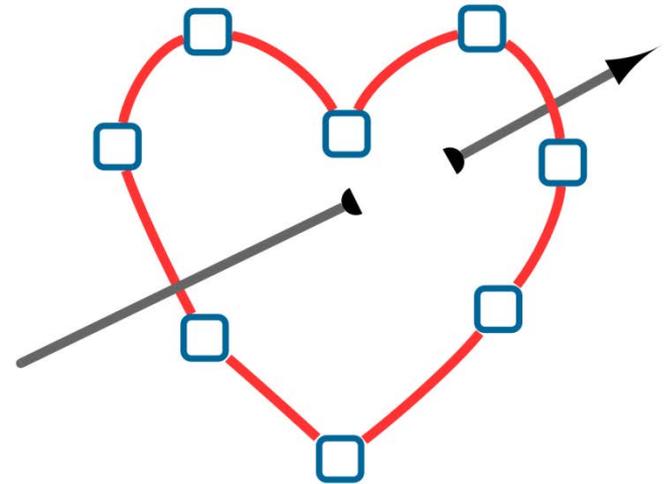




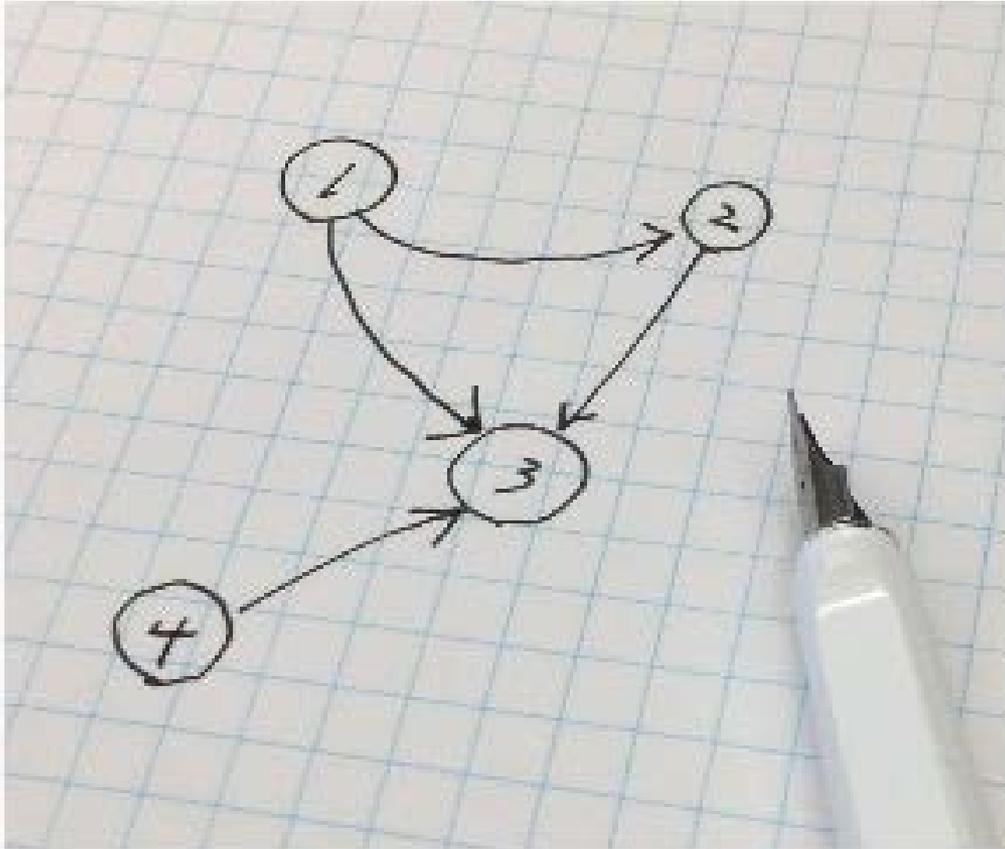
Cytoscape et moi, une histoire qui dure!

Ludovic Cottret

Work4Graph, le 7 décembre 2020



➤ Qu'est ce qu'un réseau?



- **Réseau** = ensemble de noeuds et d'arêtes
- **Noeud** = n'importe quel objet
- **Arête** = Relation entre 2 noeuds

➤ Pourquoi dessiner un réseau ?

- Une fois que les noeuds et les arêtes sont définis, le dessin d'un réseau est compréhensible par n'importe quelle communauté
- Utile pour résumer un ensemble de données
- Facilite l'interprétation des données



INRAE



Cytoscape et moi, une histoire qui dure !

07/12/2020 / Work4Graph / Ludovic Cottret

➤ Les réseaux sont largement utilisés

- Réseaux de transport
- Réseaux sociaux
- Réseaux biologiques : réseau d'interactions protéines-protéines, réseaux de régulation génétique, réseaux métaboliques, réseaux trophiques, etc...



➤ Histoire de Cytoscape

- 2002 : naissance à l'Institute for Systems Biology aux USA
- Mars 2003 : Version 1.0
- Novembre 2003 : papier dans Genome Research
- 2004 : version 2.0
- 2013 : version 3.0
- Aujourd'hui : version 3.8.2

➤ Genome Res. 2003 Nov;13(11):2498-504. doi: 10.1101/gr.1239303.

Cytoscape: a software environment for integrated models of biomolecular interaction networks

Paul Shannon¹, Andrew Markiel, Owen Ozier, Nitin S Baliga, Jonathan T Wang, Daniel Ramage, Nada Amin, Benno Schwikowski, Trey Ideker

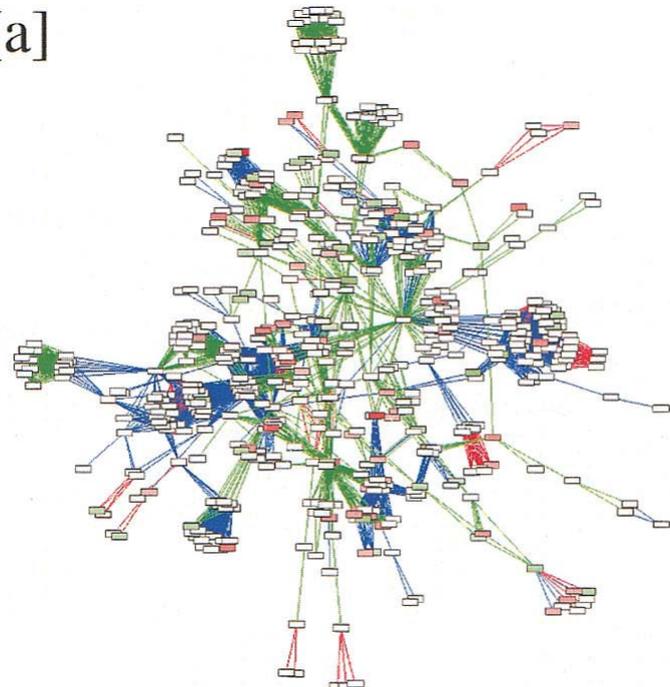
Affiliations – collapse

Affiliation

¹ Institute for Systems Biology, Seattle, Washington 98103, USA.

PMID: 14597658 PMCID: PMC403769 DOI: 10.1101/gr.1239303

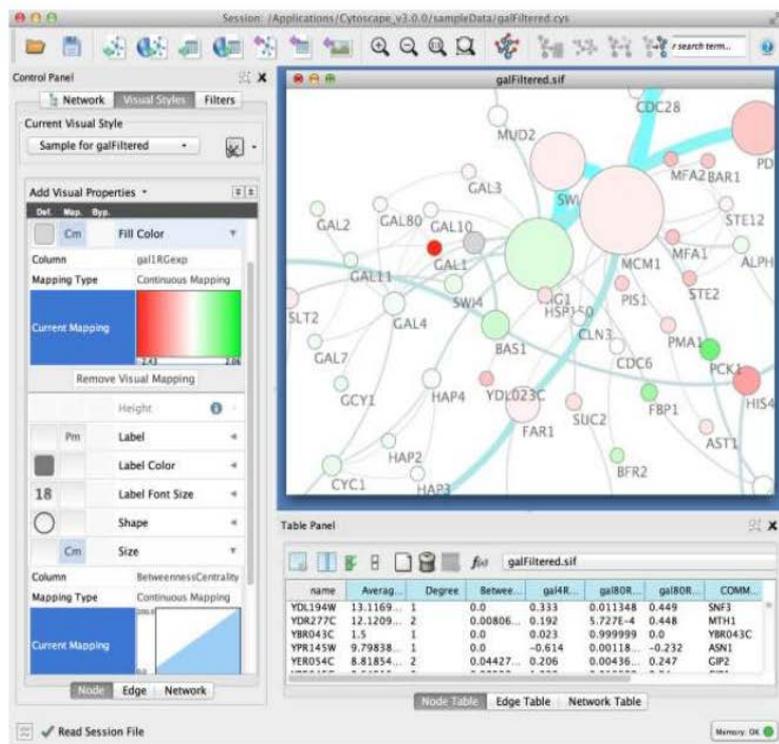
[a]



➤ Aujourd'hui...



Cytoscape



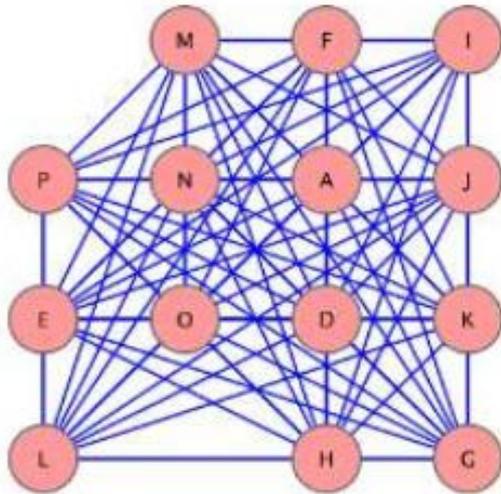
- Open source
- Cross platform
- Consortium



Cytoscape et moi, une histoire qui dure !

07/12/2020 / Work4Graph / Ludovic Cottret

➤ Principe de Cytoscape



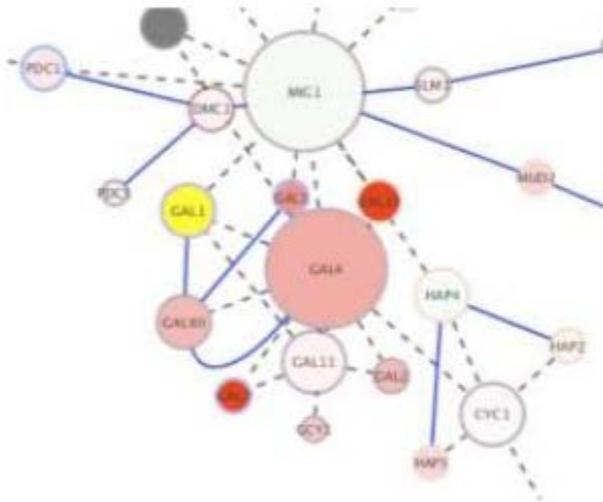
Réseau

GI_Number	Gene	HPMS_ID	RPMSQ	Entrez_ID	OMIM_ID	SR2558	Gene_Description	Architecture	Arch
g13792993	SOX1	11942	NF_012441.1	113675		SOX1	SOX1	SP	Modif
g14591020	SEC23B	11543	NF_110781.1	10483		Q15437	Protein transport protein SEC23B	QBL	Domain
g14620492	SCMH1	11540	AAH06752.1	22665		Q96D03	SCMH1	SAM	Domain
g121351025	SKDC2	11544	NF_060723.3	55770		Q96P01	Sec53 like 1	PT	Domain
g11990951	SERPINF1L2	11548	NF_030722.1	85777		Q96P01	Serpin B12	SERPIN	Domain
g163267713	SPZAA4	10447	NF_0010332	440275	606380		Similar to GON2 at23ata kinase	S_T_Y_Kinase	Domain
g14517755	TYROBP	4999	NF_001325.1	7305	804142	Q43M14	DAP12	ITAM	Domain
g15903086	IL24	4993	NF_006841.1	11009	804136	Q12007	Interleukin 24	IL10	Domain
g121265034	ADAMTS13	4994	NF_020584.1	11003	604134		ADAMTS 13	TSP1	Domain
g11743873	PPP54A	4470	AAO39911.1	8394	603279	Q96759	Phosphatidylinositol-4-phosphate 5 P4Pase	PP4C	Domain
g11450082	SVIL	4992	NF_068506.1	8640	604120	Q65425	Superoxide	yHP	Domain
g11563386	SLIT2B1	4991	NF_014444.1	8420	604123		Sulfotransferase family 2B, member 1		
g121040399	RBMS9	4990	AAH32940.1	3932	604124	Q96708	CTP	LE	Modif
g14718278	PCP	2102	NF_004480.1	2277	300091	Q43613	VGFP D	PCGF	Domain
g11321517	OPYSL4	7463	NF_000417.1	10570	606407	Q14531	Colocalin response mediator protein 3		
g18923032	TASR1	7460	NF_060184.1	95617	606270	Q96899	Troponin receptor 1		
g15454090	SSRA	2101	NF_006271.1	6748	300090	P51571	Signal sequence receptor delta	SP	Modif
g10710045	SIC16A2	2106	NF_006506.1	6567	300095	P36021	K linked FRET containing transporin TM		Domain
g11968027	PTS	7467	NF_019121.1	94400	606483	Q96870	FTS	UBC	Domain
g14820885	PKAC1	2103	NF_002732.3	5584	600529	P41743	Protein kinase C, iota type	S_T_Kinase	Domain
g19116731	RFAA	6591	NF_037479.1	29935			RFAA	TANNA	Domain
g11379812	PPP1A2P9	6593	NF_079486.1	80316			Type 1 domain phosphatase inhibitor		
g113820832		6595	NF_290379.1	90560			PH1 protein	CC	Modif
g17681844	CCDC32	6594	NF_054727.1	28922			PH1 protein	CC	Domain
g19029794	HLAF1	6596	NF_009144.1	11230			PH1 protein	TM	Domain

Tableau de metadonnées



➤ Principe de Cytoscape



Chromosome	Gene	HPRO_ID	REFSEQ	Entrez_ID	OMIM_ID	SWISS	Gene_Description	Architecture	Arch
1	SOX1	11542	NF_012441.1	113675		SOX1	SOX1	SP	Home
2	SOC2	11543	NF_118781.1	10483		Q15437	Protein transport protein SOC2	QSL	Home
3	SOC3	11544	AAH06752.1	22655		Q96003	SOC3	SP	Home
4	SOC4	11544	NF_060773.3	55770		Q96003	SOC3	SP	Home
5	SOC5	11548	NF_330722.1	83777		Q96003	SOC3	SP	Home
6	SOC6	8947	NF_00101372	44275		600380	Similar to GON2 alpha kinase	S_T_Y_Kinase	Home
7	SOC7	4996	NF_003323.1	7305	804142	043914	DAPL2	ITAM	Home
8	SOC8	4993	NF_006841.1	11009	804136	Q12007	Interleukin 24	IL10	Home
9	SOC9	4994	NF_020594.1	11093	604134		ADAMTS 13	TSP1	Home
10	SOC10	4470	AAH30911.1	8394	603273	Q96733	Protein tyrosine kinase 4-ghostlike 5	PTK	Home
11	SOC11	4992	NF_008356.1	8840	604126	Q95423	Suzorubin	VHF	Home
12	SOC12	4991	NF_014444.1	9420	604125		Sulfotransferase family 2B, member 1	ST	Home
13	SOC13	4990	AAH35560.1	3652	604124	Q96708	CYP	CYP	Home
14	SOC14	2102	NF_004460.1	2277	300091	043913	VEGF D	VEGF	Home
15	SOC15	3463	NF_006417.1	10570	606407	Q54531	Colipase response mediator protein 3	SR	Home
16	SOC16	3460	NF_060184.1	55617	606270	Q96870	Tetraspanin 9	SP	Home
17	SOC17	2101	NF_006271.1	6748	300090	P51571	Signal sequence receptor delta	SP	Home
18	SOC18	2100	NF_006508.1	6567	300089	P36021	X-linked REST-containing transmembrane protein	TM	Home
19	SOC19	7467	NF_071921.1	94400	606403	Q96870	FTS	VEG	Home
20	SOC20	3105	NF_002773.3	2569	600339	P61743	Protein kinase C, zeta type	S_T_Kinase	Home
21	SOC21	6581	NF_007479.1	29035			PKA	TANAA	Home
22	SOC22	6593	NF_079496.1	80316			Type 1 protein phosphatase inhibitor 1	SP	Home
23	SOC23	6592	NF_206175.1	80660			PH1 protein	CC	Home
24	SOC24	6294	NF_004727.1	28952			PH1 protein	CC	Home
25	SOC25	6594	NF_009144.1	11230			PH1 protein	TM	Home

Réseau

Tableau de metadonnées

Style visuel



Cytoscape et moi, une histoire qui dure !
07/12/2020 / Work4Graph / Ludovic Cottret

➤ Style visuel

The screenshot displays a network diagram and its corresponding style settings in a software application. The diagram shows nodes representing genes (E, P, I, B), proteins (Y, Z, i), and metabolites (glc, lac, AMPc) connected by edges. The style settings are organized into two main sections: 'Network' and 'Dynamic Network'.

Network Style Settings:

- Style: default
- Border Paint: [Color swatch]
- Border Width: 4.0
- Fill Color: [Color swatch]

Dynamic Network Style Settings:

- Height: [Value]
- Label: [Value]
- Label Color: [Color swatch]
- Label Font Size: 12
- Shape: [Shape swatch]

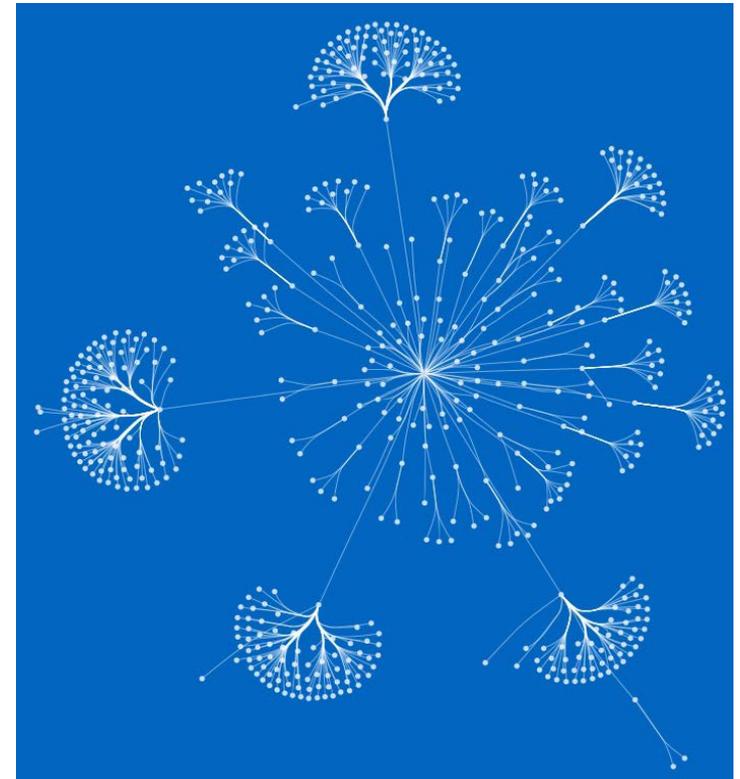
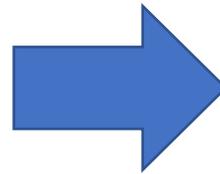
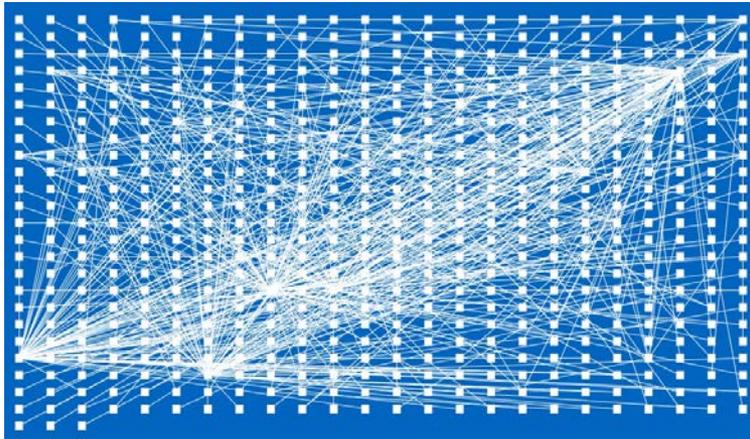
Node Table:

shared name	name	Type
i	i	gene
B	B	protein
P	P	protein
Y	Y	gene
Z	Z	gene
AMPc	AMPc	metabolite
I	I	protein
lac	lac	metabolite
E	E	protein
glc	glc	metabolite

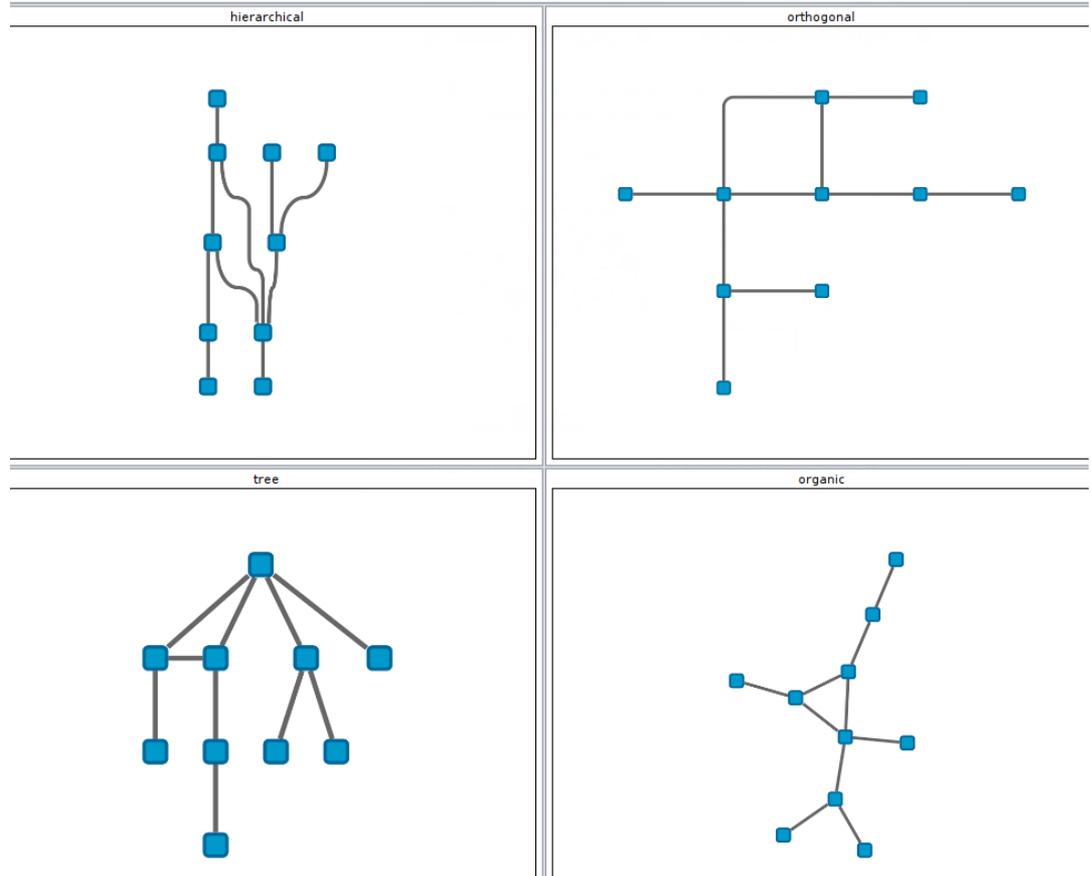
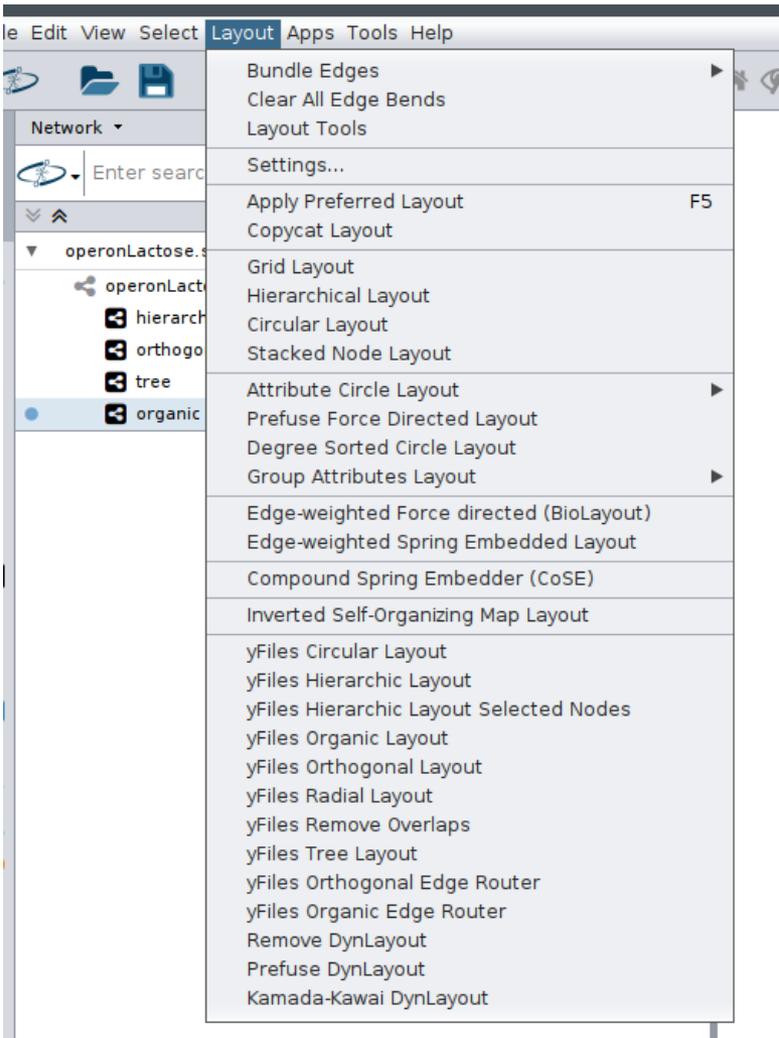
Style visuel en fonction des colonnes de la table des noeuds ou des arêtes

Le style visuel peut également être réglé pour un noeud ou une arête en particulier (bypass)

➤ Organisation du dessin : le layout



➤ Organisation du dessin : le layout



➤ Format des réseaux

- **Le plus simple : .sif : liste d'arêtes**
- **Spécifique des graphes : xggml**
- **Spécifique des réseaux biologiques : BioPax et SBML (styles visuels adaptés)**



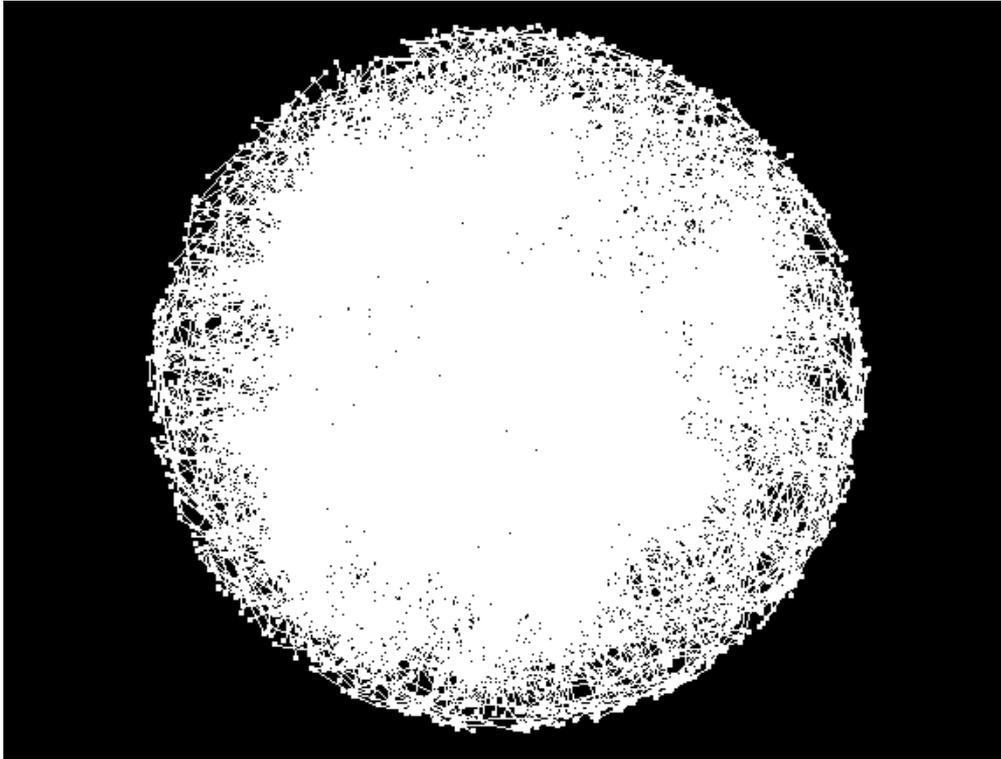
INRAE



Cytoscape et moi, une histoire qui dure !

07/12/2020 / Work4Graph / Ludovic Cottret

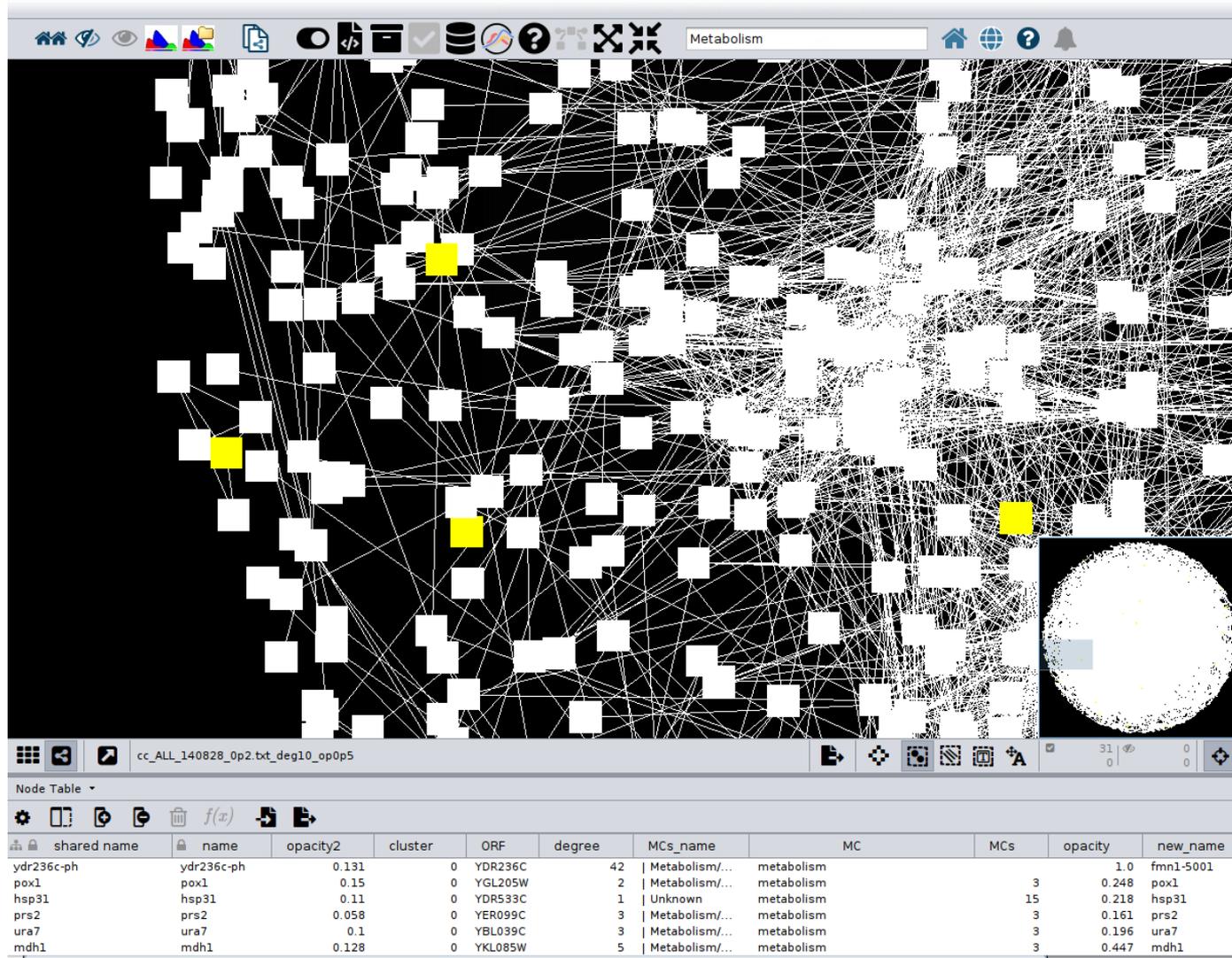
➤ Extraction de sous-réseaux



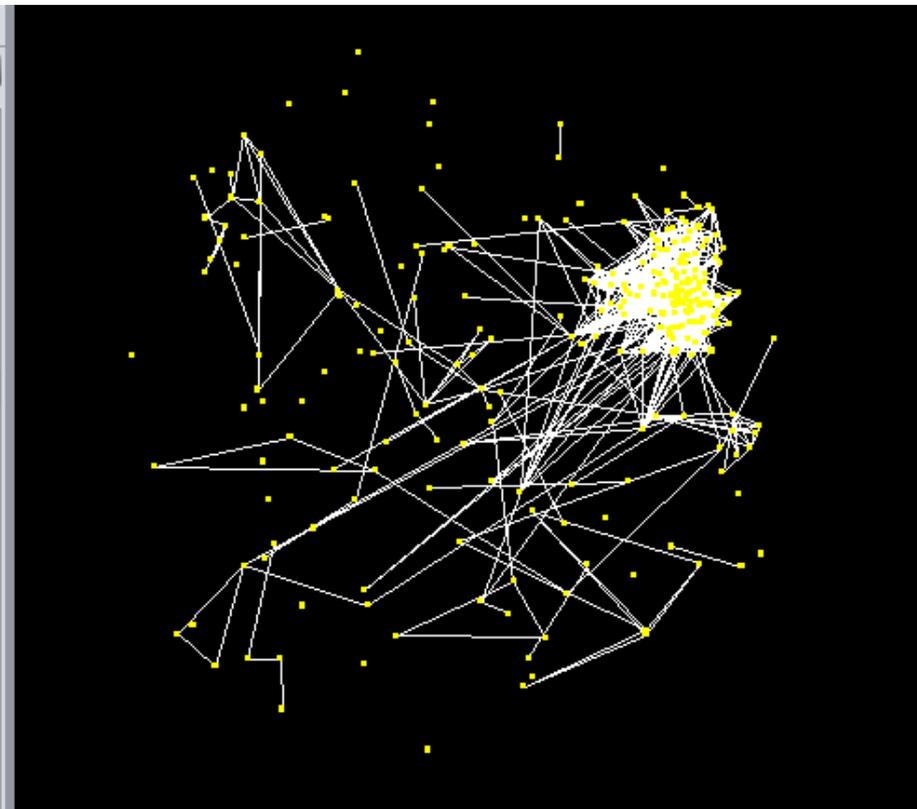
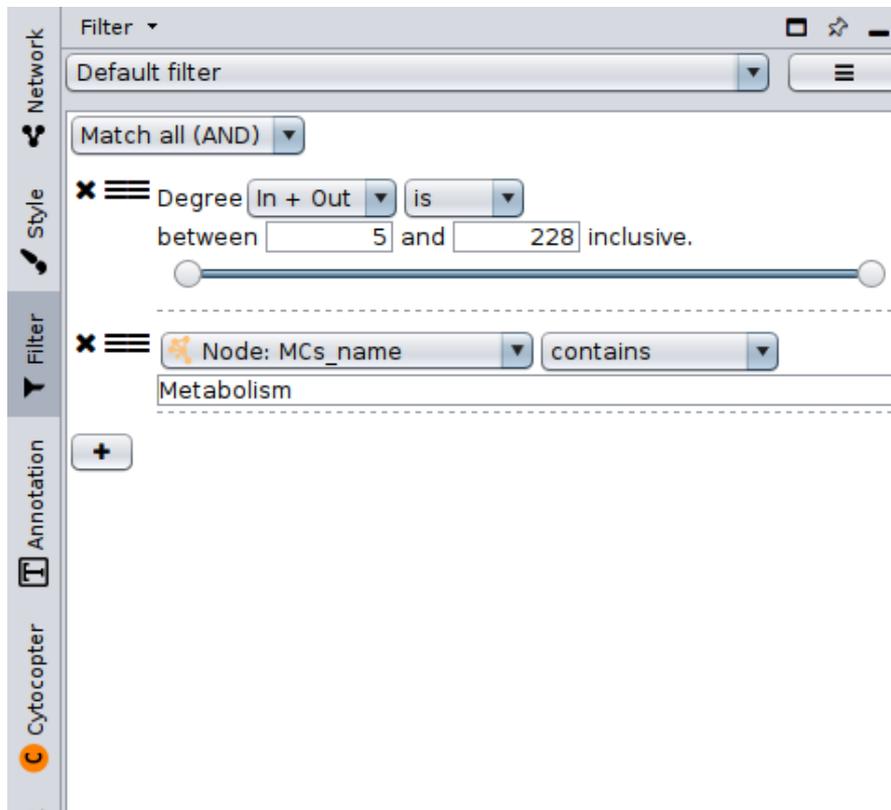
**Très difficile d'explorer
un tel réseau !**

Réseau d'interactions de gènes de la
levure (3971 noeuds et 28202 arêtes)

➤ Recherche de noeuds ou d'arêtes dans le réseau



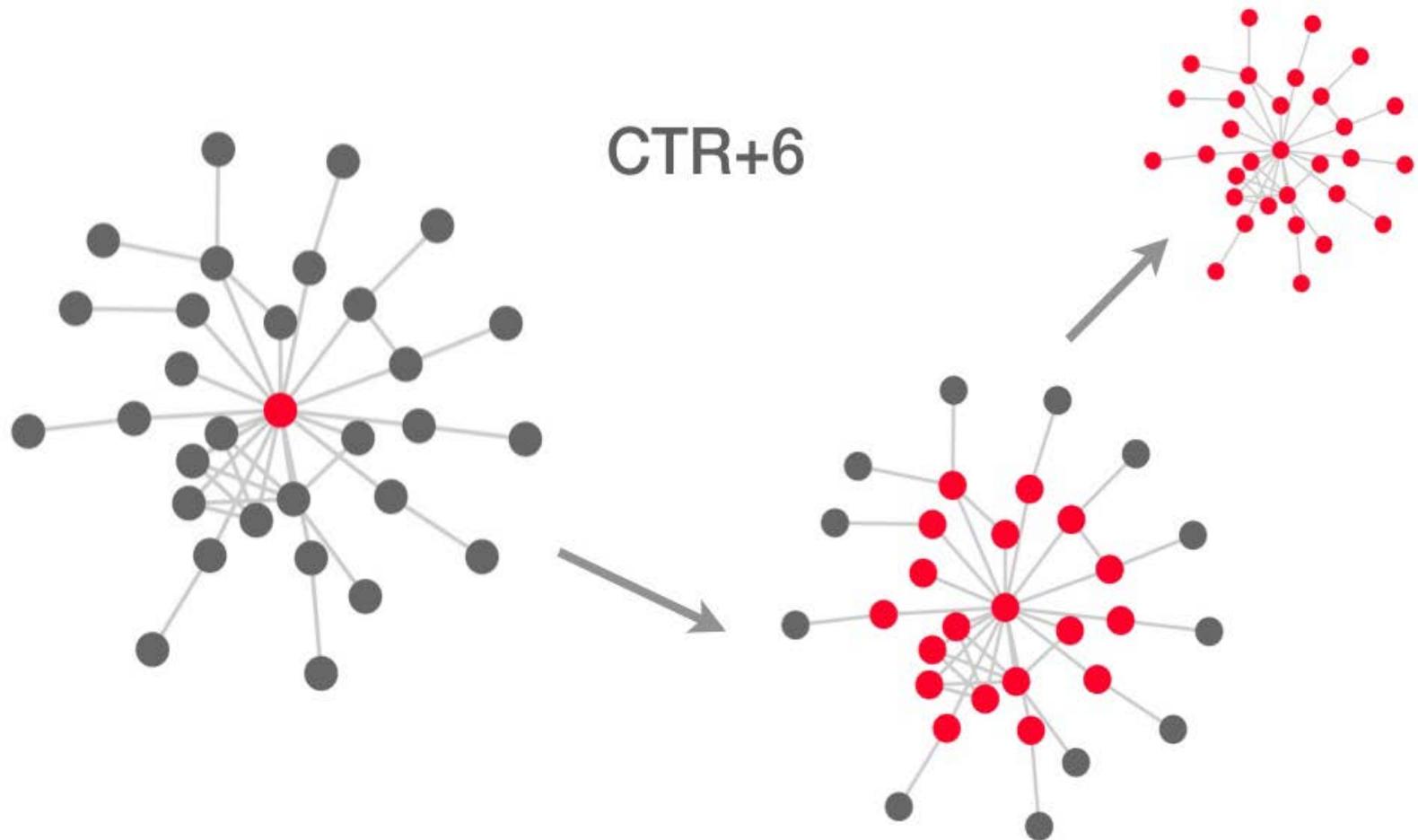
➤ Filtres



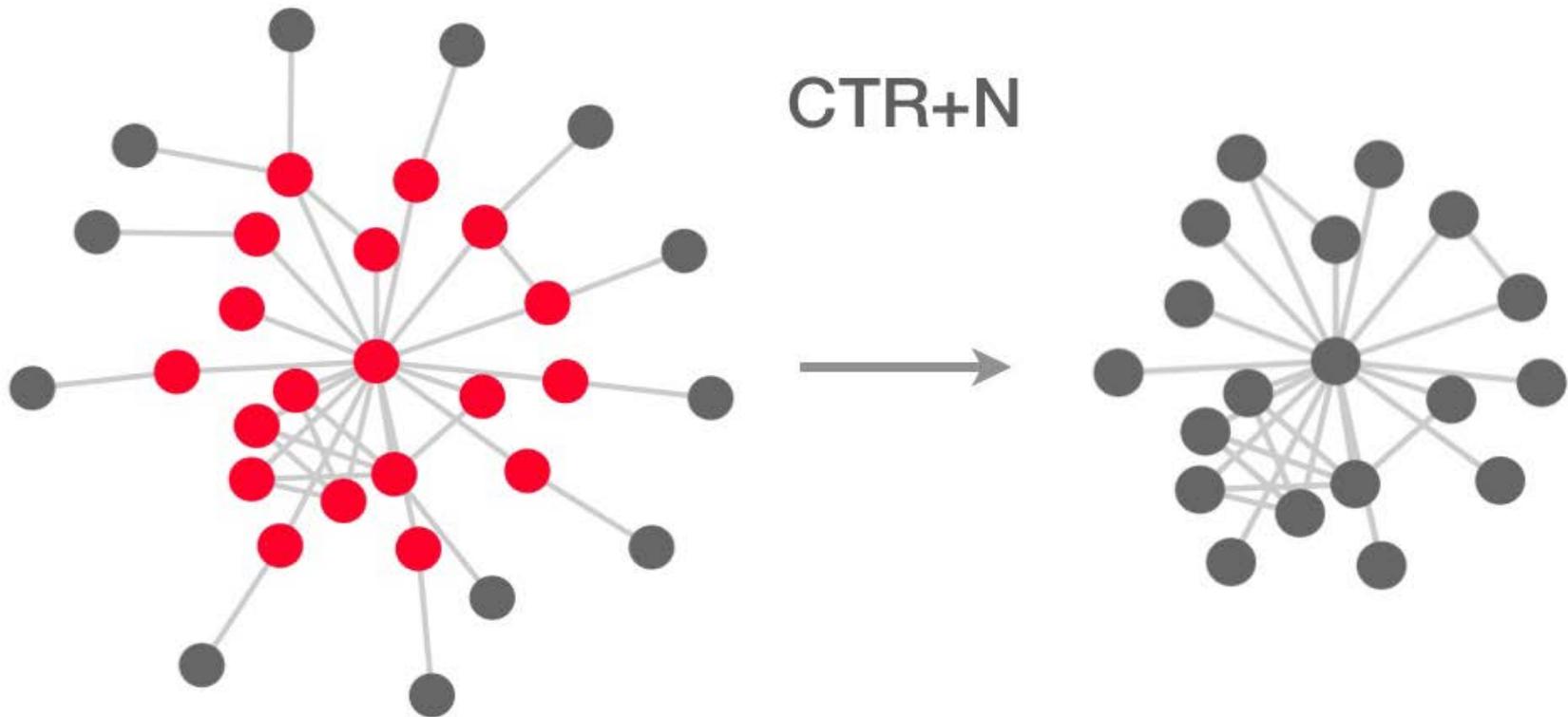
Exemple : recherche des noeuds avec plus de 5 connections et impliqués dans le métabolisme



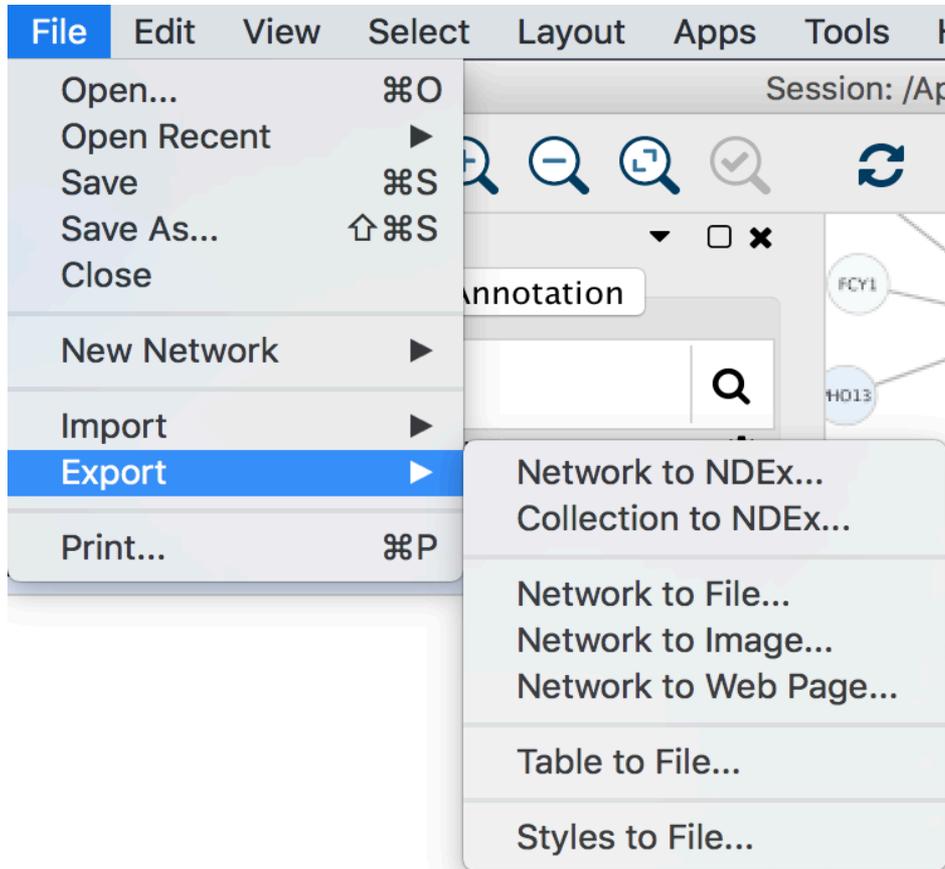
➤ Sélection par voisinage



➤ Sélection -> nouveau réseau



➤ Export du réseau



- **Session**
- **Image**
- **Styles**
- **NDEx : collection of shared networks**
- **Page web !**



➤ Une grande communauté et de nombreux plugins

Wall of Apps 366 total

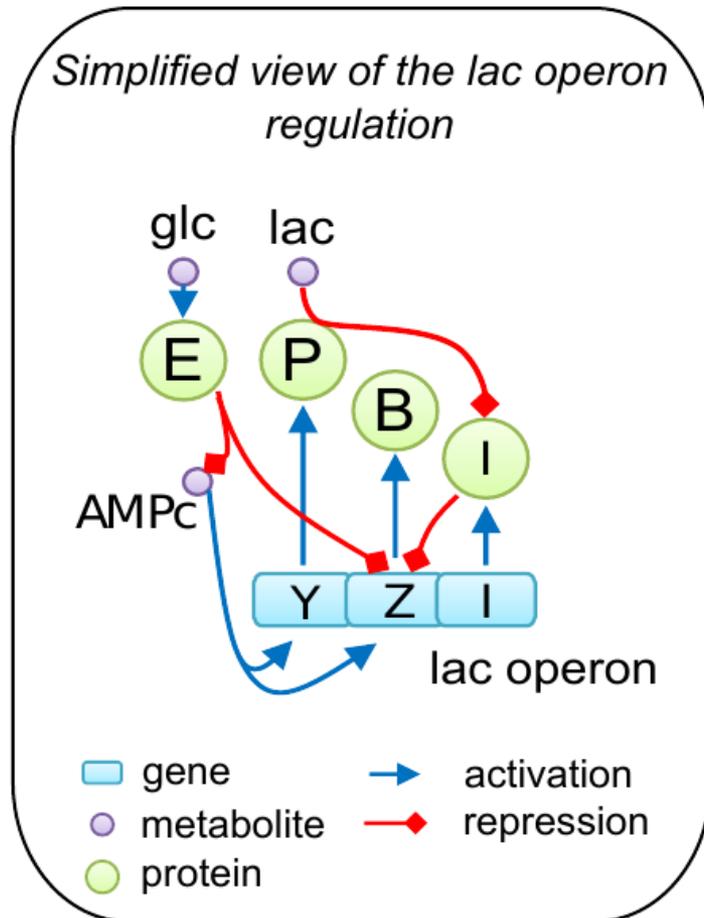
collections

The 'Wall of Apps' displays 366 total applications, categorized into three main sections: 'collections', 'data visualization', and 'network generation'. Each section contains a grid of application icons. The 'collections' section shows various icons for data management and analysis. The 'data visualization' section features a wide variety of icons, including network visualizers, charts, and data analysis tools. The 'network generation' section includes icons for network construction, simulation, and analysis. Some icons are highlighted in color, while others are faded.

data visualization

network generation

> Démo

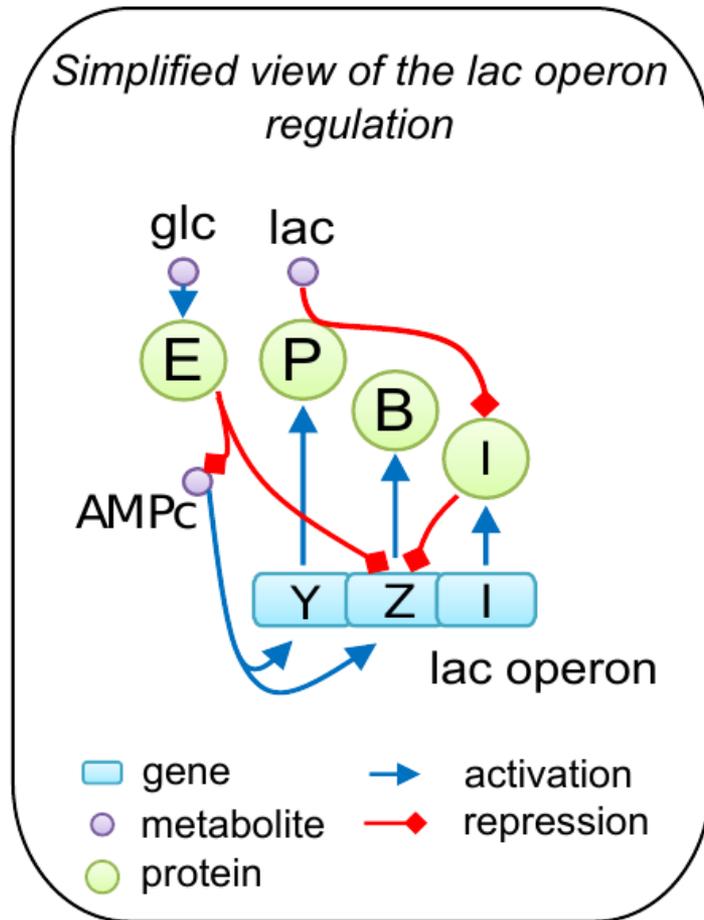


Source : R. Peyraud

Objectif: reproduire ce dessin dans Cytoscape



➤ Démo (1) : Création du fichier sif

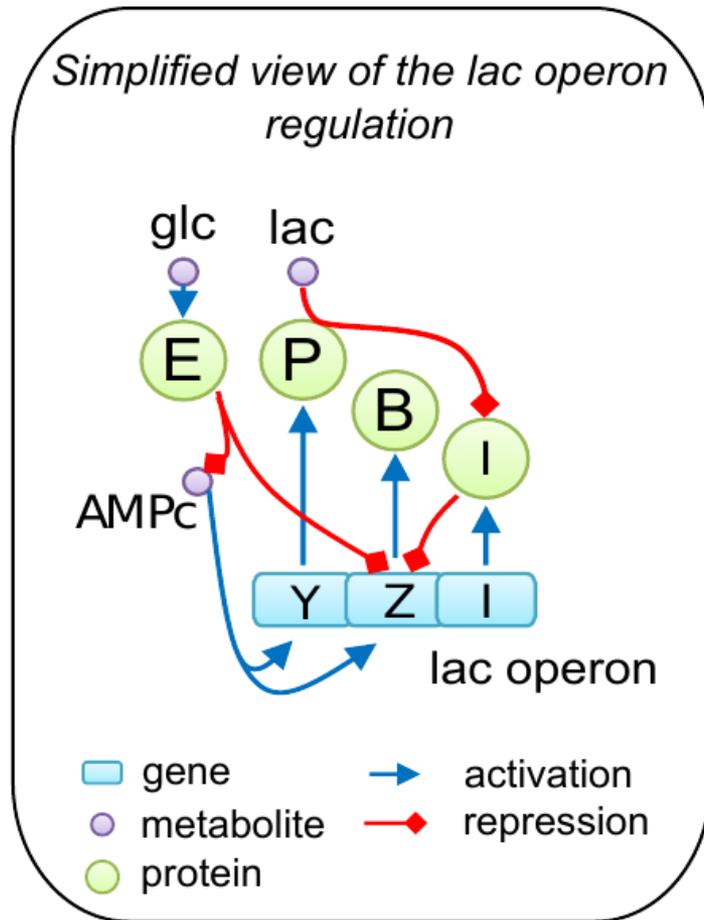


```
operonLactose.sif
glc activates E
lac inhibits I
E inhibits AMPc
E inhibits Z
I inhibits Z
AMPc activates Y
AMPc activates Z
Y activates P
Z activates B
i activates I
```

Source : R. Peyraud



➤ Démo (2) : Création du fichier des attributs des noeuds

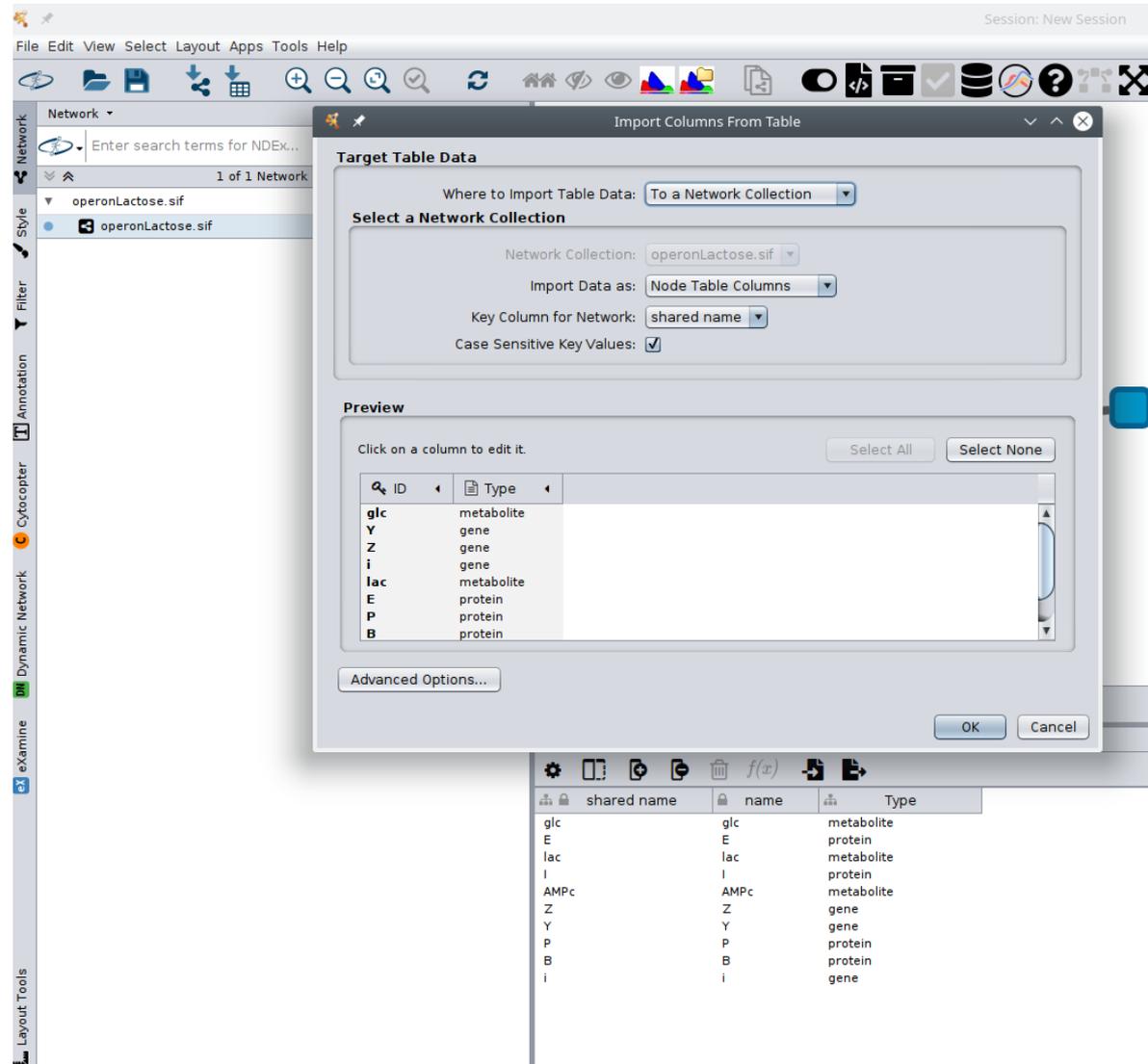
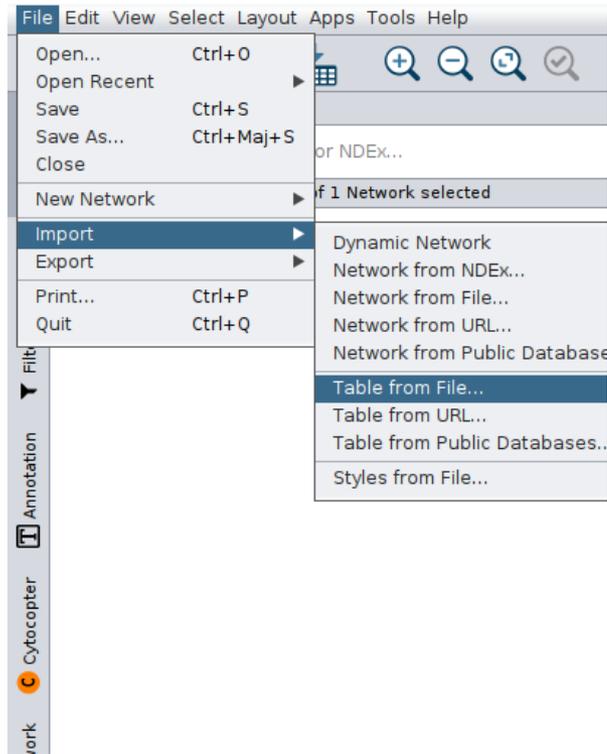


```
nodes.csv
ID,Type
glc,metabolite
Y,gene
Z,gene
i,gene
lac,metabolite
E,protein
P,protein
B,protein
I,protein
AMPC,metabolite
```

Source : R. Peyraud



➤ Démo (2) : Création du fichier des attributs des noeuds



➤ Demo (3) Création du style visuel des noeuds en fonction des attributs chargés

The screenshot displays the Cytoscape interface with a network diagram and its style configuration. The network diagram shows nodes labeled P, Y, AMPc, Z, E, glc, I, B, and i, connected by edges. The nodes are styled based on their type: proteins (P, Y, AMPc, I, B, i) are blue circles, genes (Z, E) are green circles, and metabolites (glc) are light blue circles. The style configuration panel on the left shows the following settings:

- Style:** default
- Properties:** Def. Map. Byp.
- Style:** Border Paint, Border Width (4.0), Fill Color
- Filter:** Column, Mapping Type (Discrete Mapping)
- Annotation:** gene (R:153 G:255 B:255 - #99F...), metabolite (R:204 G:204 B:255 - #CCC...), protein (R:51 G:204 B:0 - #33CC00)
- Cytopainter:** Height, Label
- Dynamic Network:** Column (name), Mapping Type (Passthrough Mapping)
- Examine:** Label Color, Label Font Size (12), Label Position, Shape (Rectangle)
- Layout Tools:** Size (15.0), Transparency (255)

The Node Table at the bottom shows the following data:

shared name	name	Type
glc	glc	metabolite
E	E	protein
lac	lac	metabolite
I	I	protein
AMPc	AMPc	metabolite
Z	Z	gene
Y	Y	gene
P	P	protein
B	B	protein
i	i	gene

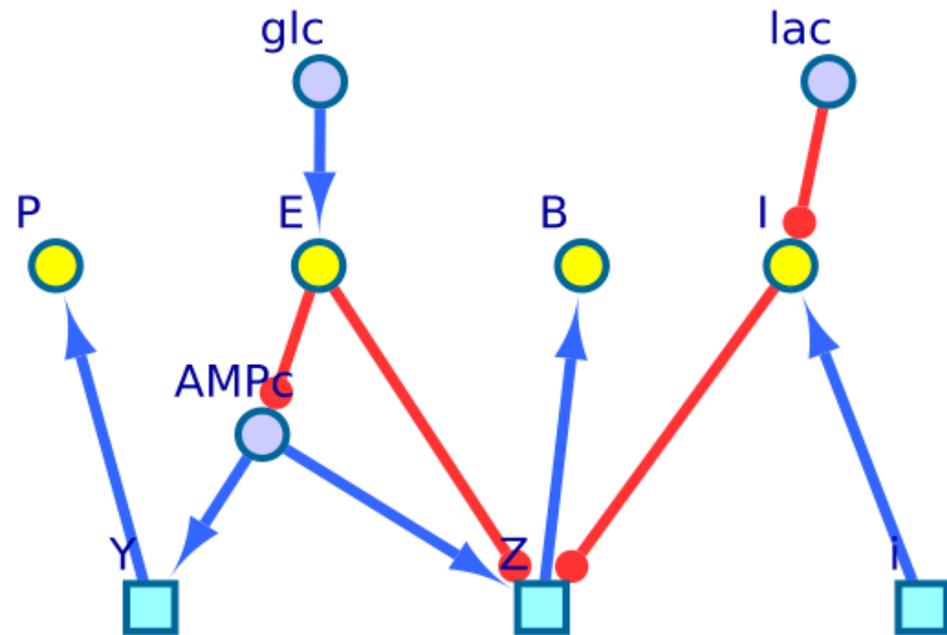
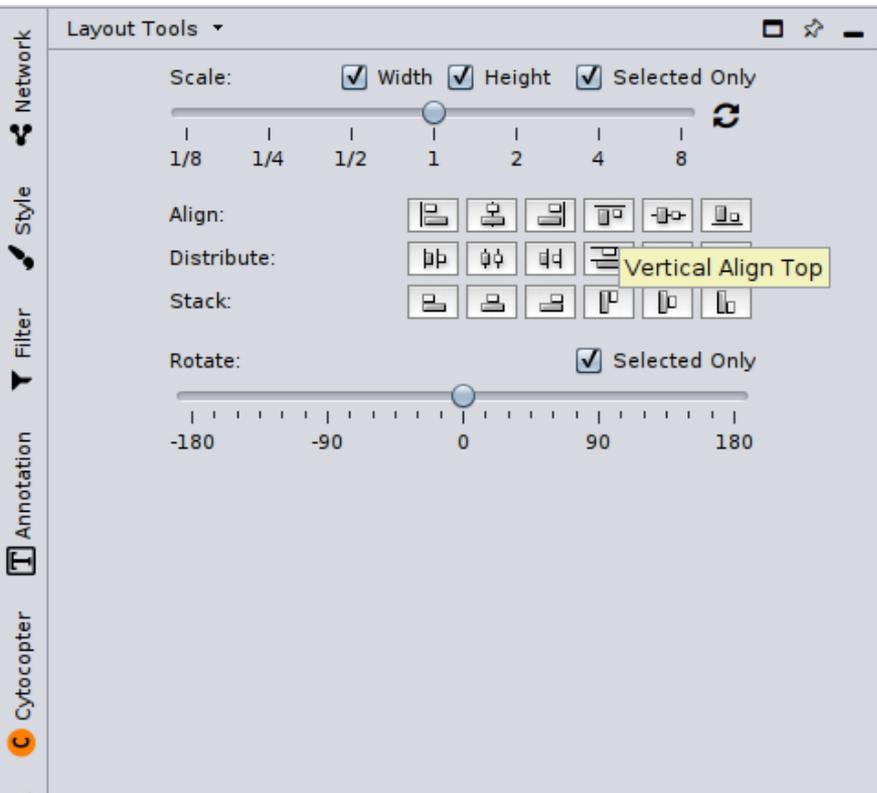
➤ Demo (4) Création du style visuel des arêtes en fonction du type d'interactions

The screenshot displays the Cytoscape interface with the 'Style' panel on the left and a network diagram on the right. The 'Style' panel is configured for 'interaction' mapping with 'Discrete Mapping'. The 'Stroke Color (Unselected)' is set to a blue color (R:51 G:102 B:255 - #3366FF) for 'activates' and a red color (R:255 G:51 B:51 - #FF3333) for 'inhibits'. The 'Target Arrow Shape' is set to 'None'. The network diagram shows nodes (circles and squares) and edges (arrows) colored according to these settings. The nodes are labeled P, Y, AMPc, B, Z, E, I, lac, and i. The edges are colored blue for activation and red for inhibition. The diagram is titled 'operonLactose.sif'.

Property	Value
Column	interaction
Mapping Type	Discrete Mapping
activates	Blue (R:51 G:102 B:255 - #3366FF)
inhibits	Red (R:255 G:51 B:51 - #FF3333)



➤ Demo (5) Layout manuel



➤ Demo (6) : arêtes incurvées

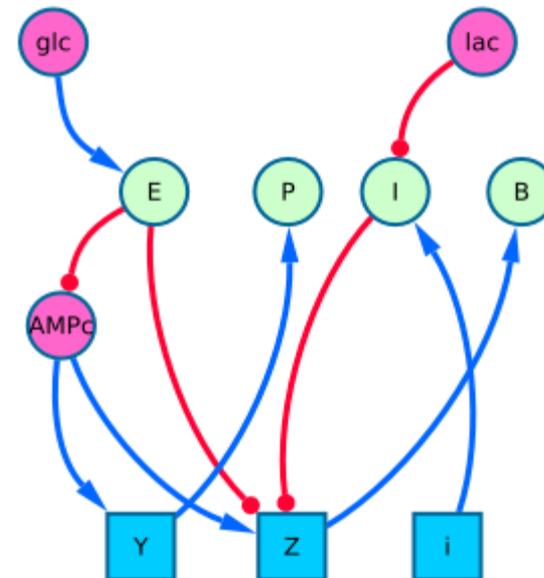
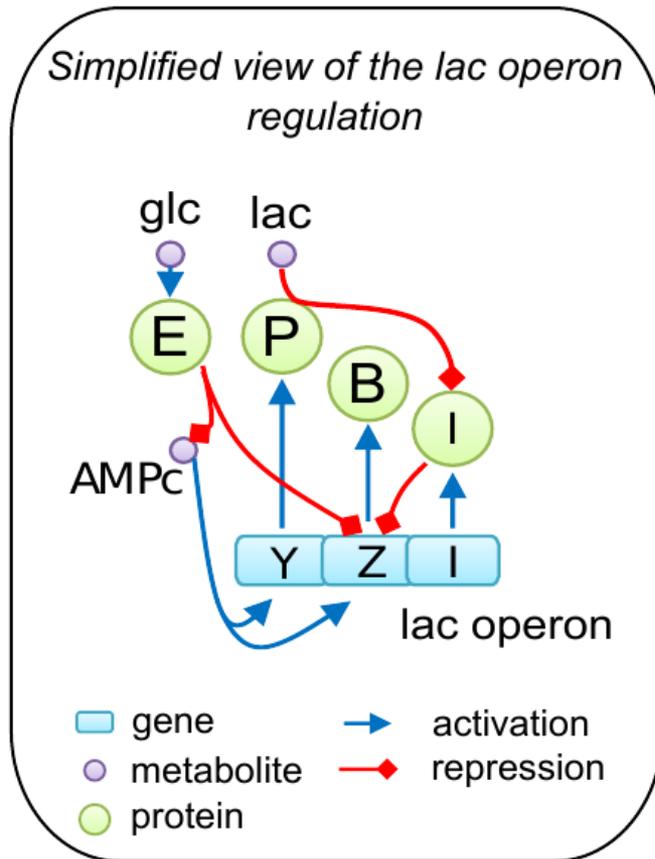
The screenshot shows the Cytoscape interface with the Edge Bend Editor dialog open. The dialog contains the following instructions:

1. **Ctrl-Alt-click** the edge to add a new handle.
2. Drag handles to bend (select the edge first).

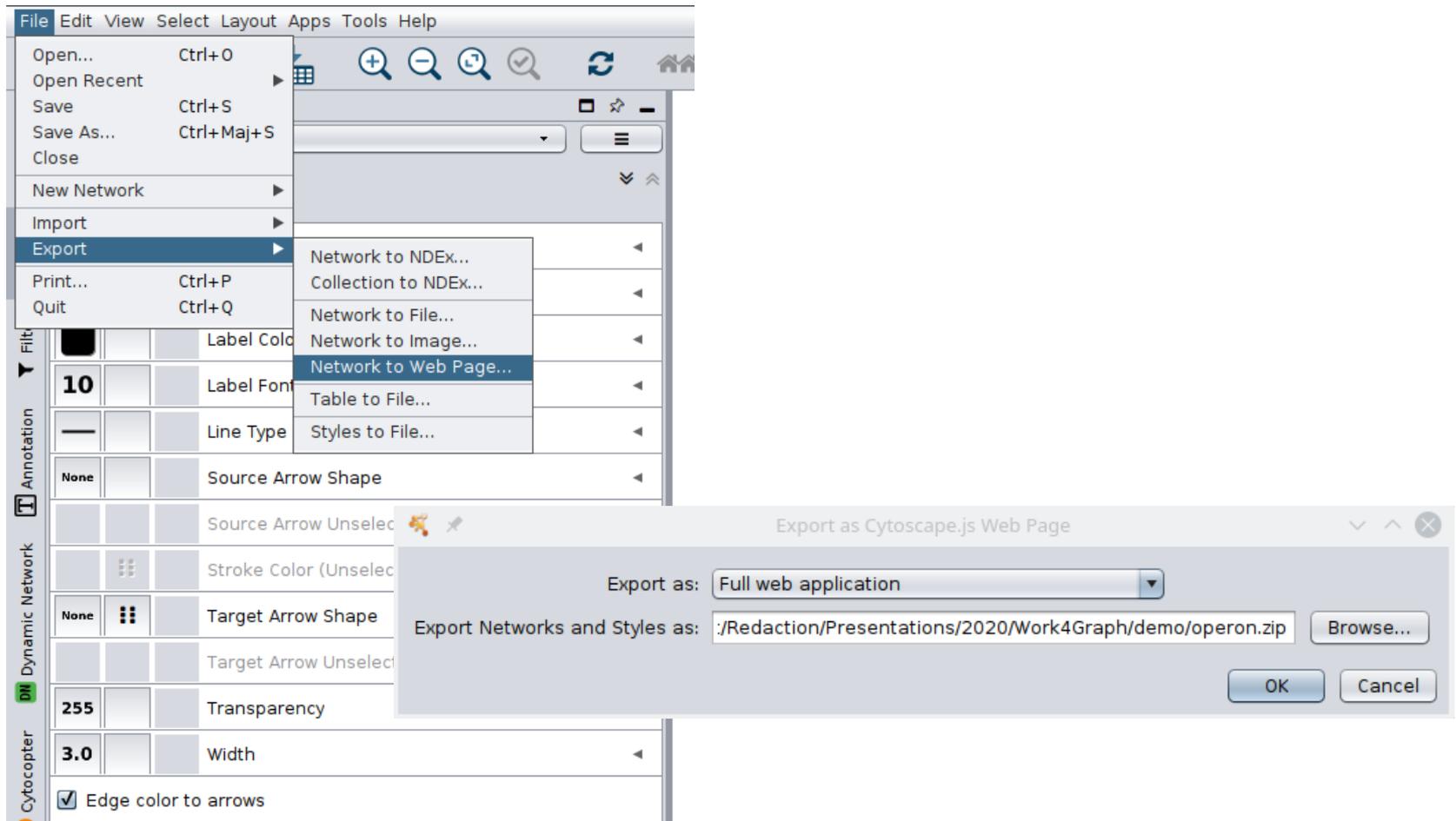
The network diagram shows nodes labeled P, E, B, I, glc, lac, and i. Edges are colored blue and red, with some being curved. A table at the bottom right lists interactions:

shared interaction	name	interaction
activates	glc (activate...	activates
inhibits	lac (inhibits) I	inhibits
inhibits	E (inhibits) ...	inhibits
inhibits	E (inhibits) Z	inhibits
inhibits	I (inhibits) Z	inhibits
activates	AMPC (activ...	activates

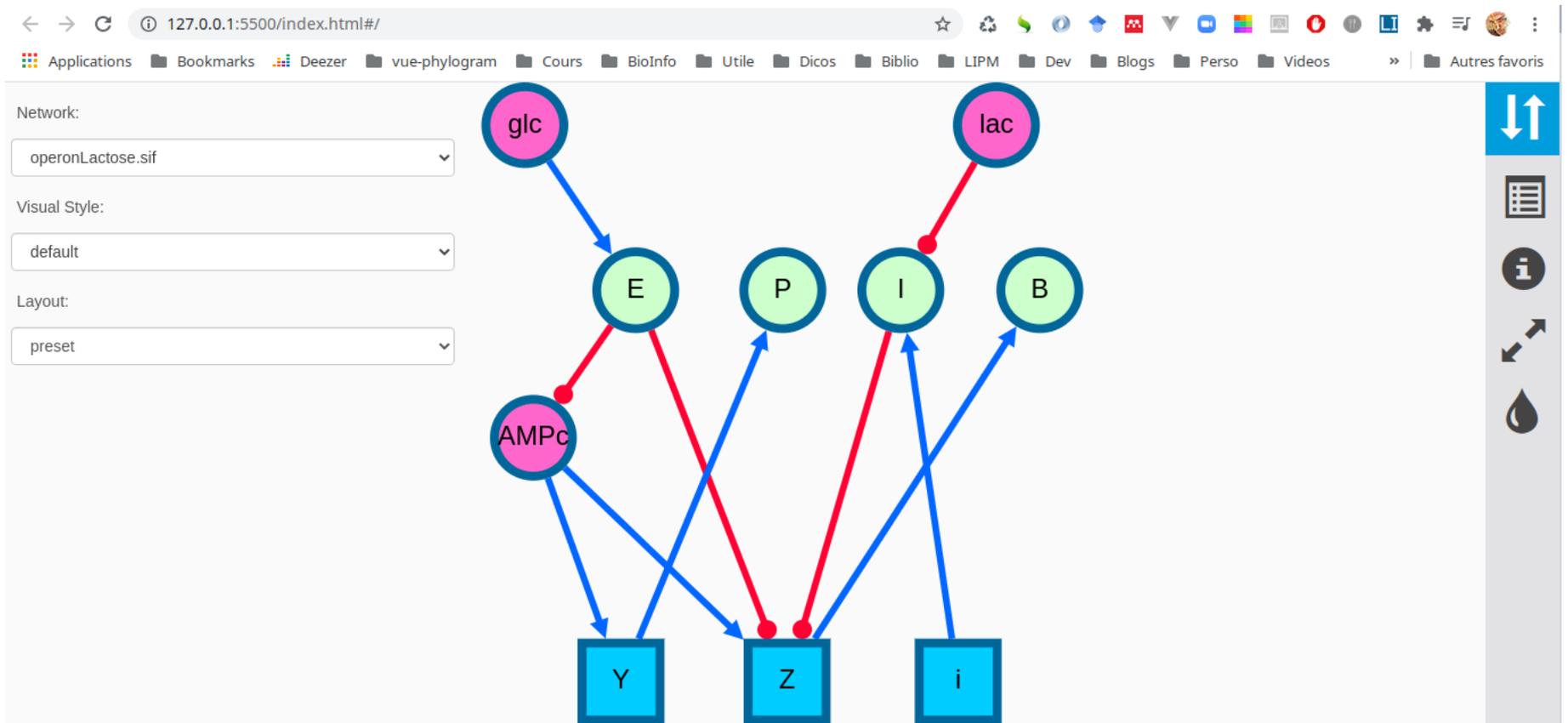
➤ Demo (7) Résultat final



➤ Demo (8) Construction d'un site web en un clic



➤ Demo (8) Construction d'un site web en un clic



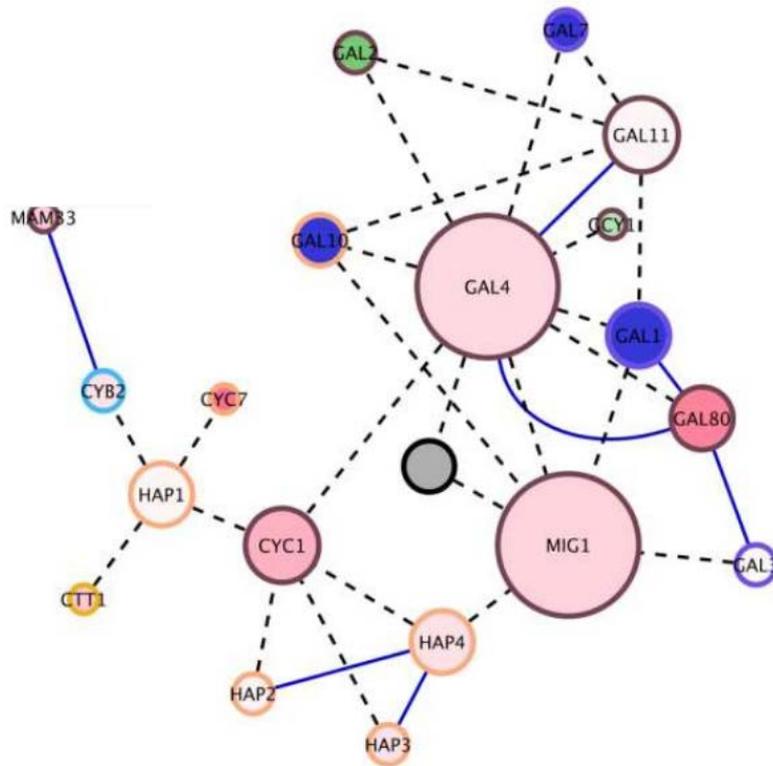
INRAE



Cytoscape et moi, une histoire qui dure !

07/12/2020 / Work4Graph / Ludovic Cottret

➤ Attention à la surcharge d'informations visuelles !



- Un message = un style visuel particulier
- Attention, l'oeil humain peut seulement discerner un nombre limité de couleurs
- Attention aux choix esthétiques malheureux !

➤ Mon expérience avec Cytoscape

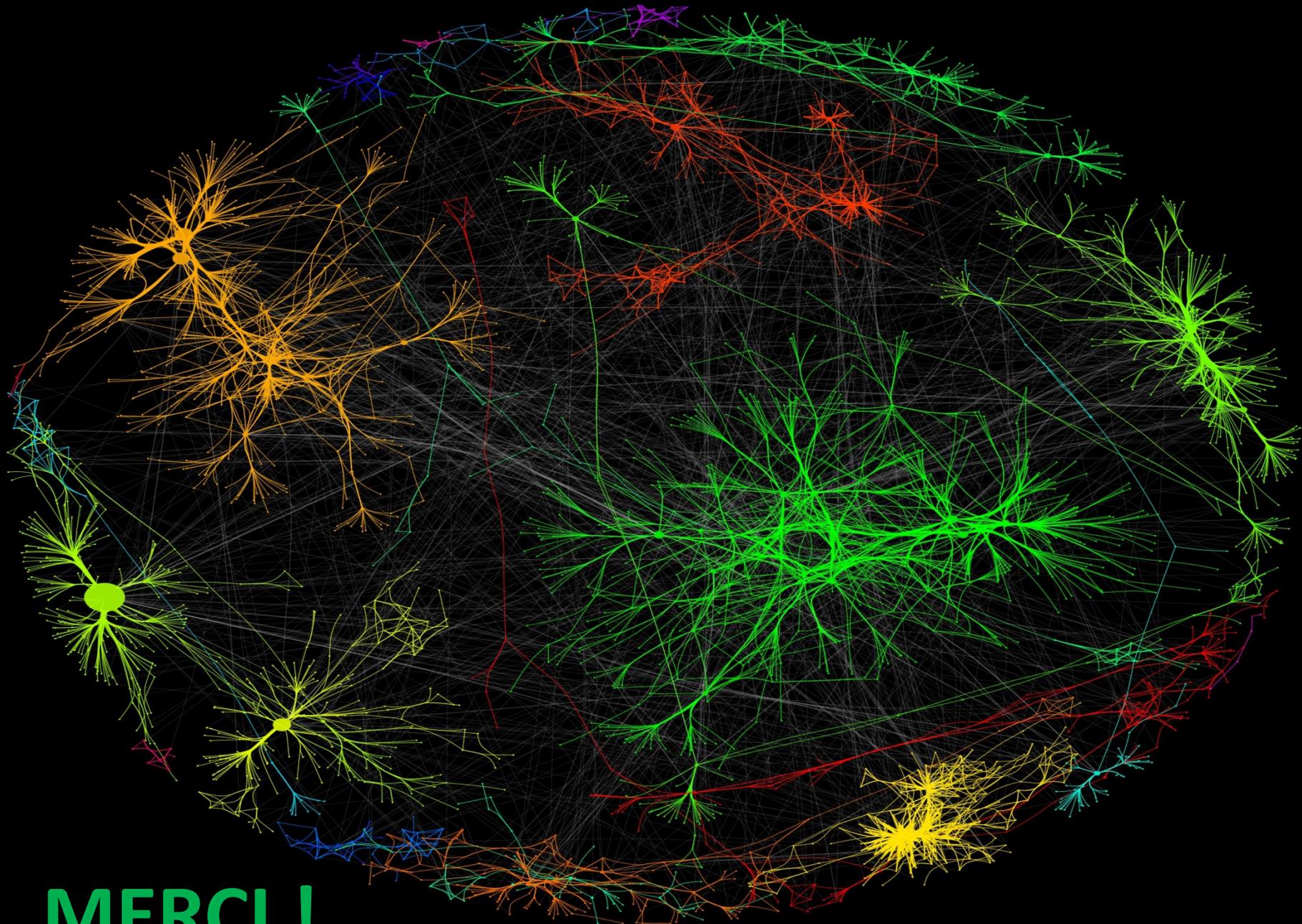
- Depuis 2005, visualisation de réseaux métaboliques
- Première visualisation dans MetExplore par jnlp
- Génération de sites web pour des analyses spécifiques.
Exemple : réseau de domaines partagés par des protéines :
<https://iant.toulouse.inra.fr/EFFECTOORES/webapp/data/clustering>
- Enseignements
- Formations CATI et LIPM pour les biologistes



➤ Pourquoi je continue à utiliser Cytoscape ?

- Pas de bugs majeurs
- Continue à être maintenu depuis maintenant 18 ans (glups...)
- Toujours de nouvelles fonctionnalités
- Simple à utiliser pour les biologistes
- Documentation exhaustive, tutoriaux
- Pas de réelle concurrence gratuite à ma connaissance pour la visualisation (pour les calculs sur les graphes, iGraph est très pratique)
- La seule concurrence vient des librairies Javascript Cytoscape.js et d3.js mais qui ne répondent pas aux mêmes besoins





MERCI !