
Projet ABC
Analyse et Base de Connaissances

Problème d'interprétation des données

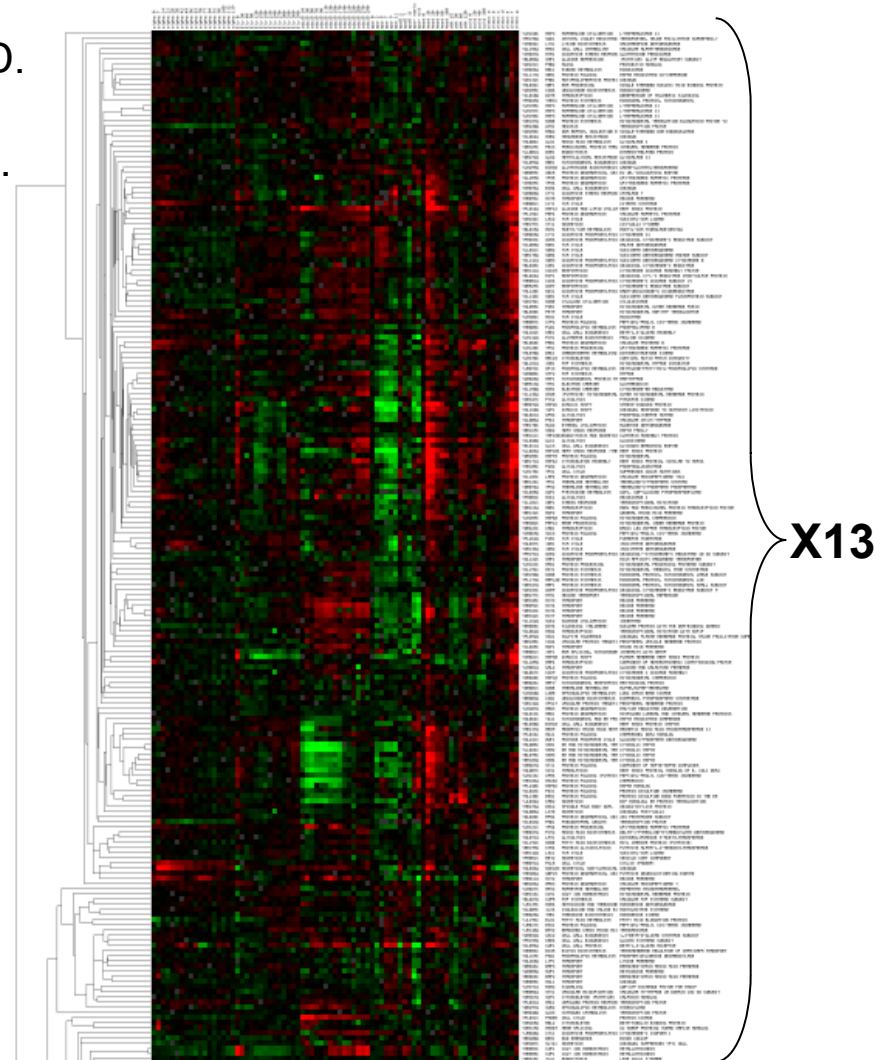
Données extraites de :

Eisen MB, Spellman PT, Brown PO, Botstein D.
Cluster analysis and display of genome-wide
expression patterns. Proc Natl Acad Sci U S A.
1998 Dec 8;95(25):14863-8.

6221 gènes de *Saccharomyces cerevisiae*

80 expériences

2467 gènes ont une fonction répertoriée



Modélisation hiérarchique des concepts dans GO

Top Documentation [Gene Ontology](#) [GO Links](#) Terms Gene Products

- [-] [GO:0003673 : Gene Ontology \(30674\)](#)
 - [+] [GO:0008150 : biological process \(23487\)](#)
 - [+] [GO:0008151 : cell growth and/or maintenance \(15526\)](#)
 - [+] [GO:0008152 : metabolism \(10826\)](#)
 - [+] [GO:0009058 : biosynthesis \(2327\)](#)
 - [+] [GO:0009059 : macromolecule biosynthesis \(1684\)](#)
 - [+] [GO:0006412 : protein biosynthesis \(1572\)](#)

- [+] [GO:0006411 : protein metabolism and modification \(4159\)](#)
- [+] [GO:0006412 : protein biosynthesis \(1572\)](#)
- [+] [GO:0006461 : protein complex assembly \(111\)](#)
- [+] [GO:0007582 : physiological processes \(695\)](#)
- [+] [GO:0007585 : respiration \(26\)](#)
- [+] [GO:0005575 : cellular component \(14402\)](#)
- [+] [GO:0005623 : cell \(11523\)](#)
 - [+] [GO:0005622 : intracellular \(10239\)](#)
 - [+] [GO:0005737 : cytoplasm \(5373\)](#)
 - [+] [GO:0005739 : mitochondrion \(1433\)](#)
- [+] [GO:0003674 : molecular function \(21459\)](#)
- [+] [GO:0005215 : transporter \(1920\)](#)
 - [+] [GO:0005372 : water transporter \(5\)](#)

Get this GO tree as [RDF XML](#).

1

2

3

Objectifs du projet

1. utiliser les termes biologiques pour qualifier les résultats de puces à ADN
 2. utiliser les résultats des puces pour étendre une base de connaissances
 3. exploiter les données recueillies et faire du data-mining
-

ABC v1 : version consultation

Basé sur DAG-EDIT (éditeur d'ontologie au format GO)

+ modules de consultation

associations termes $\leftarrow \rightarrow$ produits de gènes

Localisation chromosomique

Classifications hiérarchiques

+ modules de data-mining

Choix d'annotations caractérisant des groupements de transcrits

Détection des gènes covariants et co-localisés

- fonctions d'édition

Une des constatations de Eisen et al.

« les gènes responsables de la synthèse des protéines de la mitochondrie sont exprimés de manière concordante avec de nombreux gènes impliqués dans la respiration »

Visualisation de transcrits associés à un terme

The screenshot displays the ABC version 1.311 software interface, which is used for visualizing gene products associated with specific terms. The interface is divided into several panels:

- Terms Viewer:** A list of biological terms, with 'aerobic respiration' and 'anaerobic respiration' highlighted.
- Search Panel:** A search interface where the term '*respiration*' was entered, resulting in 11 terms found.
- DAG Viewer:** A Directed Acyclic Graph (DAG) showing the hierarchical structure of the terms, starting from 'root' and branching into 'Gene_Ontology', 'biological_process', 'cellular_process', 'cell_growth', 'metabolism', 'energy_derivation', and 'energy_reserve_metabolism'.
- Classification Tree Viewer:** A tree view showing the classification of gene products, with '6221 leaves' and 'depth=0' indicated.
- Gene Displayer:** A panel showing the search results for 'aerobic respiration', listing 69 associations found for 59 gene products. The list includes gene IDs and names such as SGD:S0002225 cytochrome c oxidase, SGD:S0002333 D-lactate ferredoxin, SGD:S0002486 cytochrome c oxidase, SGD:S0002522 YDR115W (2), SGD:S0002523 MRPL1 [YDR115W], and SGD:S0002582 mitochondrial ribosomal protein.



Annotation automatique des partition

ABC version 1.311

Terms Viewer

- response to stress
 - DNA protection
 - SOS response
- response to DNA damage
- response to ER-overload
- response to hypoxia
- response to osmotic stress
- response to oxidative stress
- response to pest/pathogen
- response to starvation
 - fruiting body formation ()
 - fruiting body formation ()
 - response to glucose star

Classification Caption

Color	Query
Red	TERM=response to starvation
Blue	TERM=response to biotic stimulus
Green	TERM=response to oxidative stress
Purple	TERM=response to stress

Search Panel

find TERMS with

ASP3-2/YLR157C - Netscape

Standard Name	ASP3-2 (s
Alias	ASP3
Systematic Name	YLR157C
Feature Type	ORF
GO Annotations	ASP3-2 GO evidence and references

SGD Curated Paper - Netscape

Bon EP, Carvajal E, Stanbrough M, Rowen D, Magasa
Asparaginase II of *Saccharomyces cerevisiae*. GLN3/URF
a periplasmic enzyme. *Appl Biochem Biotechnol* 63-65():

Abstract: The production of some extracellular enzymes is known to be affected by readily metabolized nitrogen sources such as NH₄⁺ and a consensus regarding the involved mechanisms. Asparaginase II is a protein of *Saccharomyces cerevisiae* encoded by the ASP3 gene. The enzyme is secreted by cells grown in either ammonia, glutamine, or glutamate, but it is not secreted by cells subjected to nitrogen starvation or have been grown on a poor nitrogen source.

Gene Displayer

search collapse expand

response to starvation
41 association(s) found
for 38 gene products

inherited terms' gene products

- SGD:S0002264 homeobox tran
- SGD:S0002702 dihydrosphingo
- SGD:S0002885 serine/threonin
- SGD:S0004145 nitrogen catabo
- SGD:S0004147 nitrogen catabo
- TAS SGD_REF:8664

Hierarchical Tree Diagram

- response to stress
 - response to biotic stimulus (+)
 - response to stress (+)
 - response to oxidative stress (+)
 - YPL154C
 - YIR037W
 - YEL060C
 - YIL033C
 - YLR287C-A
 - YGR088W
 - YMR169C
 - YBR006W
 - YPRO26W
 - YDR171W
 - YCL035C
 - YJ513W
 - YDR074W
 - YGR209C
 - YIL101C
 - YBR072W
 - YNL160W
 - YOR178C
 - YBR126C
 - YCR021C
 - YDR258C
 - YLL026W
 - YCR069W
 - YFLO31W
 - YLR157C
 - YLR155C
 - YLR158C
 - YLR160C
 - YLR043C
 - YDR160W
 - response to stress
 - protein folding
 - response to nitrogen starvation (+)

Utilisation de l'ontologie anatomique de FB

The screenshot displays the ABC version 1.311 software interface, which is used for analyzing gene products based on ontological terms. The interface is divided into several panels:

- Terms Viewer:** A hierarchical tree of ontological terms. The 'cellular_component' and 'cell' categories are expanded, with 'axon' selected.
- Search Panel:** A search interface with tabs for 'Terms', 'Gene products', and 'Expert'. The search criteria are set to 'Name = axon'. A 'Find' button is visible.
- Classification Caption:** A table showing the classification of the search results. It lists terms and their associated counts.
- Gene Displayer:** A section for displaying gene products. It shows 'axon' with 95 associations found. A list of related genes is provided, including 'Abl', 'App1', 'arm', 'eya', 'Lar', and 'emc'.
- DAG Viewer:** A Directed Acyclic Graph (DAG) showing the hierarchical structure of the search results. The root is 'Gene_Ontology', which branches into 'cellular_component' and then 'cell', with 'axon' at the bottom.
- Main Display:** A large tree diagram showing the distribution of 5089 leaves across various anatomical terms. The terms are color-coded: pink for 'adult head (+)', 'adult head (+) ocellus (+)', 'spermatid (+)', 'male accessory gland secondary cell (+)', 'testis (+)', 'testis (+)', 'testis', 'nervous system (+)', 'embryonic/larval midgut', 'indirect flight muscle (+)', and 'adult muscle system (+)'. A vertical bar on the right side of the tree shows the relative frequency of each term, with a color scale from green to red.

Color	Query	Hits/Tree	Hits/Subtree
Orange	TERM=(axon axon)	40/5089	40/5089
Red	TERM=ribosome	85/5089	85/5089
Blue	TERM=mitochondrial electron transport chain...	40/5089	40/5089
Dark Blue	TERM=(muscle fiber muscle fiber)	20/5089	20/5089

5089 leaves
depth=0

central nervous system (+)
adult head (+)
adult head (+) ocellus (+)
spermatid (+)
male accessory gland secondary cell (+)
testis (+)
testis (+)
testis
nervous system (+)
embryonic/larval midgut
indirect flight muscle (+)
adult muscle system (+)

Identification des transcrits co-localisés

The screenshot displays the ABC version 1.311 software interface. The main window is divided into several panels:

- Terms Viewer:** A hierarchical tree of biological terms. The term "response to starvation" is selected and highlighted in blue.
- Search Panel:** Contains search filters and results. The "Terms" tab is active, showing "find TERMS with" and "Name = response to starvation". It indicates "1 term(s) founds".
- DAG Viewer:** A Directed Acyclic Graph (DAG) showing the relationship between terms. The "Gene_Ontology" root is visible, with "biological_process" and "cellular_process" as children.
- Classification Caption:** A pop-up window showing a table of classification results for the selected term.
- Gene Displayer:** A search bar and a list of gene products associated with the selected term. The search results show "41 association(s) found for 38 gene products".
- Main Display:** A large visualization of gene products, likely a heatmap or dendrogram, showing the co-localization of transcripts. The visualization is labeled "6221 leaves" and "depth=0".

The Classification Caption table is as follows:

Color	Query	Hits/Tree	Hits/Subtree
Blue	CO-LOCALIZED	337/6221	337/6221

The Gene Displayer search results are as follows:

- SGD:S0002264 homeobox tran
- SGD:S0002702 dihydrosphingo
- SGD:S0002885 serine/threonin
- SGD:S0004145 nitrogen catabo
- SGD:S0004147 nitrogen catabo
- TAS SGD_REF:8664

Détail des gènes co-localisés sur le même brin

The screenshot displays the ABC version 1.311 software interface, which is used for analyzing gene co-localization. The interface is divided into several panels:

- Terms Viewer:** A hierarchical tree of Gene Ontology (GO) terms. The term "response to starvation" is selected and highlighted in blue.
- Search Panel:** Contains search criteria for "response to starvation". It shows results for "FIND Terms WITH Name = 'response to starvation'" (1 term(s) founds) and "FIND Terms WITH Name = 'response to biotic stimulus'" (1 term(s) founds).
- DAG Viewer:** A Directed Acyclic Graph (DAG) showing the relationship between terms. The path from "Gene_Ontology" to "biological_process" to "cellular_process" to "cell_growth" to "response" to "response_to_starvation" is visible.
- Classification Tree Viewer:** A tree diagram showing the classification of 53 leaves (depth=43). The tree is zoomed horizontally and vertically. The leaves are labeled with gene IDs and their associated chromosome and coordinates, such as "YDR241W [IV:945142]", "YIL009C-A [IXr:336209]", "YER066C-A [Vr:292201]", "YNL242W [XIV:191323]", "YDR235W [IV:933495]", "YOR304C-A [XVr:888743]", "YFL013W-A [VI:109804]", "YDR238C [IVr:940807]", "YML058C-A [XIIIr:160412]", "YAL058C-A [Ir:39047]", "YMR324C [XIIIr:922443]", "YMR254C [I:]", "YAL068C [Ir:2169]", "YMR175W [XIII:611015]", "YIL071C [IXr:229991]", "YFL035C-A [I:]", "YMR084W [XIII:436627]", "YDR243C [IVr:950275]", "YDR239C [IVr:943411]", "YAL069W [I:335]", "YCL001W [III:111910]", and "YDR238C [IVr:940807]".

Annotations on the right side of the Classification Tree Viewer indicate chromosome locations:

- An orange bracket labeled "Ch IV" points to the top group of genes, including YDR241W, YIL009C-A, YER066C-A, YNL242W, YDR235W, YOR304C-A, YFL013W-A, and YDR238C.
- A green bracket labeled "Ch IVr" points to the bottom group of genes, including YAL058C-A, YMR324C, YMR254C, YAL068C, YMR175W, YIL071C, YFL035C-A, YMR084W, YDR243C, YDR239C, YAL069W, YCL001W, and YDR238C.

The Gene Displayer panel at the bottom left shows the "response to starvation" term with 41 associations found for 38 gene products. The "inherited terms' gene products" checkbox is checked, and a list of gene products is displayed, including "SGD:S0002264 homeobox tran", "SGD:S0002702 dihydrosphingo", "SGD:S0002885 serine/threonin", "SGD:S0004145 nitrogen catabo", "SGD:S0004147 nitrogen catabo", and "TAS SGD_REF:8664".

Localisation des gènes sur le chromosome IV

Coordonnées 920000 à 960000

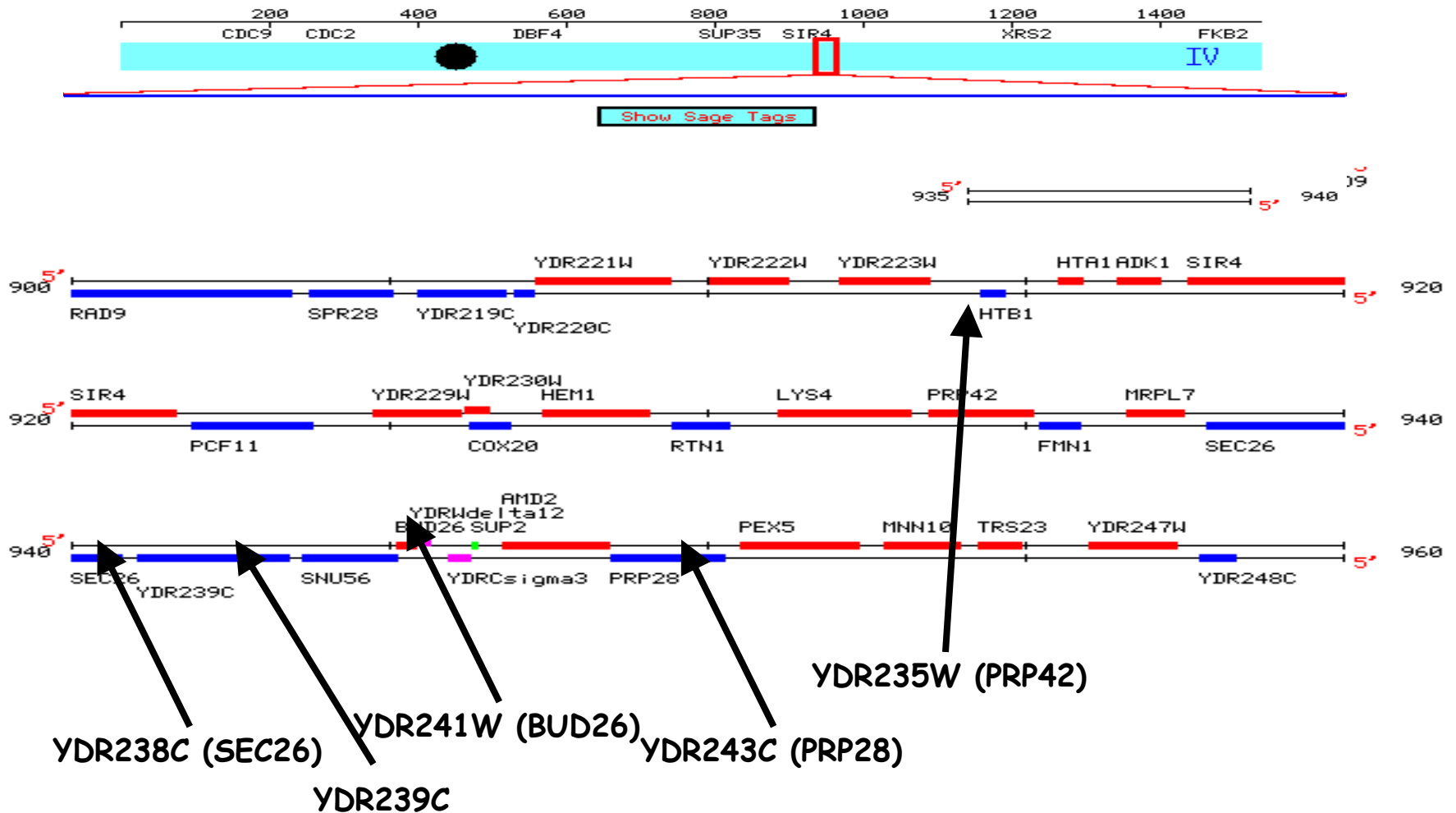
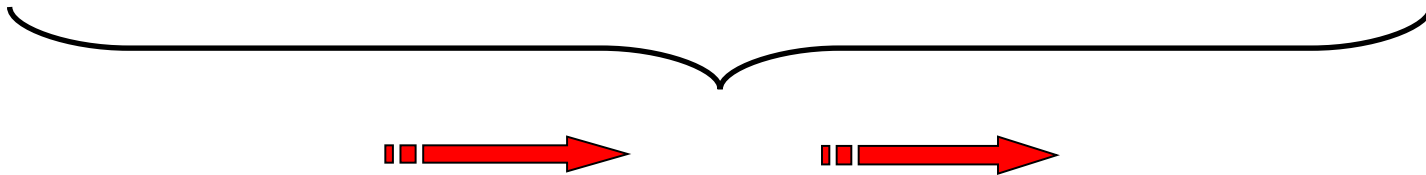
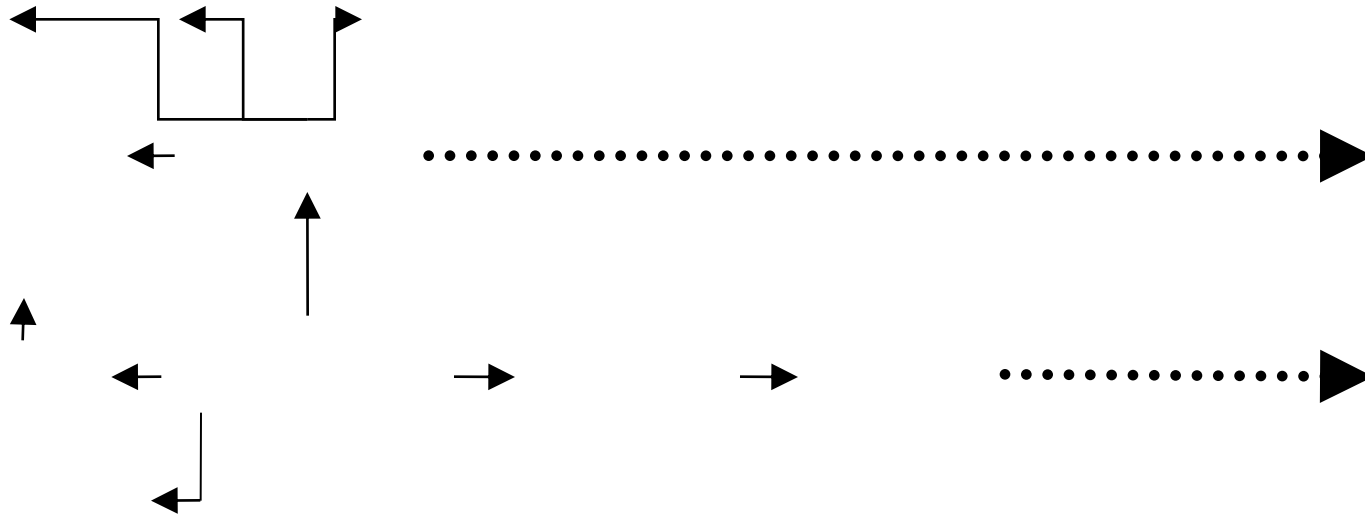
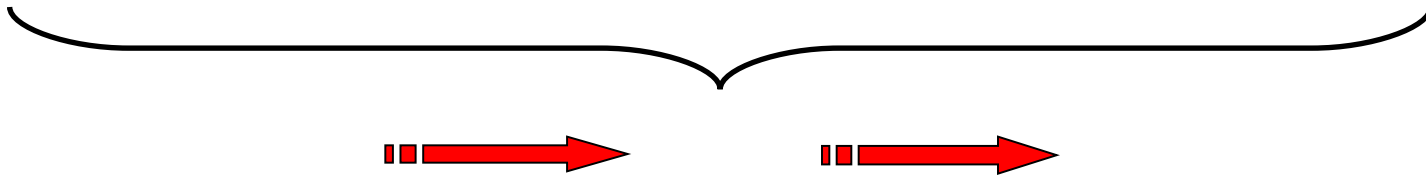
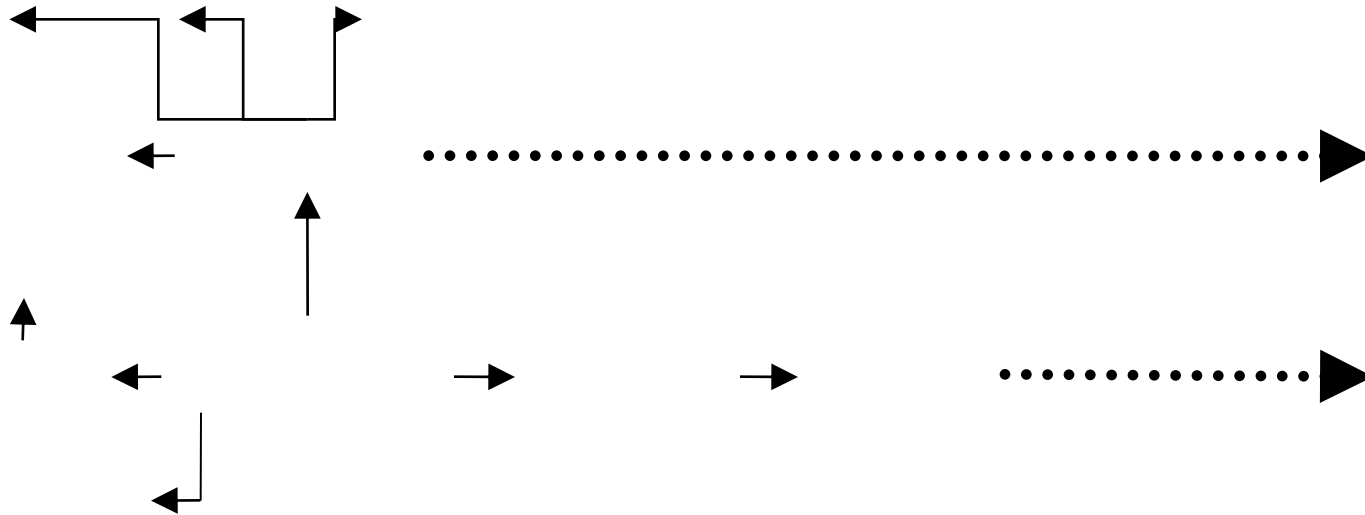


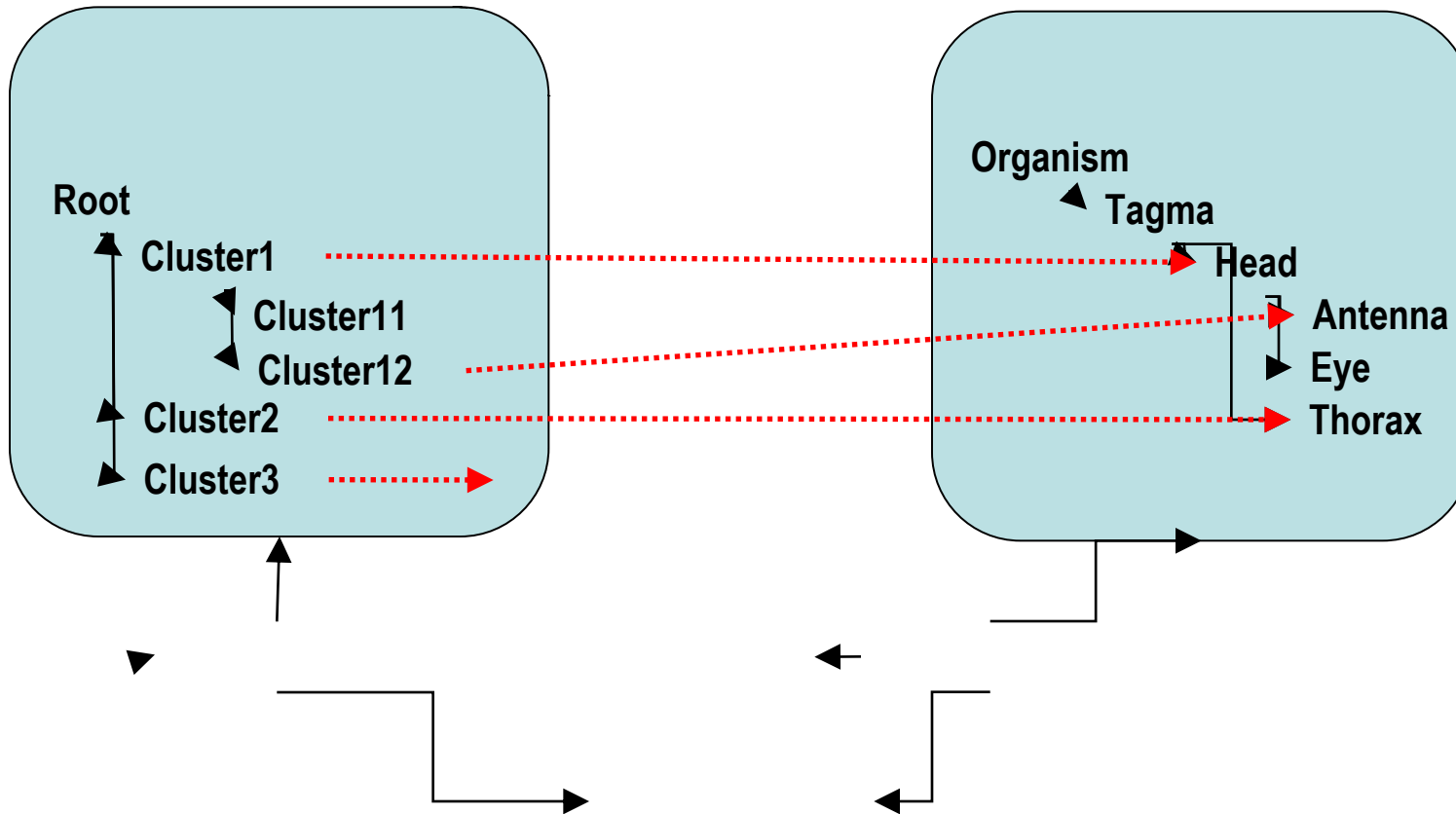
Schéma actuel



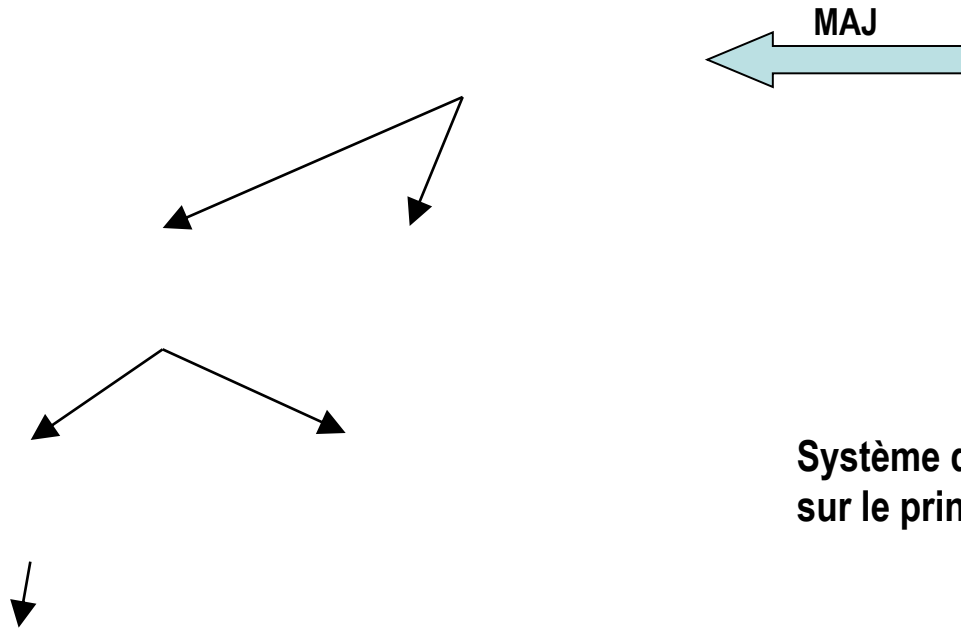
Evolution futures



Réutilisation des annotations



Arborescence de vues



**Systeme de réécriture des requêtes
sur le principe du « local-as View »**

Cadre du projet



UMR 6543 Institut de signalisation, biologie du développement et cancer

Université de Nice - Sophia Antipolis

CENTRE DE BIOCHIMIE

Parc Valrose

06108 NICE CEDEX 2
