

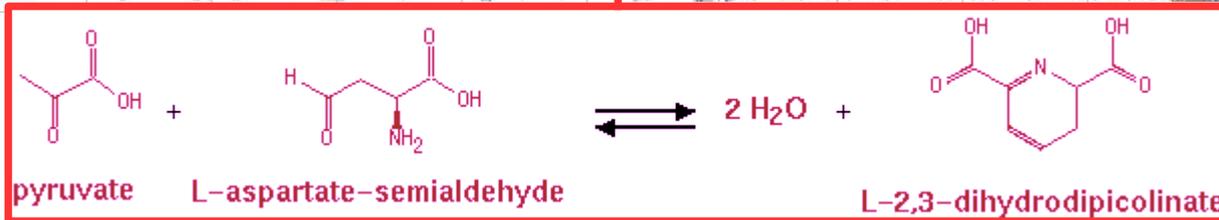
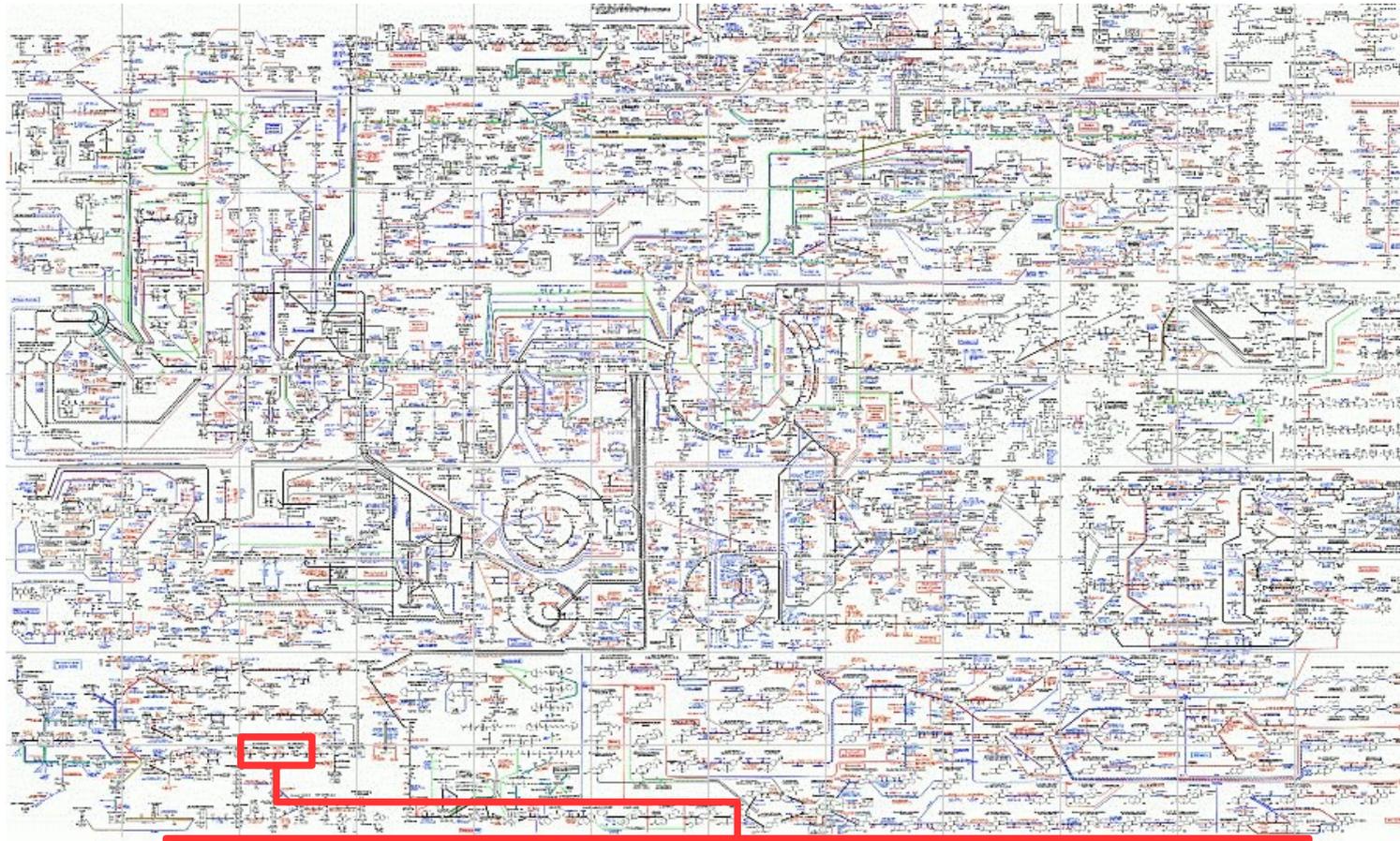
Metabolic network visualisation

Ludovic Cottret
LIPM
CATI BBRIC

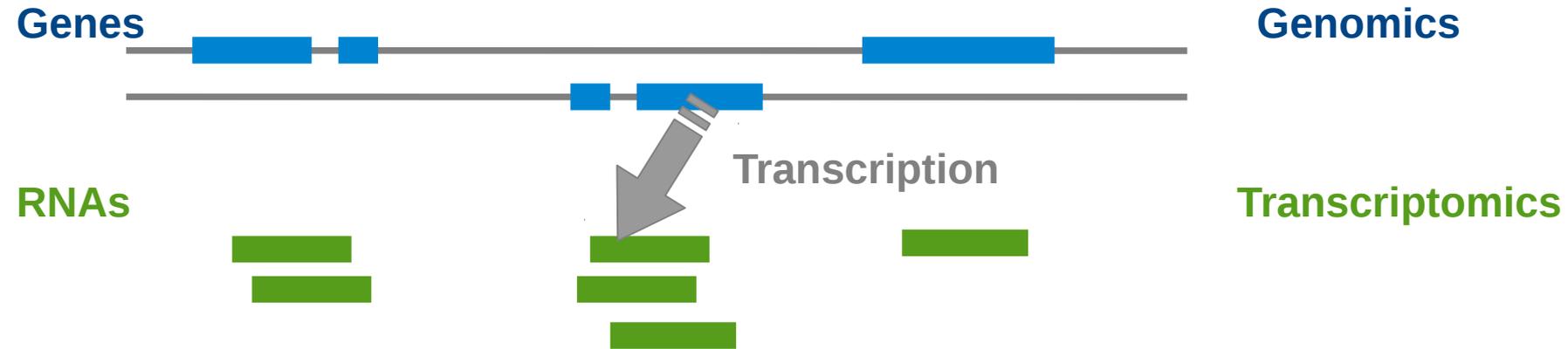
Séminaire Méthodes et outils pour l'Open Data
Notre-Dame de Londres

18 décembre 2014

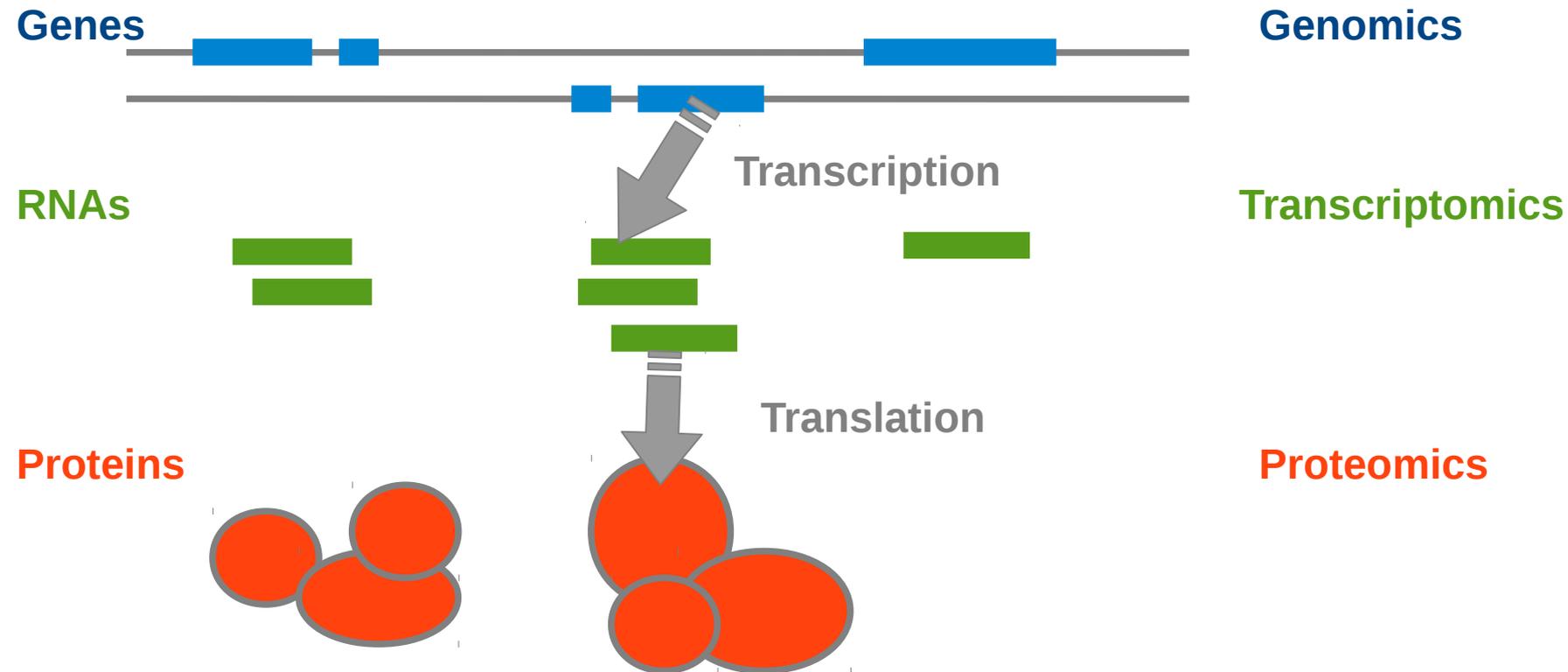
What is a metabolic network?



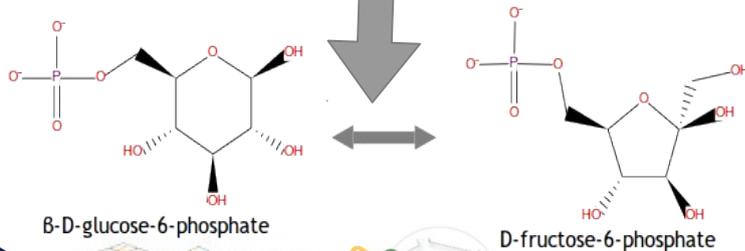
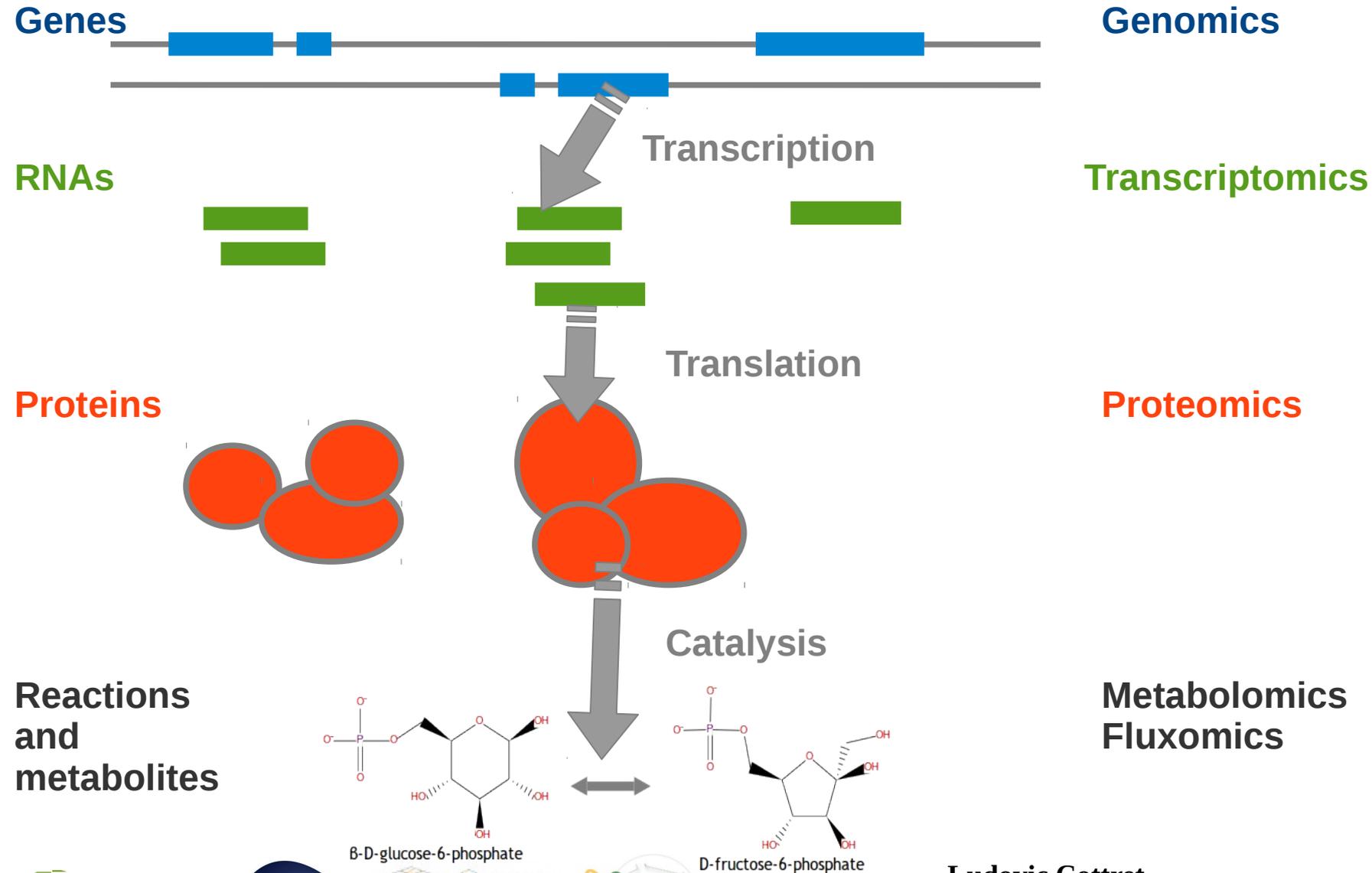
From genes to reactions



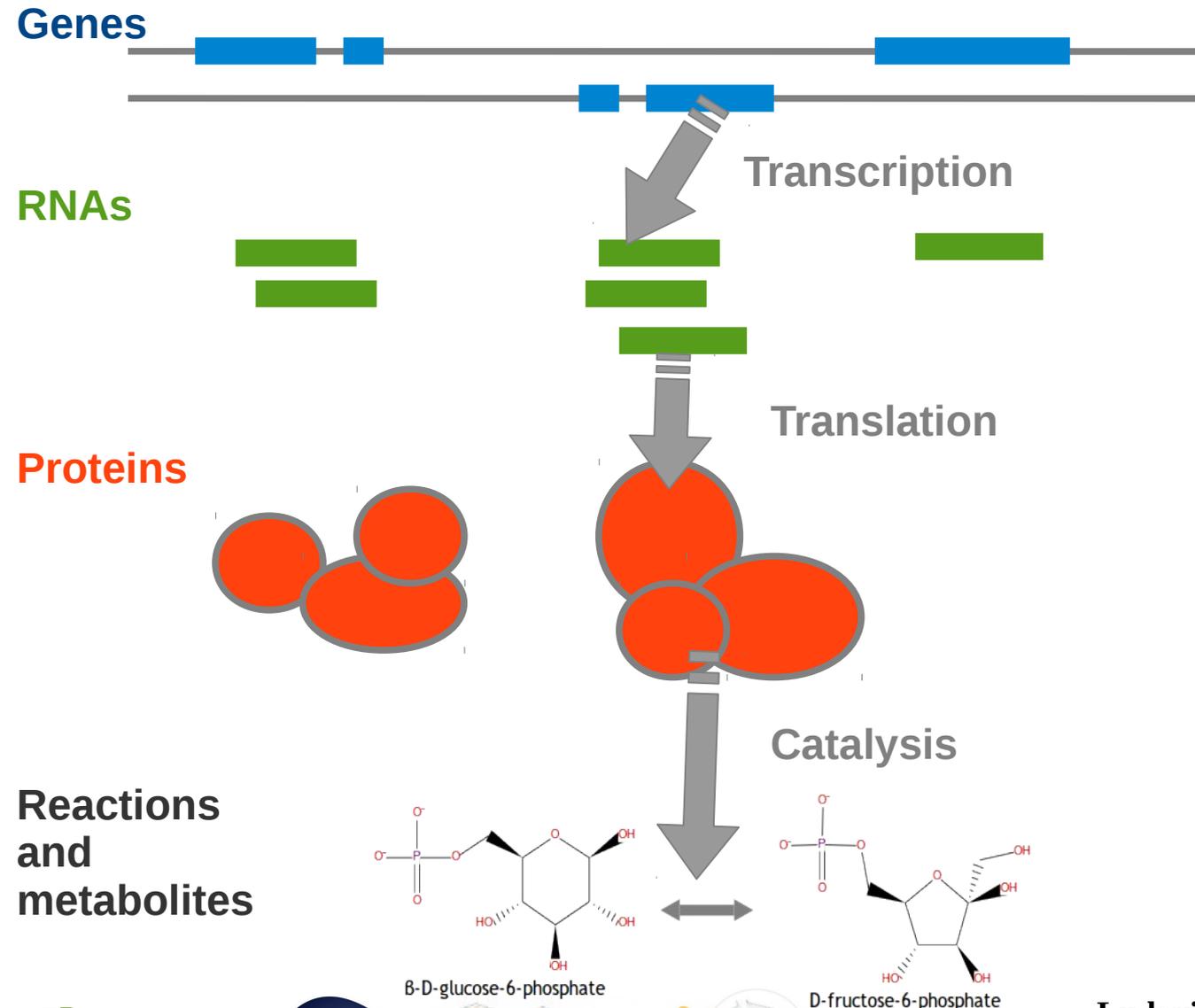
From genes to reactions



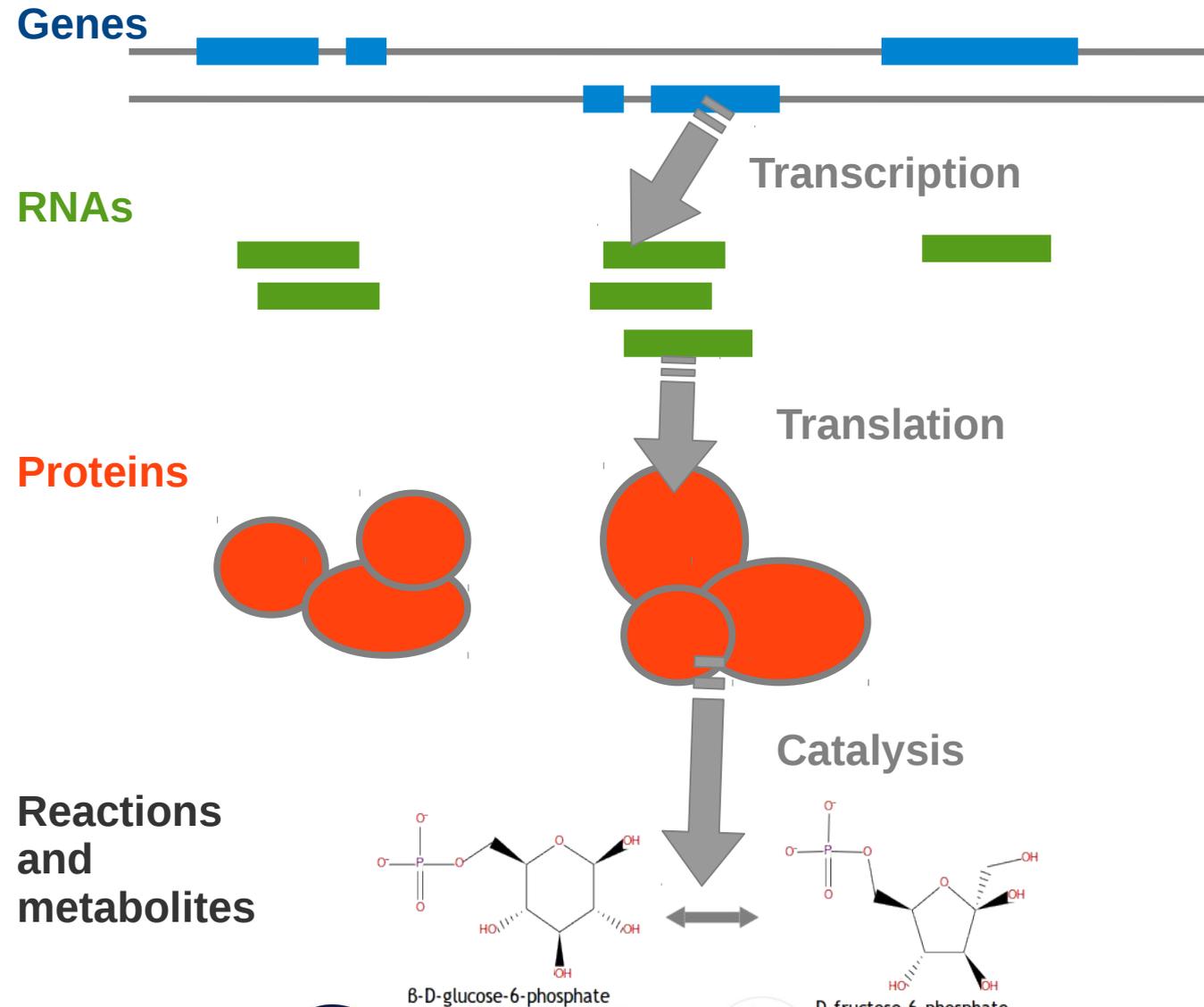
From genes to reactions



From genes to reactions



From genes to reactions

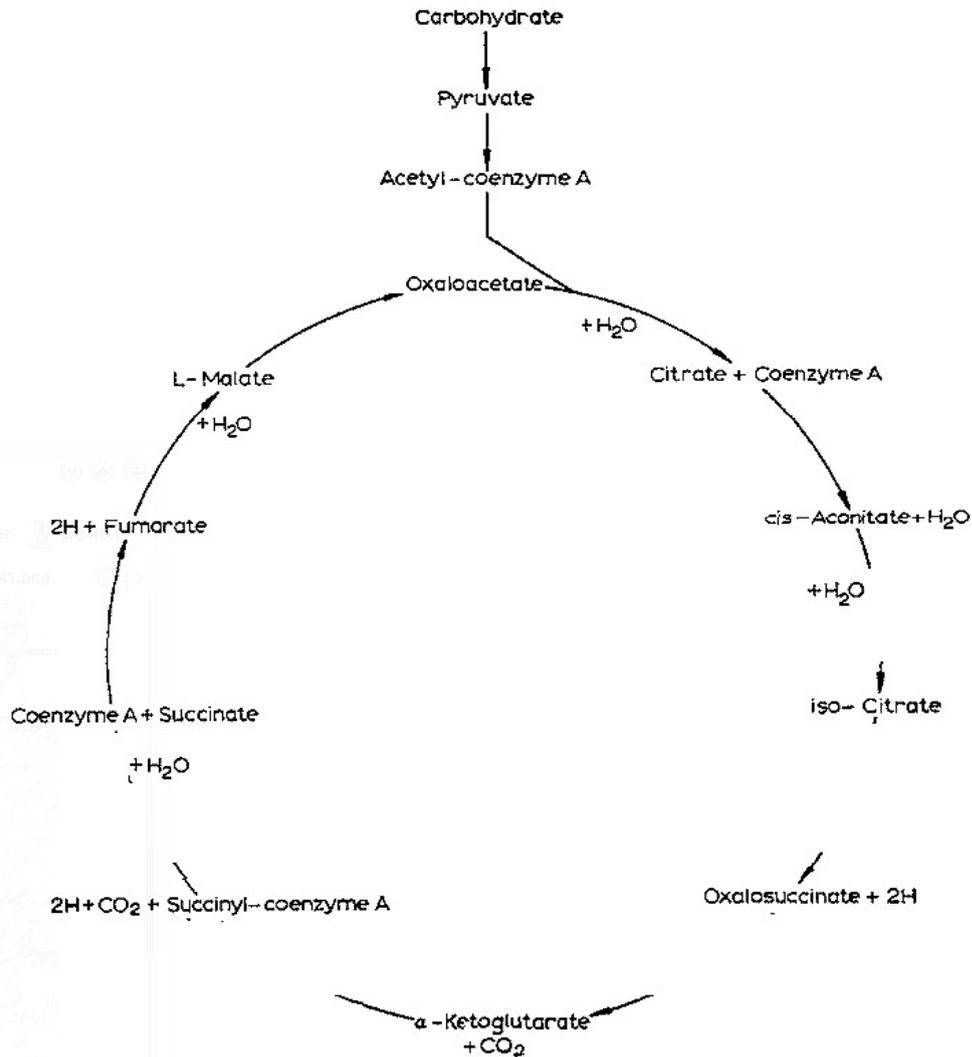


Metabolic network is a useful context for omics data interpretation...

... and visualisation facilitates this interpretation!

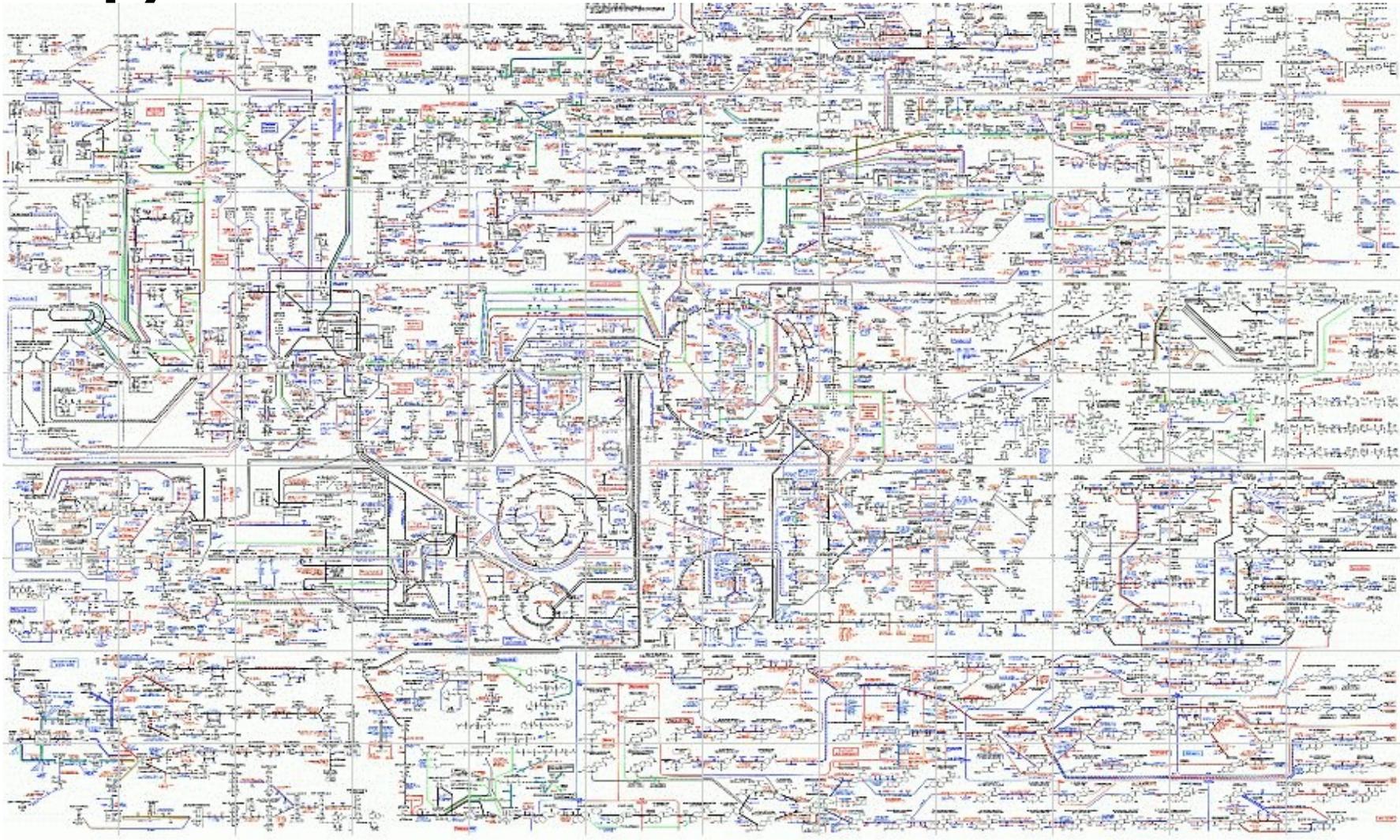
Metabolic network static visualisation

Before computers...

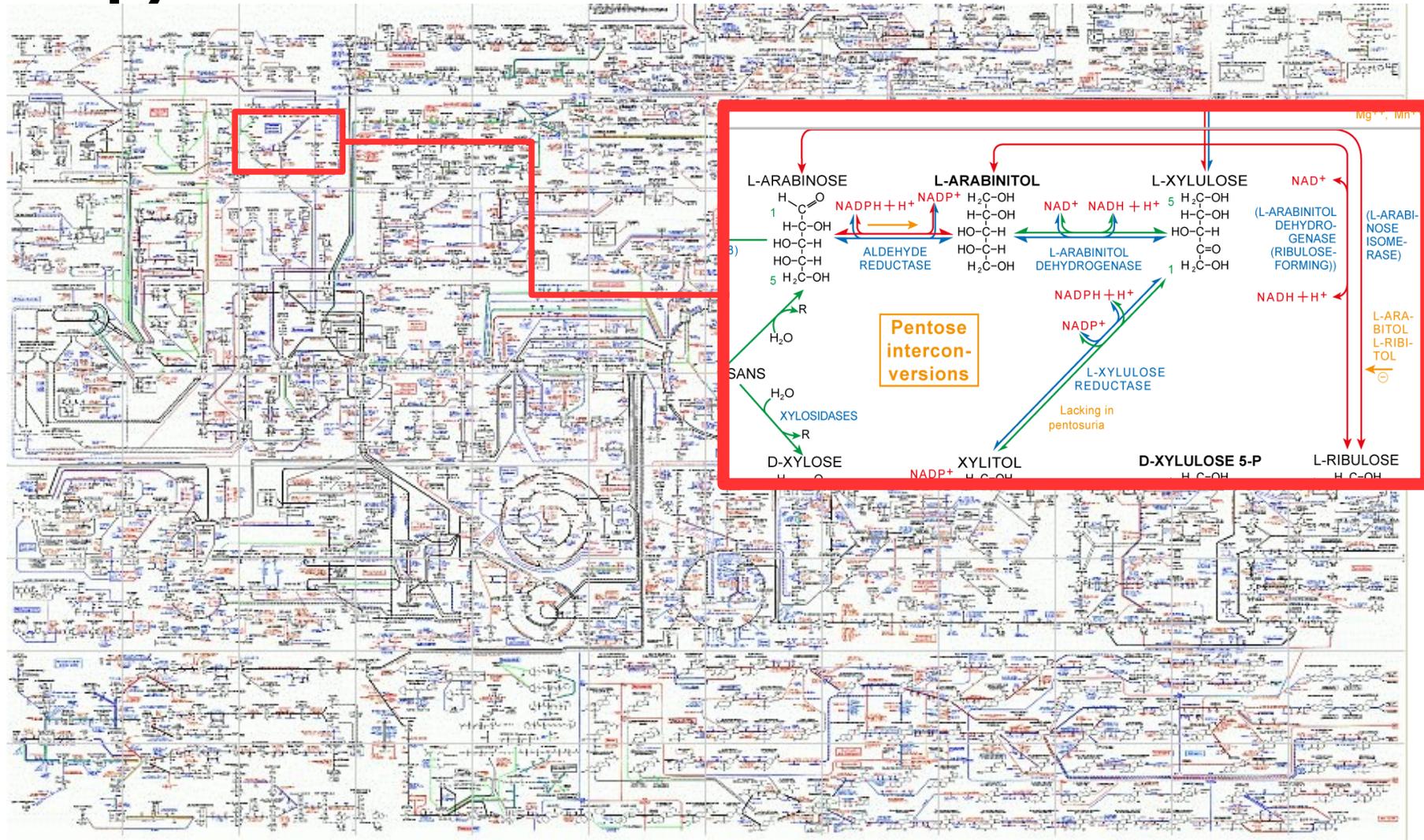


HANS A. KREBS
 The citric acid cycle
 Nobel Lecture, December 11, 1953

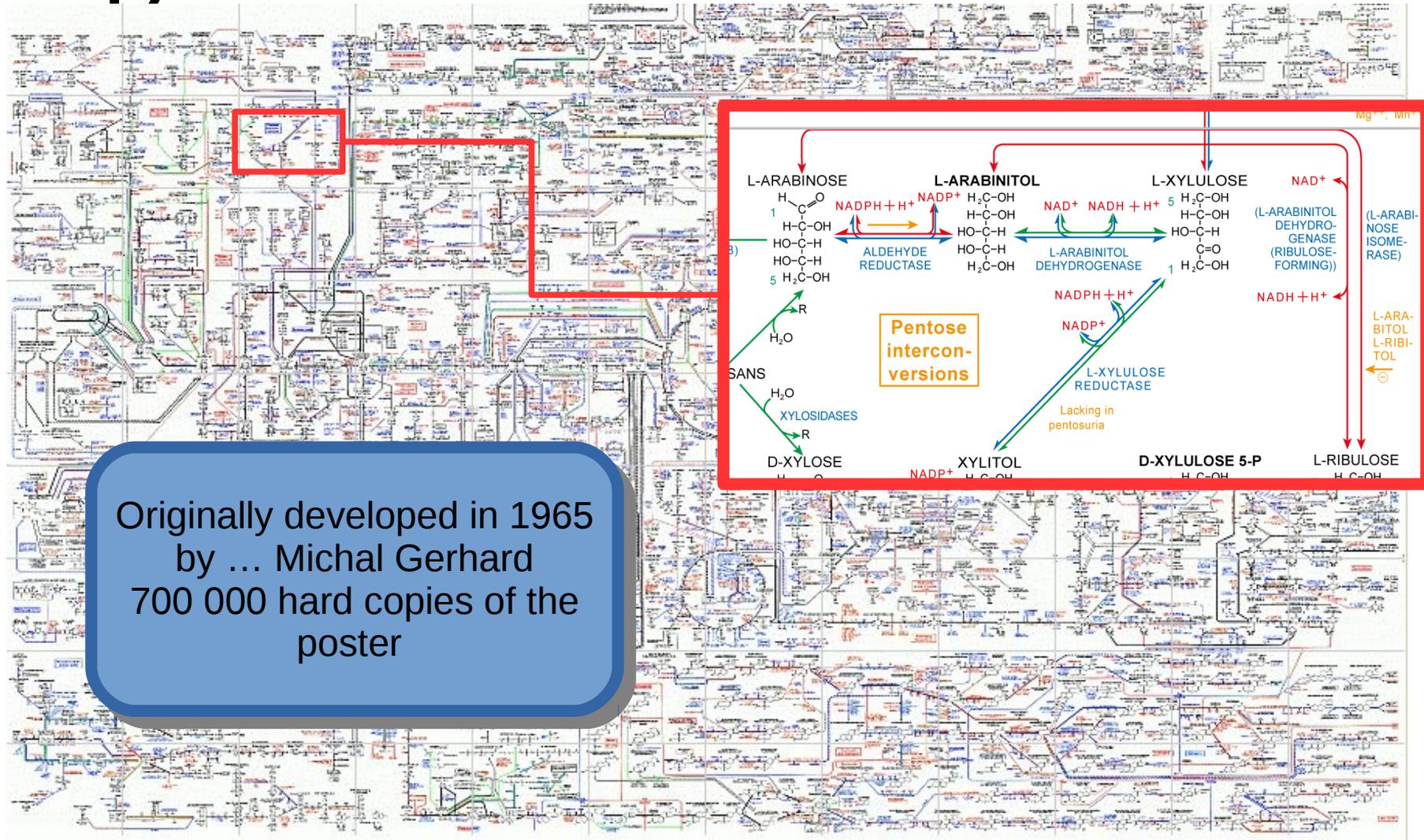
Roche BioChemical Pathways (Boehringer map)



Roche BioChemical Pathways (Boehringer map)



Roche BioChemical Pathways (Boehringer map)



Originally developed in 1965
by ... Michal Gerhard
700 000 hard copies of the
poster

Interactive Boehringer map

Biochemical Pathways Roche.com Contact Share

Part 1: Metabolic Pathways Part 2: Cellular and Molecular Processes

Search Filter Focus Legend

glycolysis

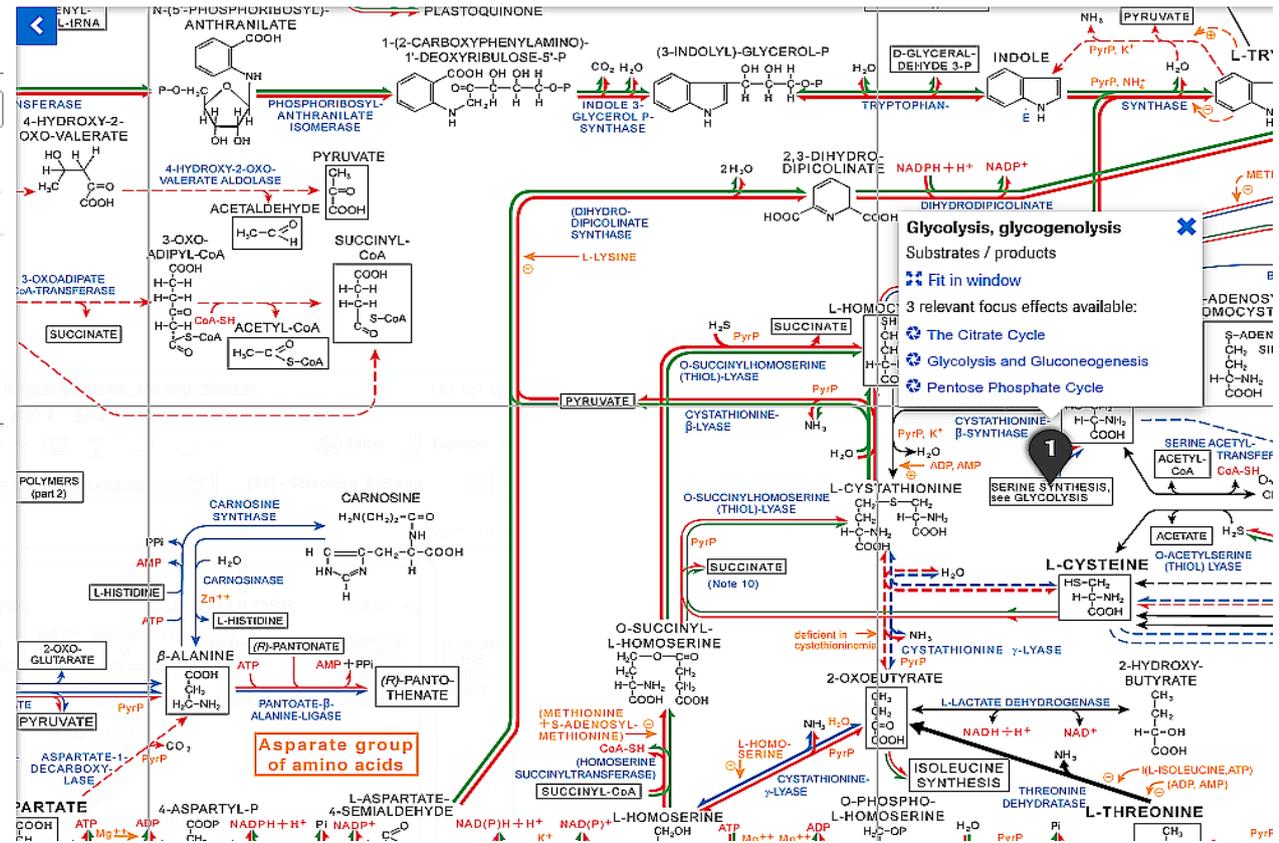
Show 1 matches in Part 2

1 matches in Part 1

1 Glycolysis, glycogenolysis
Substrates / products

3 relevant focus effects available:

- The Citrate Cycle
- Glycolysis and Gluconeogenesis
- Pentose Phosphate Cycle



<http://biochemical-pathways.com/>

Ludovic Cottret
Séminaire “Méthodes et outils pour l'Open Data”
Saint-Martin de Londres, 18 décembre 2014

Metabolic pathways in KEGG



Lysine biosynthesis - Reference pathway

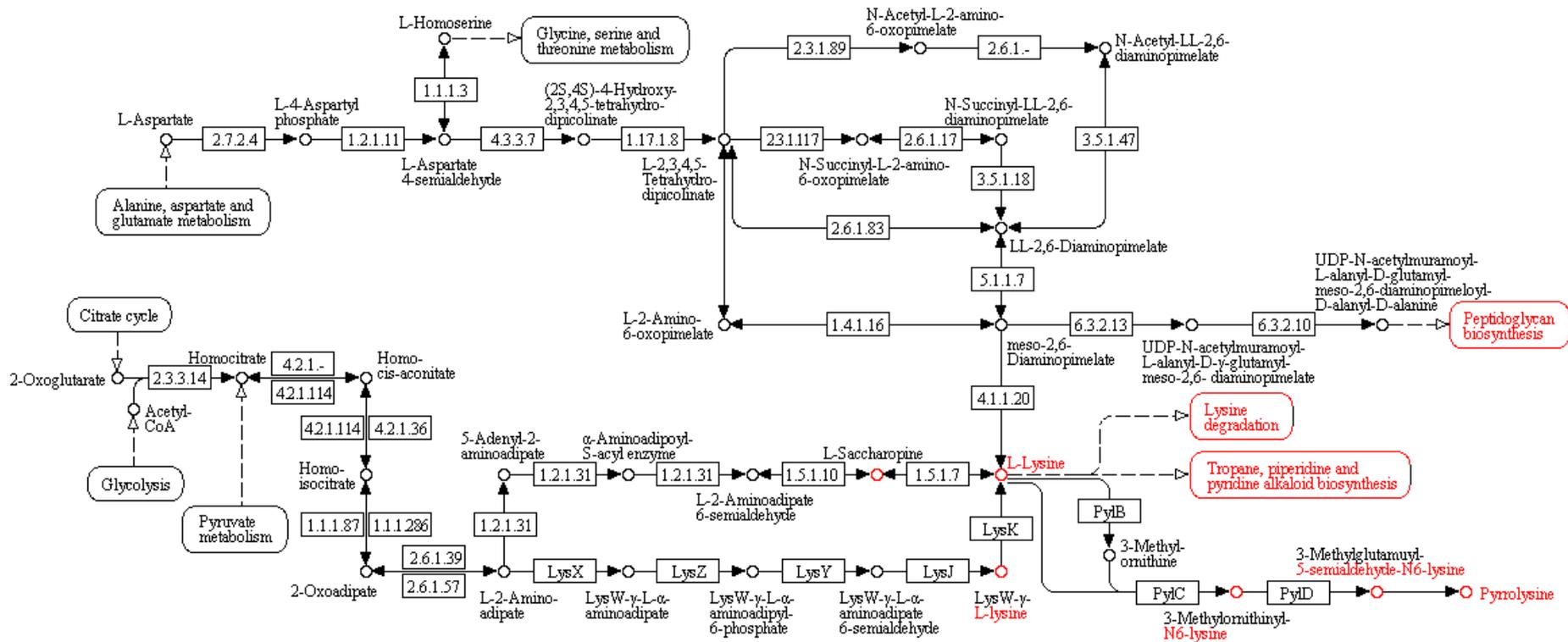
[Pathway menu | Organism menu | Pathway entry | User data mapping]

Reference pathway

Go

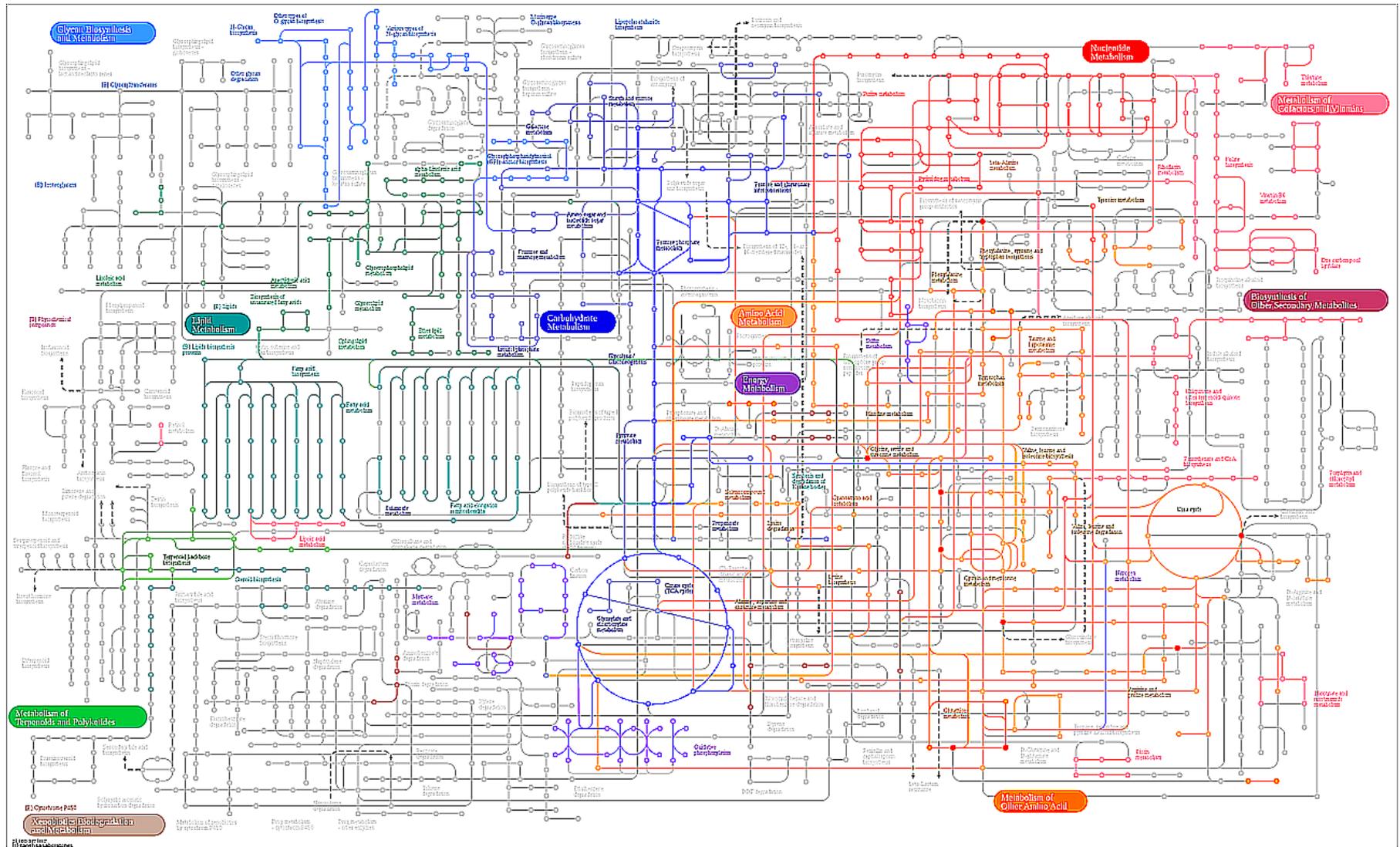
100%

LYSINE BIOSYNTHESIS



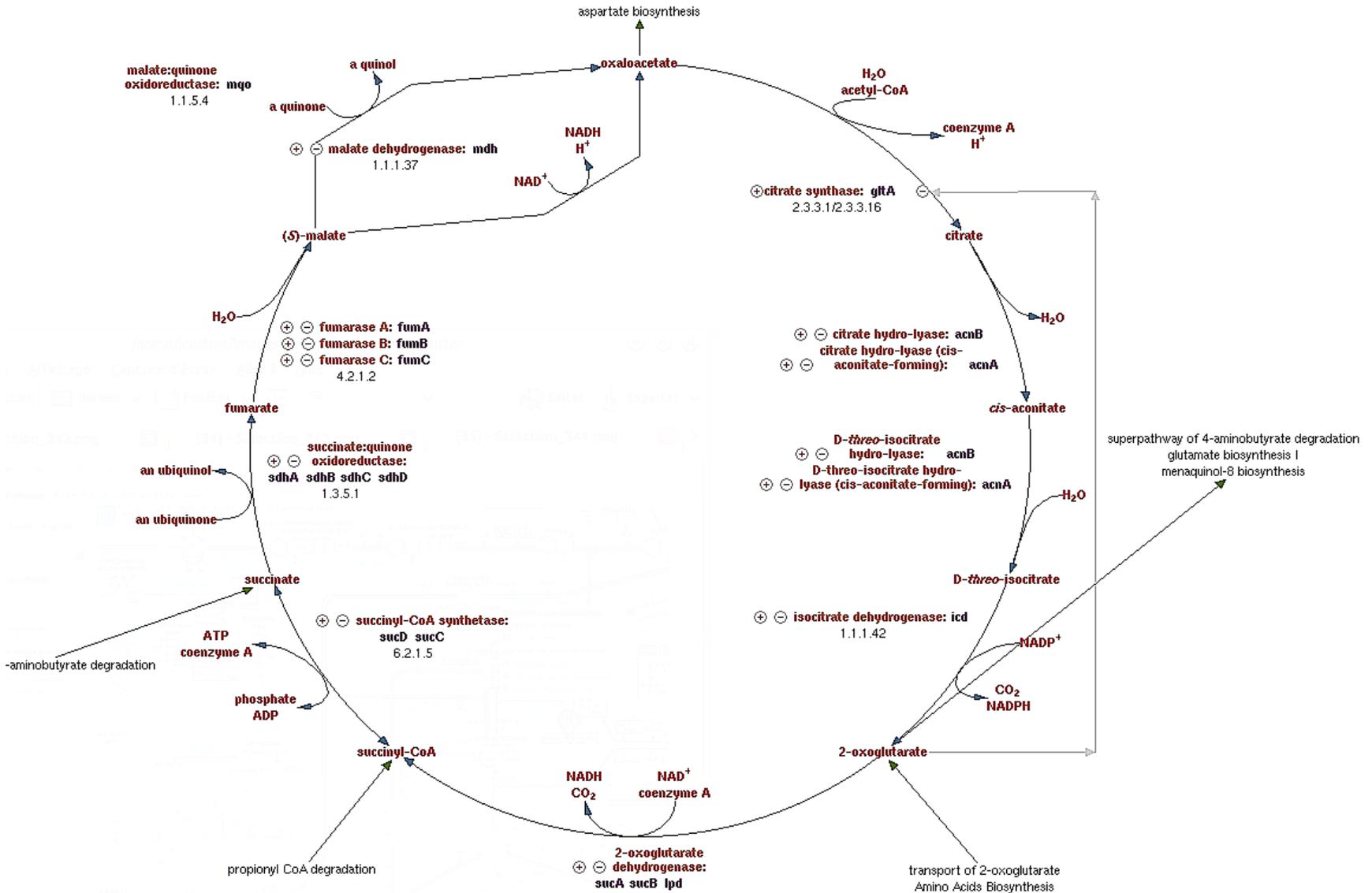
00300 11/19/13
(c) Kanehisa Laboratories

Metabolic pathways in KEGG



Metabolic pathways in BioCyc

More Detail Less Detail

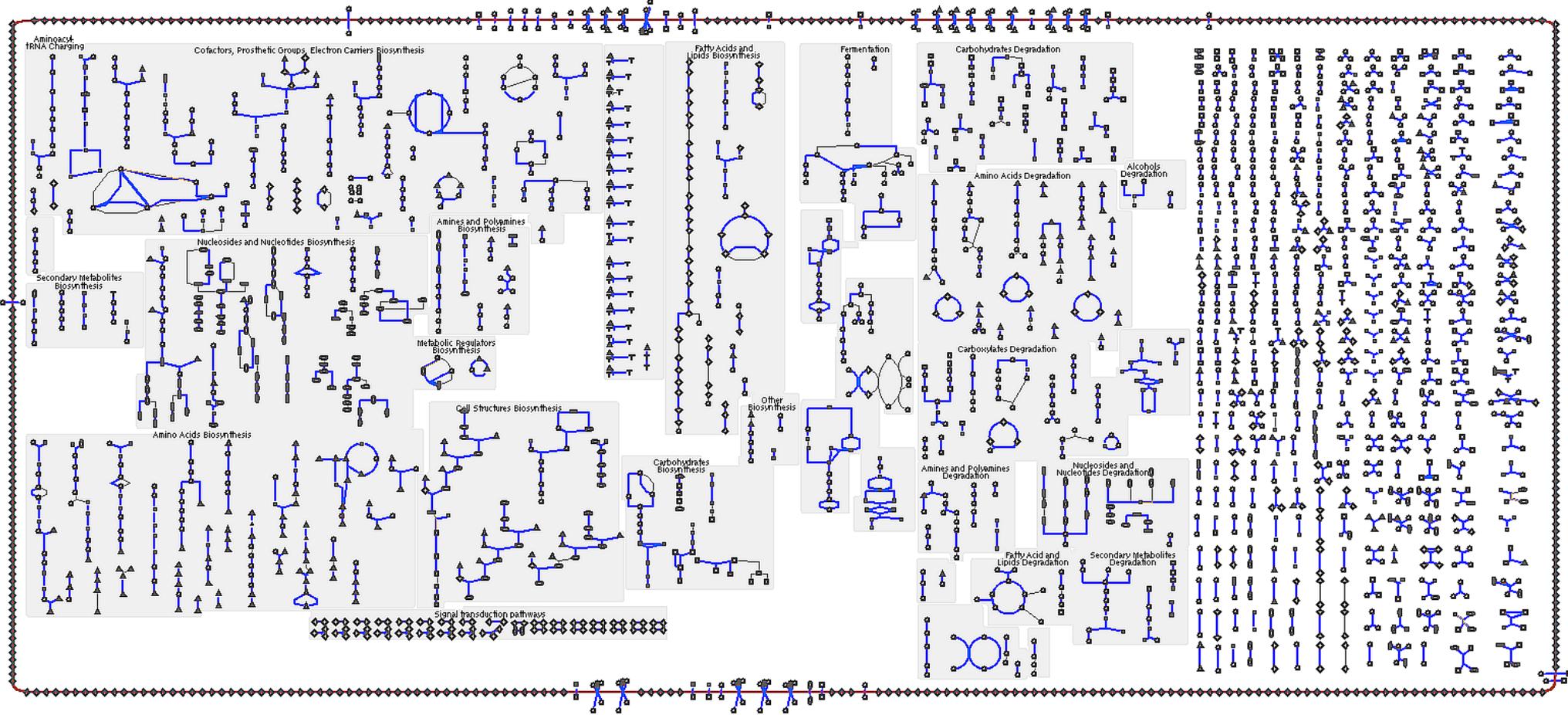
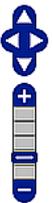


Metabolic pathways in BioCyc

Cellular Overview of *Bacillus subtilis subtilis* 168

Pan left/right/up/down the entire diagram by holding the left mouse button, click on an object for more info, right-click (ctrl-click for Mac) for menu

Cellular Overview



Creation and edition of metabolic pathways

Pathway editor in the pathway-tools (BioCyc)

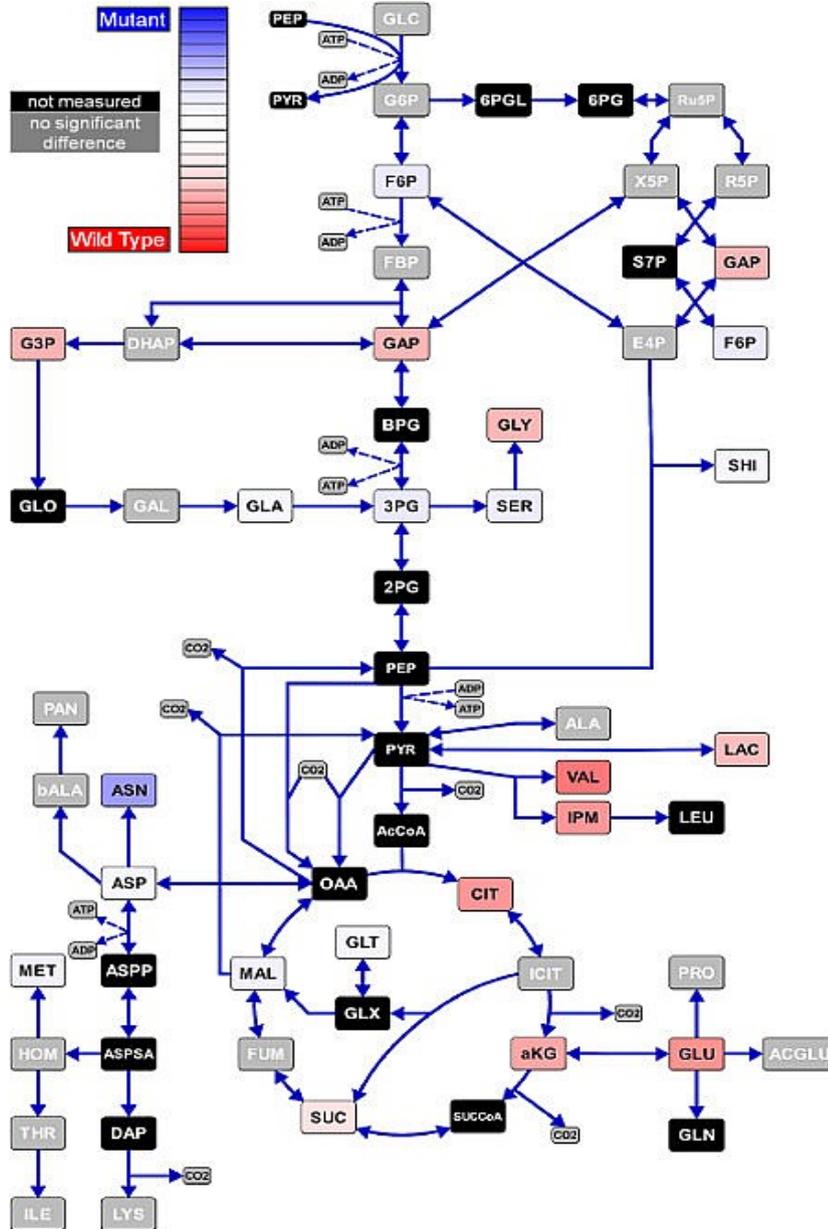
The screenshot displays the BioCyc pathway editor interface. The top menu bar includes "Exit", "Reaction", and "Pathway". Below it, the title bar reads "Editing Pathway test".

The left panel shows a reaction box with the following text:

```
AMINEPHEN-RXN  
1.4.3.21  
phenylethylamine + oxygen +  
H2O -> phenylacetaldehyde +  
ammonia + hydrogen peroxide + H+
```

The right panel shows a pathway diagram. At the top, "phenylacetaldehyde" is written in red. Two green arrows branch out from it. The left arrow points down to "phenylacetate". To the left of this arrow, the text "NAD⁺" and "H₂O" is written in red. Below the arrow, "2 H⁺" and "NADH" is written in red. The right arrow points down to "2-phenylethanol". To the right of this arrow, "NADH" and "H⁺" is written in red. Below the arrow, "NAD⁺" is written in red.

Omix



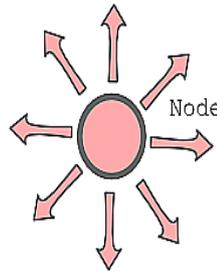
<http://www.omix-visualization.com/>

Metabolic graph visualisation

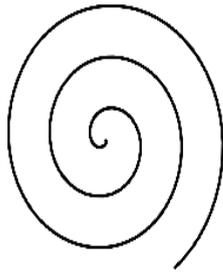
Principles

- One entity (reaction or metabolite) = one node
- Nodes are linked by interactions (edges)
- Layout algorithms for optimising the visualisation

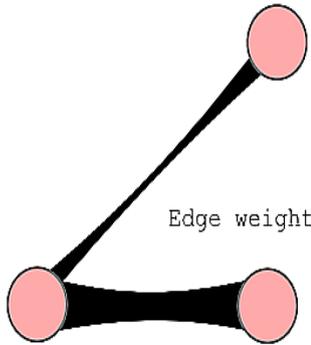
Force-directed layout



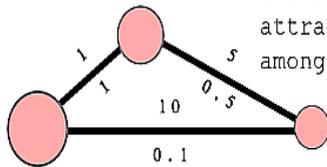
Node size exerts repulsion



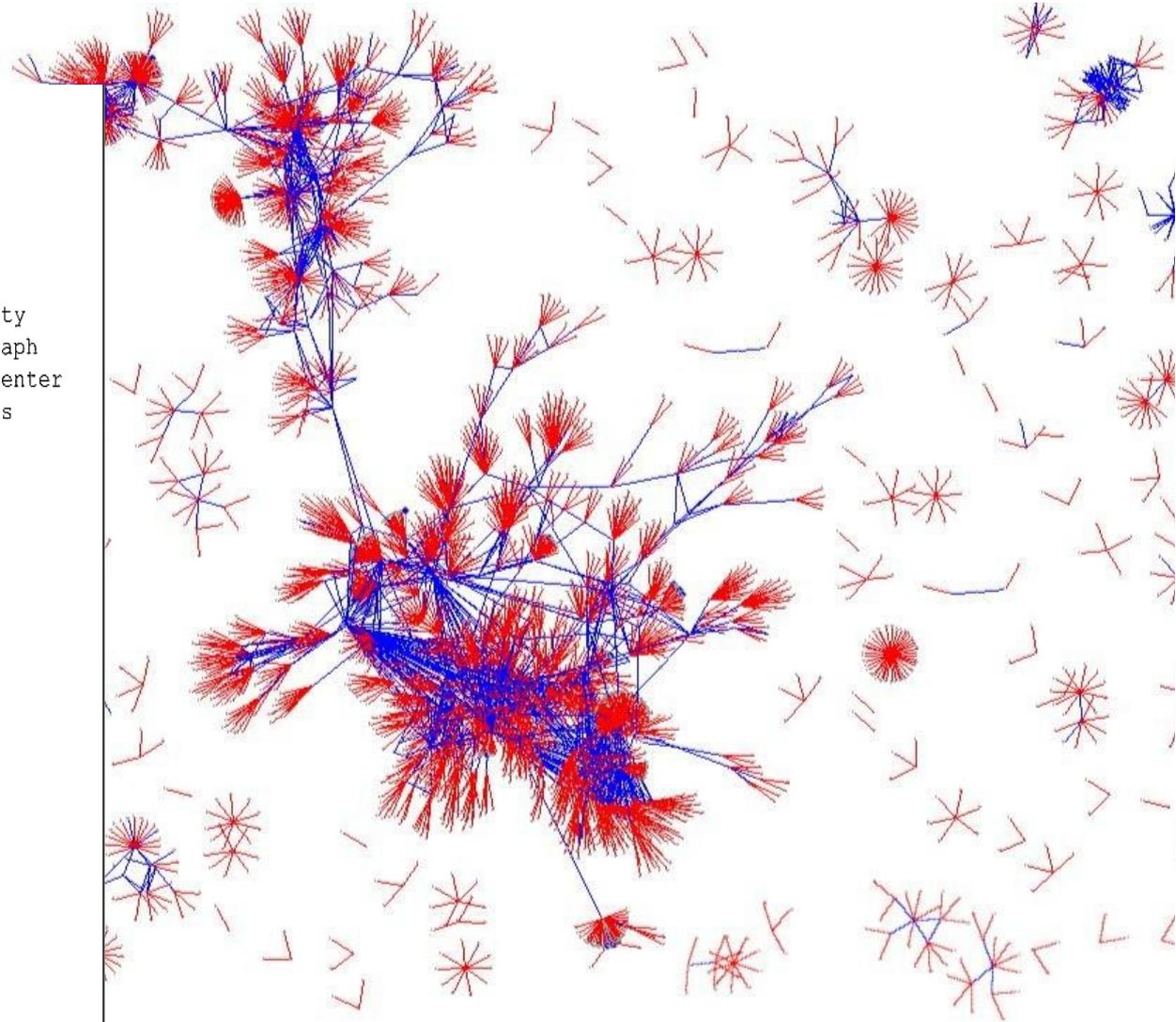
Canvas Gravity
Draws the graph
toward the center
of the canvas



Edge weight exerts attraction



Force-Directed Algorithm
Shows the results of
attraction and repulsion
among all nodes



<https://dhs.stanford.edu/visualization/more-networks/>



INRA
SCIENCE & IMPACT



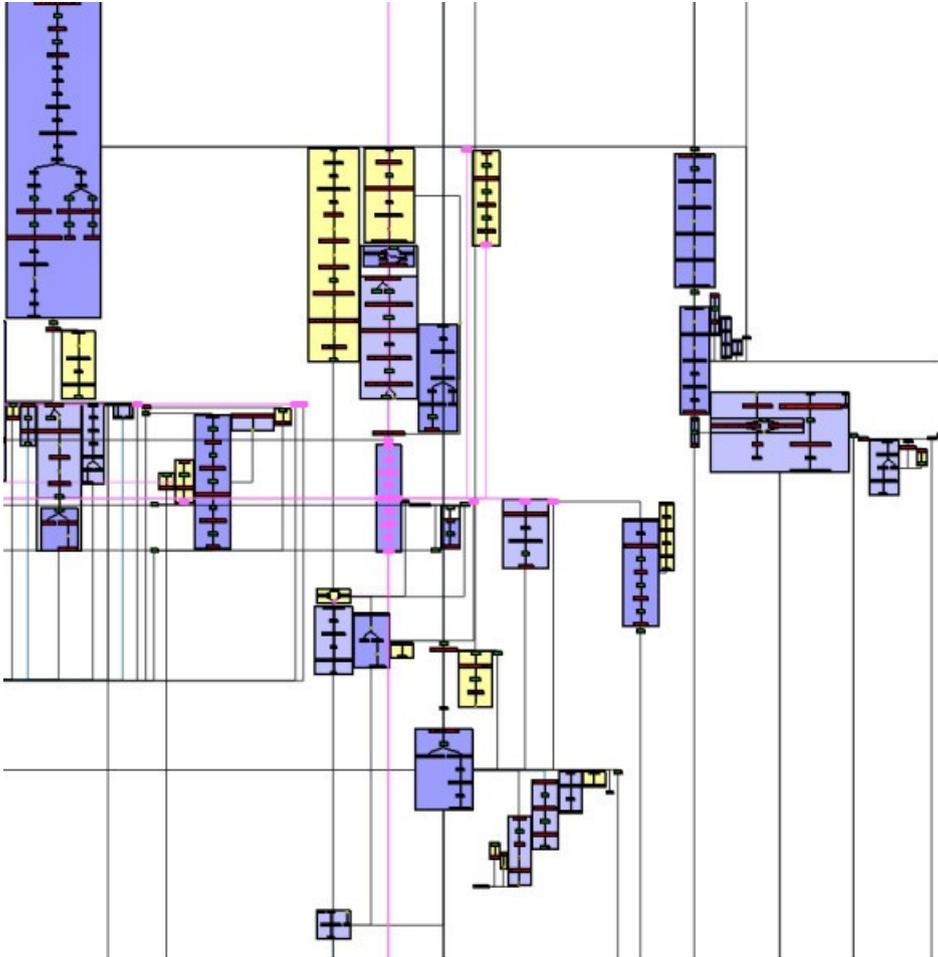
Bioinformatique
Biodiversité
Représentation
& Intégration
des Connaissances



Laboratoire Interactions Plantes Micro-organismes

Ludovic Cottret
Séminaire “Méthodes et outils pour l’Open Data”
Saint-Martin de Londres, 18 décembre 2014

Layout designed for metabolic networks

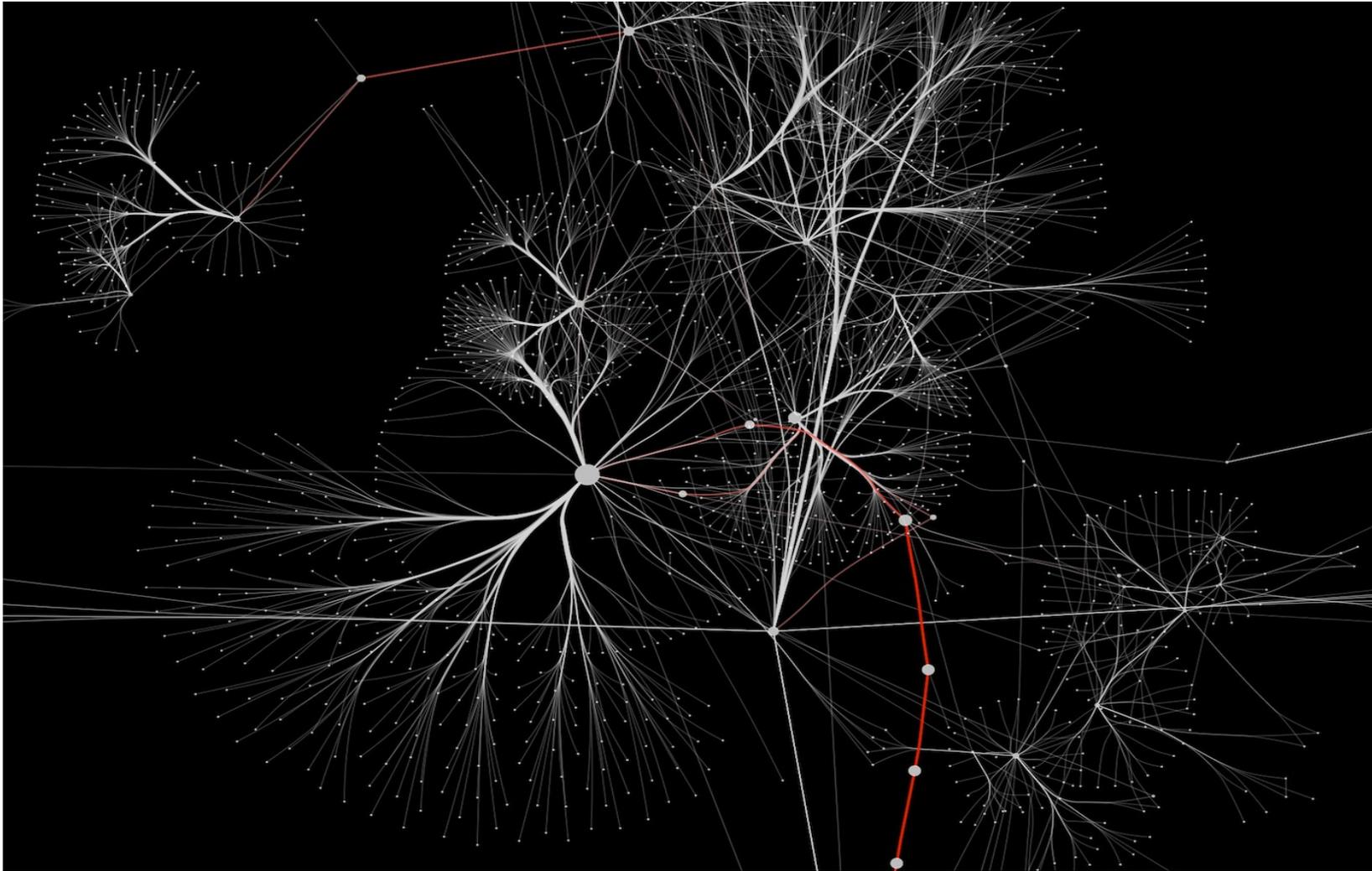


Objectives

- when possible, preserves the classification of the reactions in metabolic pathways
- takes into account the metabolic pathway drawing conventions

Bourqui R, Cottret L, Lacroix V, et al. Metabolic network visualization eliminating node redundancy and preserving metabolic pathways. BMC Systems Biology 2007;1:29.

Cytoscape Java



Cytoscape JAVA

- user friendly
 - visual styles
 - a lot of plugins
 - may be linked to Web pages by Java Web start
- Version 3.0 :
 - V2 plugins can not be used
 - Web start not possible?



Tulip / Systrip

<http://tulip.labri.fr/TulipDrupal/?q=systrip>

Trypanosoma brucei
Nodes : 1168 Edges : 1294

Id	Name	Type	Nodes	Select
▶ Unclasi...			0000326	00000
▶ superpa...	superpath...	Pathway	0000058	00000
▶ superpa...	superpath...	Pathway	0000042	00000
▶ purine n...	purine nuc...	Pathway	0000039	00000
▶ superpa...	superpath...	Pathway	0000035	00000
▶ superpa...	superpath...	Pathway	0000034	00000
▶ glycolysi...	glycolysis III	Pathway	0000034	00000
▶ superpa...	superpath...	Pathway	0000033	00000
▼ peptido...	peptidogly...	Pathway	0000033	00000
UDP...	UDP-N-ace...	Reaction		
UTP	UTP	Compound		
UDP...	UDP-N-ace...	Reaction		
NAC...	NA	Reaction		
UDP...	UDP-N-ace...	Reaction		
UDP...	UDP-N-ace...	Reaction		
N-AC...	N-acetyl-gl...	Compound		
D-GLT	D-glutamate	Compound		
L-GL	Glucosami...	Reaction		
D-GL	D-glucosa...	Compound		
UDP...	myo-inosit...	Compound		

Views

- Sub-view
 - threonine degradation I
 - CYS
 - AMINO-OXOBUT
 - GLYOXIL-RXN
 - S-LACTOYL-GLUTATHIONE
 - METHYL-GLYOXAL
 - AKBLIG-RXN
 - THREODEHYD-RXN
 - THREOSON-RXN
 - AMINO-ACETONE
 - CYSTACLY-RXN
 - THS
 - D-LACTATE
 - 2-OXOBUTANOATE
 - GLYOXIL-RXN

Display pathways Display compounds Display reactions
 All Elements Selected Elements

Network Navigation | Element properties | Data Sets

Trypanosoma brucei
TTR-DUTP 5.3.3.5
FARNESYL-PP GLYCOLALDEHYD
DOLICHOL LYS 3.8.1.2
2.4.1.- GUANINE D-LACTATE
[Fatty-Aldehydes] 1.3.99.3 null 4.2.1.18
2.7.1.68 2.7.8.5 2.7.9.1 4.3.2.2 CPD-61
3.6.1.7 2.6.1.1 3-KETO-ADIPATE
1.1.99.1 null CPD-650
2.1.1.103 CPD-1301 CPD-207
6.3.2.13 NA
CPD1V8-181 1.1.1.4 HOMO-SER
CTAPRENYL-4-HYDROXYBENZOATE

Sub-view
5-P-RIBOSYL-N-FORMYLGLYC
2.1.2.2 ALA
THR DGDP 3.5.4.1
DGLACTATE DGTP
1.1.1.1 PYRUVATE
1.1.1.1 UDP-AA GLUTAMATE
C6 GLT
COSAMINE-6-P
LPOAM
4.1.1.3
1.3.1.6

Histogram view : data_TrypanoSanger_3ppm.csv
number of nodes
5
4
3
2
1
0
-sd m +sd

Parallel Coordinates view : data_TrypanoSanger_...

Information View
Go Back

REACTION: R06417

Entry	R06417	Reaction
Name	(+) -cis-isopulegone:NADP+ oxidoreductase	
Definition	(-) -Isopiperitenone + NADPH + H+ <=> (+) -cis-	
Equation	C02485 + C00005 + C00080 <=> C11951 + C00006	

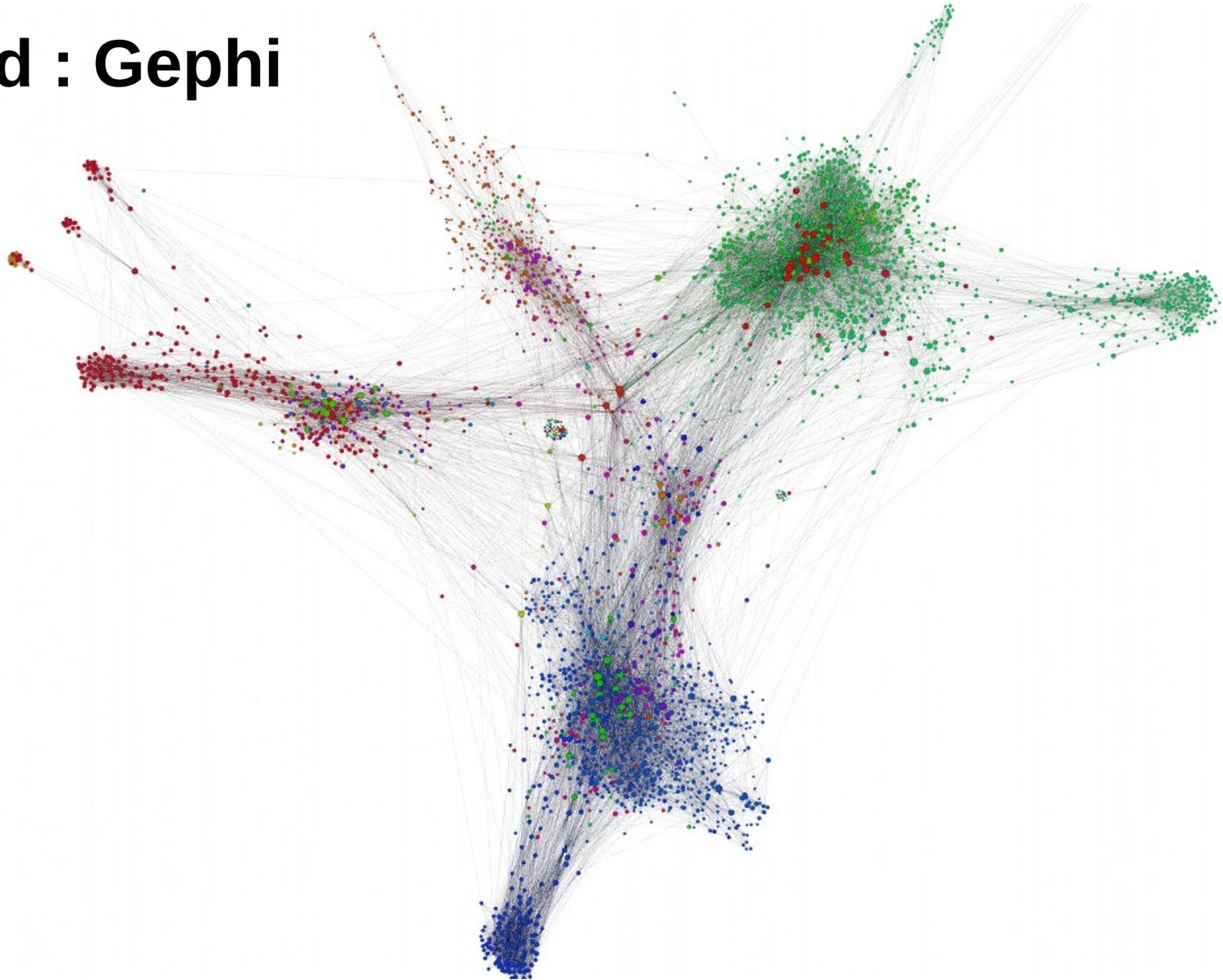
pyruvate

Tulip et Systryp

- Powerful (OpenGL)
- Many graph measures
- Not very user friendly
- Nobody to maintain !!!

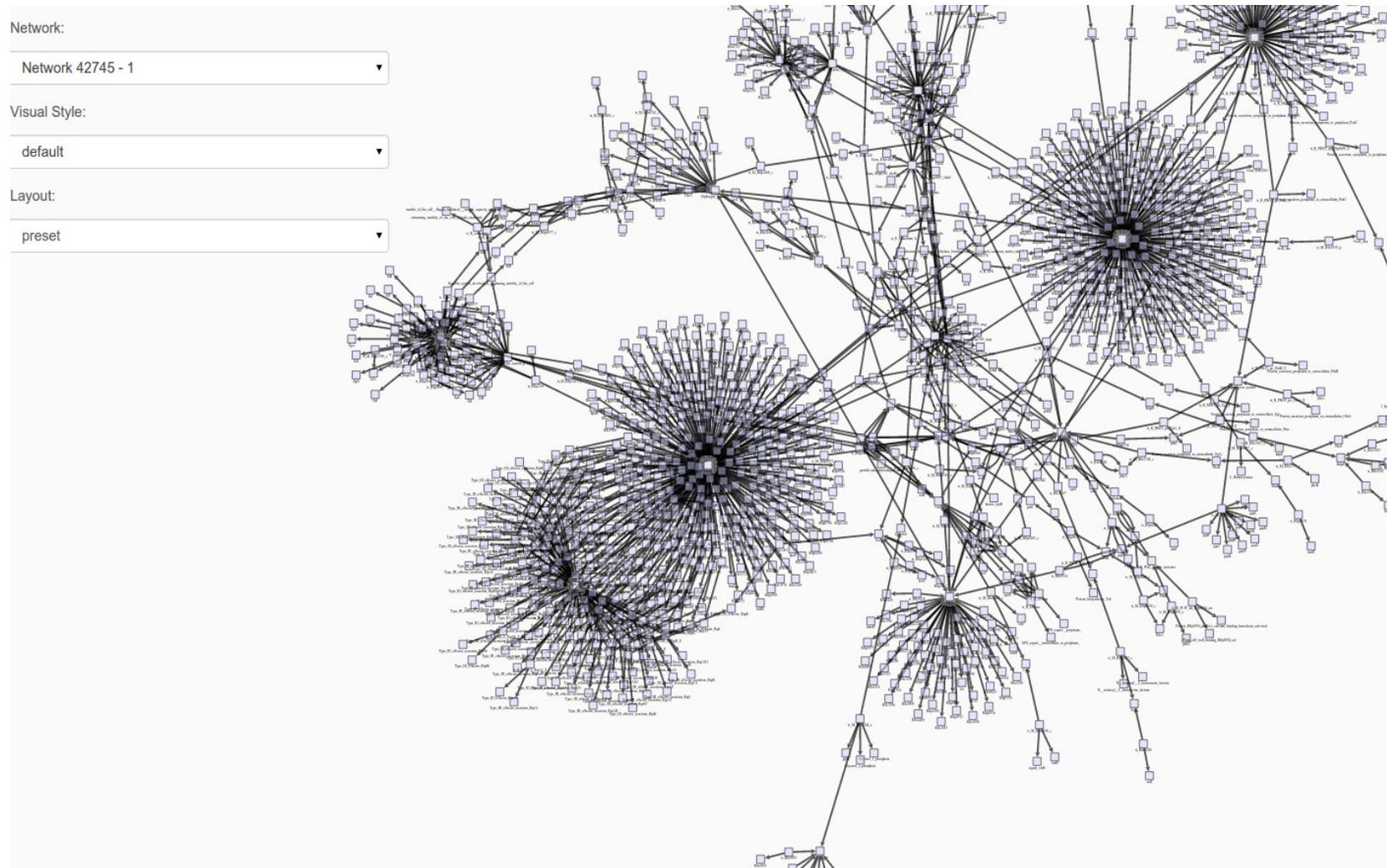


Trend : Gephi



Metabolic network web visualisation

Cytoscape.js



Cytoscape.js

- Large community
- Links with Cytoscape Java
- Lags with large networks



d3.js

- Powerful
 - More generic
 - Large community
- Interactions more difficult to handle (→ SVG)



Future : Network visualisation with WebGL ?

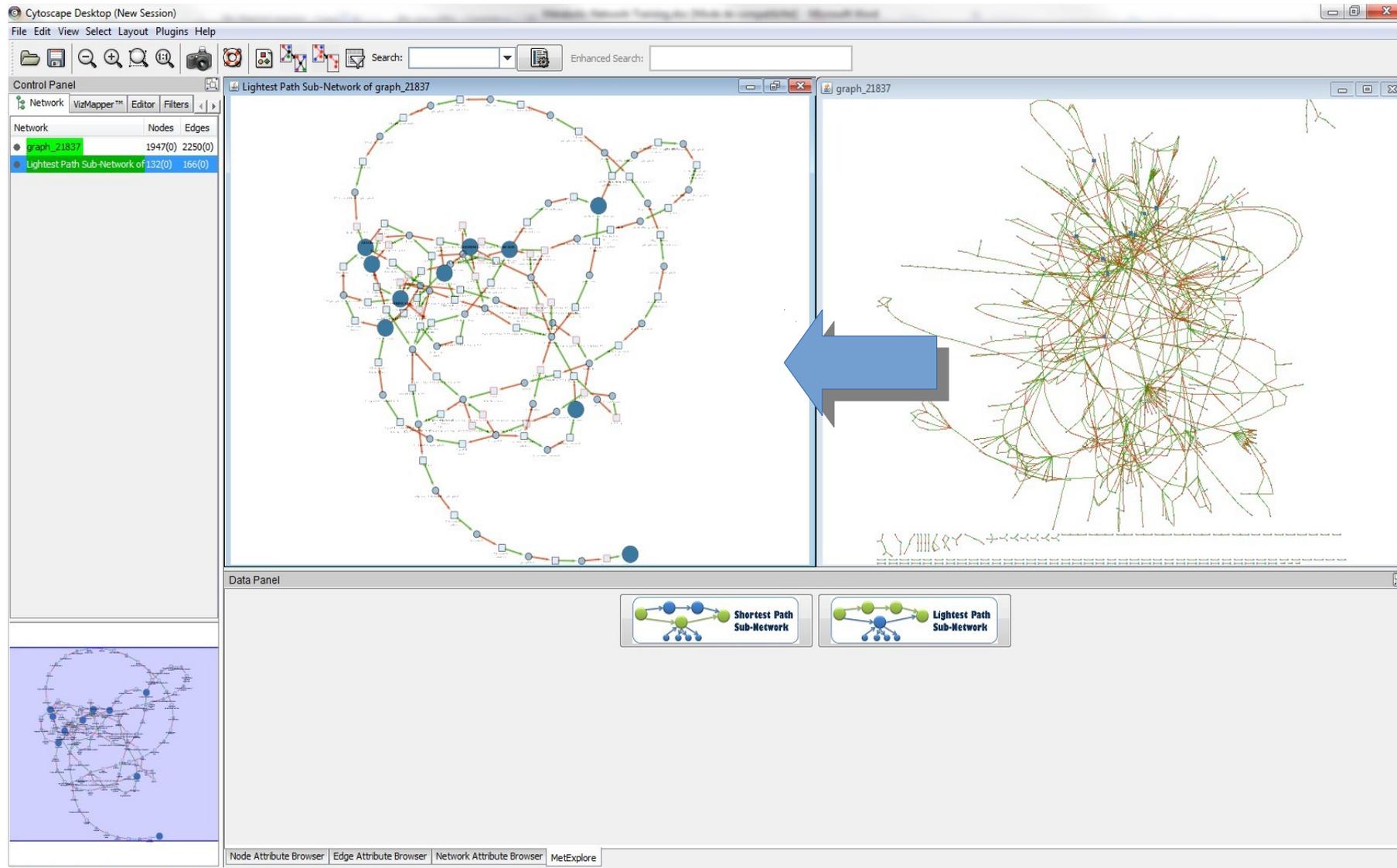
Ex: sigmaajs

- Uses the graphic cards

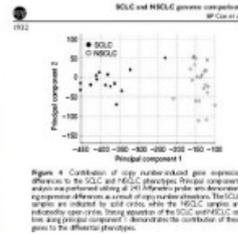
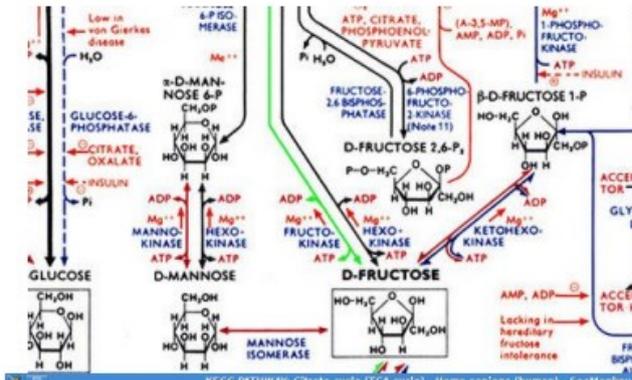


<http://sigmaajs.org/>

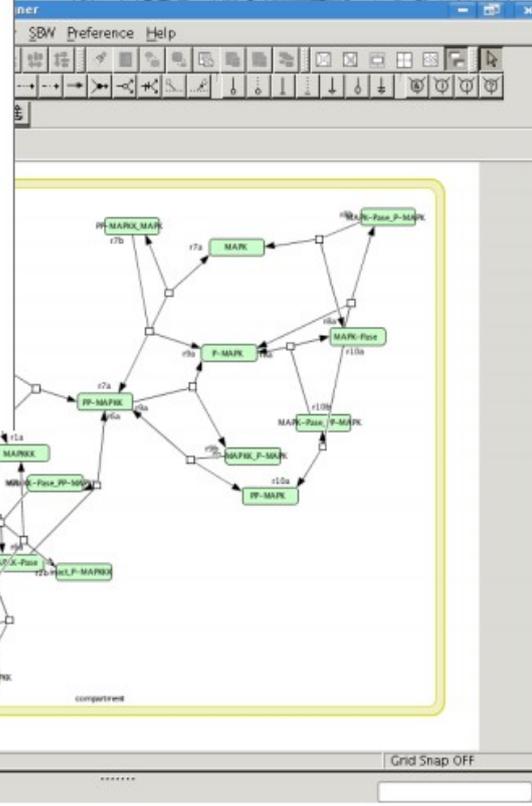
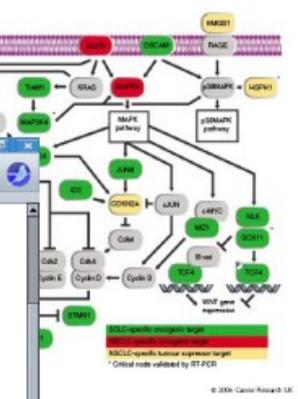
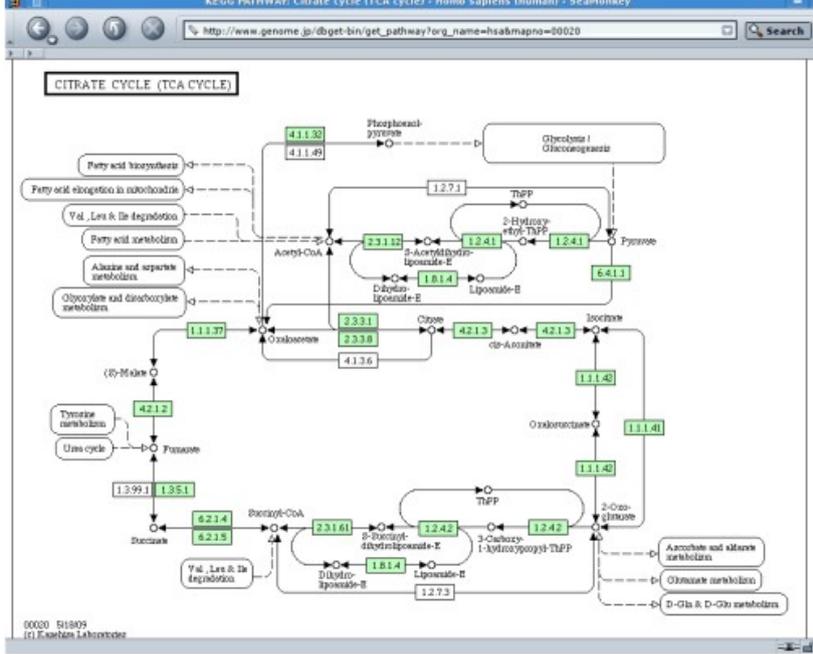
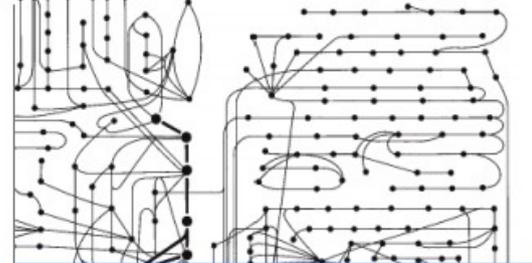
Sub network identification helps the visualisation



How to share network views?



implication in the SGLC complex as well suggests that this gene may play an essential role in the development of long neurons (Santoro et al. 2006).
 It is noteworthy that a subset of the genomic alterations between the SGLC and HSLC cell lines could be the result of adaptation to culturing conditions. Testing this hypothesis through analysis of different tissues than their counterparts in genomic alterations and gene regulation.
 Regions of difference
 Through our analysis, numerous regions throughout the genome were identified to be differentially altered between the SGLC and HSLC samples. This difference-based approach compares the regions of copy number variants and identifies the regions that are strongly linked to clinical disease. These regions suggest to use from whole genome sequencing (WGS) to identify genetic alterations in the SGLC and HSLC cell lines. In addition, the SGLC and HSLC cell lines were followed by additional sequencing, we also identified several regions that differed strongly in their alternative splicing patterns. We refer to these as phenotype-specific copy number alterations (PS-CNA). These included 19q13.31-19q13.42.



From N. Lenovère's slides

Ludovic Cottret
 Séminaire "Méthodes et outils pour l'Open Data"
 Saint-Martin de Londres, 18 décembre 2014



INRA
 SCIENCE & IMPACT



Bioinformatique
 Biodiversité
 Représentation
 & Intégration
 des Connaissances

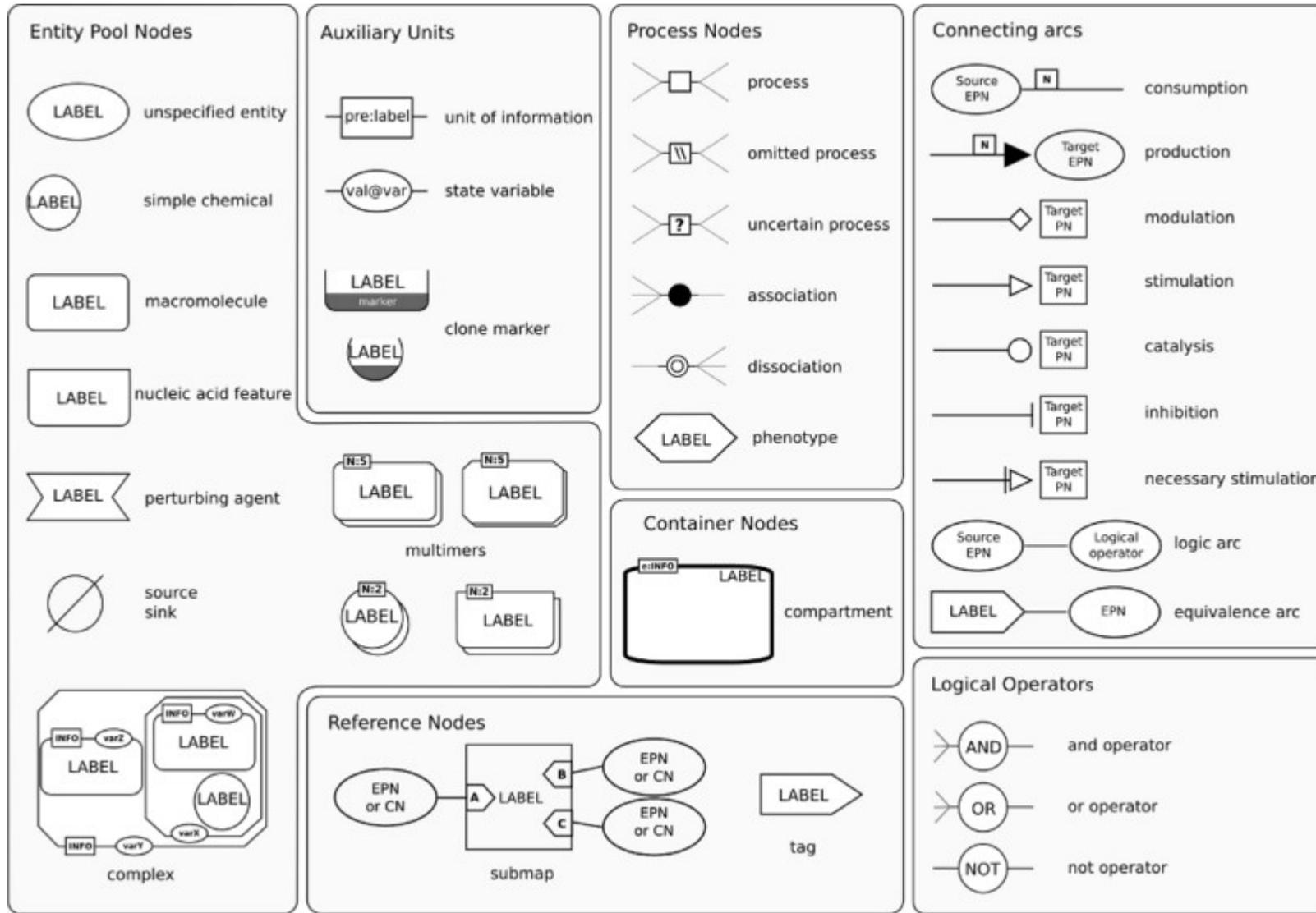


Laboratoire Interactions Plantes Micro-organismes

Challenges

- Numerous visual styles
- Layouts not always deterministic

Systems Biology Graphical Notation (SBGN)



<http://www.sbgn.org/>

Ludovic Cottret
 Séminaire “Méthodes et outils pour l'Open Data”
 Saint-Martin de Londres, 18 décembre 2014

Take home messages

- Network visualisation is now interactive
- Large network visualisation is still a challenge
- Sub network identification helps the interpretation
- Network visualisation in web is now possible
- Standardisation initiatives allows sharing biological network views



INRA
SCIENCE & IMPACT



Bioinformatique
Biodiversité
& Représentation
des Connaissances



Laboratoire Interactions Plantes Micro-organismes

Ludovic Cottret
Séminaire “Méthodes et outils pour l'Open Data”
Saint-Martin de Londres, 18 décembre 2014