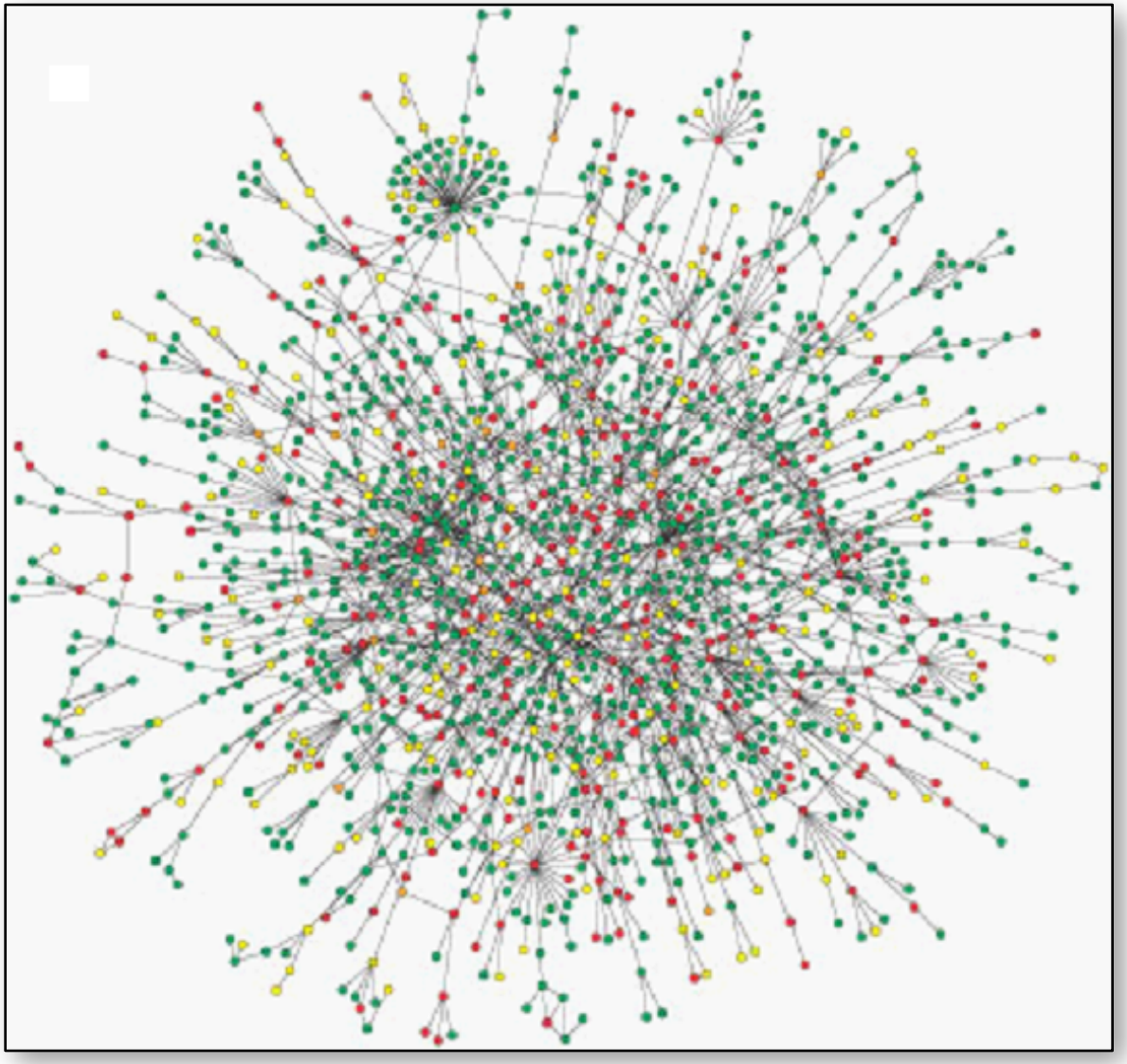
# Visualisation of Biological Networks

Bjorn Sommer

Falk Schreiber

University of Konstanz

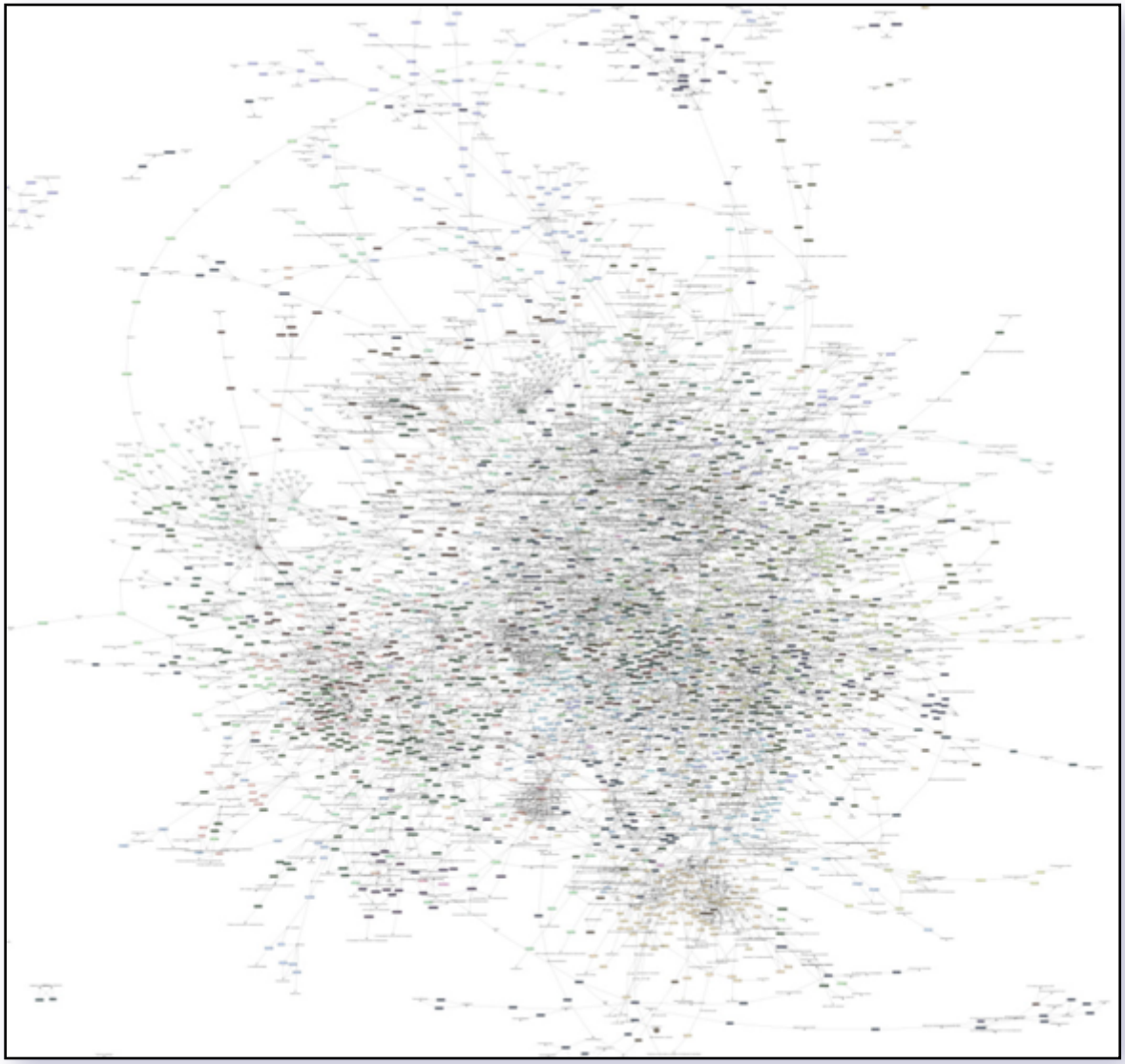
# Question 1 – Can you Read this?



A network with $10^2$ nodes

Protein interaction network, source: Jeong *et al. Nature*, 2001

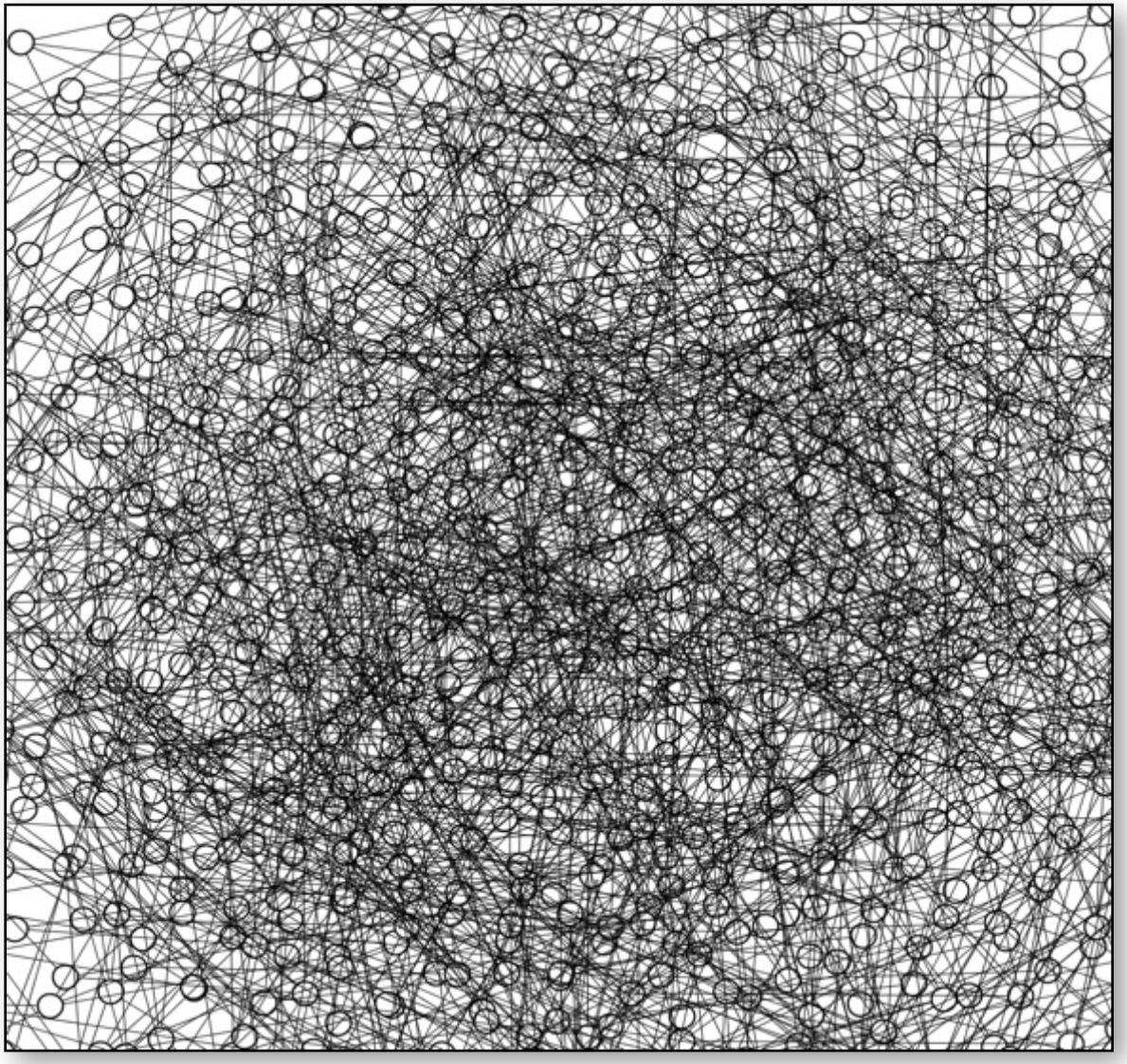
# Question 1 – Can you Read this?



A network with $10^3$ nodes

Metabolic network,
source: KEGG, 2012

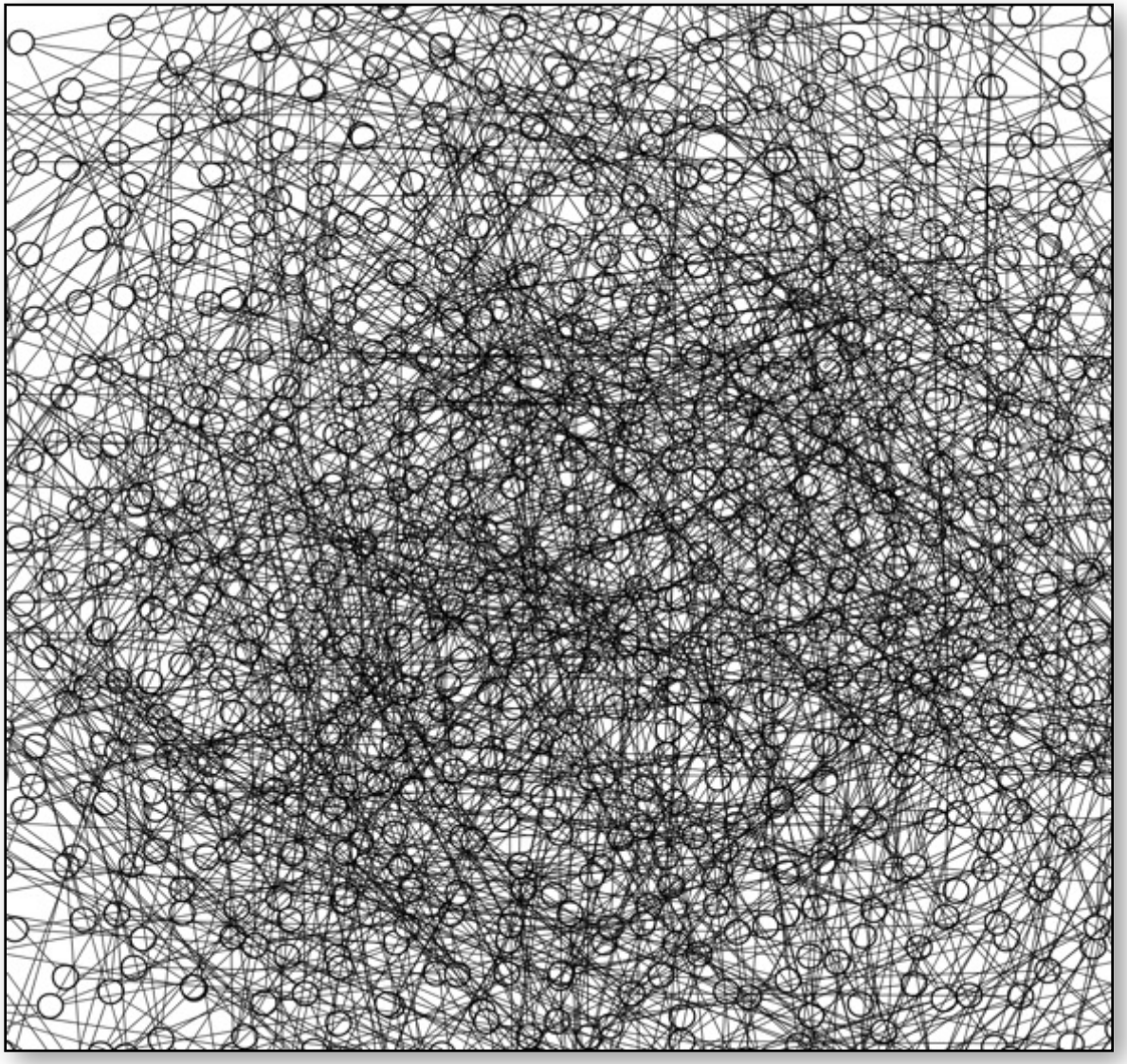
# Question 1 – Can you Read this?



A network with $10^4$ nodes

Protein interaction network, source: DIP, 2013

# Part 1

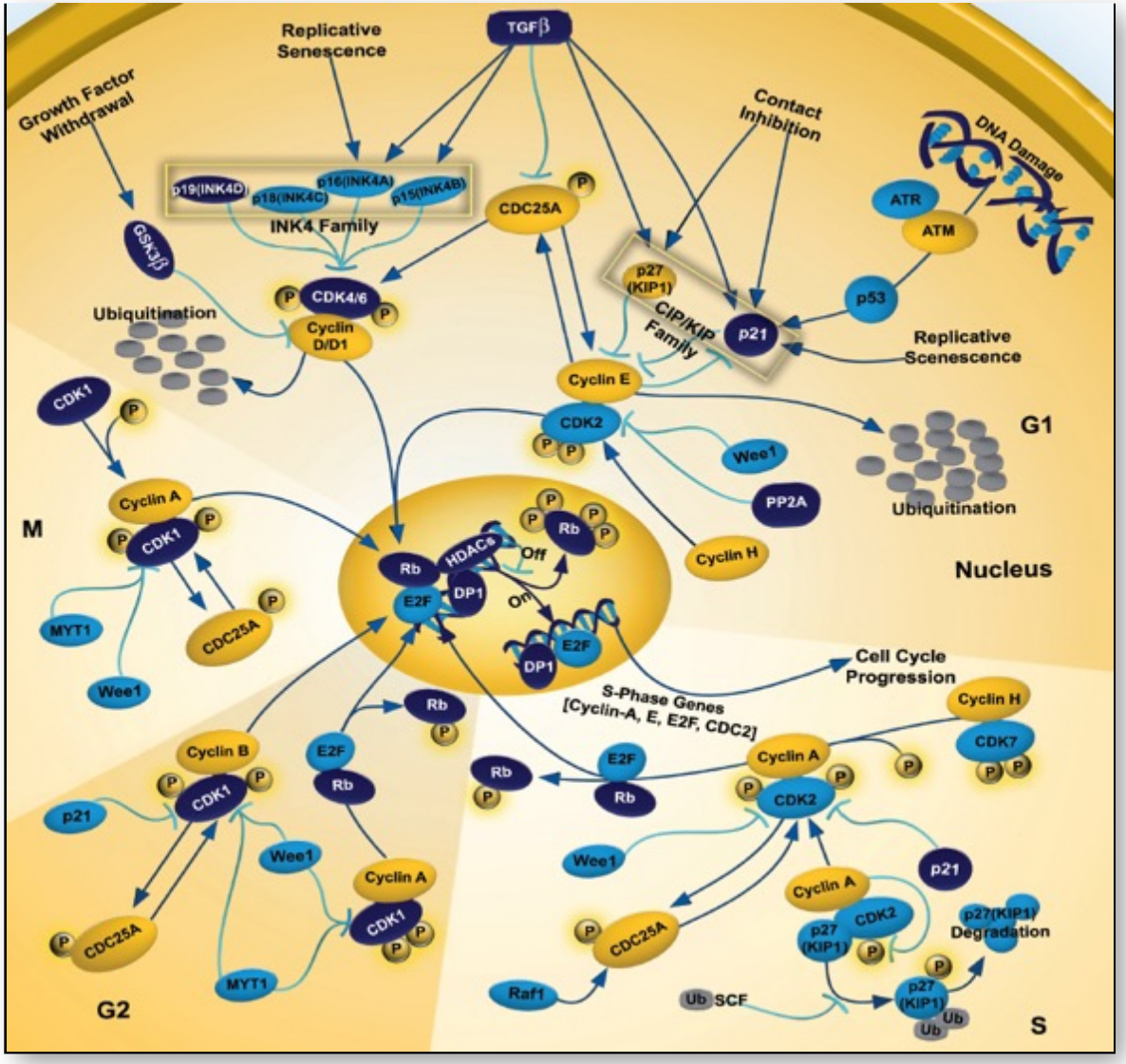


A network with $10^4$ nodes

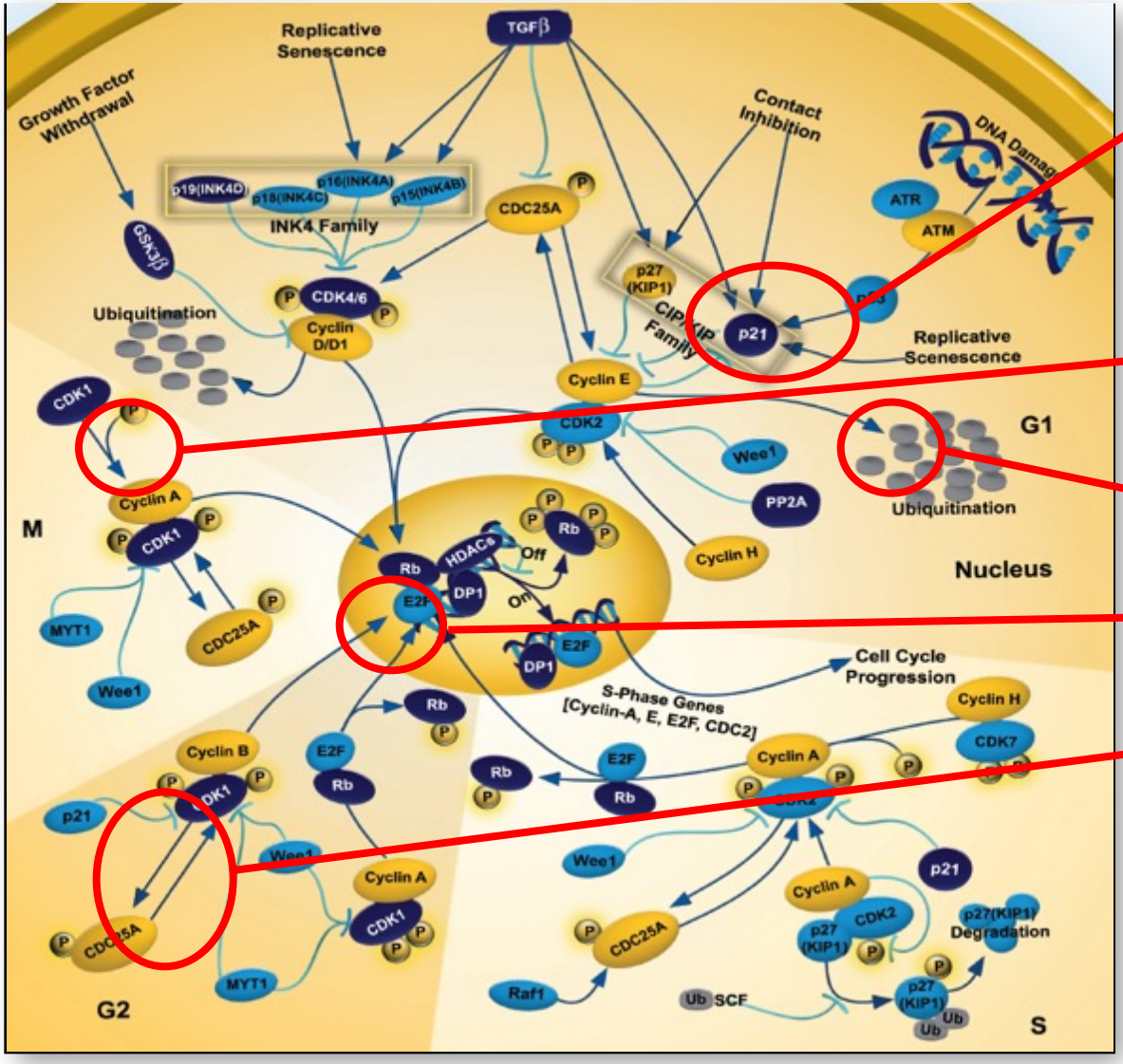
Protein interaction network, source: DIP, 2013

→ Automatic layout of large networks and interactive exploration of networks

# Question 2 – Can you Understand this?



# Question 2 – Can you Understand this?




Stimulates gene transcription?

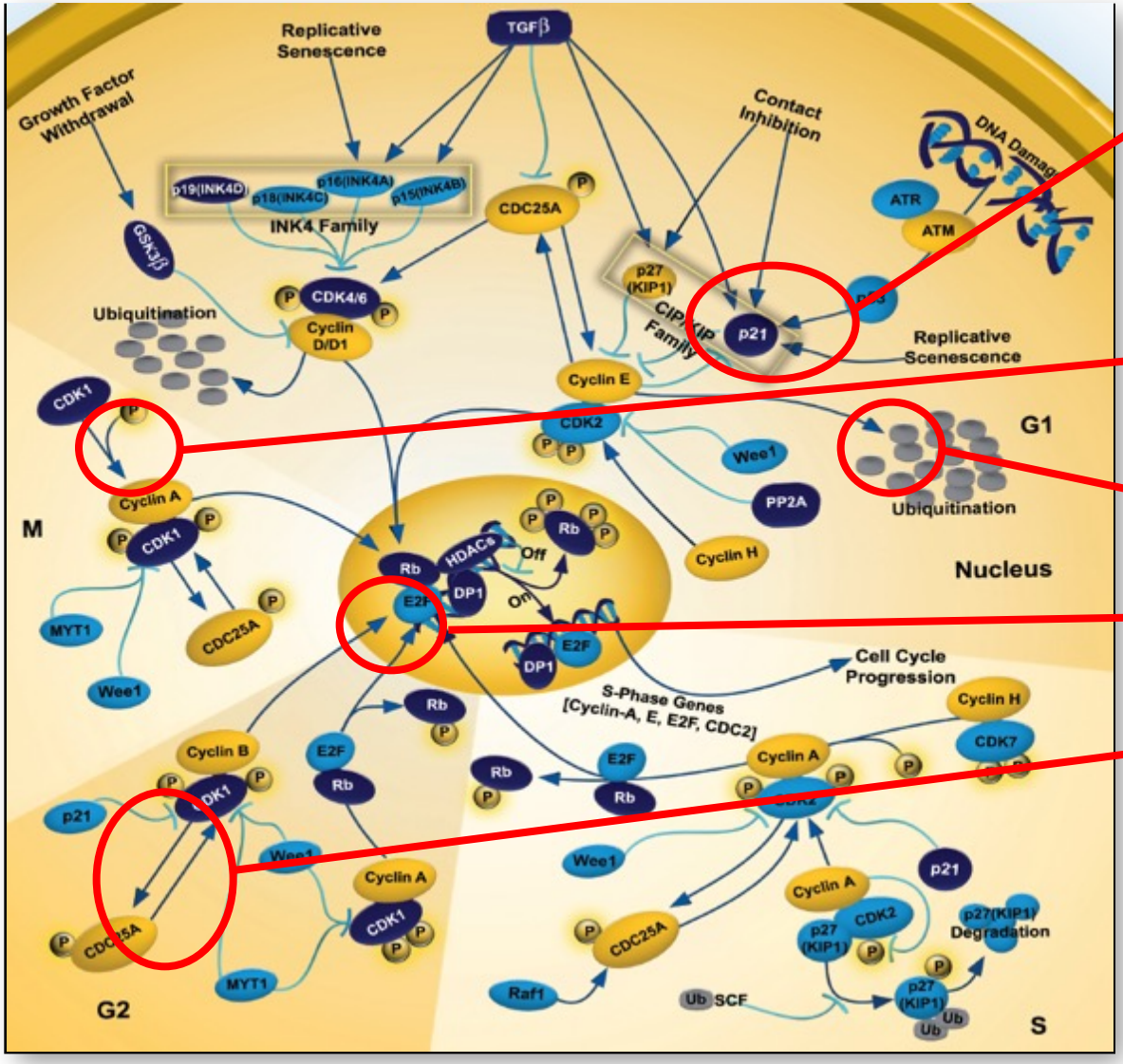
Associates into?

Is degraded?

Translocates?

Reciprocal stimulation?

# Part 2




Stimulates gene transcription?

Associates into?

Is degraded?

Translocates?

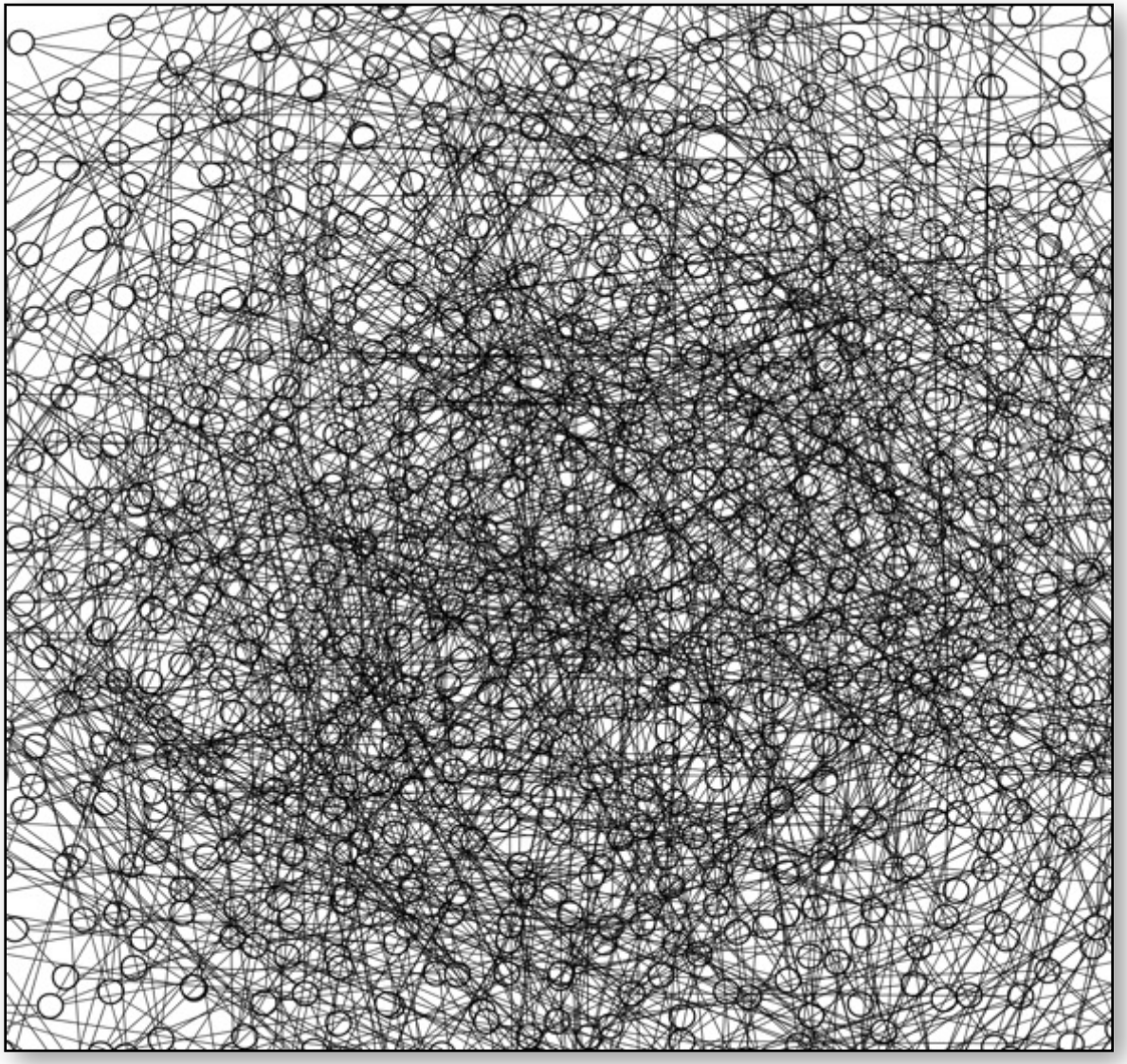
Reciprocal stimulation?

→Standardisation of graphical representation

# Part 1

- Automatic layout of large networks and interactive exploration of networks

# Part 1



A network with $10^4$ nodes

Protein interaction network, source: DIP, 2013

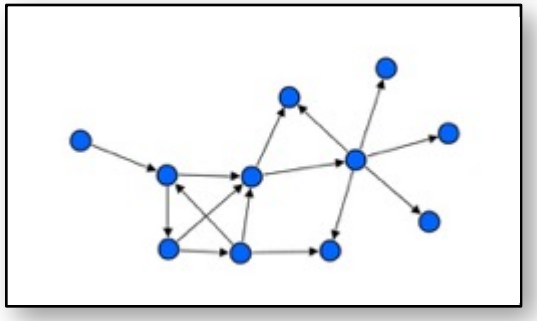
→ Automatic layout of large networks and interactive exploration of networks

# Automatic Layout of Networks

- Force-based approaches
  - Simulate a system of physical forces

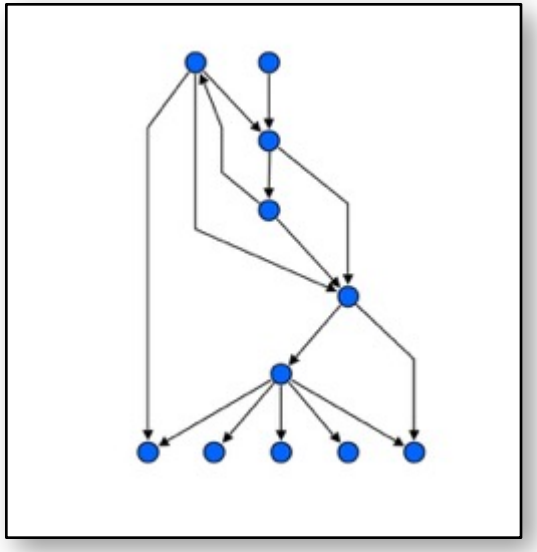
Eades. *Congressus Numerantium*, 1984.

Fruchterman & Reingold. *Software - Practice and Experience*,1991.



- Layered approaches
  - Decycling - layering - crossing reduction - coordinate assignment

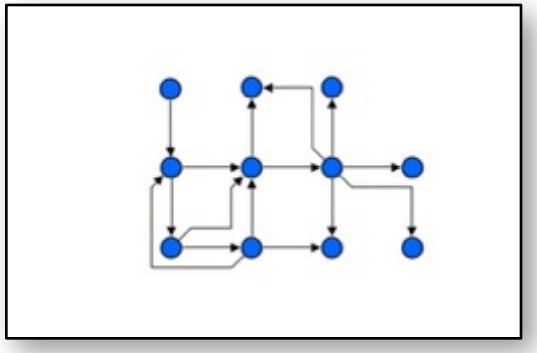
Sugiyama et al. *IEEE Transactions on Systems, Man and Cybernetics*, 1991.



- Orthogonal / grid-based approaches

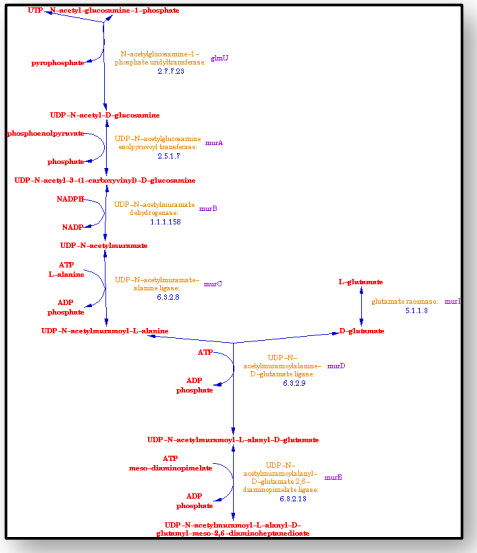
Tamassia. *SIAM Journal on Computing*, 1987.

Biedl *et al. Graph Drawing, LNCS 1353*, 1998.

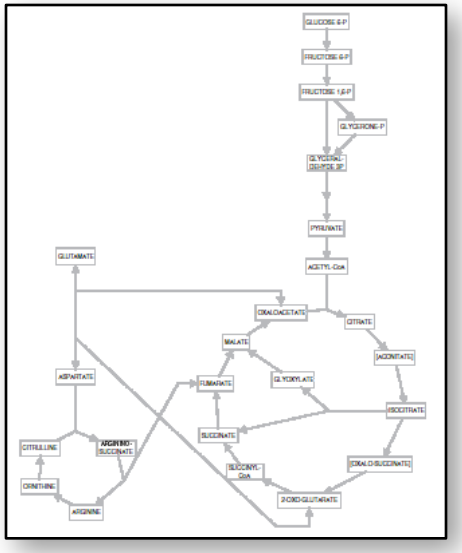


# Many Special Layout Algorithms

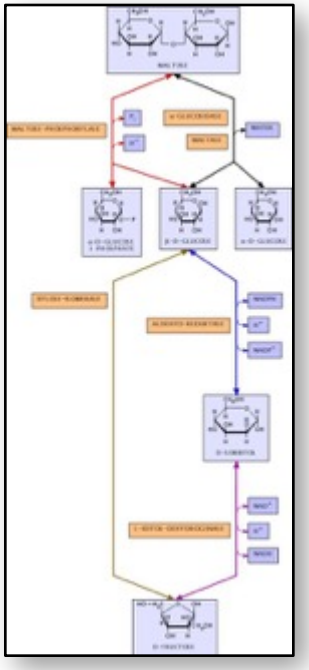
- Commonly extensions of the three classes of layout algorithms
  - Force-based
  - Layered
  - Orthogonal / grid-based
- Examples



Source: Karp & Paley. *Conf. Bioinformatics and Genome Research*, 1994.



Source: Becker & Rojas. *Bioinformatics*, 2001.



Source: Schreiber. *In Silico Biology*, 2002.

Source: Genc & Dogrusoz. *Graph Drawing LNCS 2912,* 2004.

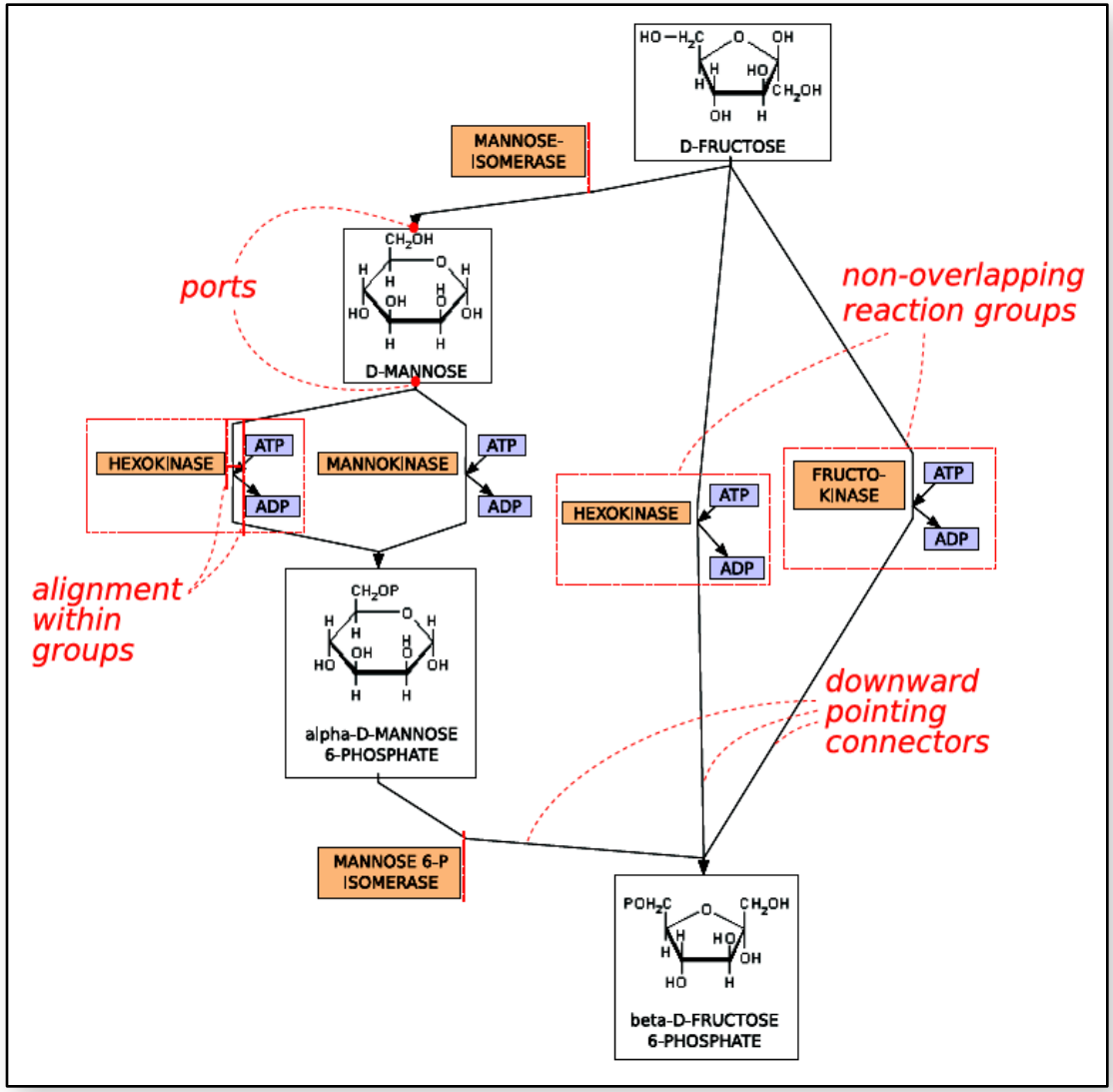
# Good Network Layout

- Better layouts have
  - Fewer edge crossings
  - Large crossing angles
  - Straighter edges
  - Horizontal and vertical edges
  - Symmetrical parts shown symmetrically
  - …

→ Network visualization methods as optimization methods

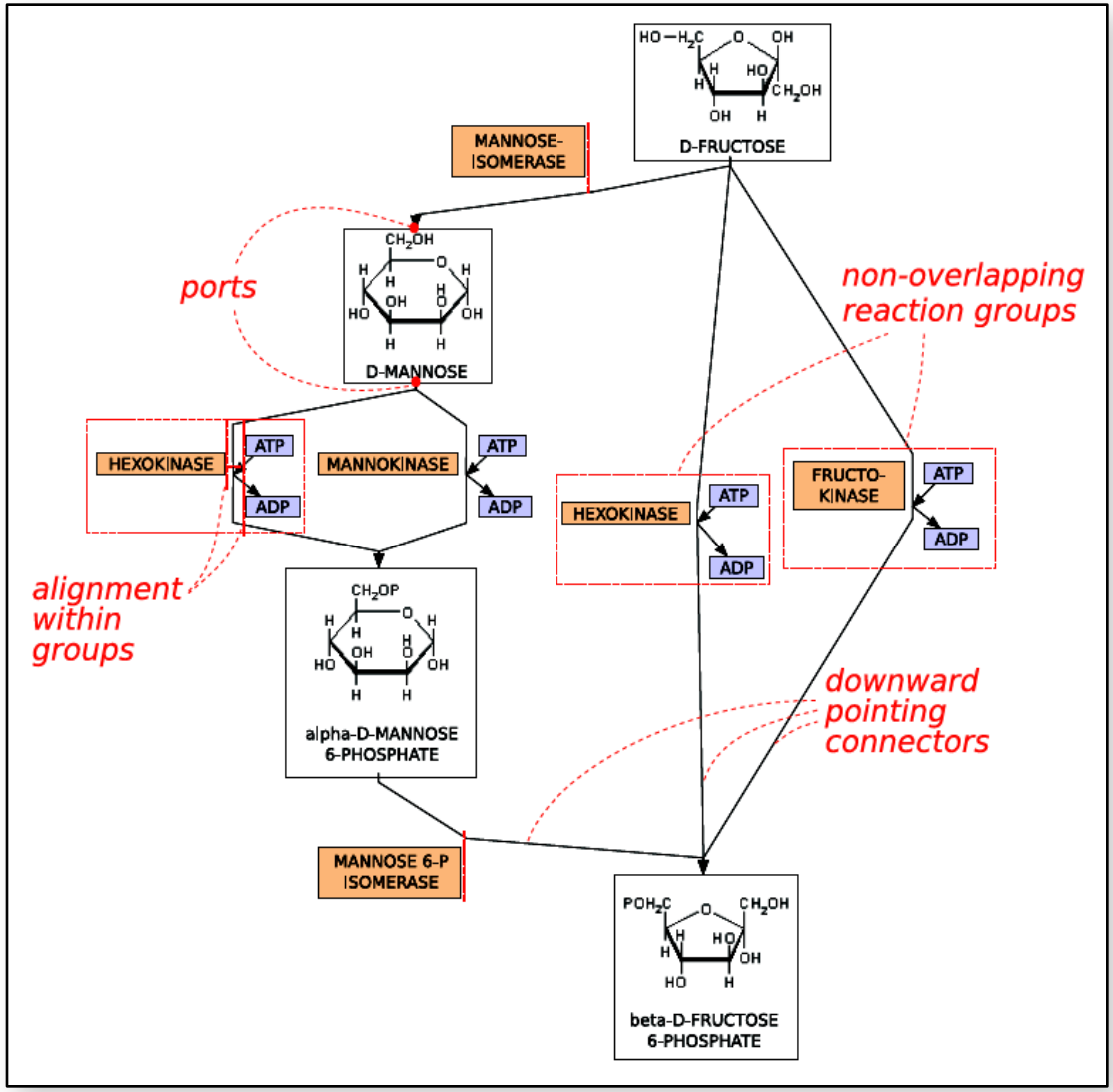
- Constraint-based layout algorithms
  - Fulfils requirements
  - Adaptable to special questions (highlighting paths, …)
  - Adaptable to different biological networks

# Constraint-based layout



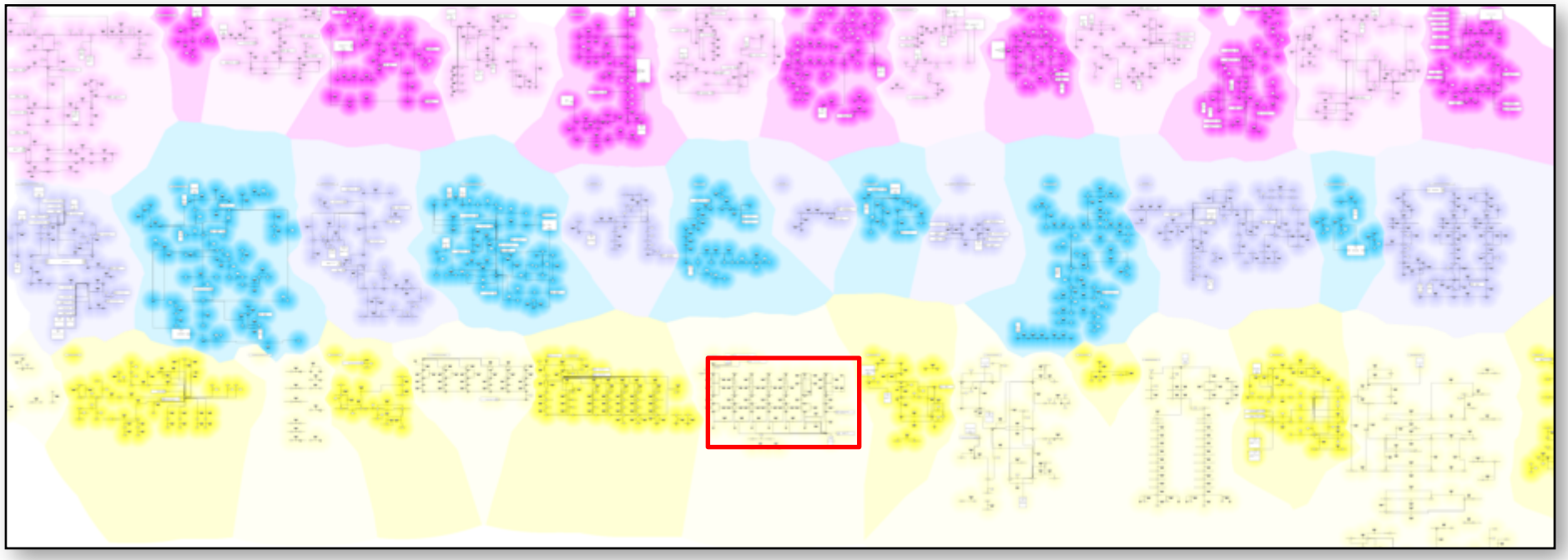
# Constraint-based layout

- Translated into separation constraints
- Constrained stress majorisation
- Instead of solving unconstrained quadratic forms we solve subject to separation constraints (i.e. Quadratic Programming)



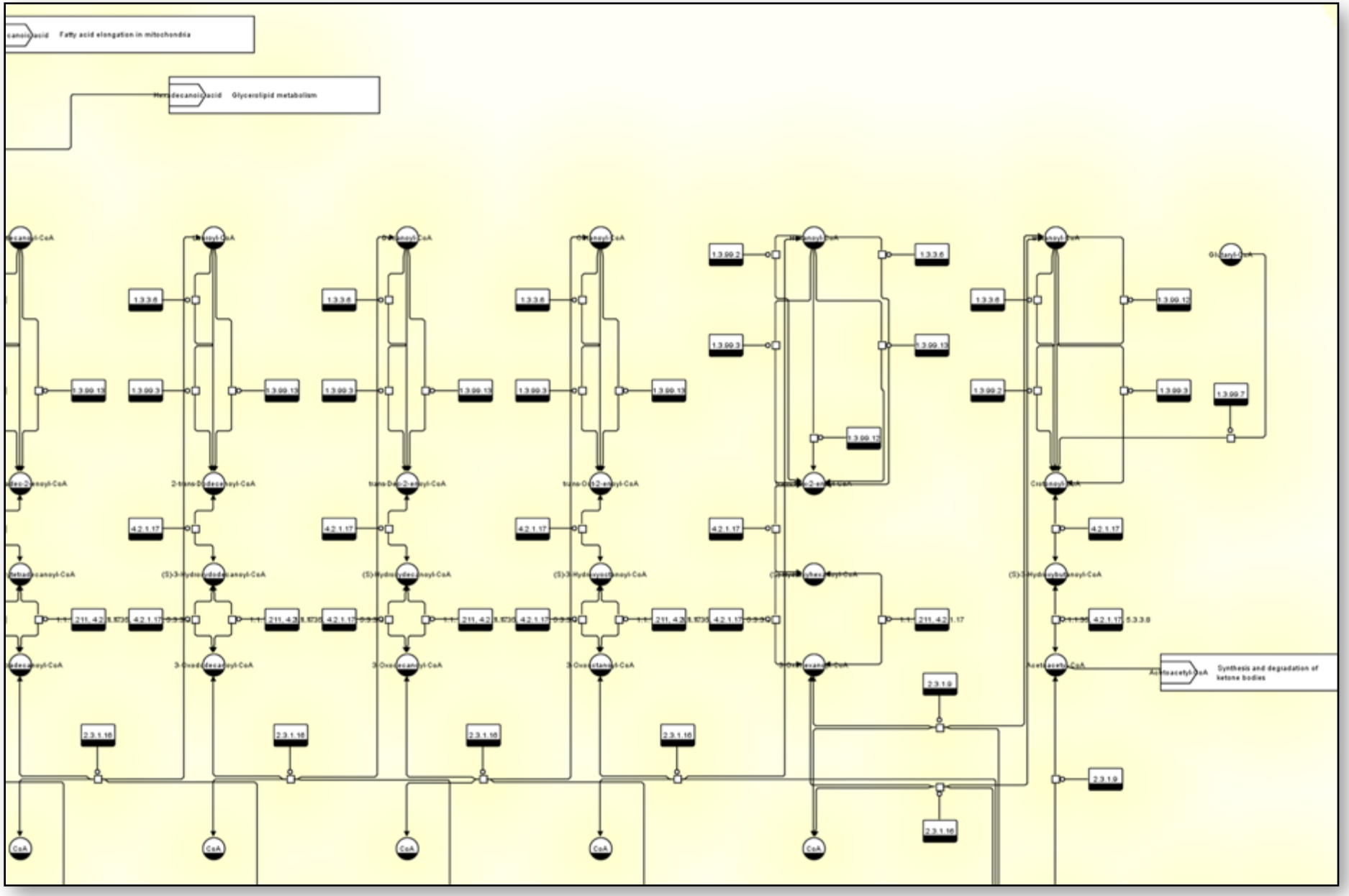

with T. Czauderna, T. Dwyer, M. Wybrow, K. Marriott
Dwyer *et al.*, *IEEE Trans. Visualization & Computer Graphics*, 2008
Schreiber *et al.*, *BMC Bioinformatics*, 2009

# Constrained KEGG Layout



with T. Czauderna, T. Dwyer, M. Wybrow, K. Marriott
Czauderna *et al.*, *BMC Bioinformatics*, 2013

# Constrained KEGG Layout



# Navigation in Large Networks

- Mantra of Information Visualisation:
  - “overview first,
  - zoom and filter,
  - then details-on-demand”
    (B. Sneiderman)

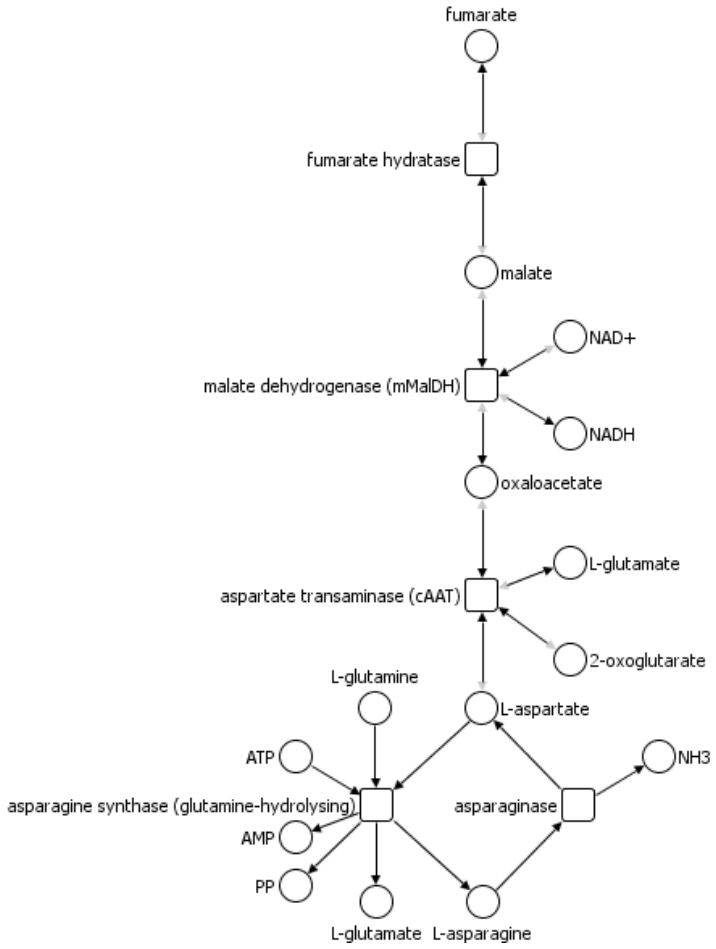
→

  - Hierarchical abstraction of large networks (e.g. nested or compound graphs)
  - Navigation within the networks

  - Example: metabolic pathways

# Navigation in Large Networks

# Separation in Pathways

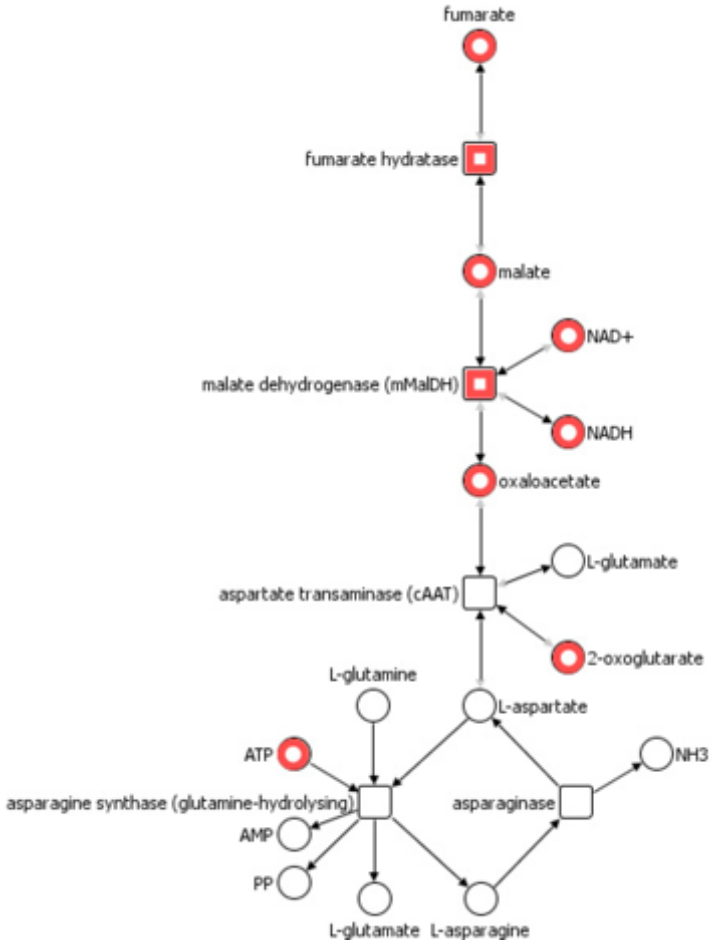



- Amino Acid Metabolism
  - Alanine, Valine, Leucine biosynthesis
  - Arginine biosynthesis
  - Asparagine biosynthesis
  - GS-GOGAT cycle
  - Histidine biosynthesis
  - Isoleucine biosynthesis
  - Lysine biosynthesis
  - Methionine biosynthesis
  - Methionine recycling
  - Phenylalanine, Tyrosine, Tryptophan biosynthesis
  - Proline biosynthesis
  - Serine, Glycine, Cysteine biosynthesis
  - Shikimate biosynthesis
  - Threonine biosynthesis
- Carbohydrate Metabolism
  - Arabinoxylan, Beta-Glucan, Cellulose biosynthesis
  - Ascorbate-Glutathione cycle
  - Ascorbate biosynthesis
  - Chlorogenic acid biosynthesis
  - Fermentation
  - Fructan biosynthesis
  - Glutathione biosynthesis
  - Glycolysis, Gluconeogenesis
  - Glyoxylate cycle
  - Pentose phosphate pathway
  - Starch metabolism (monocots)
  - Sucrose breakdown pathway (dicots)
  - Sucrose breakdown pathway (monocots)
  - Sugar metabolism
  - TCA cycle
- Cofactor Metabolism
- Energy Metabolism
- Lipid Metabolism

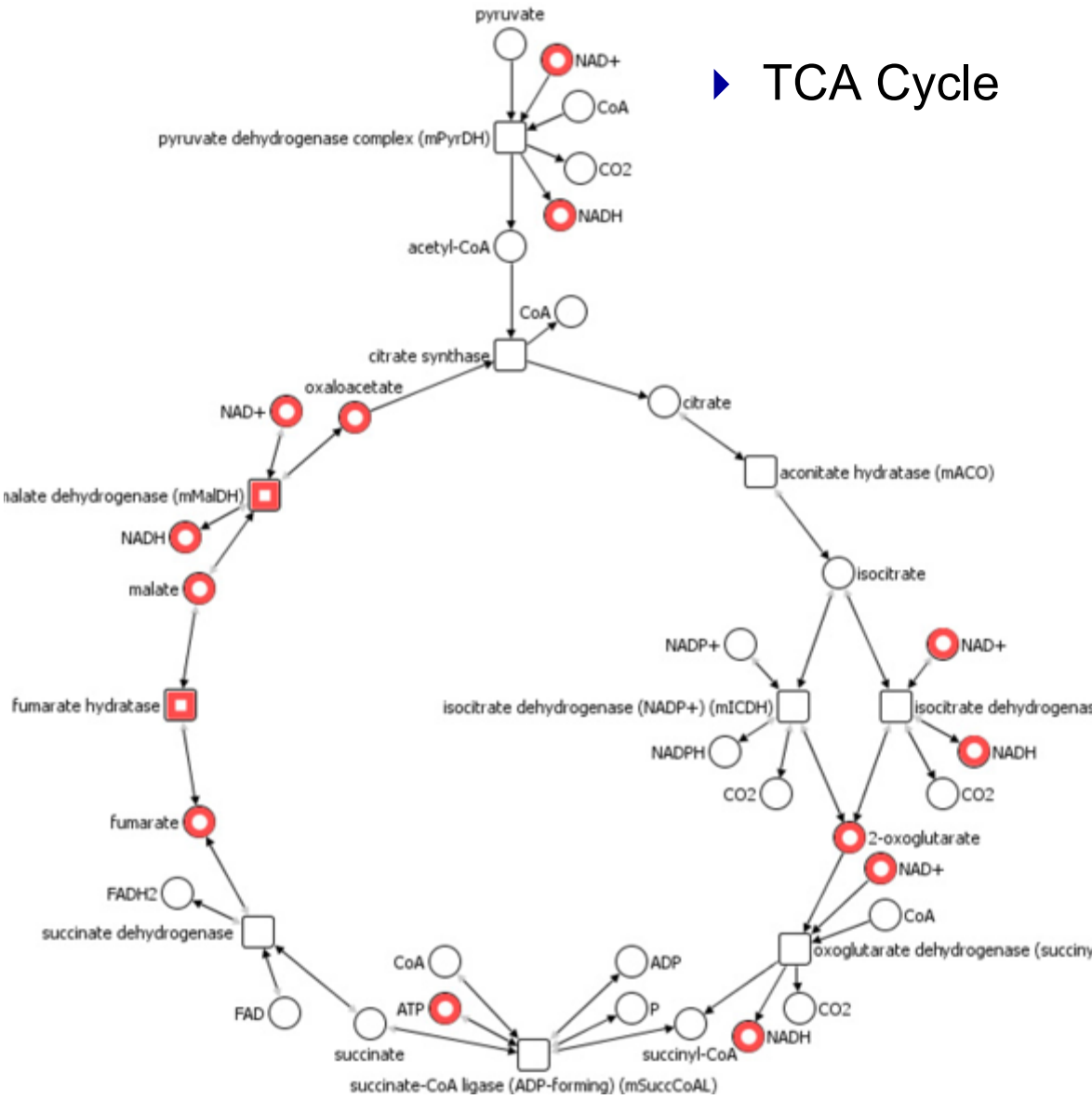
Pathway source: MetaCrop, 2013.

# Separation in Pathways

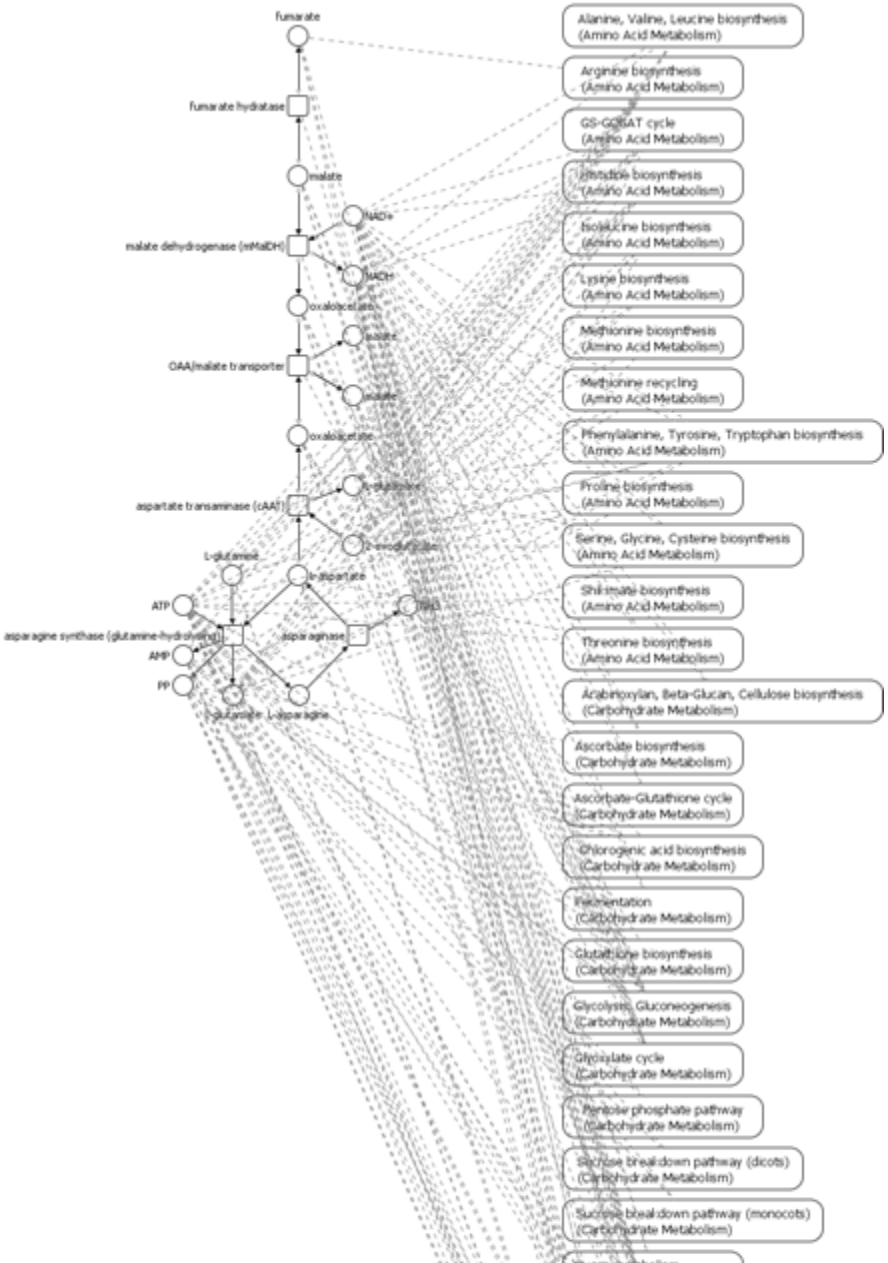
- Asparagine Pathway




- TCA Cycle

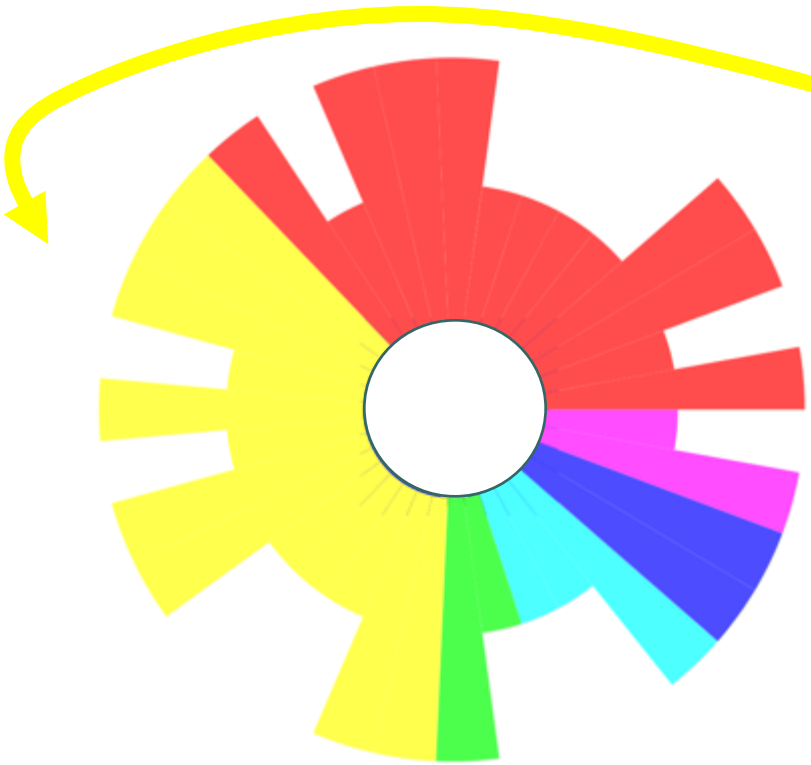


# Navigation between Pathways



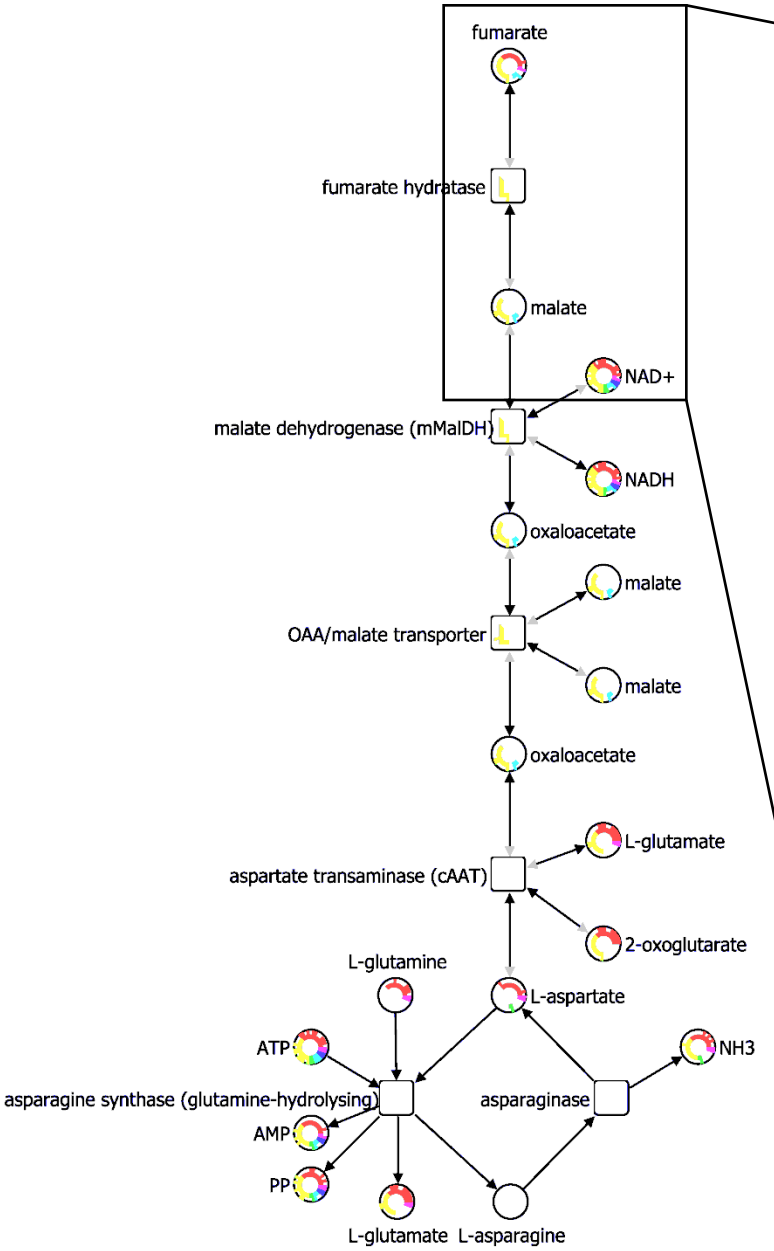
# Navigation between Pathways

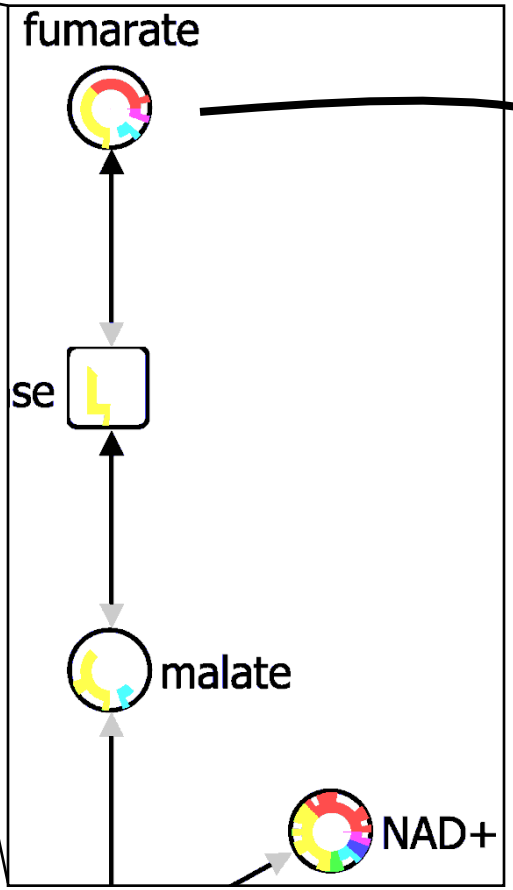
- Combination of network visualisation with glyphs



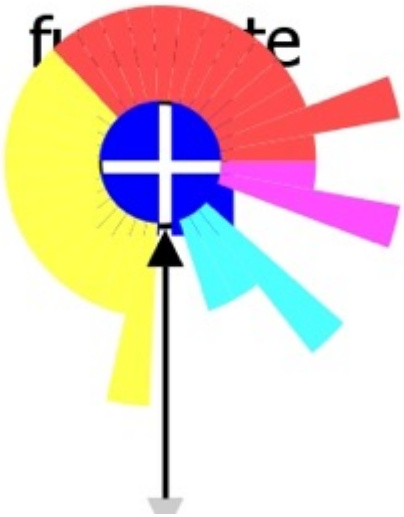
- Amino Acid Metabolism
  - Alanine, Valine, Leucine biosynthesis
  - Arginine biosynthesis
  - Asparagine biosynthesis
  - GS-GOGAT cycle
  - Histidine biosynthesis
  - Isoleucine biosynthesis
  - Lysine biosynthesis
  - Methionine biosynthesis
  - Methionine recycling
  - Phenylalanine, Tyrosine, Tryptophan biosynthesis
  - Proline biosynthesis
  - Serine, Glycine, Cysteine biosynthesis
  - Shikimate biosynthesis
  - Threonine biosynthesis
- Carbohydrate Metabolism
  - Arabinoxylan, Beta-Glucan, Cellulose biosynthesis
  - Ascorbate-Glutathione cycle
  - Ascorbate biosynthesis
  - Chlorogenic acid biosynthesis
  - Fermentation
  - Fructan biosynthesis
  - Glutathione biosynthesis
  - Glycolysis, Gluconeogenesis
  - Glyoxylate cycle
  - Pentose phosphate pathway
  - Starch metabolism (monocots)
  - Sucrose breakdown pathway (dicots)
  - Sucrose breakdown pathway (monocots)
  - Sugar metabolism
  - TCA cycle
- Cofactor Metabolism
- Energy Metabolism

# Navigation between Pathways



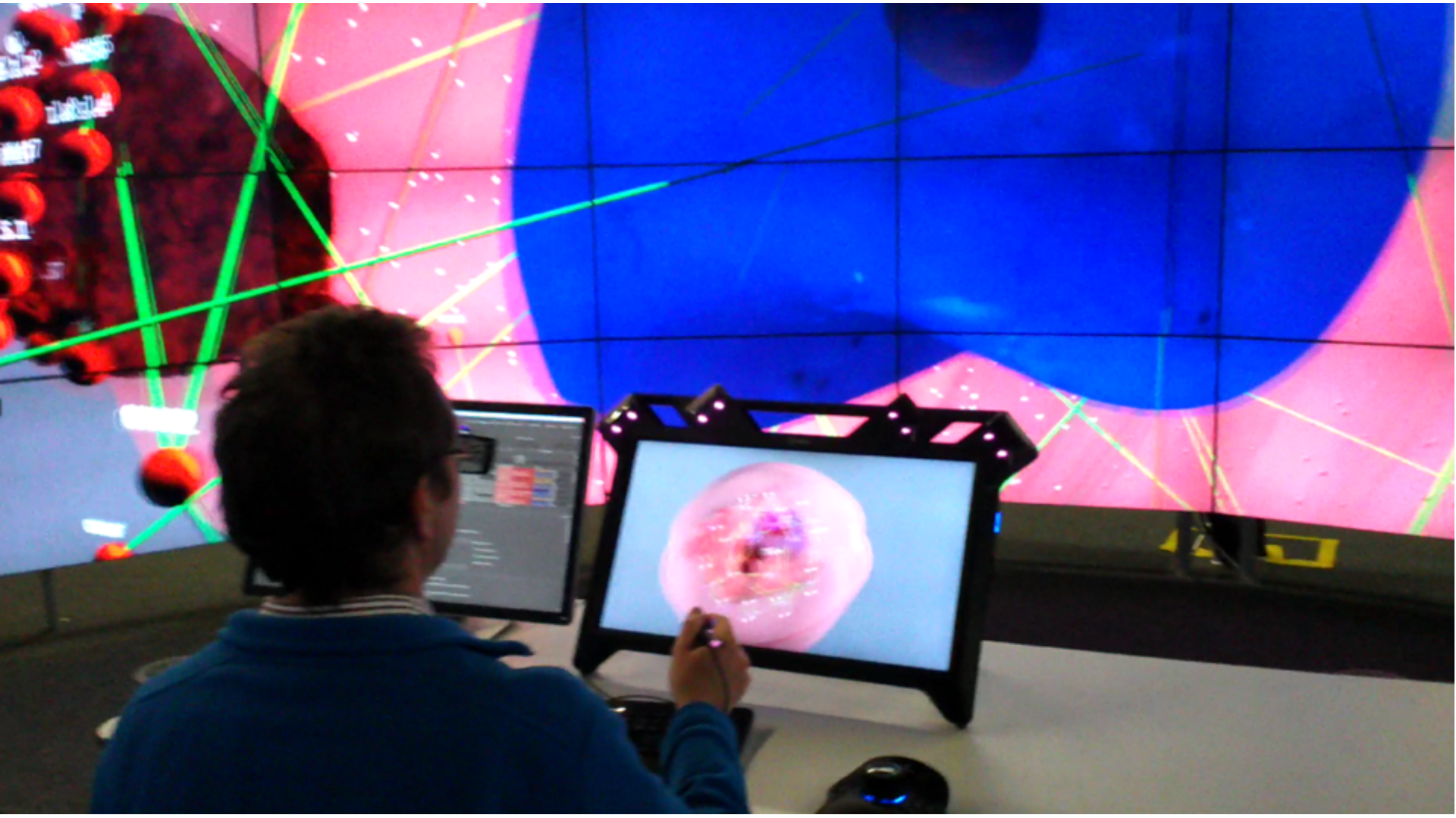




Mouse over



with C. Klukas, I. Jusufi, A. Kerren
Jusufi *et al. Information Visualization*, 2012.

# Spatially embedded Networks: Focus and Context

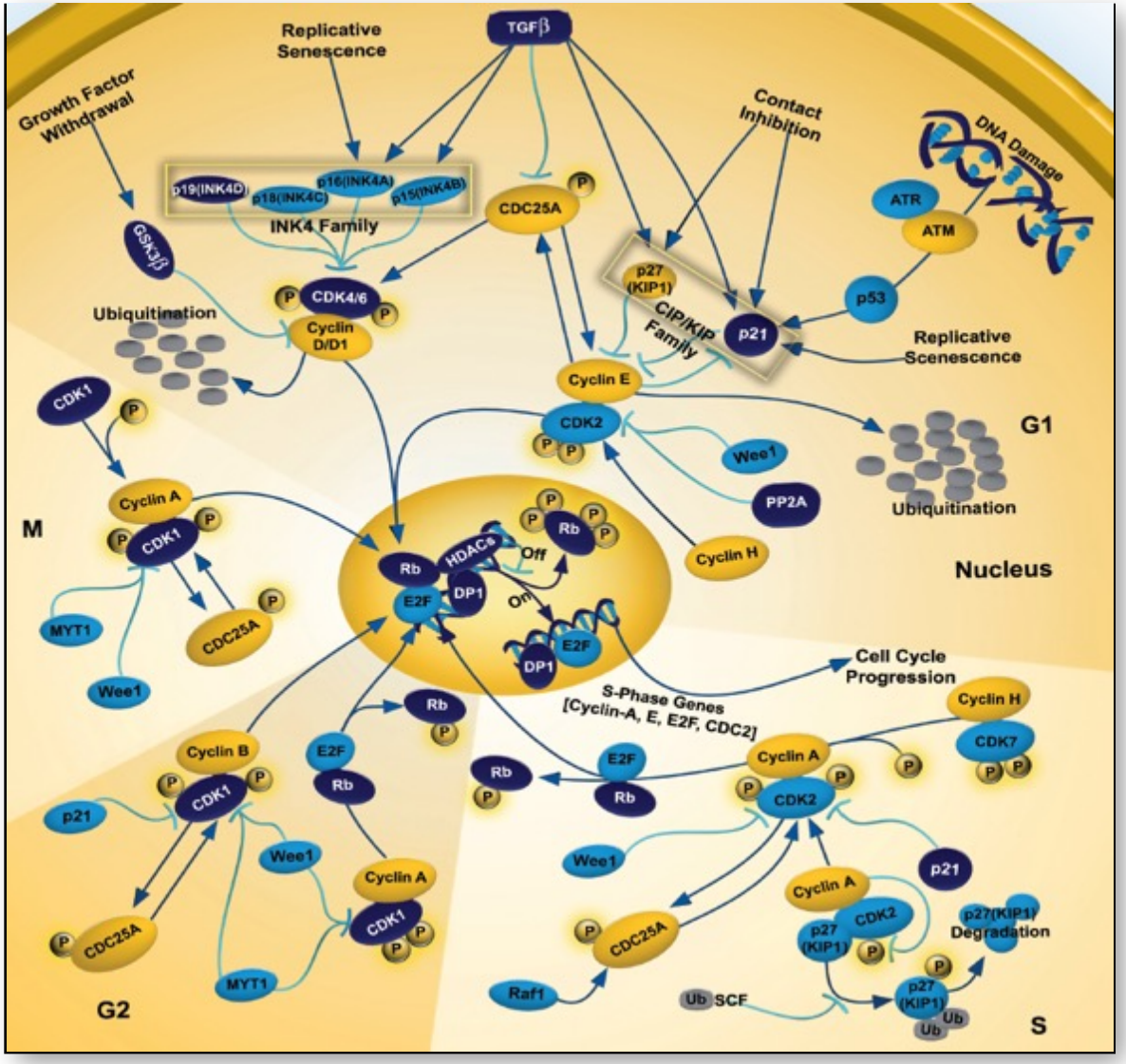


B. Sommer et al. Stereoscopic Space Map. Proceedings of Stereoscopic Displays and Applications XXVII, 2016(5):1–9, Feb. 2016.

# Part 2

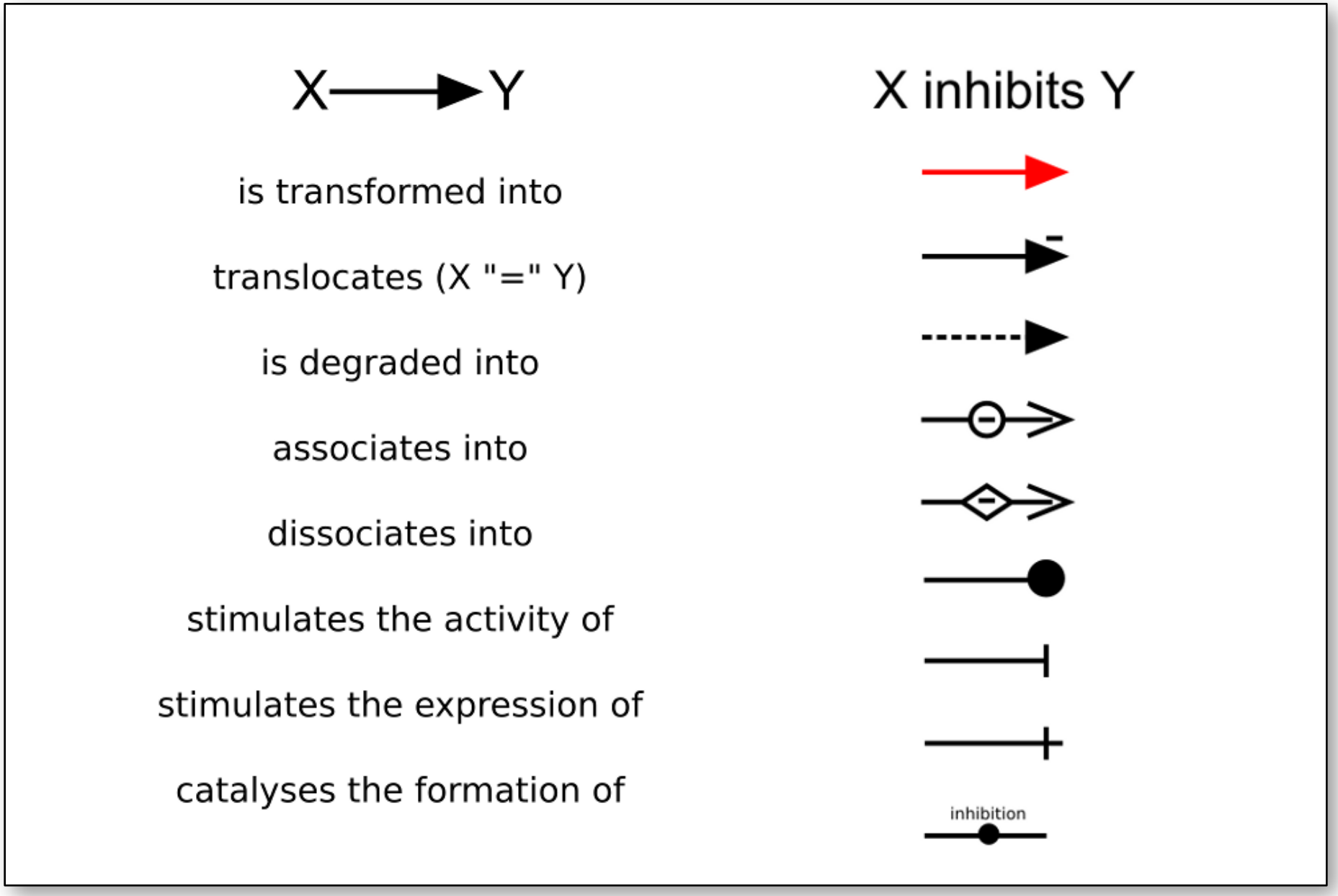
- Standardisation of graphical representation

# Part 2




→Standardisation of graphical representation

# Ambiguity in Conventional Representation



# Standardised Symbols are Important


Most English speaking country


Quebec


Iran


China

Israel

Singapore

Norway

Poland

USA and Canada

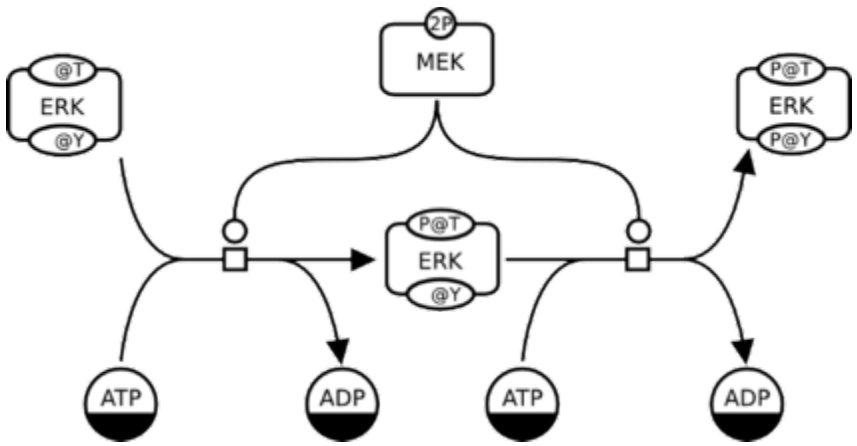
# What is SBGN?

- A way to unambiguously describe biochemical and cellular events in graphs
- Limited amount of symbols (~30) → Smooth learning curve
- Can graphically represent quantitative models, biochemical pathways, at different levels of granularity
- Developed since 2006 by a growing community, part of COMBINE

- Three languages
  - Process Descriptions → one state = one glyph
  - Entity Relationships → one entity = one glyph
  - Activity Flow → conceptual level

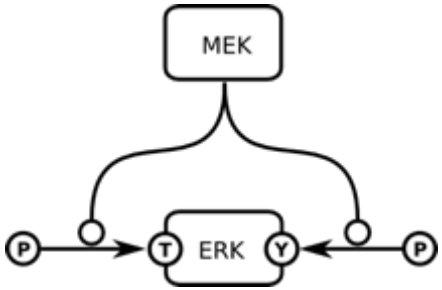
# Graph Trinity: Three Languages in One

## Process Description maps



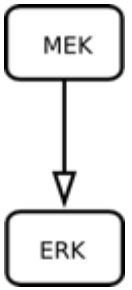

- Unambiguous
- Mechanistic
- Sequential
- Combinatorial explosion

## Entity Relationships maps




- Unambiguous
- Mechanistic
- Non-Sequential

## Activity Flow maps

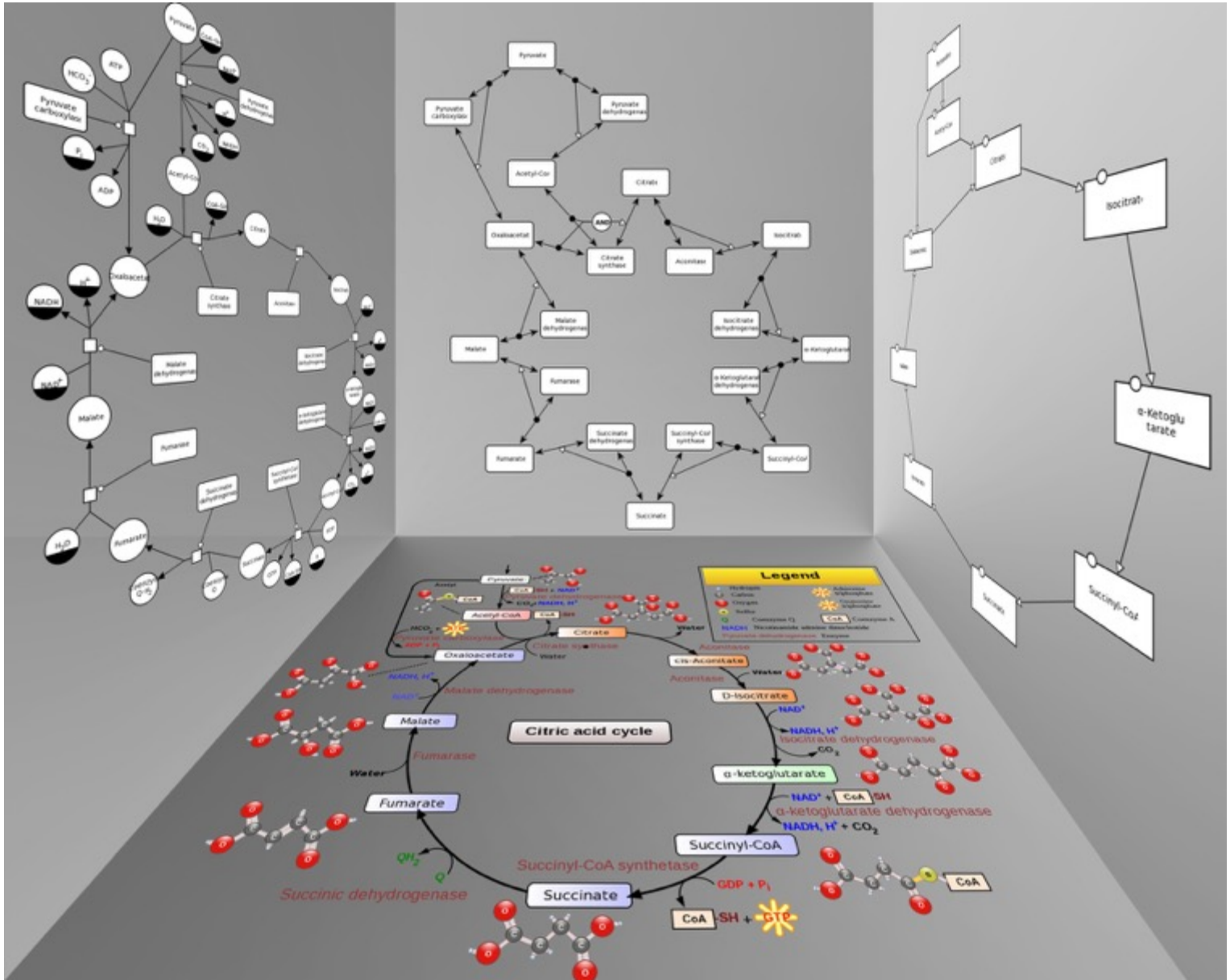



- Ambiguous
- Conceptual
- Sequential

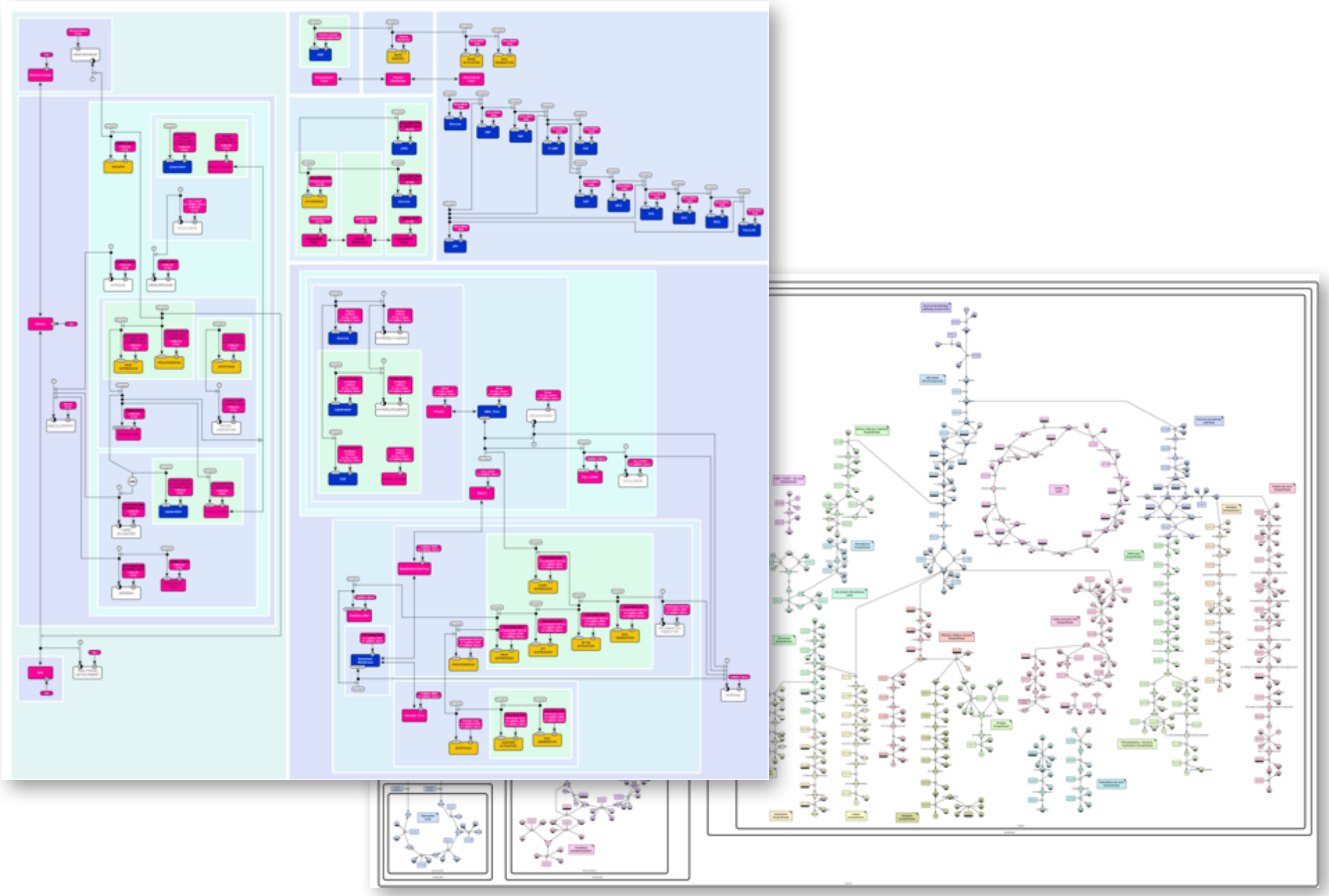
Le Novère *et al. Nature Biotechnology*, 2009

# Graph Trinity: Three Languages in One

Process Description  Entity Relationships  Activity Flow



# Systems Biology Graphical Notation (SBGN)



# Some Links

http://www.sbgn.org



http://co.mbine.org

http://www.sbgn-ed.org


http://www.vanted.org

DE GRUYTER

# JIB

Journal of Integrative Bioinformatics

- Editors-in-chief
  - Ralf Hofestädt
  - Falk Schreiber

- COMBINE Topics
  - Tool integration and workflow systems
  - Network simulation and analysis
  - Prediction and integration of metabolic and regulatory networks
  - Standards in Systems Biology
  - Visualization and animation in Bioinformatics
  - Whole cell modeling