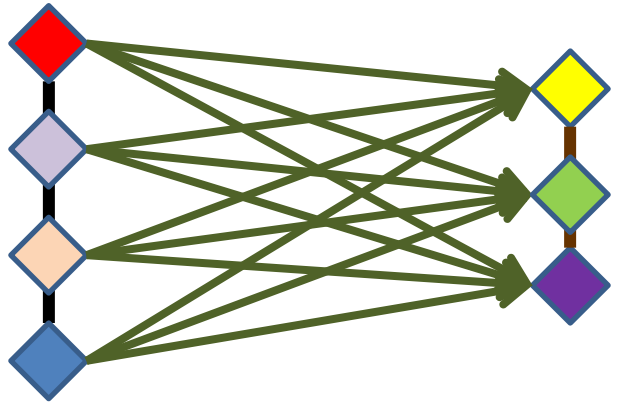


Predicted interactions

- Homology/orthology-based *(interologs)*
- Domain-motifs based (directed)
- **Domain-domain based** (interaction & **direction**)

Directed PPI from the reference database (eg., Reactome)

Domain composition as training set (eg., PFAM)

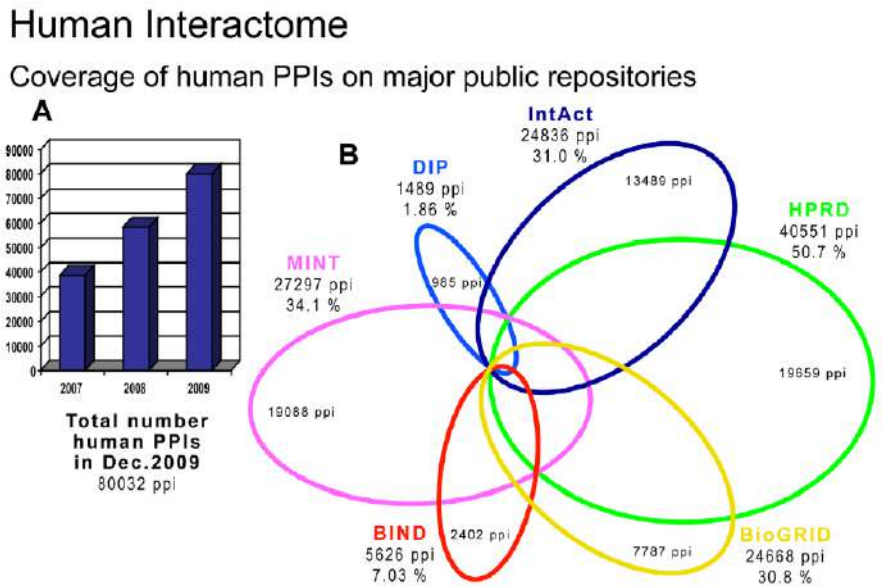


Liu *et al.*, MCP (2009) and
Rhodes *et al.*, Nature Biotechnology (2005)

Protein-protein interactions

Experimentally found interactions

- Small-scale (precise, can be directed, should be collected)
- Large-scale (limited structural details, generally undirected, found in data repositories)
- HuRI (http://interactome.baderlab.org)
- BioGRID (http://thebiogrid.org)
- STRING (https://string-db.org)
- IntAct (http://www.ebi.ac.uk/intact)
- DroID, WI8, etc.

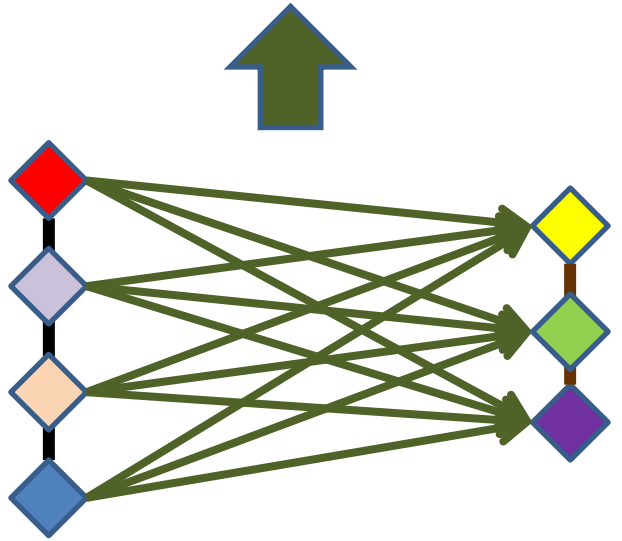


Predicted interactions

- Homology/orthology-based *(interologs)*
- Domain-motifs based (directed)
- **Domain-domain based** (interaction & **direction**)

$$F(\blacklozenge\!-\!\blacklozenge) = \frac{Pr(\blacklozenge\!\rightarrow\!\blacklozenge) - Pr(\blacklozenge\!\rightarrow\!\blacklozenge)}{Pr(\blacklozenge)\ X\ Pr(\blacklozenge)}$$

Domain composition as training set (eg., PFAM)

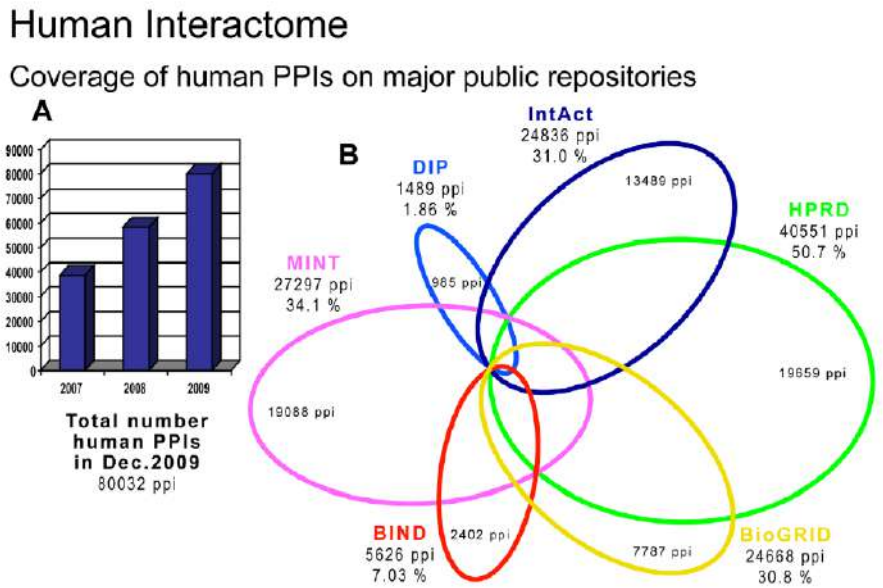


Liu *et al.*, MCP (2009) and
Rhodes *et al.*, Nature Biotechnology (2005)

Protein-protein interactions

Experimentally found interactions

- Small-scale (precise, can be directed, should be collected)
- Large-scale (limited structural details, generally undirected, found in data repositories)
- HuRI (http://interactome.baderlab.org)
- BioGRID (http://thebiogrid.org)
- STRING (https://string-db.org)
- IntAct (http://www.ebi.ac.uk/intact)
- DroID, WI8, etc.

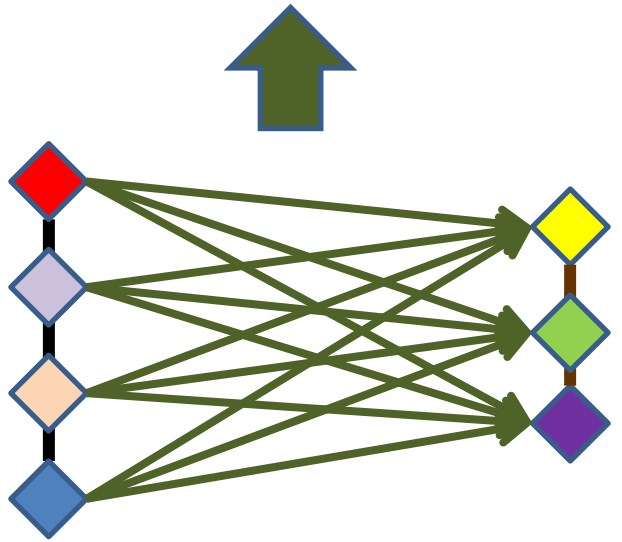


Predicted interactions

- Homology/orthology-based *(tomorrow)*
- Domain-motifs based (directed)
- **Domain-domain based** (interaction & **direction**)

$$F(d_{mn}) = \frac{\Pr(d_m \rightarrow d_n) - \Pr(d_n \rightarrow d_m)}{\Pr(d_m) \times \Pr(d_n)}$$

Domain composition as training set (eg., PFAM)



Liu *et al.*, MCP (2009) and
Rhodes *et al.*, Nature Biotechnology (2005)

Regulation of protein-protein interactions

Kinases, phosphatases, ubiquitin-ligases, peptidases, etc.

- Reversible or irreversible modulation of specific proteins
- Priming, activation, temporal de-activation, cleavage, destruction, etc.

Resources of post-translational modifications (PTMs)

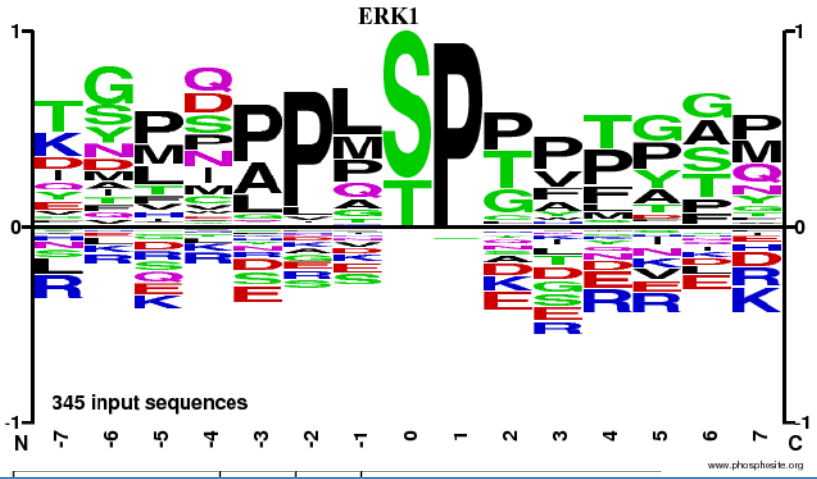
- NetworKIN (http://networkin.info)
- Phosphosite (http://phosphosite.org)
- dbPTM (http://dbptm.mbc.nctu.edu.tw)
- ELM server (http://elm.eu.org)
- …

Active site

Specificity?

Docking interactions

Target motifs



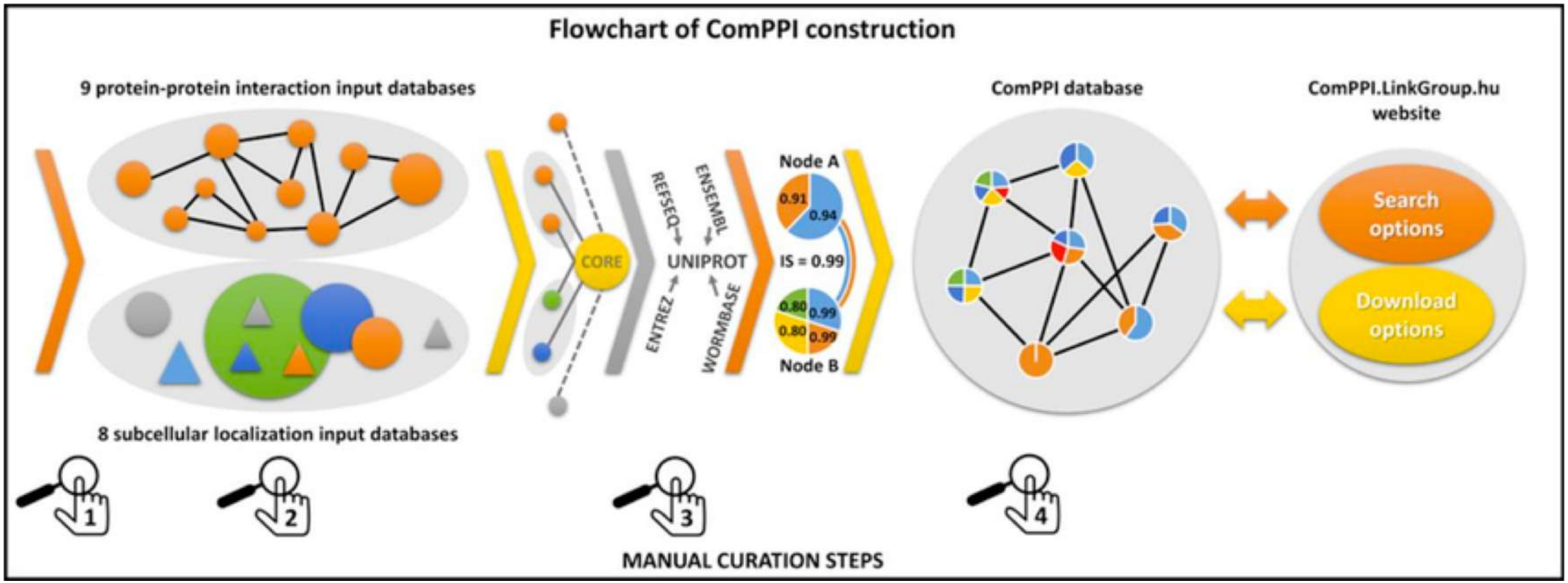
Docking motifs

Sequence	Start	End	Subsequence
NCF1_HUMAN	363	368	TQRSKPQPAVPPRPSADLIL
POLG_HCVJA	2323	2328	LPSTKAPPIPPPRRKRTVVL
FAK1_MOUSE	750	755	SGGSDEAPPKPSRPGYPSPR
DYN1_HUMAN	833	838	EPPPPQVPPRPTRAPPGVPS
P85A_HUMAN	308	313	RQPNATPPAPPKPTTVANN
P85A_HUMAN	305	310	WNERQPAPALPPKPPKPTTV
RPGF1_HUMAN	284	289	VVDNSPPPALPPKKRQSAPS
PTN22_MOUSE	614	619	RTDDEIPPPLPERTPESFIV
NEF_HV1BR	72	77	EVGFPVTPQVPLRPMTYKAA
PAK1_RAT	13	18	LDVQDKPPAPPMRNTSTMIG
SOS1_HUMAN	1152	1157	DEVPVPPPVPPRRRPESAPA

P..P.[KR]

Dinkel *et al.,* Nucleic Acids Res. (2012); Bhattacharyya *et al.,* Annu. Rev. Biochem. (2006)

Spatial localisation of protein-protein interactions
The ComPPI resource



comPPI 2.0

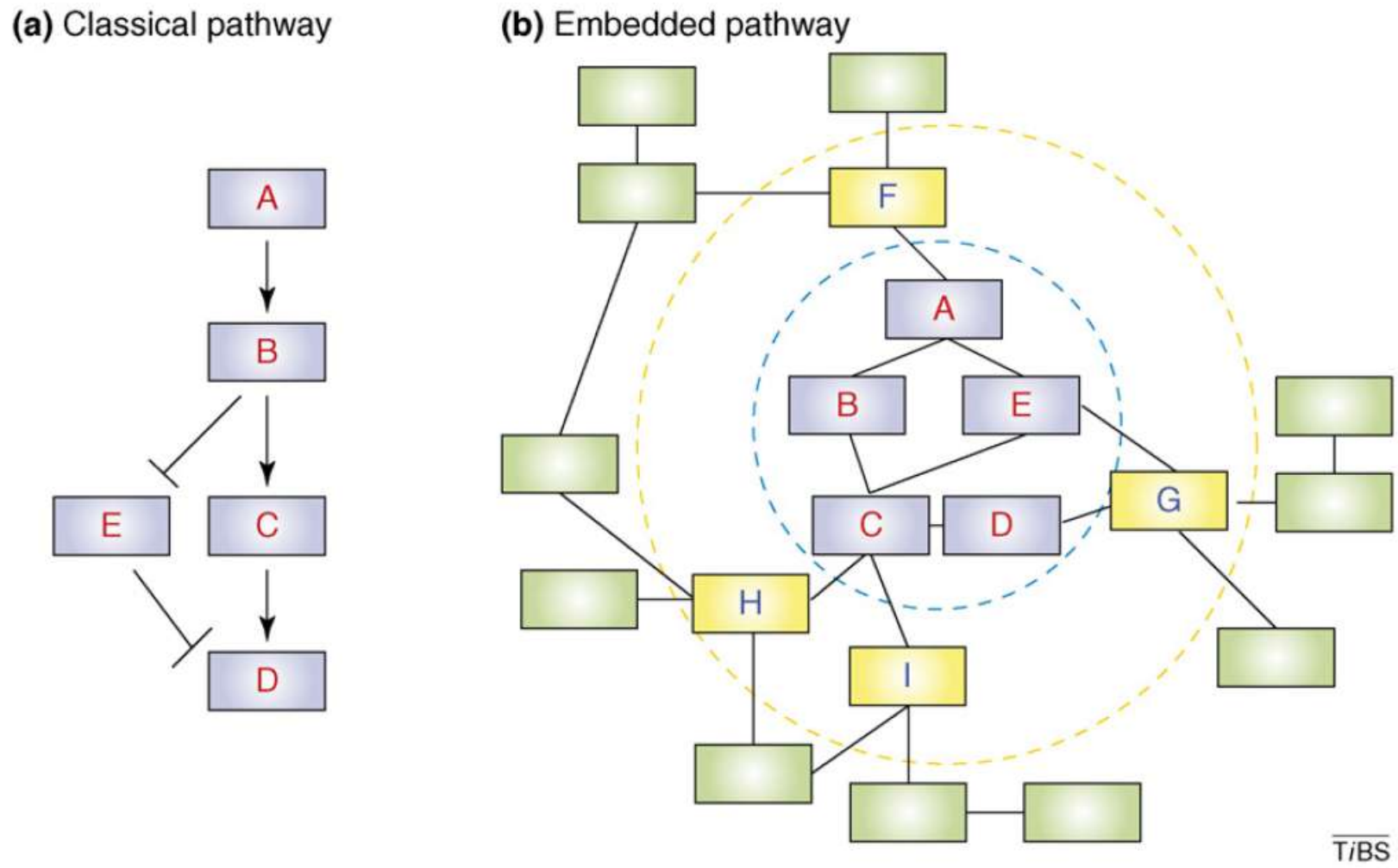
Veres *et al.,* Nucleic Acids Res. (2015)

Signalling pathways and networks

Signalling pathways and networks

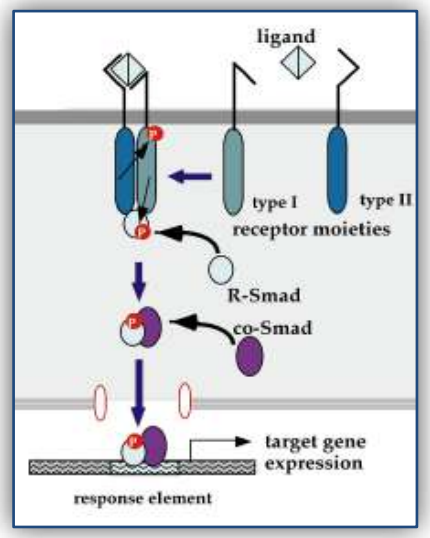
Integration of interaction data

Embedding pathways into interaction networks

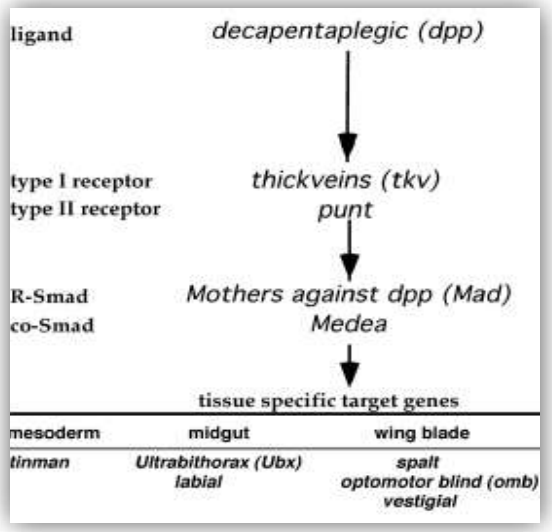


Trends Biochem Sci, **32**: 320–331 (2007)

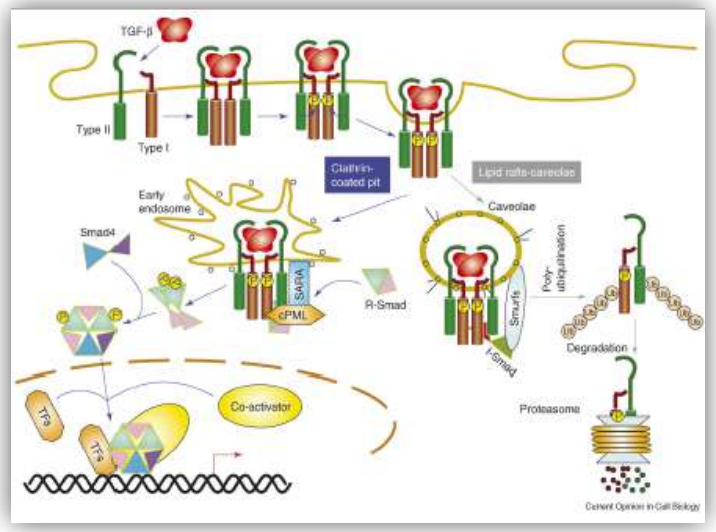
Research approaches



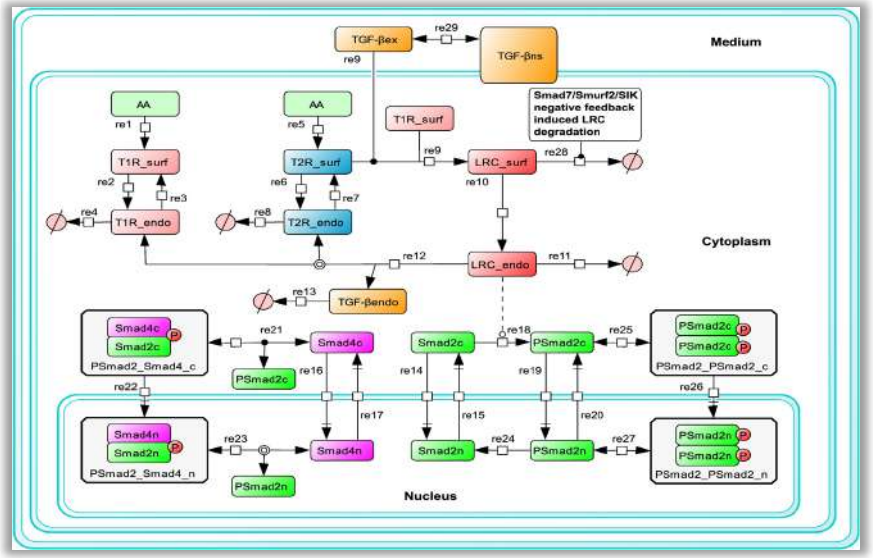
Biochemistry



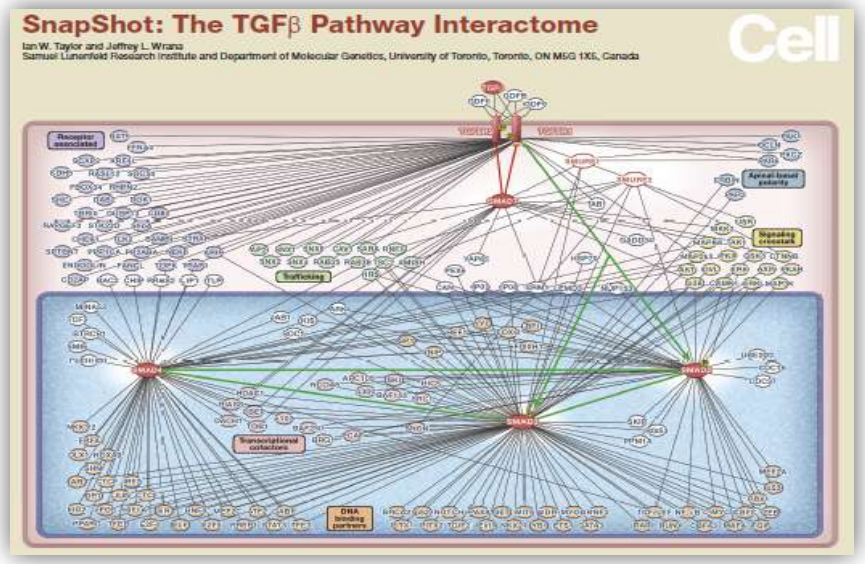
Genetics



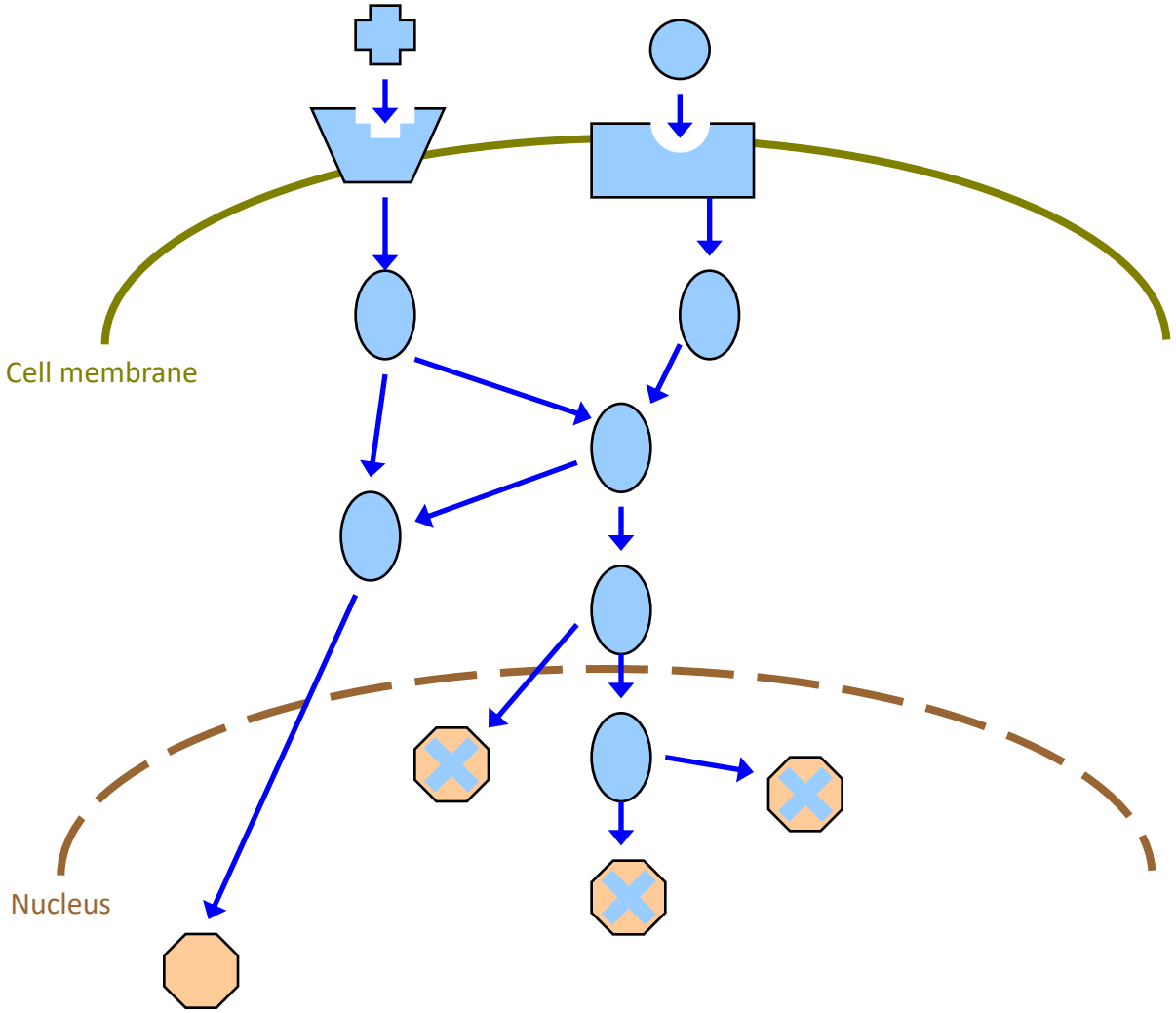
Cell biology



Bioinformatics



Network biology



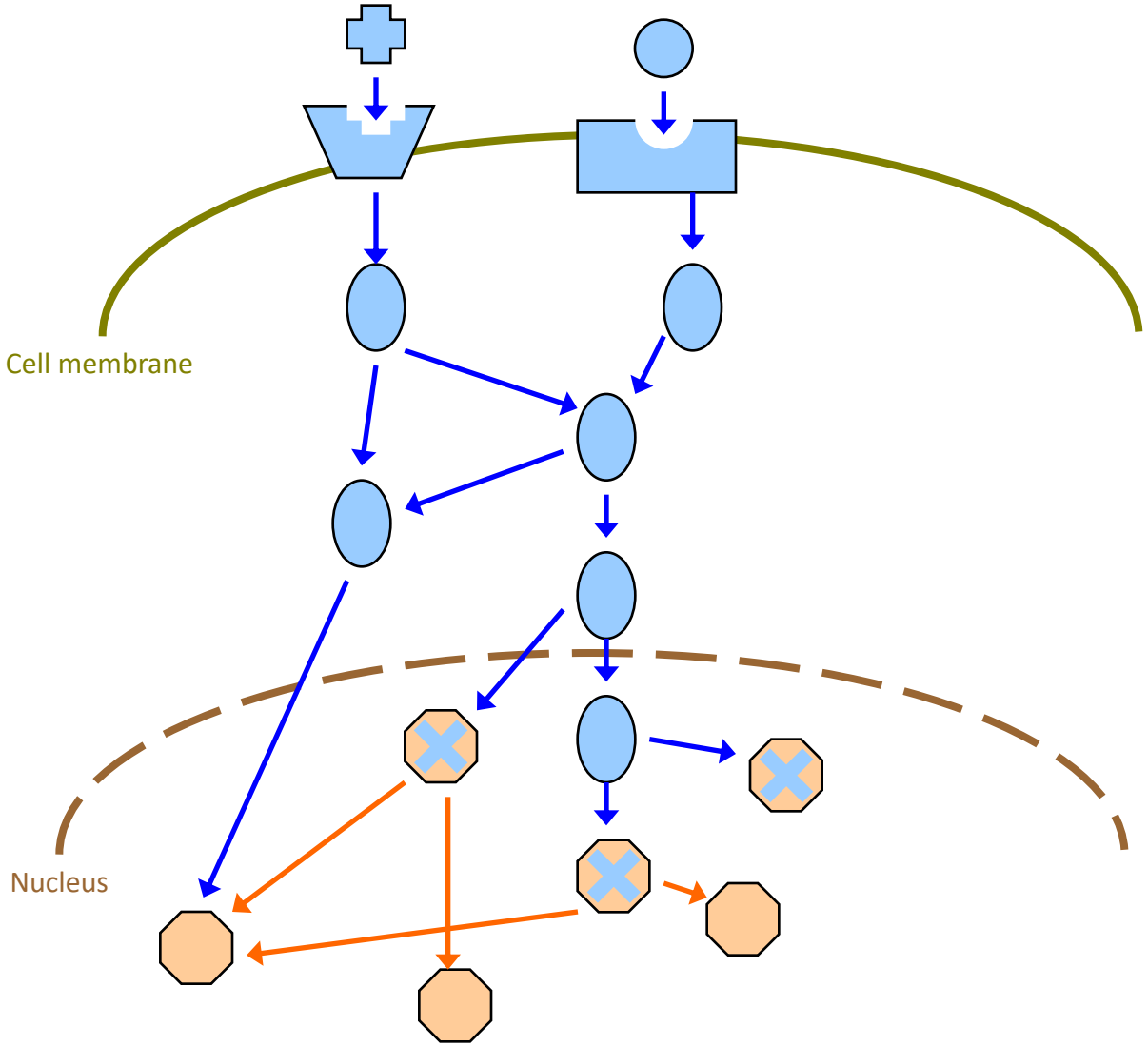
Legends:

Signalling elements

Transcription factors

Types of networks:

Network of signalling pathways



Legends:

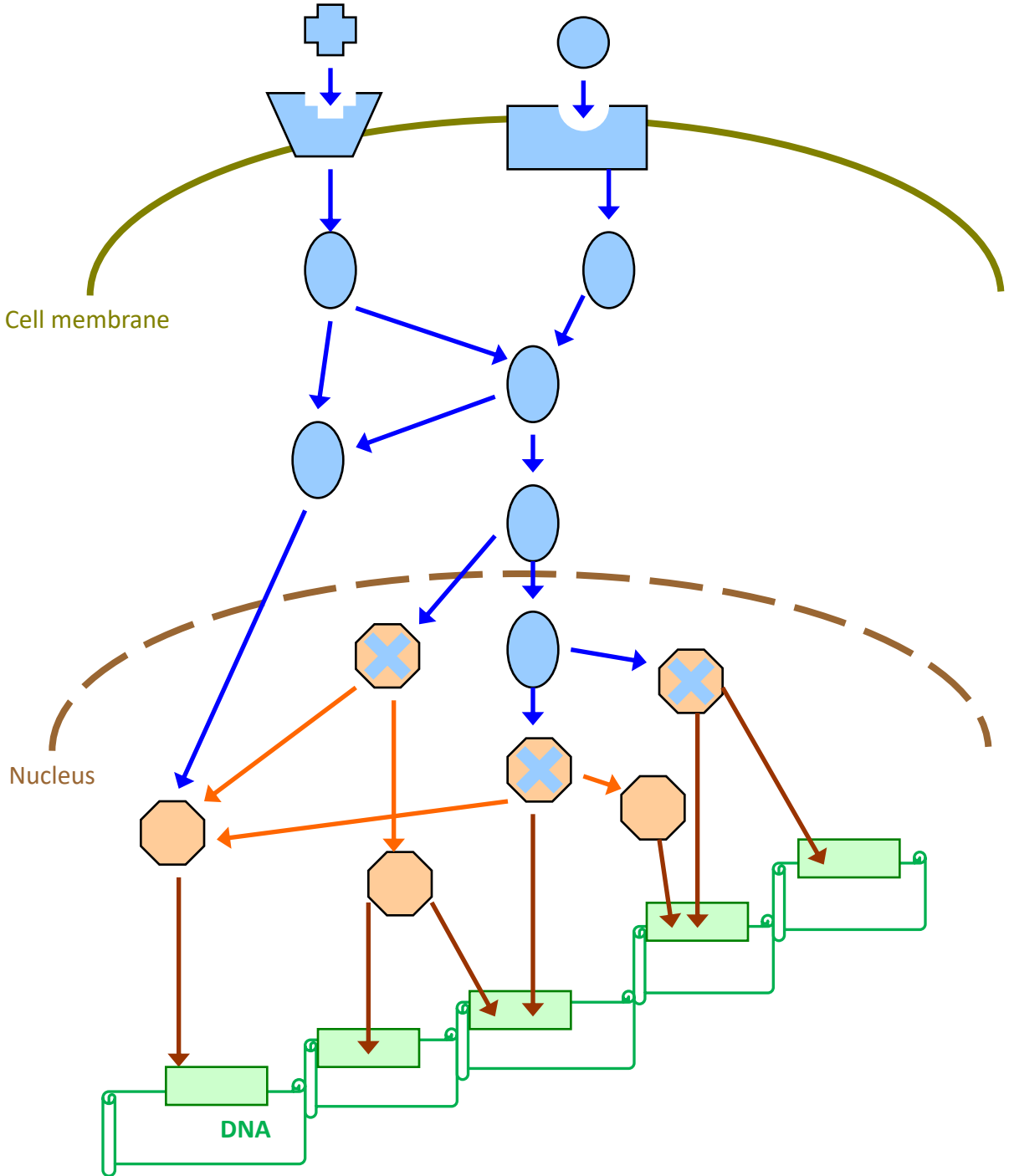
Signalling elements

Transcription factors

Types of networks:

Network of signalling pathways

Network of transcription factors

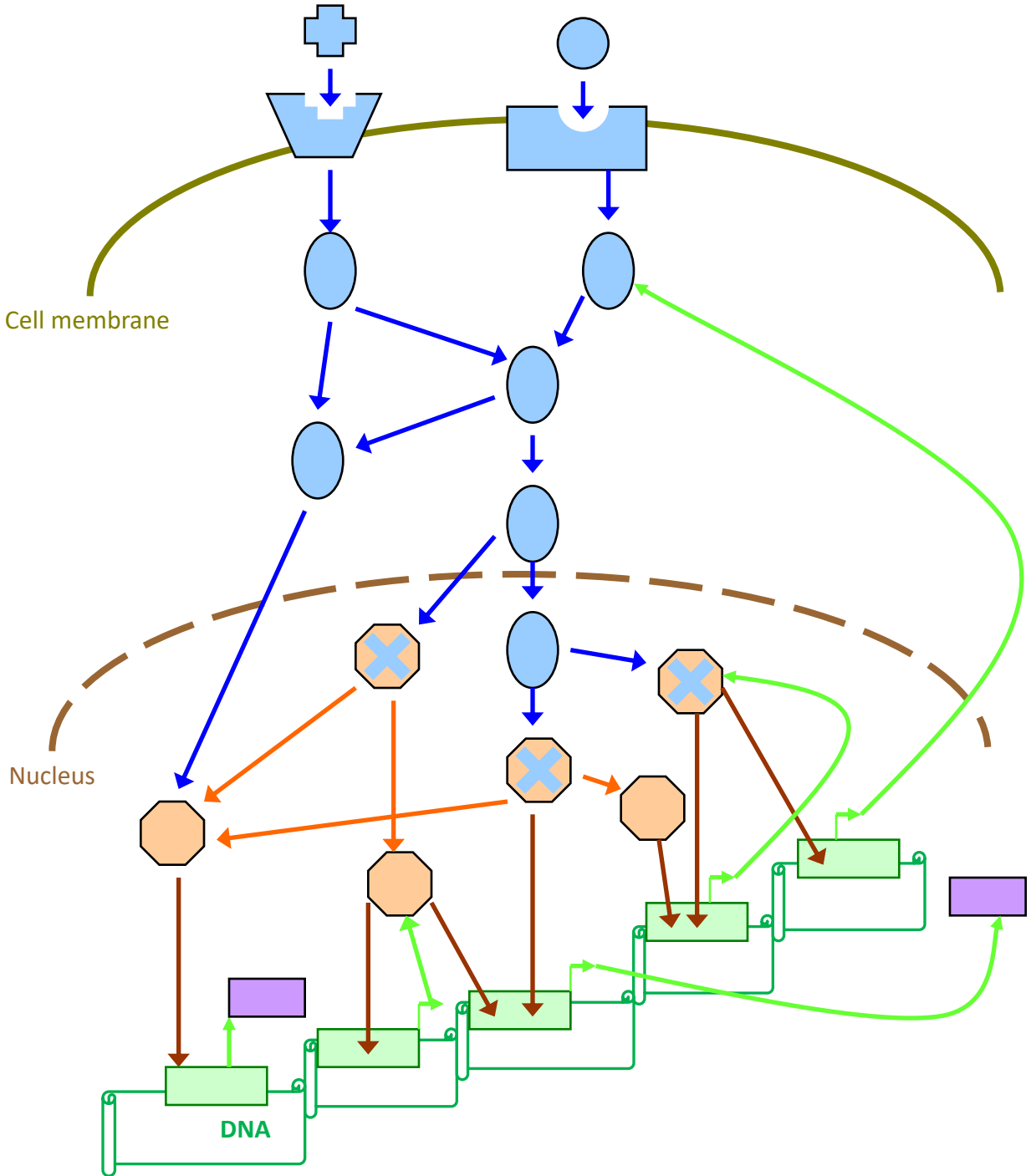


Legends:

- Signalling elements
- Transcription factors
- Promoters

Types of networks:

- Network of signalling pathways
- Network of transcription factors
- Network of transcription factors' binding sites

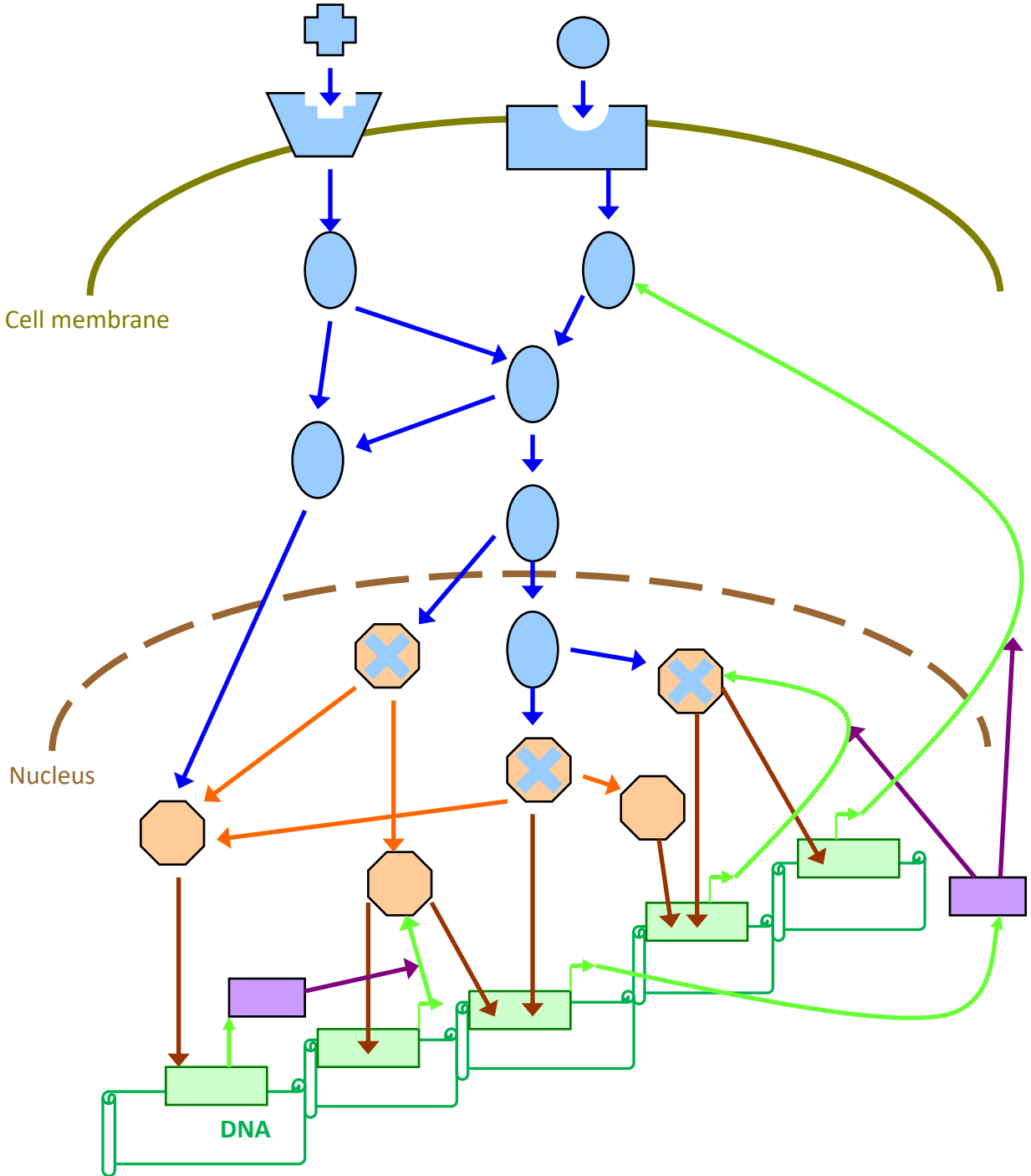


Legends:

- Signalling elements
- Transcription factors
- Promoters
- miRNAs

Types of networks:

- Network of signalling pathways
- Network of transcription factors
- Network of transcription factors' binding sites
- Network of promoters and their transcripts

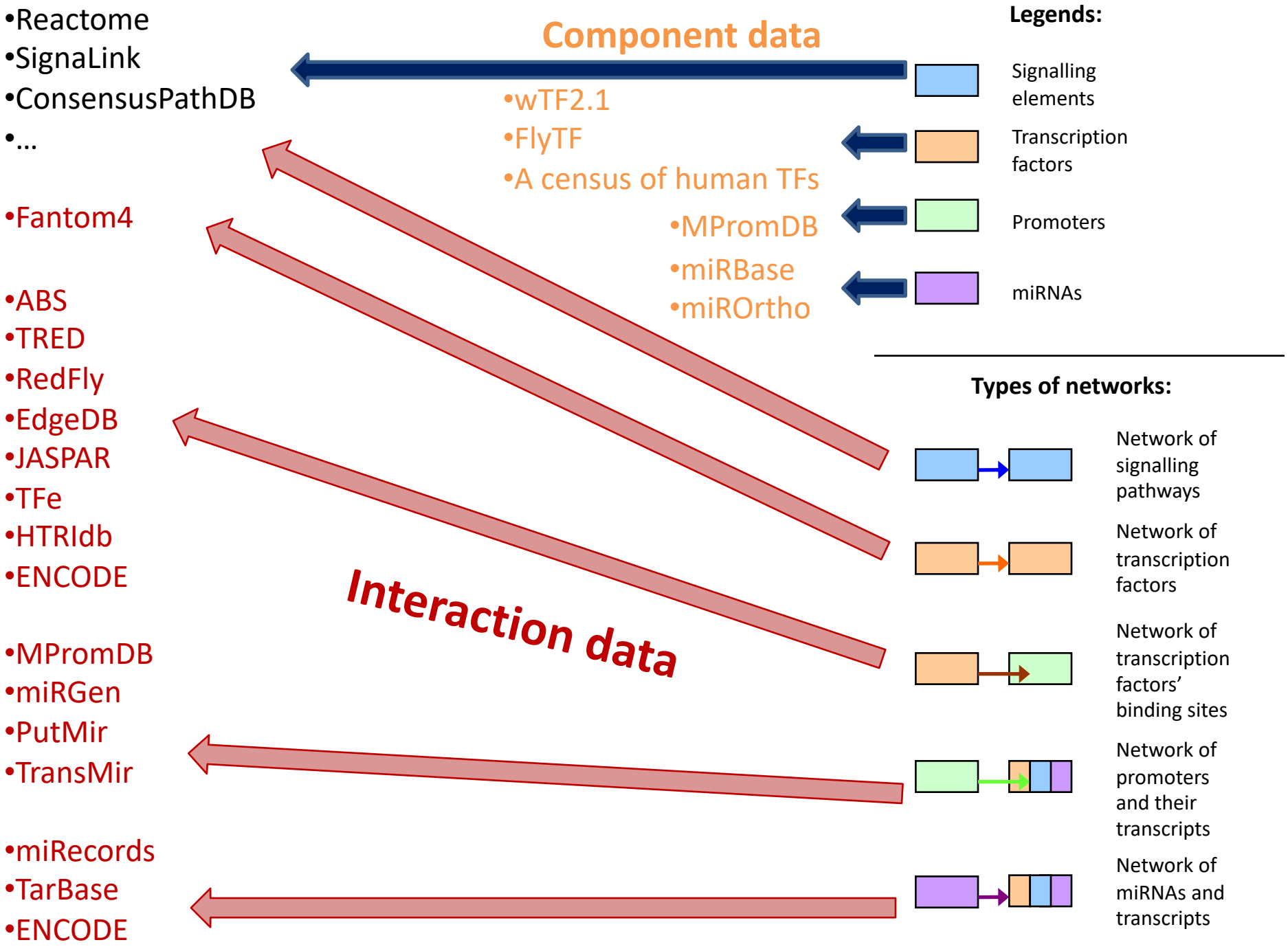


Legends:

- Signalling elements
- Transcription factors
- Promoters
- miRNAs

Types of networks:

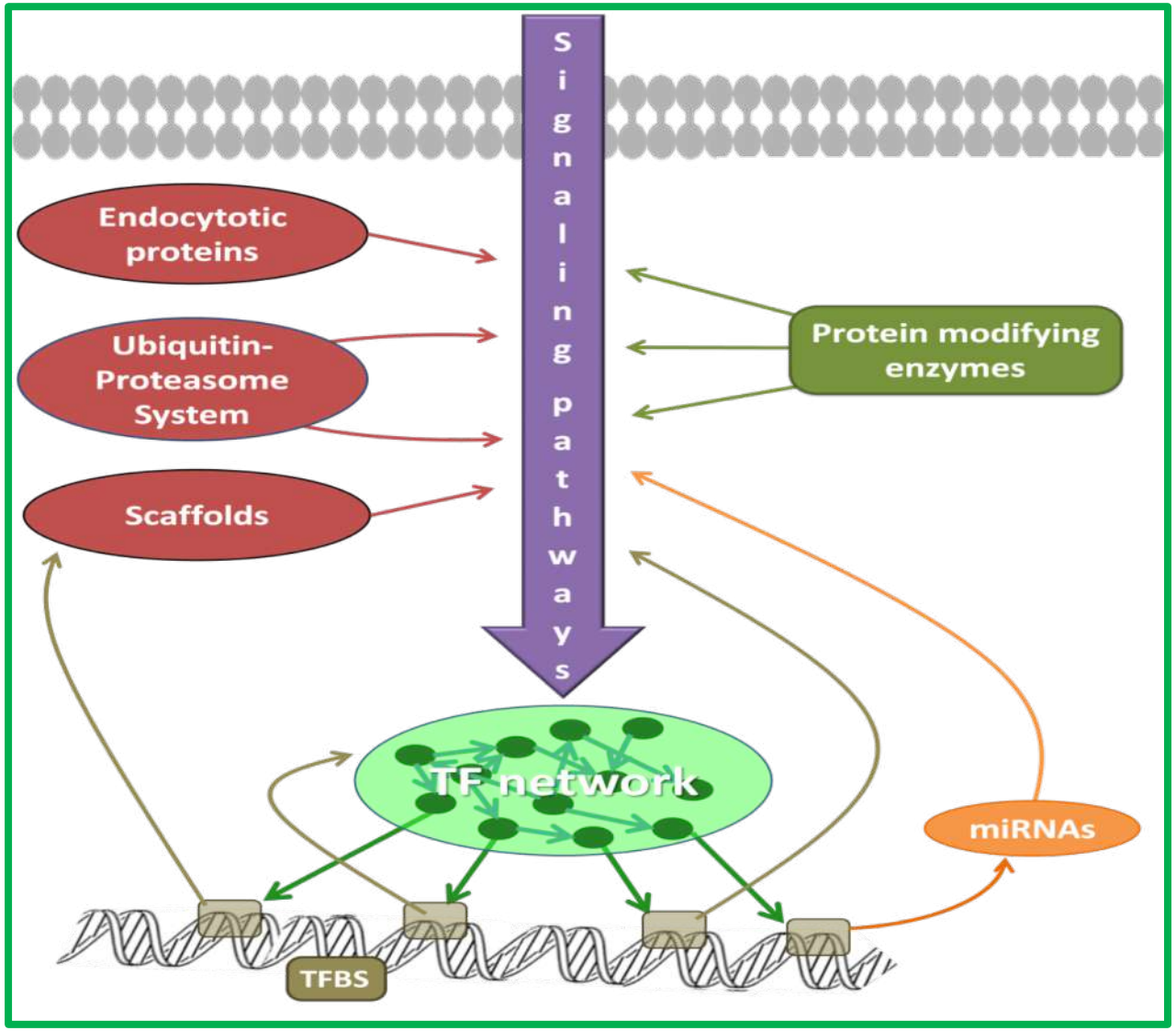
- Network of signalling pathways
- Network of transcription factors
- Network of transcription factors' binding sites
- Network of promoters and their transcripts
- Network of miRNAs and transcripts


•Reactome
•SignaLink
•ConsensusPathDB
•…
Component data
•wTF2.1
•FlyTF
•A census of human TFs
•MPromDB
•miRBase
•miROrtho
Legends:
Signalling elements
Transcription factors
Promoters
miRNAs
•Fantom4
•ABS
•TRED
•RedFly
•EdgeDB
•JASPAR
•TFe
•HTRIdb
•ENCODE
Interaction data
•MPromDB
•miRGen
•PutMir
•TransMir
•miRecords
•TarBase
•ENCODE
Types of networks:
Network of signalling pathways
Network of transcription factors
Network of transcription factors' binding sites
Network of promoters and their transcripts
Network of miRNAs and transcripts

SignaLink 2.0

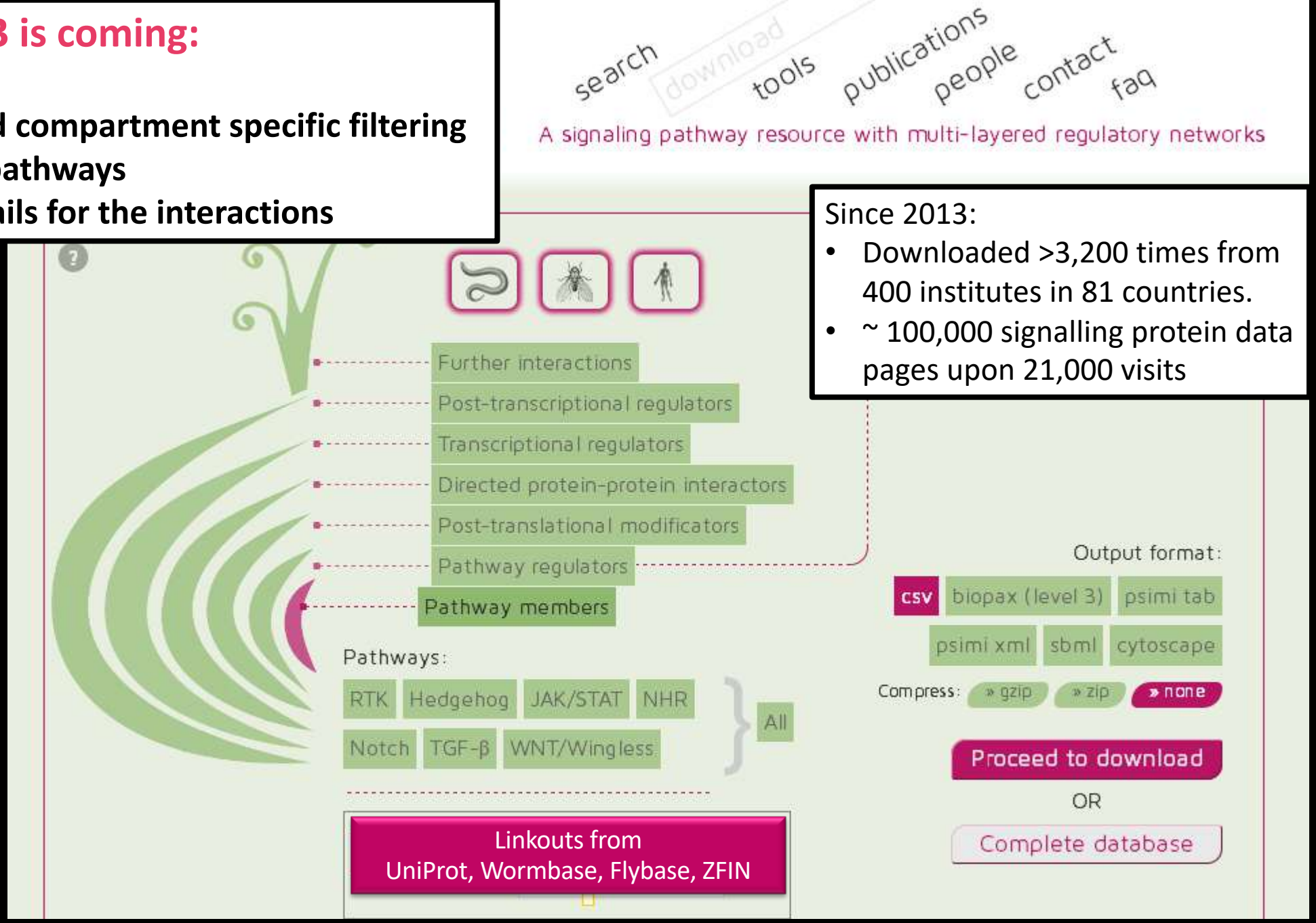
A signalling pathway resource with a multi-layered regulatory network

Fazekas *et al* BMC Syst Biol, 2013



SignaLink 3 is coming:

- lncRNAs
- Tissue and compartment specific filtering
- Immune pathways
- More details for the interactions



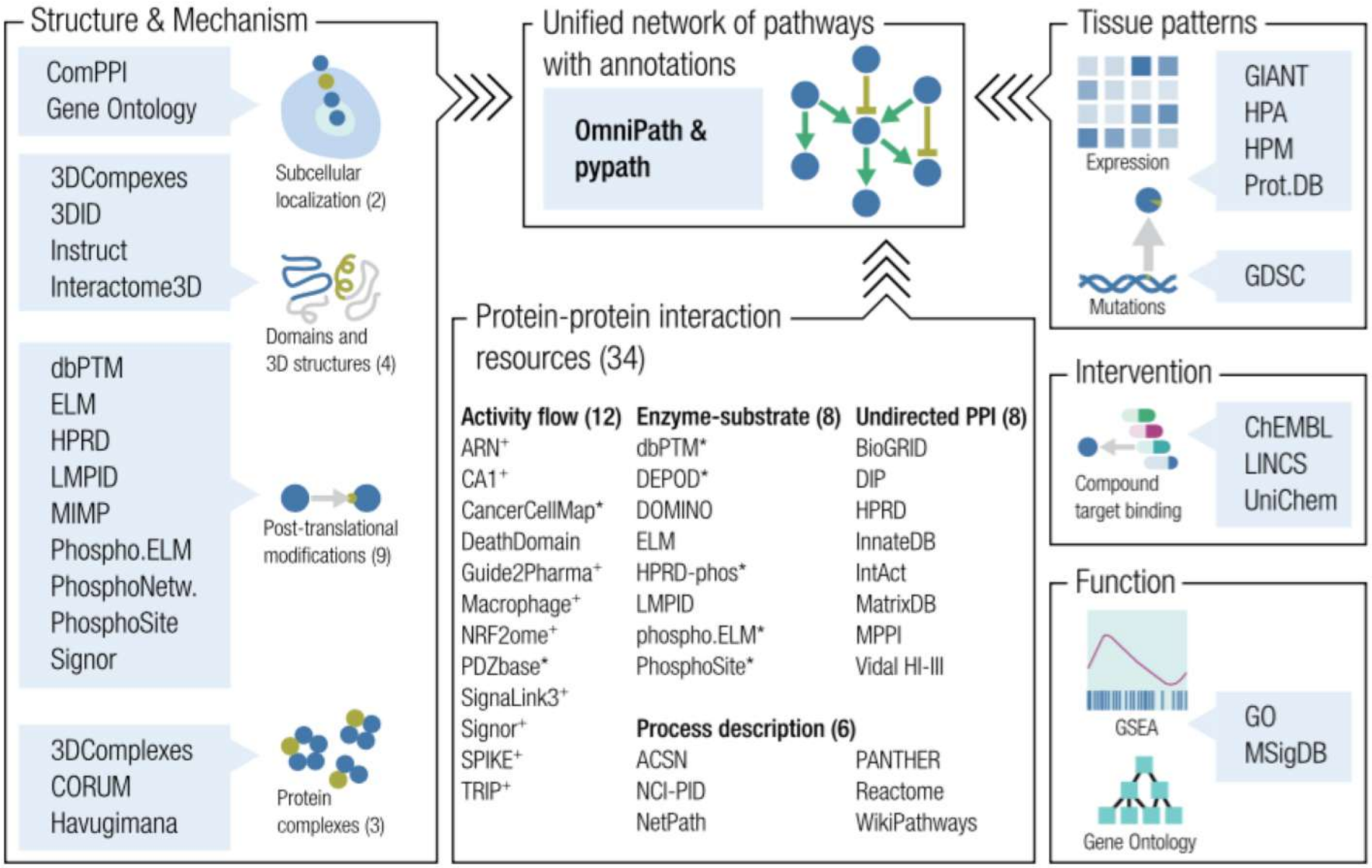
Since 2013:

- Downloaded >3,200 times from 400 institutes in 81 countries.
- ~ 100,000 signalling protein data pages upon 21,000 visits

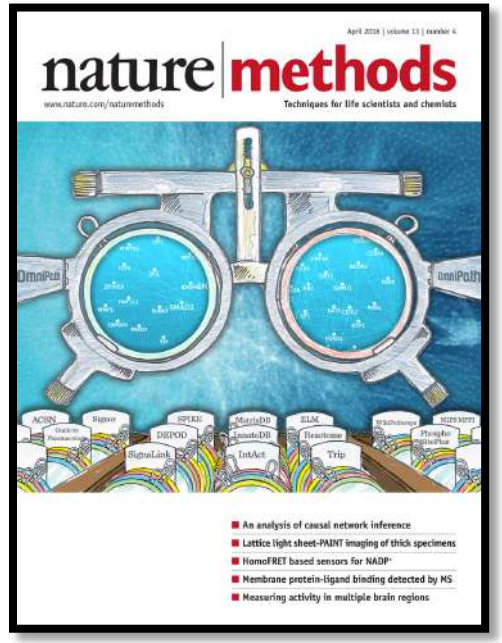
http://SignaLink.org

Main developer:
Saez-Rodriguez group

OmniPath: gateway and guide for literature curated signalling pathway resources *in human*



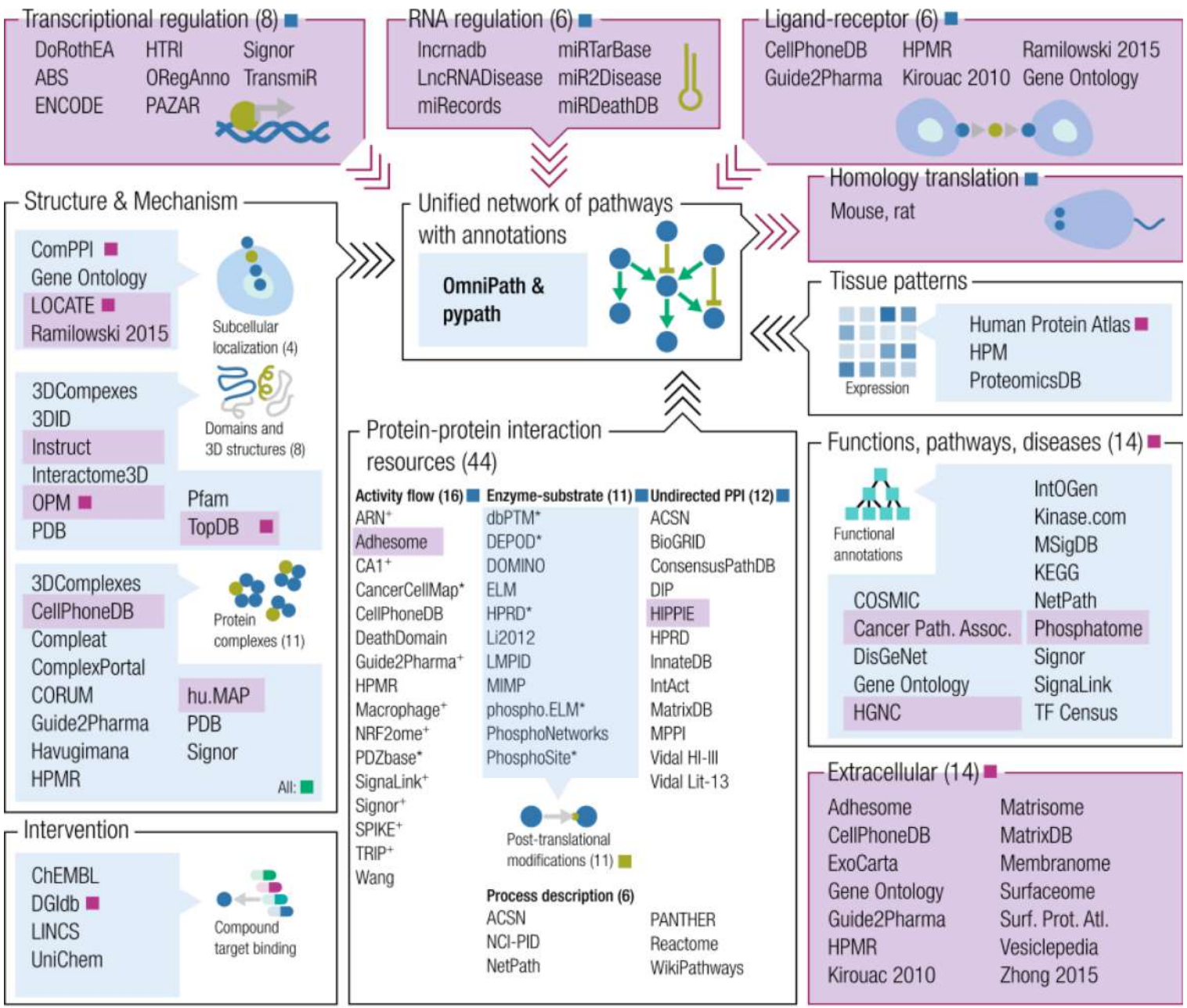
OmniPathDB and pyPath
http://omnipathdb.org



Turei *et al, Nature Methods*, 2016
(cover story)

Main developer:
Saez-Rodriguez group

OmniPath2



OmniPathDB and pyPath
http://omnipathdb.org

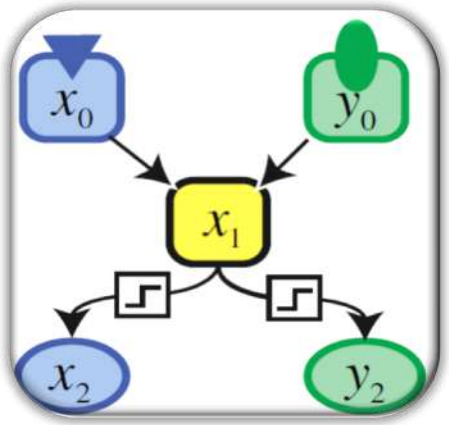
OmniPathR

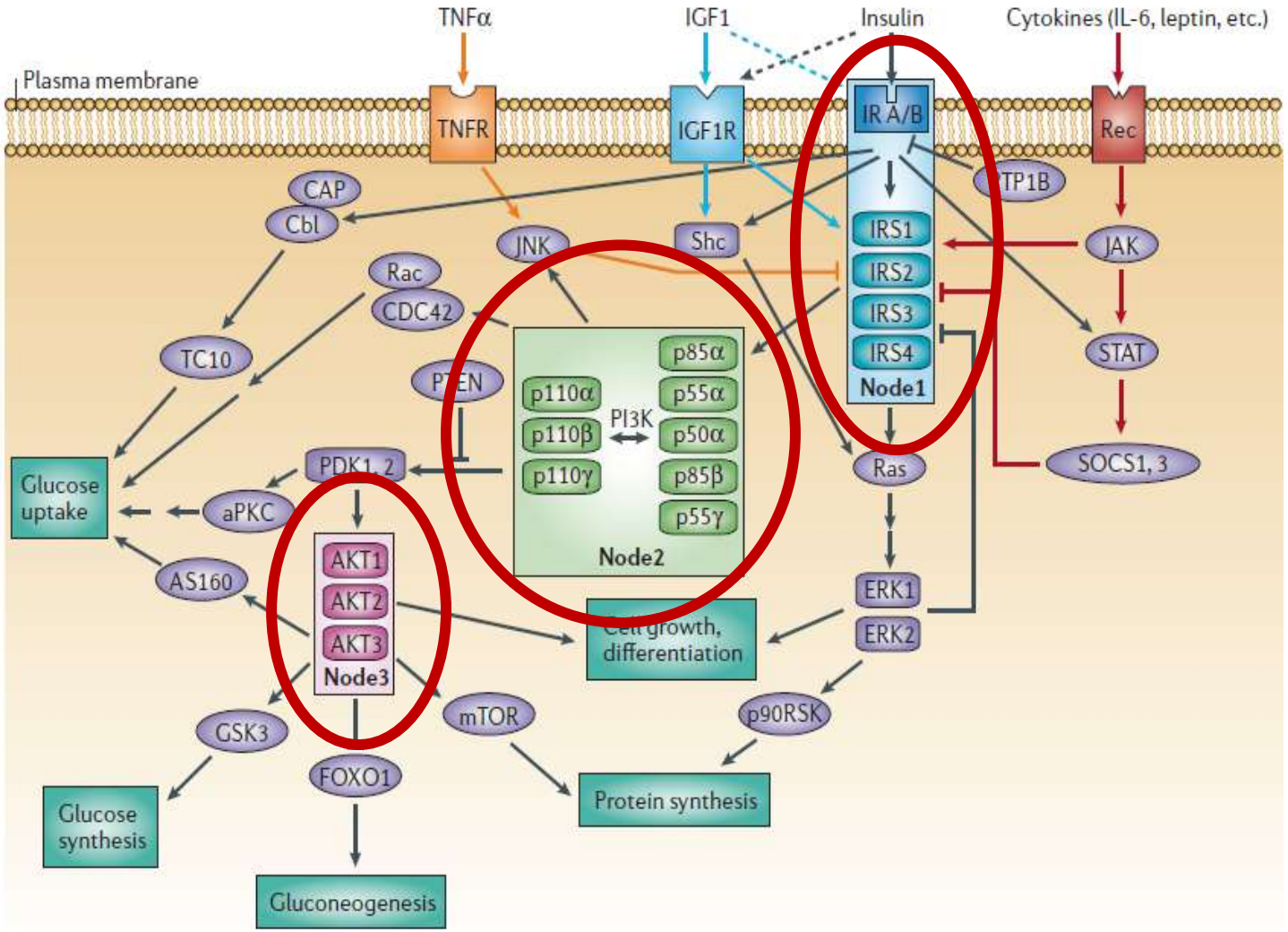
OmniPath-CytoscapeApp

Signalling pathways and networks

Regulation of the signalling flow on the PPI-level

How cross-talks and signalling flow are regulated?

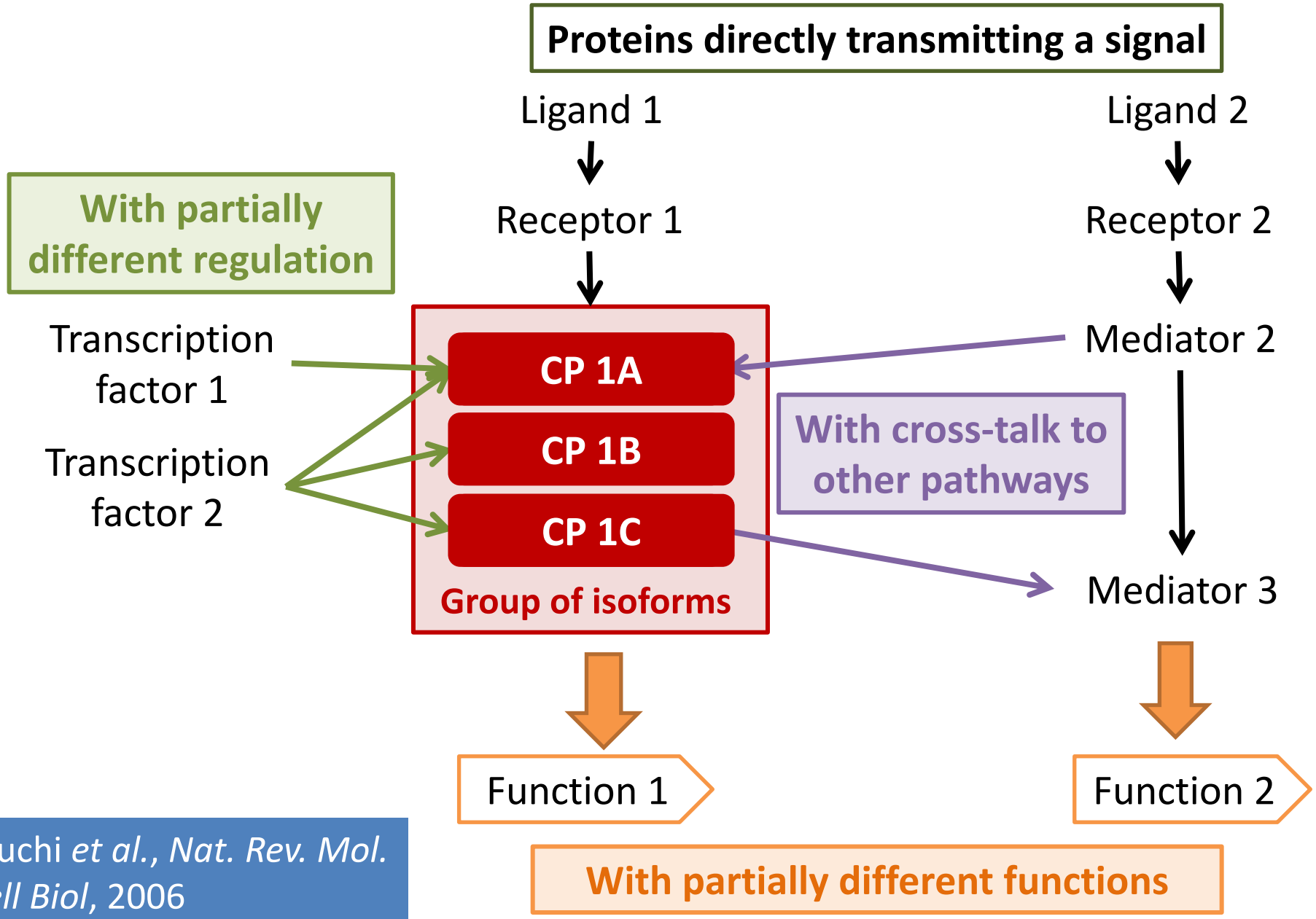




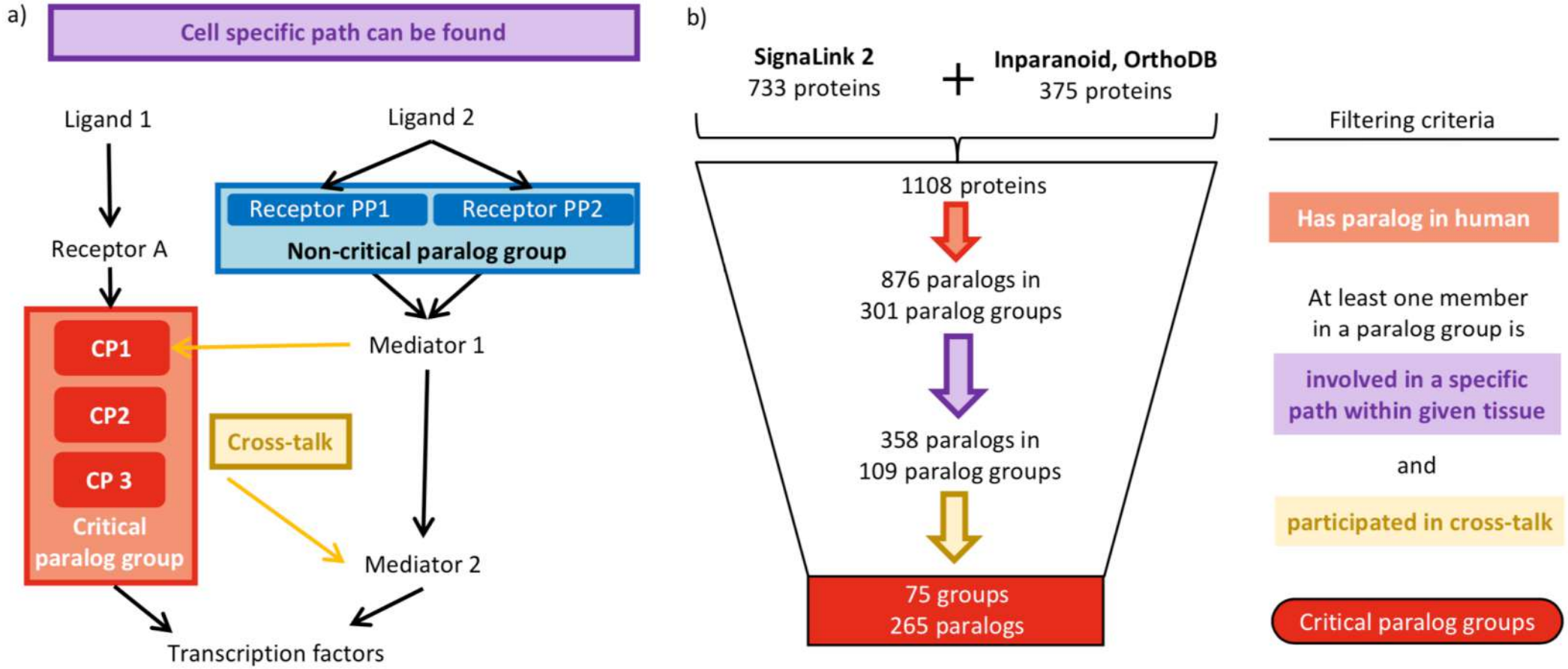
Critical nodes

Taniguchi et al, Nat. Rev. Mol. Cell Biol, 2006

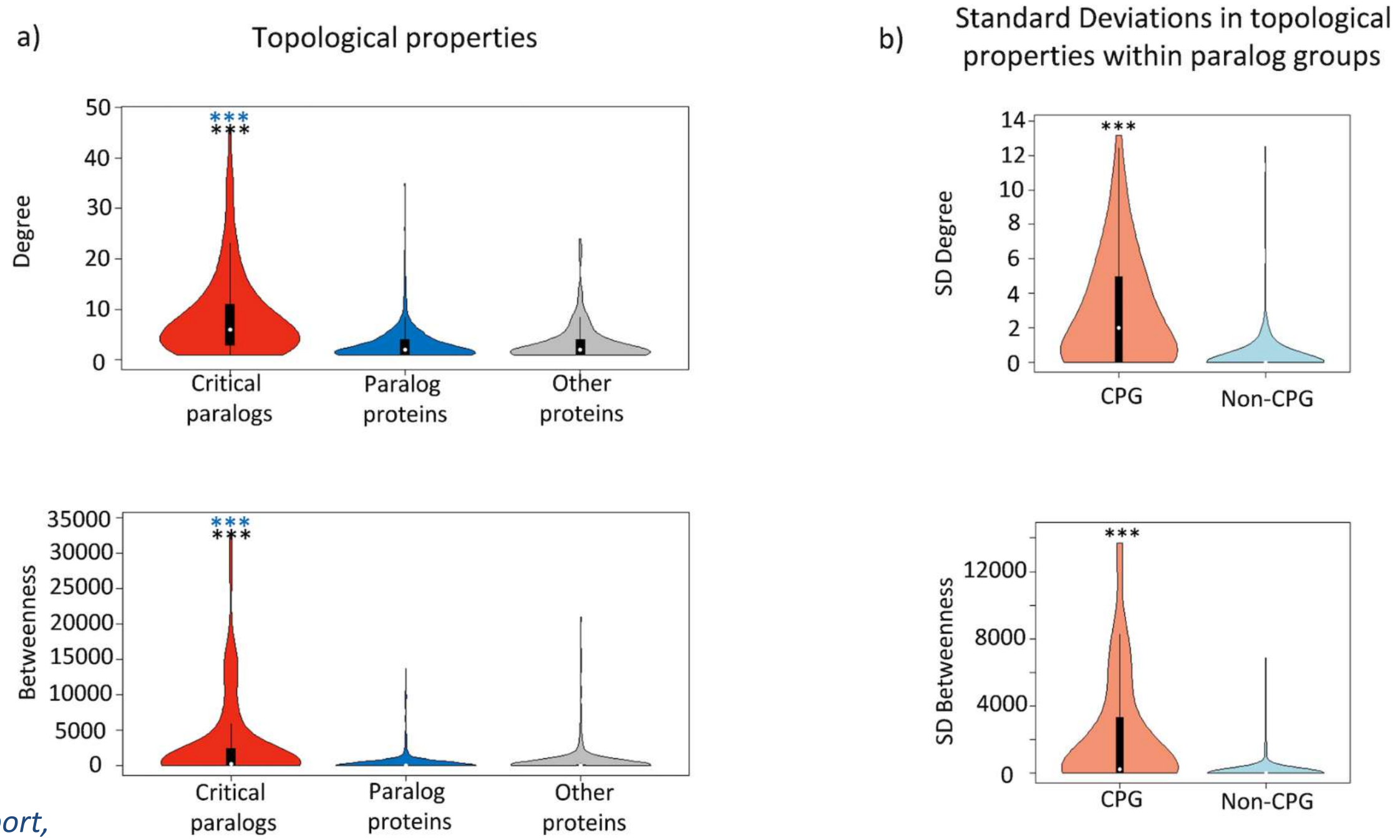
The concept of critical node proteins (CNPs)



Based on Taniguchi *et al.*, *Nat. Rev. Mol. Cell Biol*, 2006

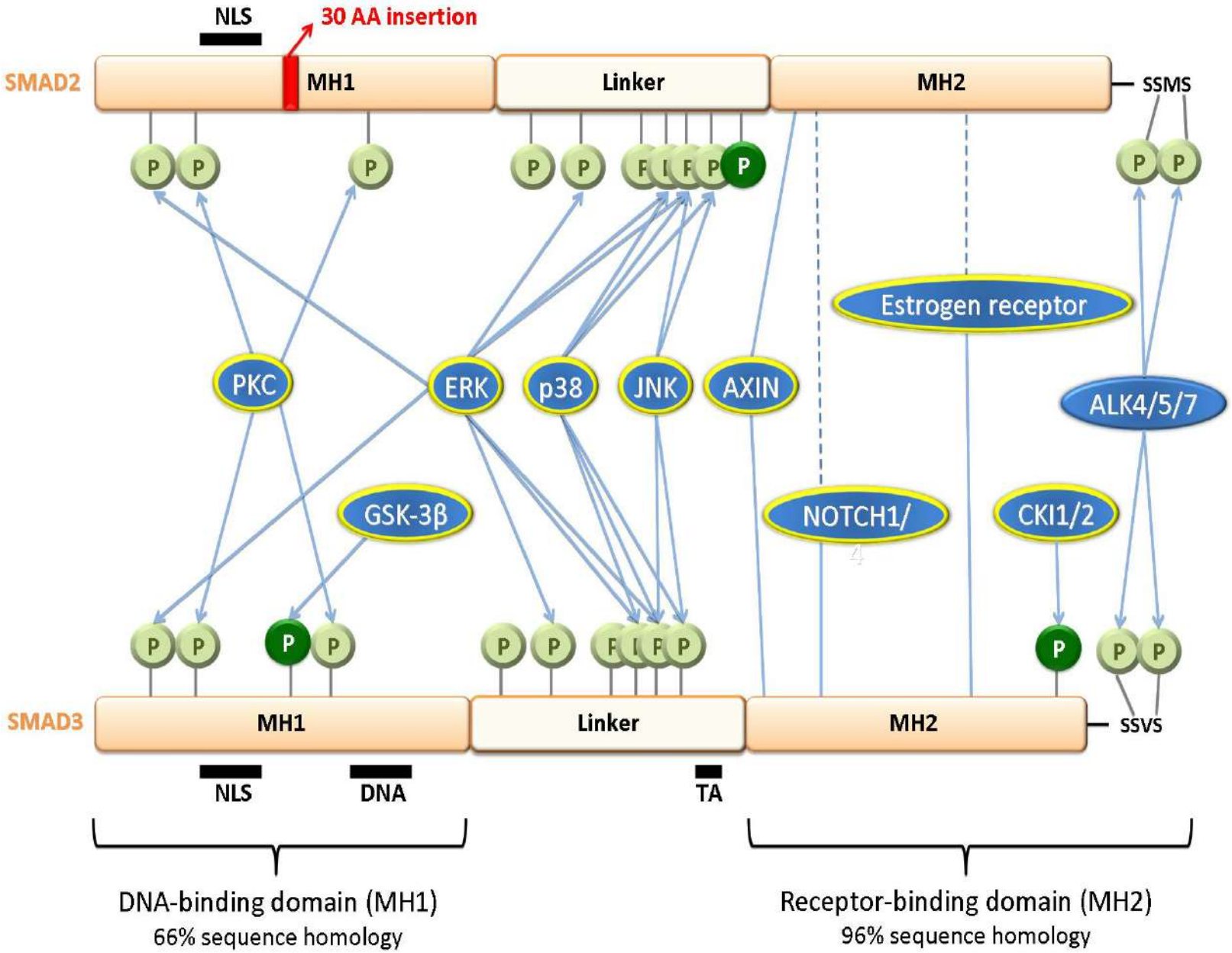


Modos *et al*, *Scientific Report*, 2016

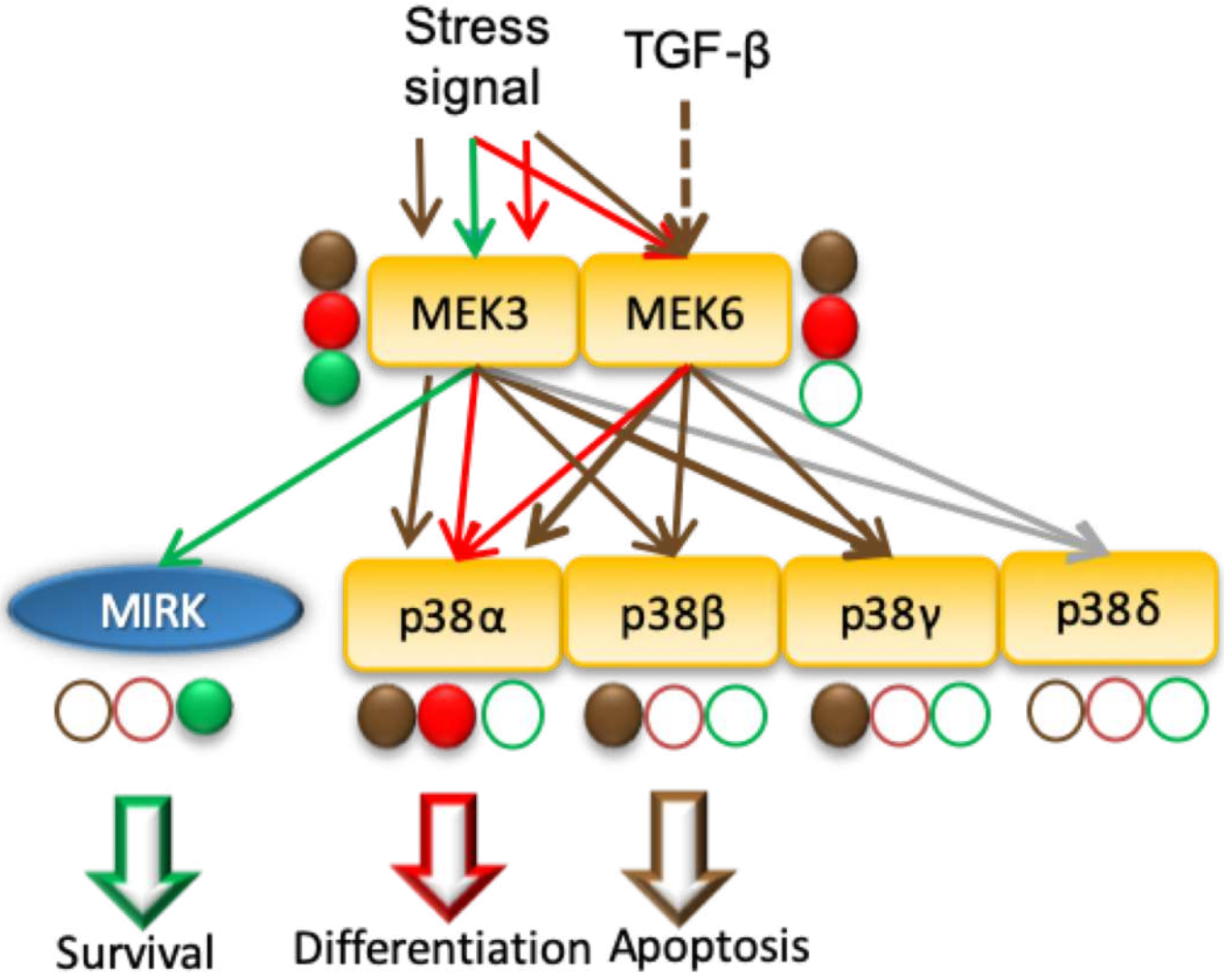


Modos *et al*, *Scientific Report*, 2016

Example 1: The SMAD2/3 critical node



Example 2: The MEK3/6 and p38 critical nodes



Peripheral Nervous System

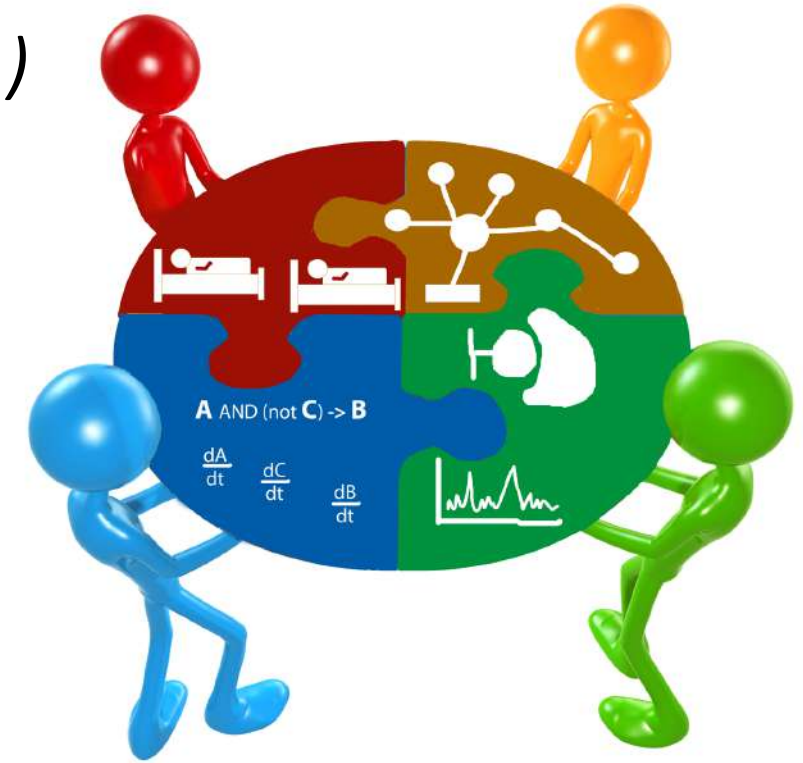
Cardiovascular

Liver and gall bladder

Protein-protein interactions in Network medicine

Network medicine – a promising history

- "Network medicine" *(Barabasi, 2007)*
- "Systems medicine is finally coming of age" *(Lemberger, 2007)*
- Network as the target *(Pawson and Linding, 2008)*
- "Think globally, act locally" *(Barabasi, Loscalzo, 2011)*
- Nowadays considered as a resource for
  - biomarker discovery
  - drug target prediction
  - drug side-effect analysis
  - drug repurposing
  - suggesting new therapies
  - patient-stratification



Korcsmaros *et al, Integrative Biology*, 2017

Which genes are mutated, which ones have differential expression in cancer?

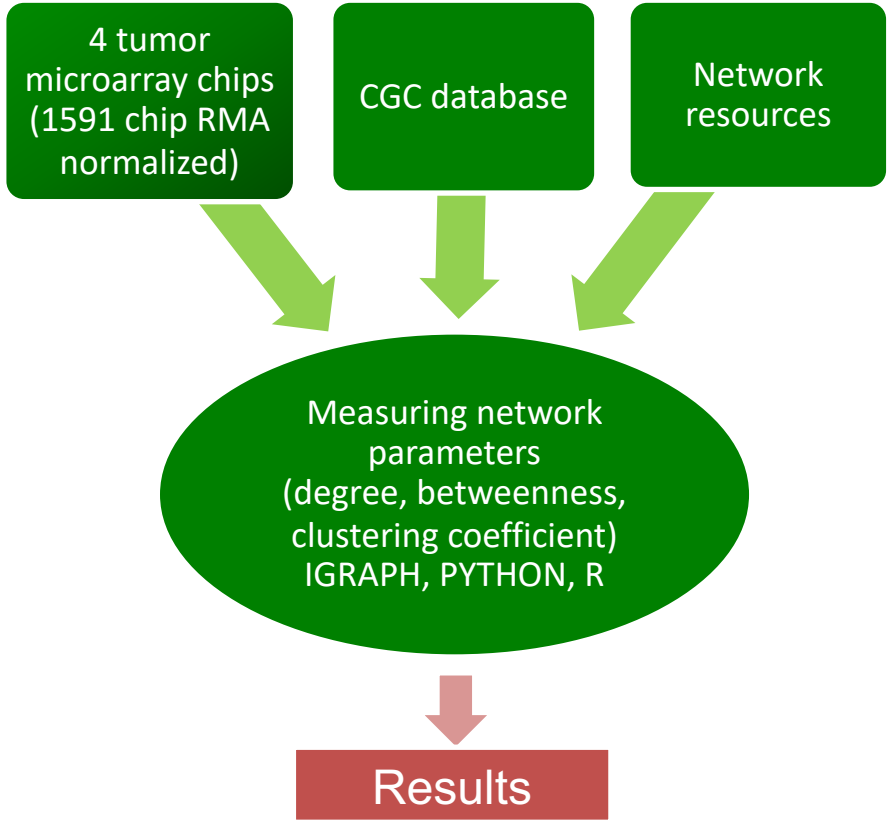
Are there hidden ‘bad guys’ among the interactors of these cancer-related proteins?

Resources & Methods

- **Cancer Gene Census:** Known driver cancer genes in 4 different cancer types separately: **colorectal carcinoma, breast carcinoma** (include all types), **non-small cell lung carcinoma, hepatocellular carcinoma**

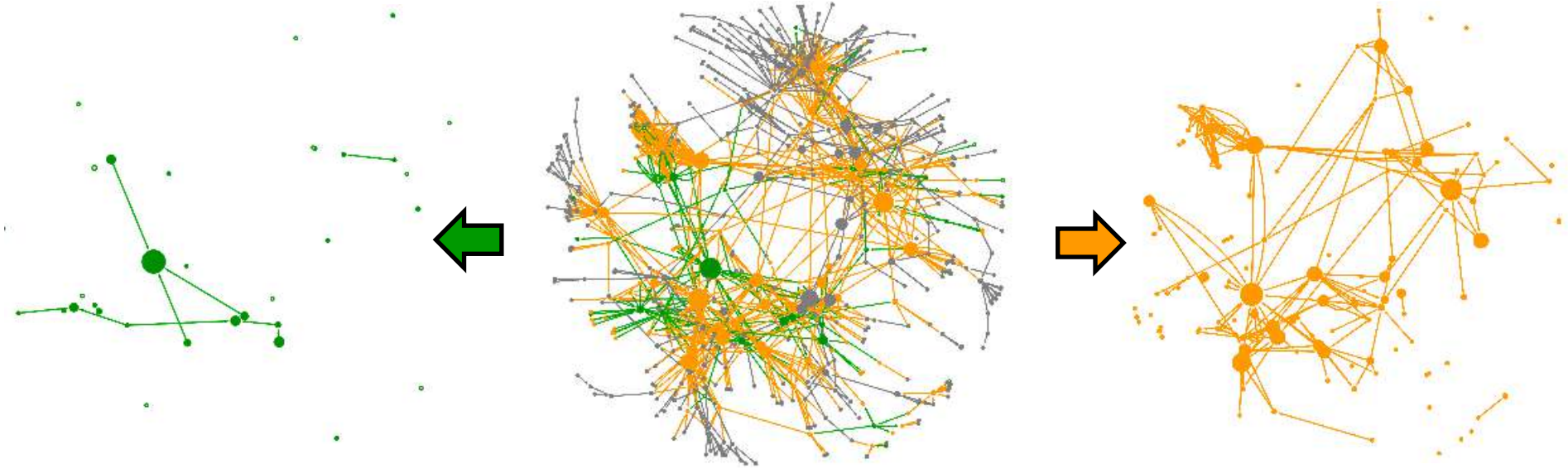
- **Gene Expression Omnibus**: Affymetrix HGU 133 plus2 microarray data from **4 different** cancer types separately.

- **Network resources - separately:**
  3 signalling networks:
  SignaLink2, Cui. et al Mol. Syst. Biol. 2007, Reactome
  2 protein-protein interaction networks
  (Biogrid+DIP+IntAct, HPRD)



➢ **Tissue- and cancer-specific networks based on expression data**

First neighbours act like glue in the colorectal carcinoma network



Cancer-related proteins
The network falls apart

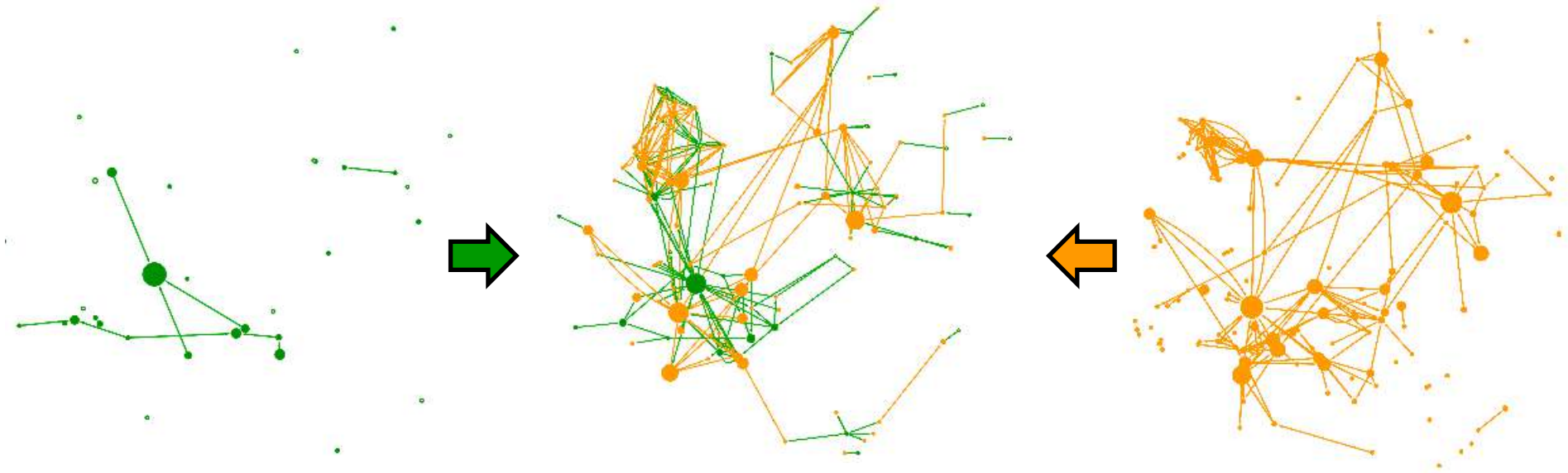
Whole manually curated interaction network of colon cancer

Colon cancer-related
First neighbour
Unaffected

First neighbours
One giant component, bigger than by chance*

*Exact test with sampling on same network

First neighbours act like glue in the colorectal carcinoma network



Cancer-related proteins
The network falls apart

One giant component
First neighbours act like glue
and connect the cancer-related proteins*

First neighbours
One giant component,
bigger than by chance*

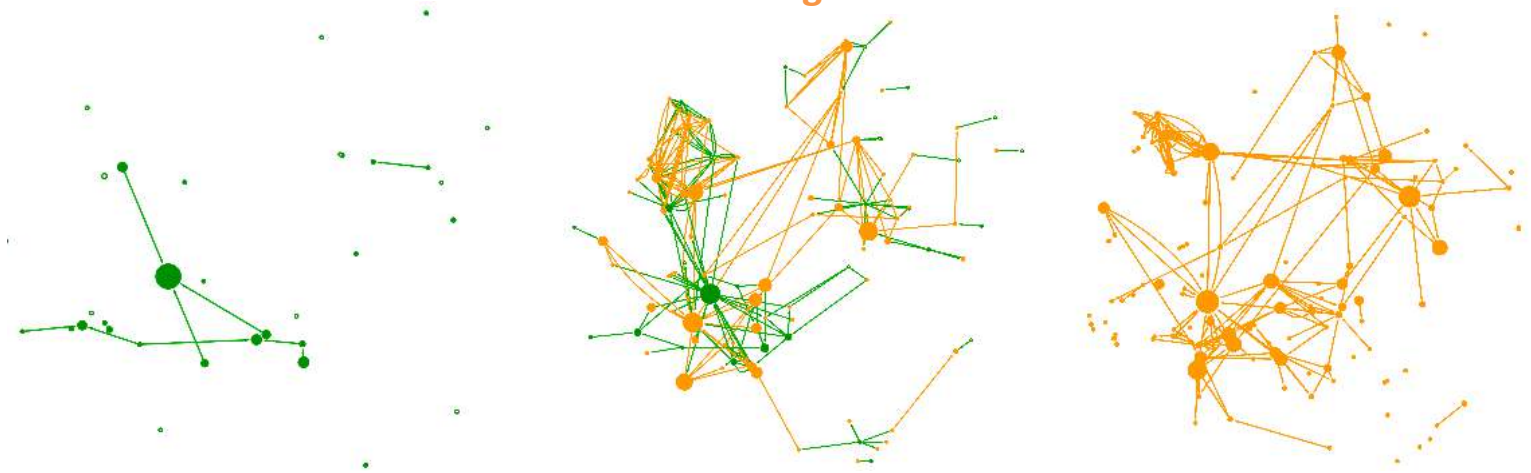
Colon cancer-related
First neighbour
Unaffected

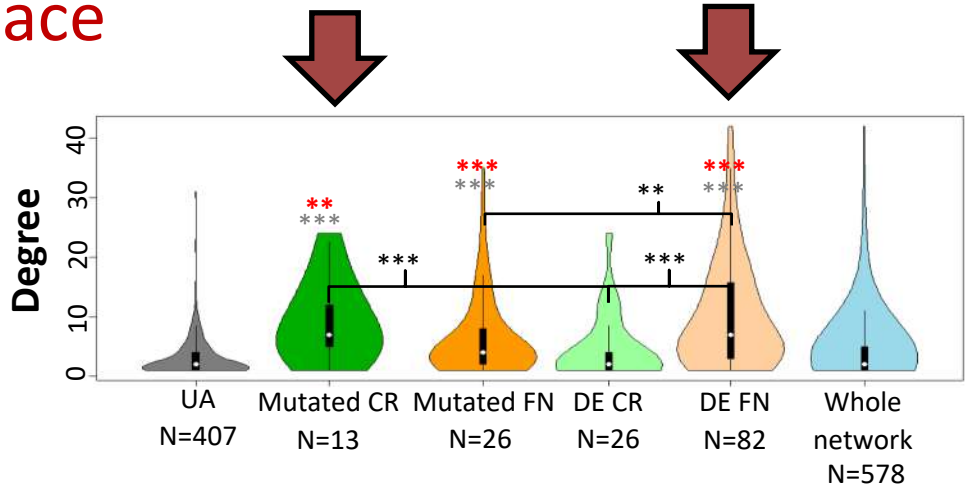
*Exact test with sampling on same network

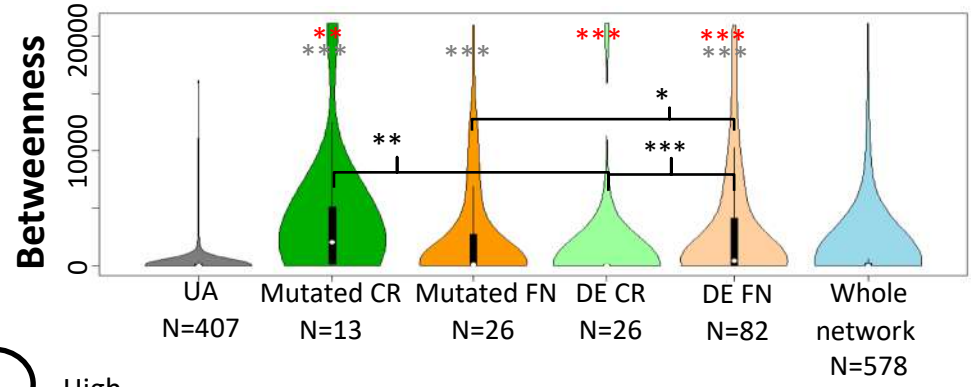
Interactors of cancer-related proteins have key influence on carcinogenesis and increase the drug target space

Colon, liver, lung and breast cancer

Cancer-related → Cancer-related and first neighbour ← First neighbour

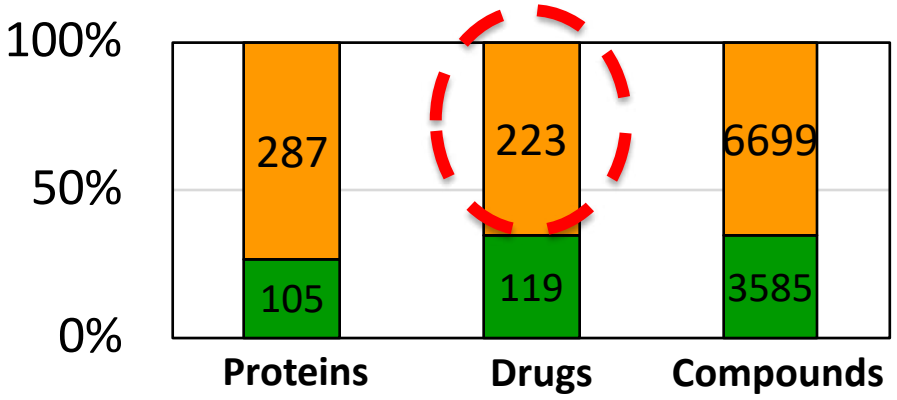




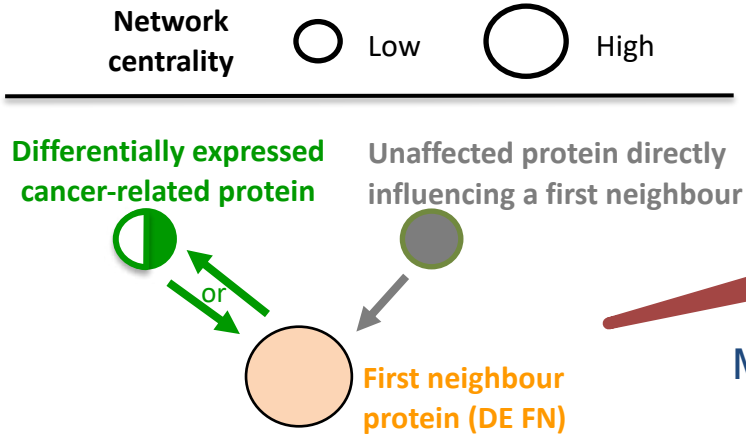


Drug repurposing

Only 27 (12%) have currently applied in oncology



Drug targeting



Cancer mimicking target selection strategy

Modos *et al*, Nature Partner Journal (npj) Systems Biology and Applications, 2017

Conclusions of the first neighbour analysis

- **First neighbours** of cancer-related proteins are at least as **important as cancer-related proteins** themselves.

- Complementary strategy in cancer to rewire signalling networks:
  - **Mutations directly** affecting the **more central** proteins
  - **Differential expression** affecting the **first neighbours of central proteins**

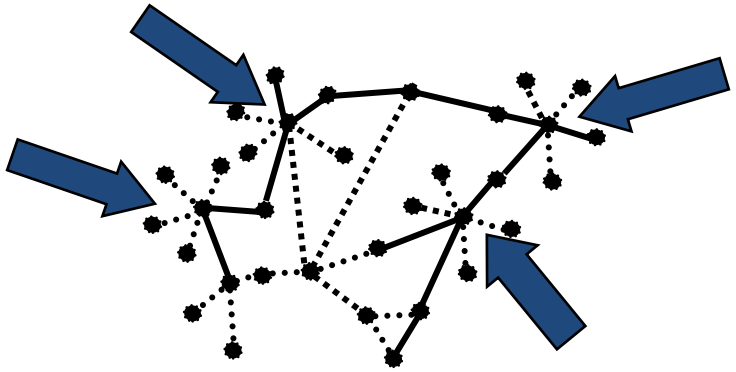
- Drug repurpose potential for some non-oncology drugs

- Cancer mimicking strategy by **targeting influencer proteins** (2nd neighbours of a cancer-related protein)

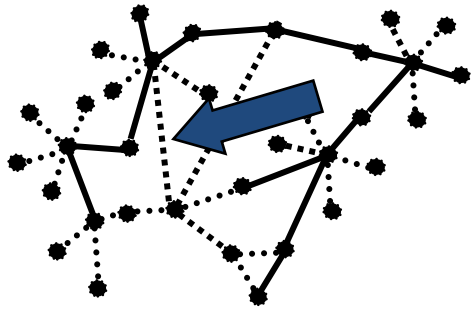
Modos *et al*, Nature (npj) Systems Biology and Applications, 2017

Protein-protein interactions in Network pharmacology

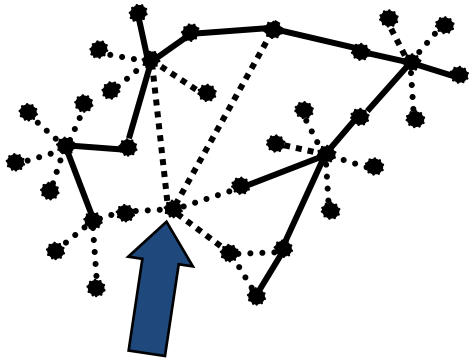
Network-based drug target options



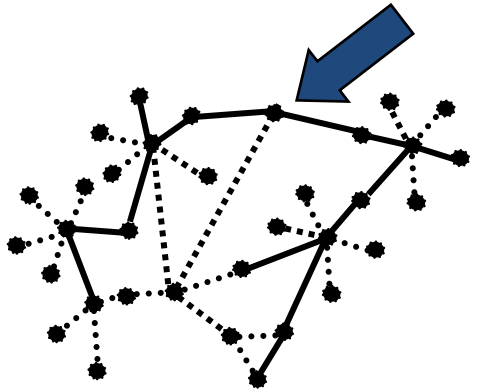
hubs



hub-connections

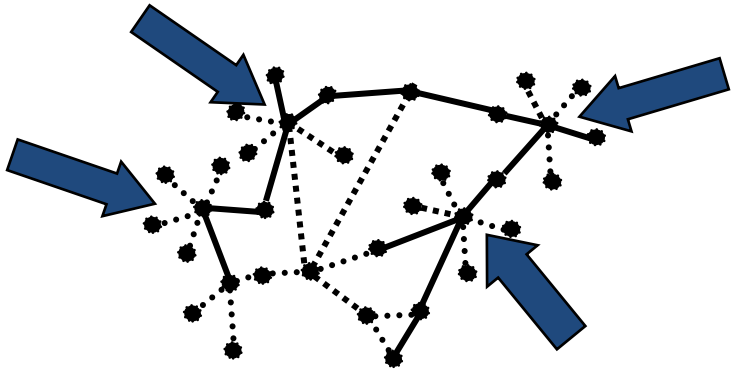


overlapping elements



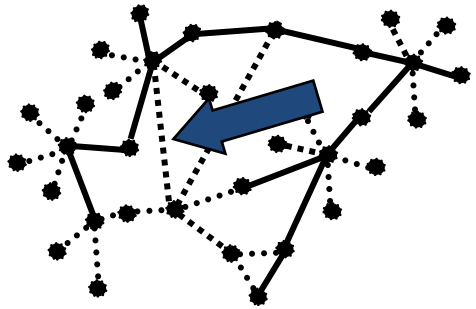
bridges

Network-based drug target options in signalling networks



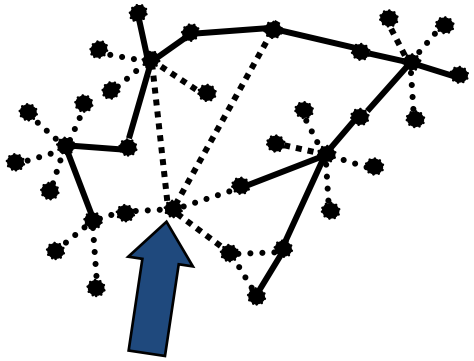
hubs

central components of pathways



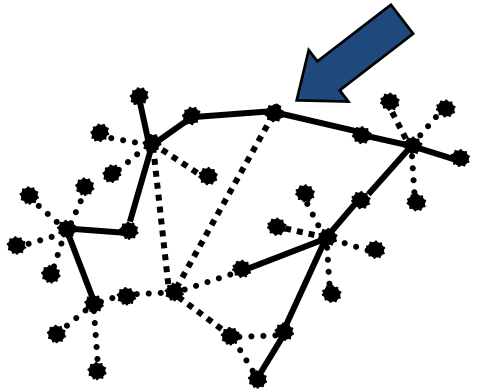
hub-connections

cross-talks



overlapping elements

multi-linker



bridges

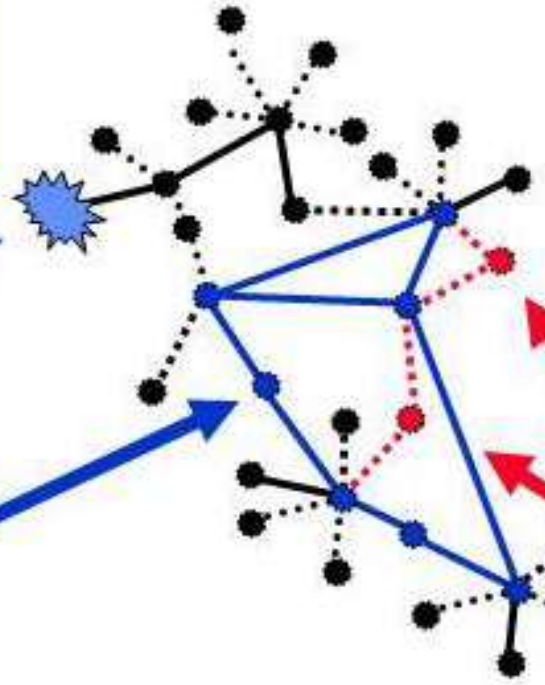
linker

multi-pathway proteins

Drug design strategies

The central hit strategy:
to kill rapidly growing cells
having flexible networks
with large dissipation

[choke point targeting]



specific, high centrality node

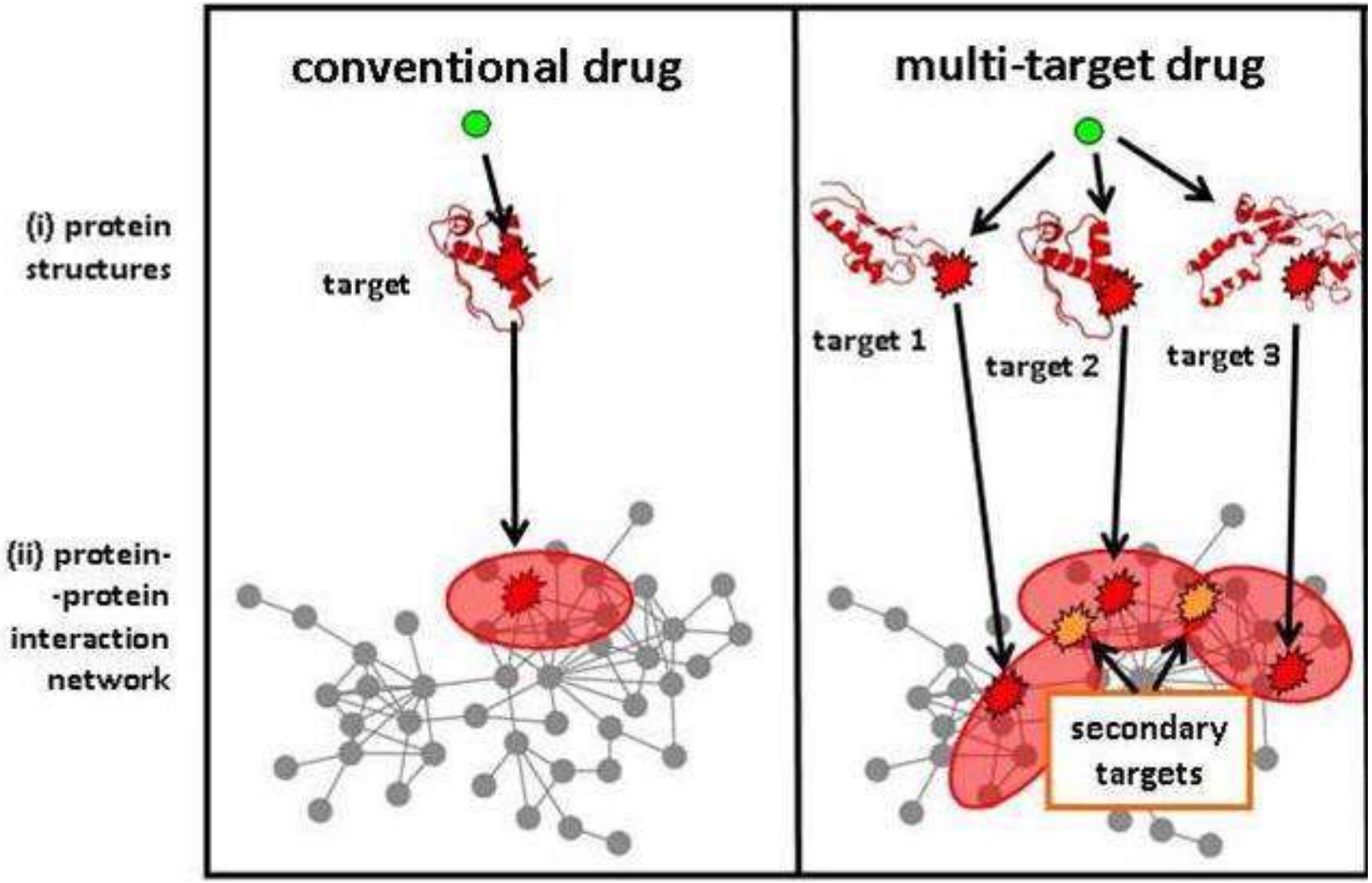
The network influence strategy:
for differentiated cells
having rigid networks
with small dissipation

[central node hit causes overload: side effects & toxicity]

- **multiple targets**
- **allo-network drugs**

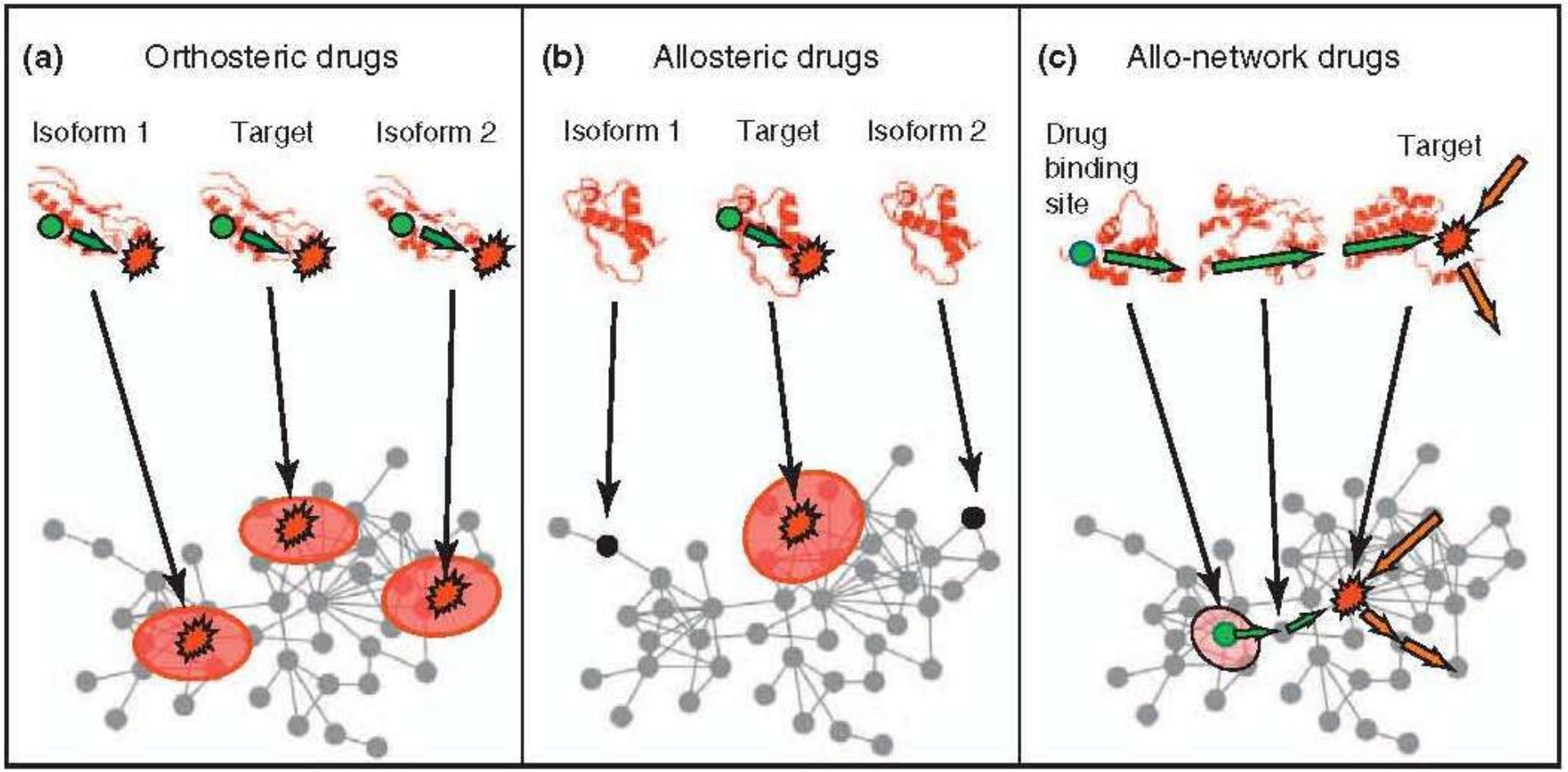
Csermely, Korcsmaros et al., *Pharmacol & Therap* **138:**333-408 (2013)

Multi-target drugs



Csermely et al, *Trends Pharmacol Sci* 26:178

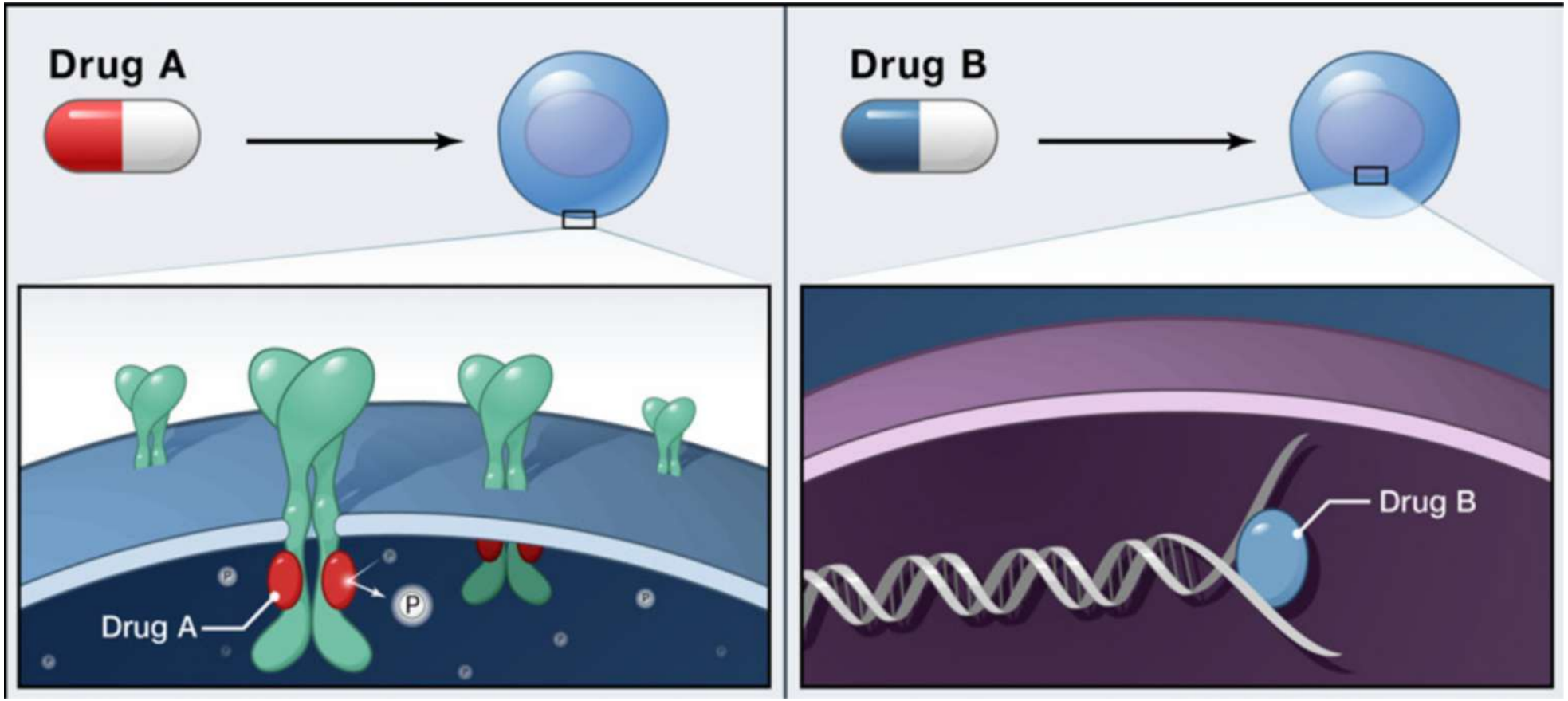
Allo-network drugs



hit of intra-cellular paths

Nussinov *et al*, *Trends Pharmacol Sci* 32:686-93

Allo-network drugs in action

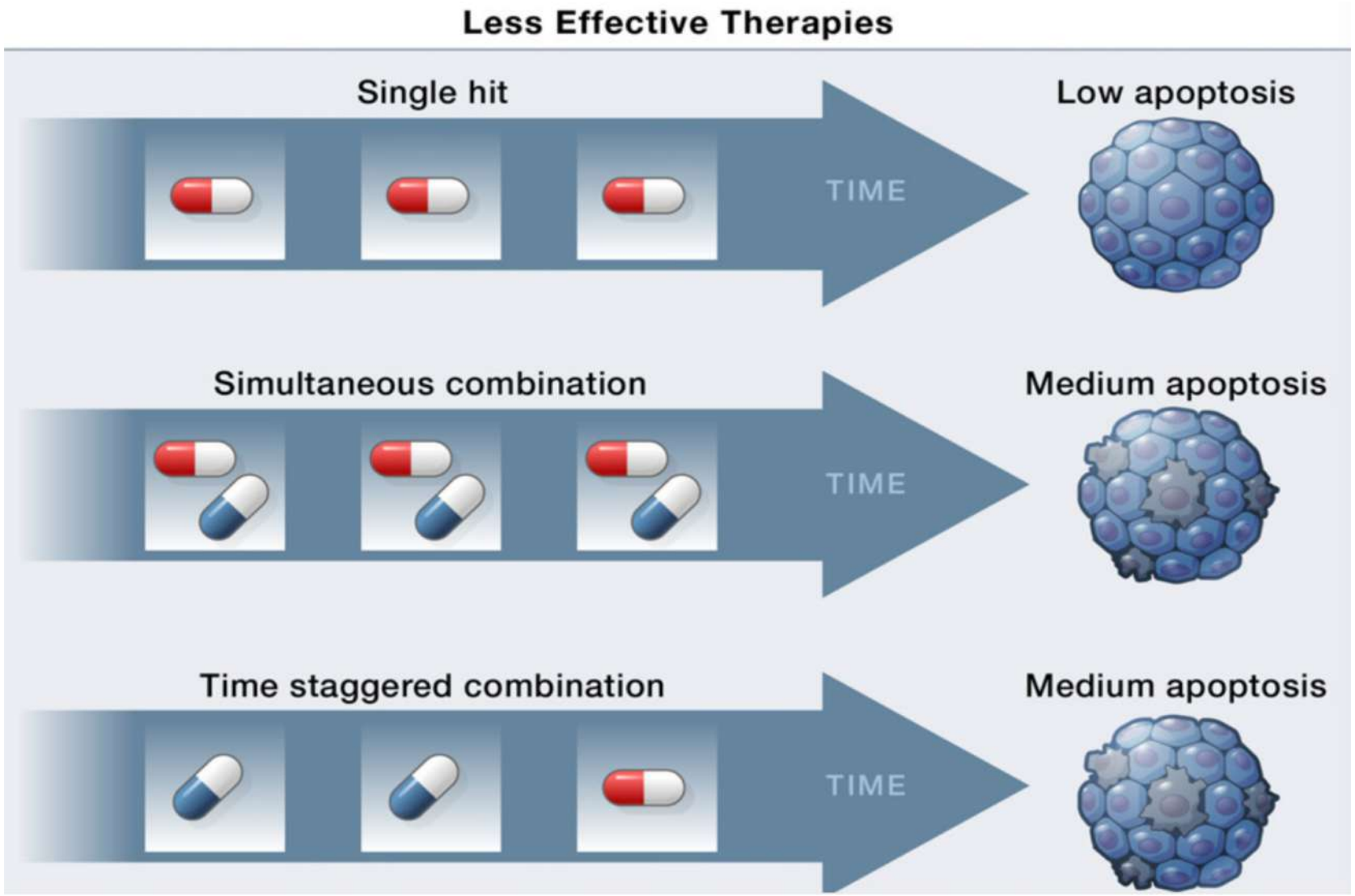


Drug A: EGFR kinase-inhibitor erlotinib

Drug B: doxorubicin (chemotherapy drug)

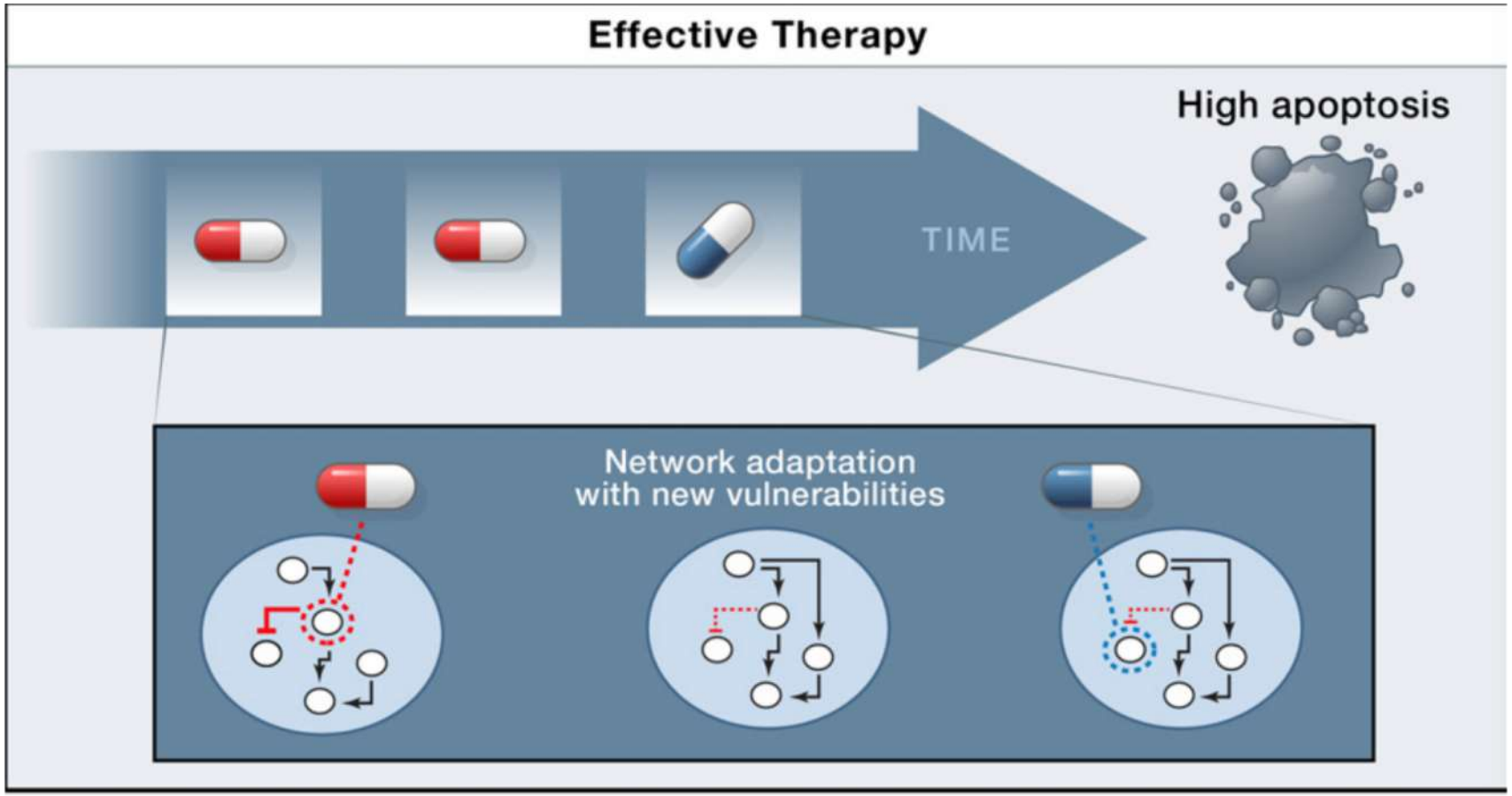
Lee *et al* (Yaffe), *Cell,* 2012; Comment and figures: Erler and Linding, *Cell, 2012*

Allo-network drugs in action



Lee *et al* (Yaffe), *Cell,* 2012; Comment and figures: Erler and Linding, *Cell, 2012*

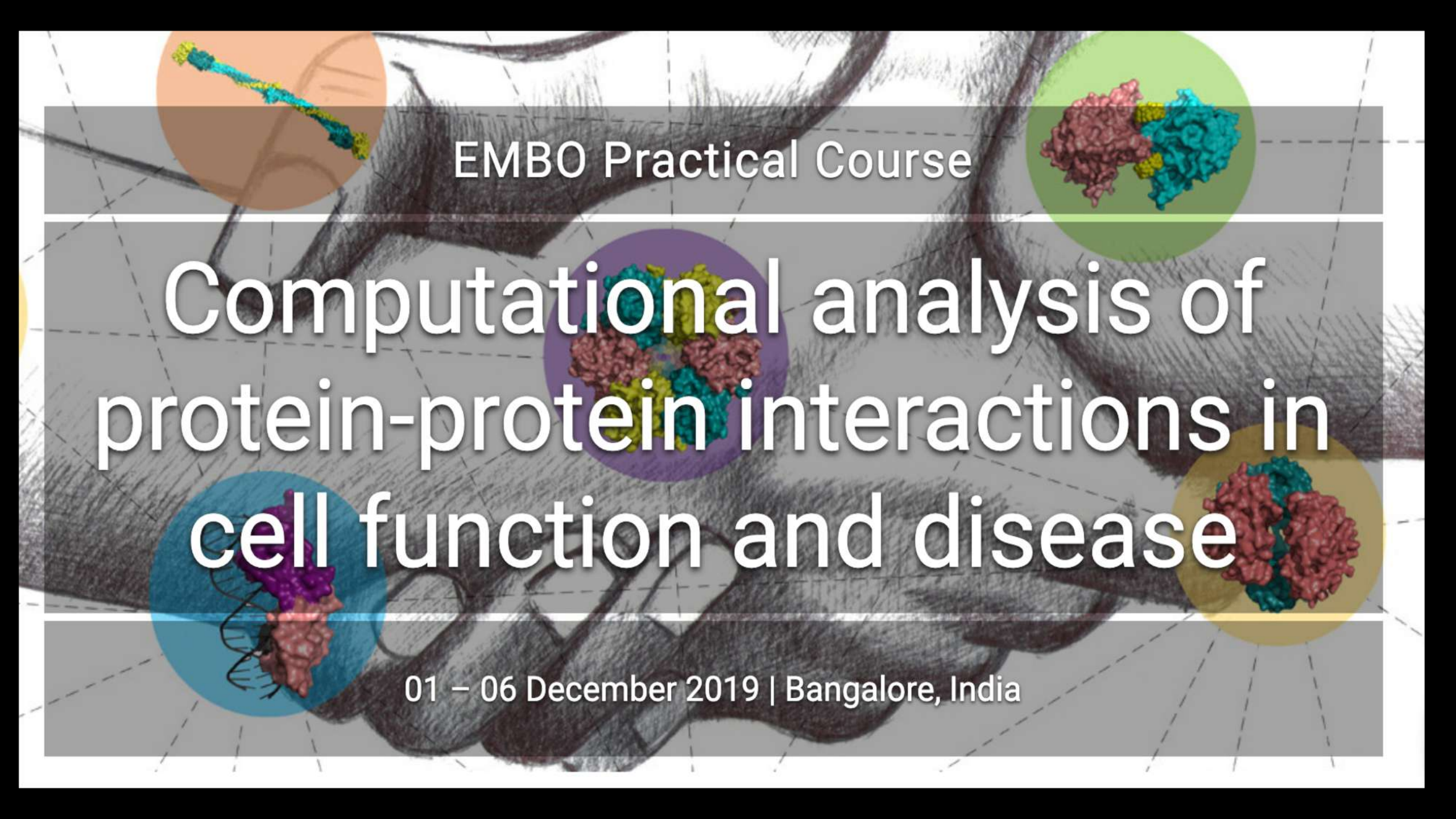
Allo-network drugs in action



Lee *et al* (Yaffe), *Cell,* 2012; Comment and figures: Erler and Linding, *Cell, 2012*

Would you like to know more?

_____



EMBO Practical Course

Computational analysis of protein-protein interactions in cell function and disease

01 – 06 December 2019 | Bangalore, India

@KorcsmarosLab

Thank you!

https://github.com/korcsmarosgroup/

Tamas.Korcsmaros@earlham.ac.uk