Projet ABC
Analyse et Base de Connaissances

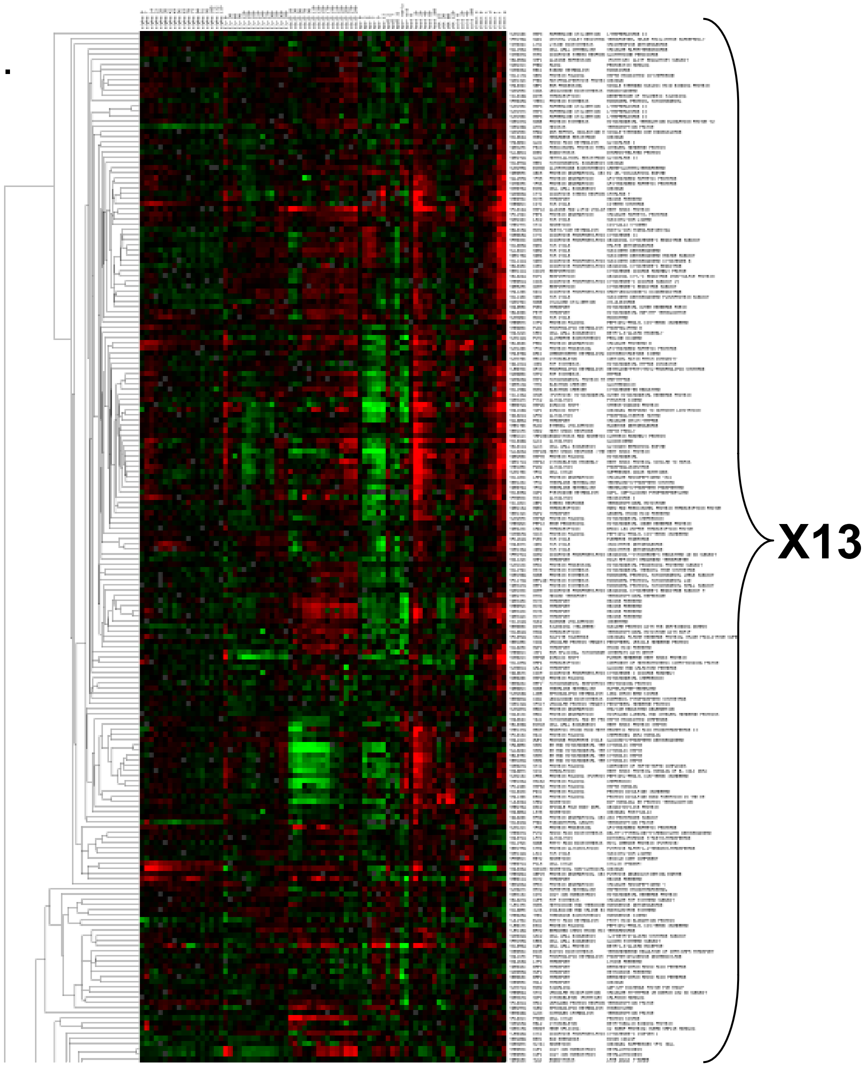
Problème d'interprétation des données

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.

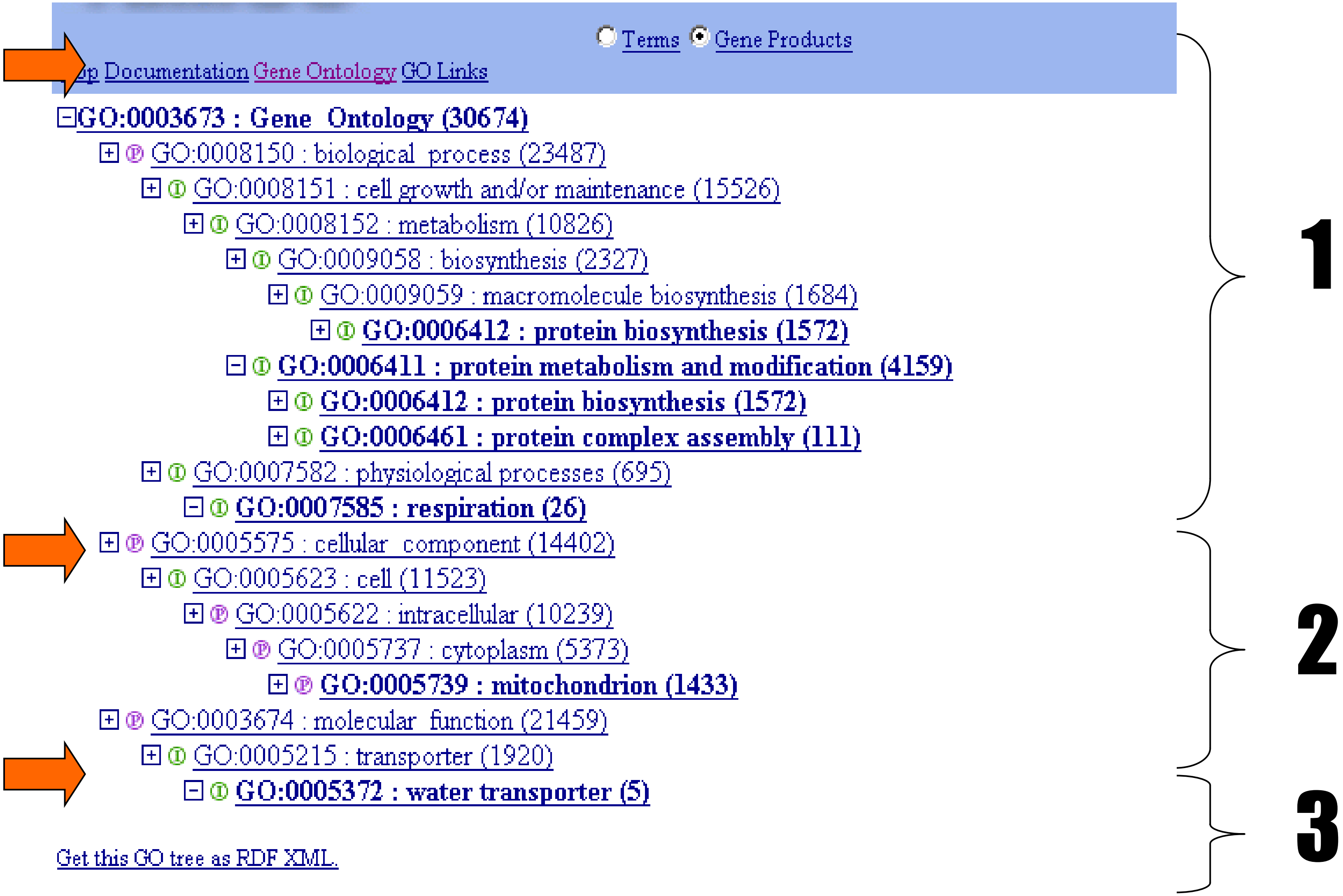
6221gènes de *Saccharomyces cerevisiae*

80 expériences

2467 gènes ont une fonction répertoriée



Modélisation hiérarchique des concepts dans GO



○ Terms ● Gene Products

op Documentation Gene Ontology GO Links

- ⊟ **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.

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←→produits de gènes
    Localisation chromosomique
    Classifications hiérarchiques

+ modules de data-mining

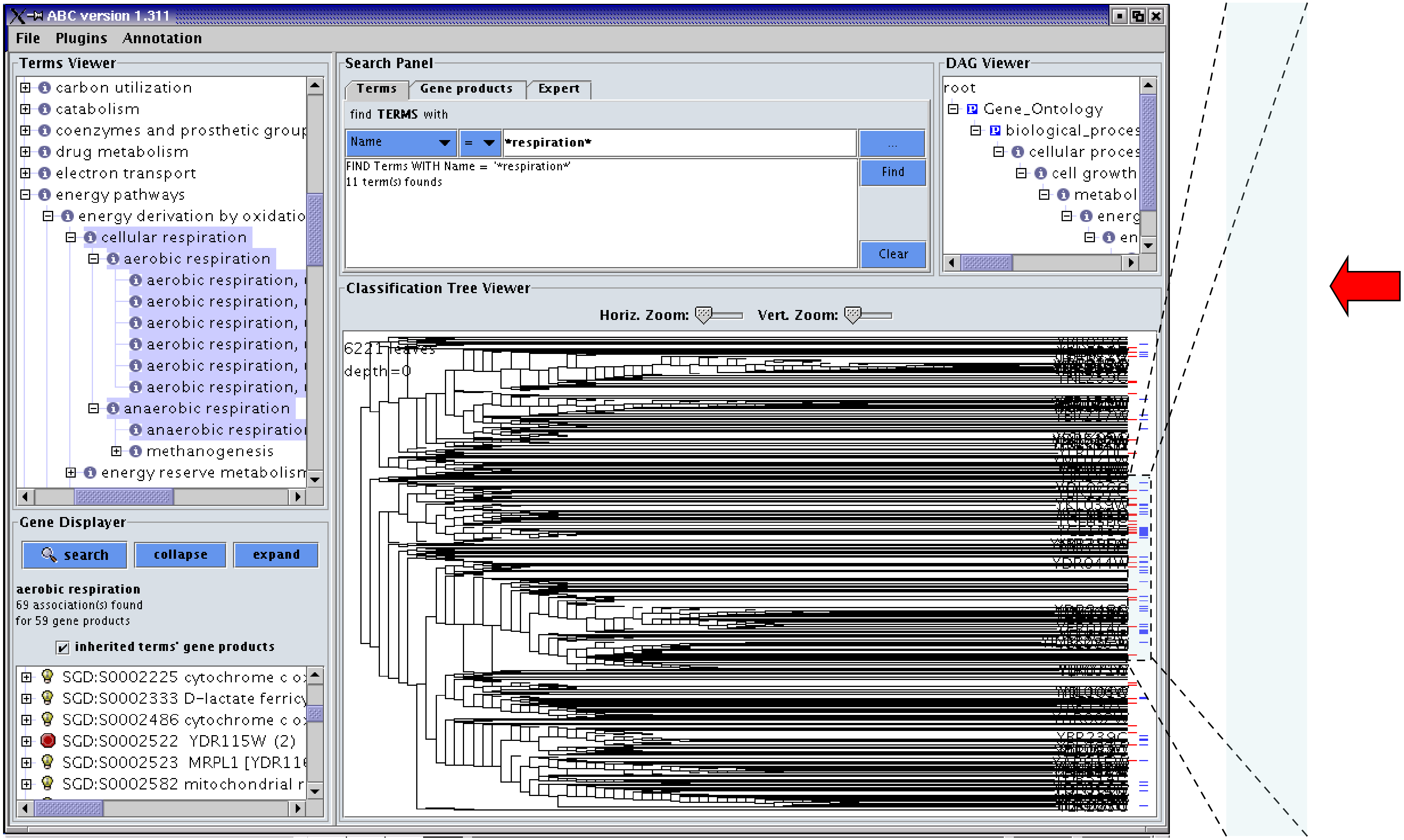
    Choix d'annotations caractérisant des groupements de transcrits
    Détection des gènes covariants et co-localisés

- 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



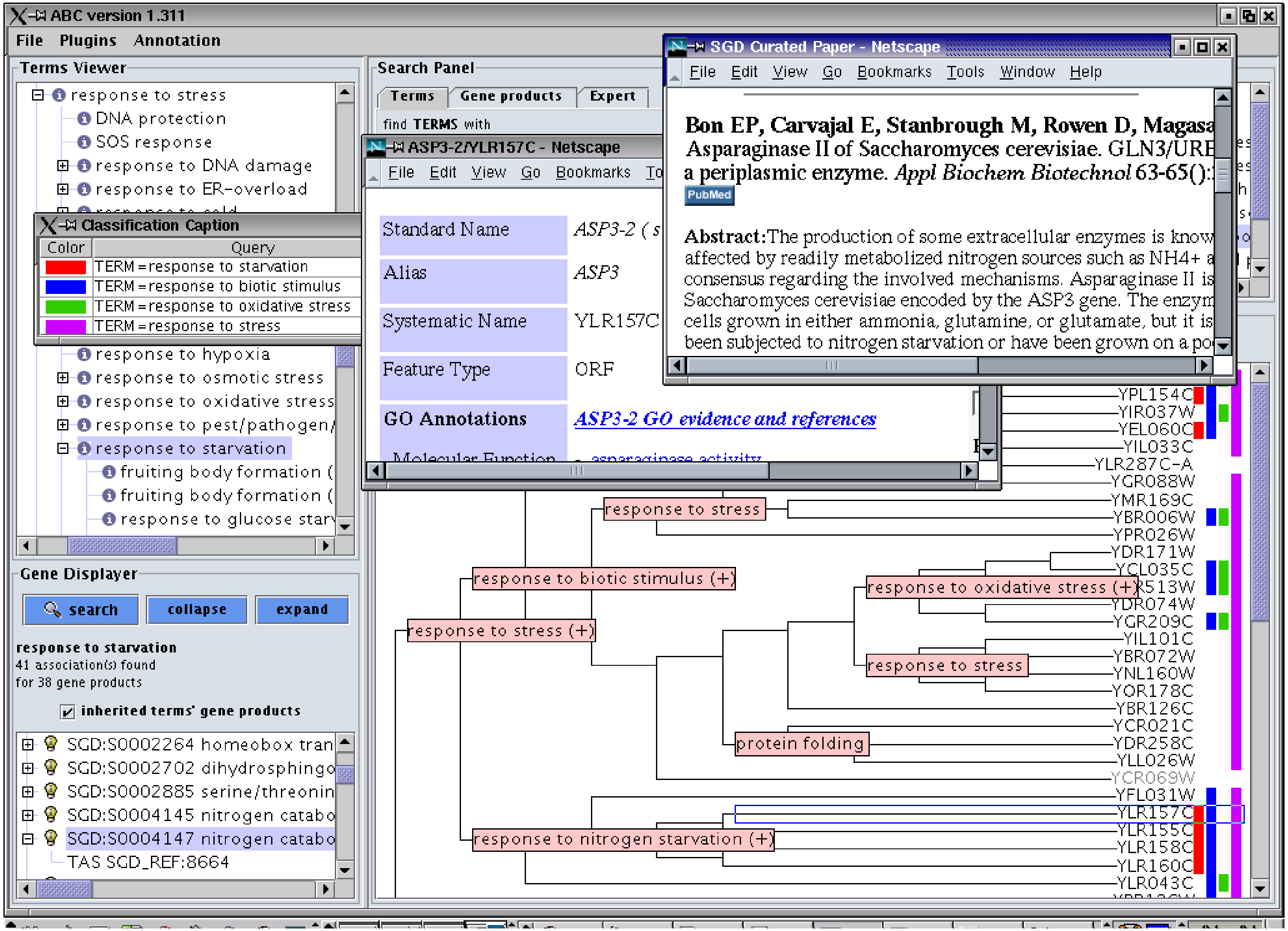
respiration

FIND Terms WITH Name = '*respiration*'
11 term(s) founds

aerobic respiration
69 association(s) found
for 59 gene products

SGD:S0002225 cytochrome c o
SGD:S0002333 D-lactate ferricy
SGD:S0002486 cytochrome c o
SGD:S0002522 YDR115W (2)
SGD:S0002523 MRPL1 [YDR11
SGD:S0002582 mitochondrial r

Annotation automatique des partition



Classification Caption

Color	Query
🟥	TERM = response to starvation
🟦	TERM = response to biotic stimulus
🟩	TERM = response to oxidative stress
🟪	TERM = response to stress

Standard Name	*ASP3-2 (s*
Alias	*ASP3*
Systematic Name	YLR157C
Feature Type	ORF
GO Annotations	*ASP3-2 GO evidence and references*
Molecular Function	asparaginase activity

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

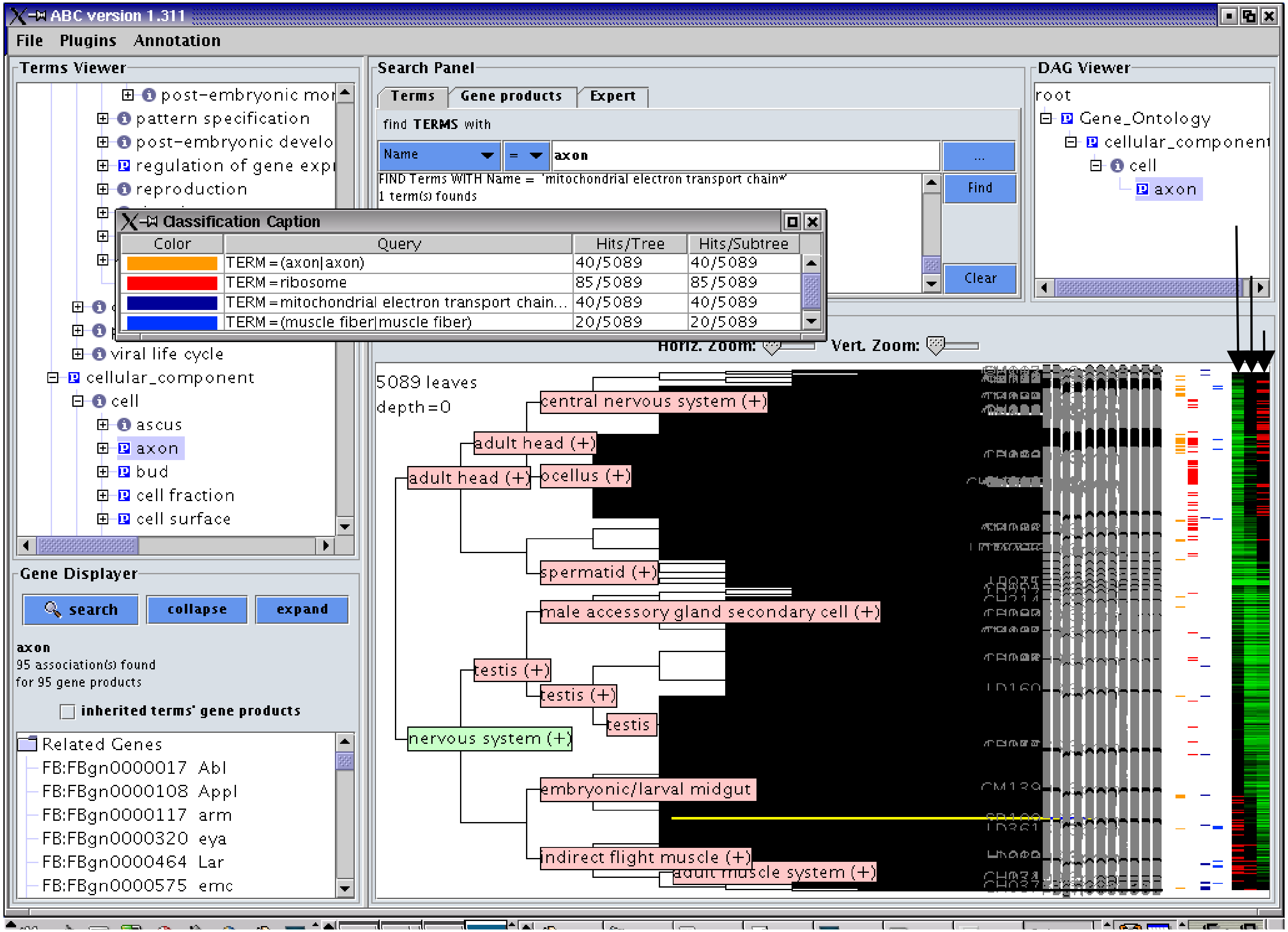
Abstract: The production of some extracellular enzymes is know
affected by readily metabolized nitrogen sources such as NH4+ a
consensus regarding the involved mechanisms. Asparaginase II is
Saccharomyces cerevisiae encoded by the ASP3 gene. The enzym
cells grown in either ammonia, glutamine, or glutamate, but it is
been subjected to nitrogen starvation or have been grown on a po

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

Utilisation de l'ontologie anatomique de FB



find **TERMS** with

Name = axon

FIND Terms WITH Name = 'mitochondrial electron transport chain*'
1 term(s) founds

Classification Caption

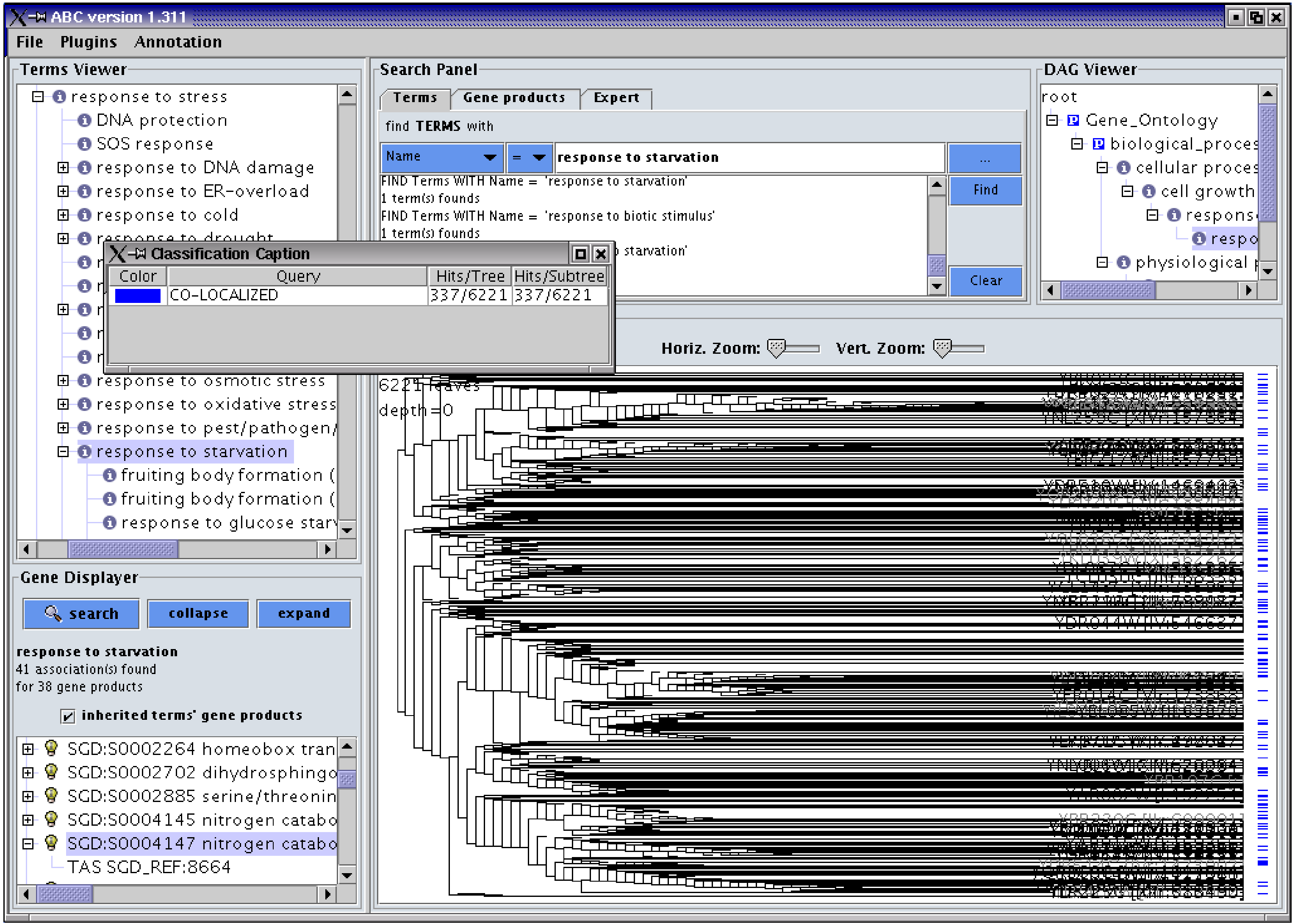
Color	Query	Hits/Tree	Hits/Subtree
🟧	TERM=(axon|axon)	40/5089	40/5089
🟥	TERM=ribosome	85/5089	85/5089
🟦	TERM=mitochondrial electron transport chain...	40/5089	40/5089
🟦	TERM=(muscle fiber|muscle fiber)	20/5089	20/5089

axon
95 association(s) found
for 95 gene products

☐ **inherited terms' gene products**

Related Genes
FB:FBgn0000017 Abl
FB:FBgn0000108 Appl
FB:FBgn0000117 arm
FB:FBgn0000320 eya
FB:FBgn0000464 Lar
FB:FBgn0000575 emc

Identification des transcrits co-localisés



response to stress
DNA protection
SOS response
response to DNA damage
response to ER-overload
response to cold
response to drought
response to osmotic stress
response to oxidative stress
response to pest/pathogen/
response to starvation
fruiting body formation (
fruiting body formation (
response to glucose starv

response to starvation

FIND Terms WITH Name = 'response to starvation'
1 term(s) founds
FIND Terms WITH Name = 'response to biotic stimulus'
1 term(s) founds
starvation'

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

root
Gene_Ontology
biological_proces
cellular proces
cell growth
respons
respo
physiological

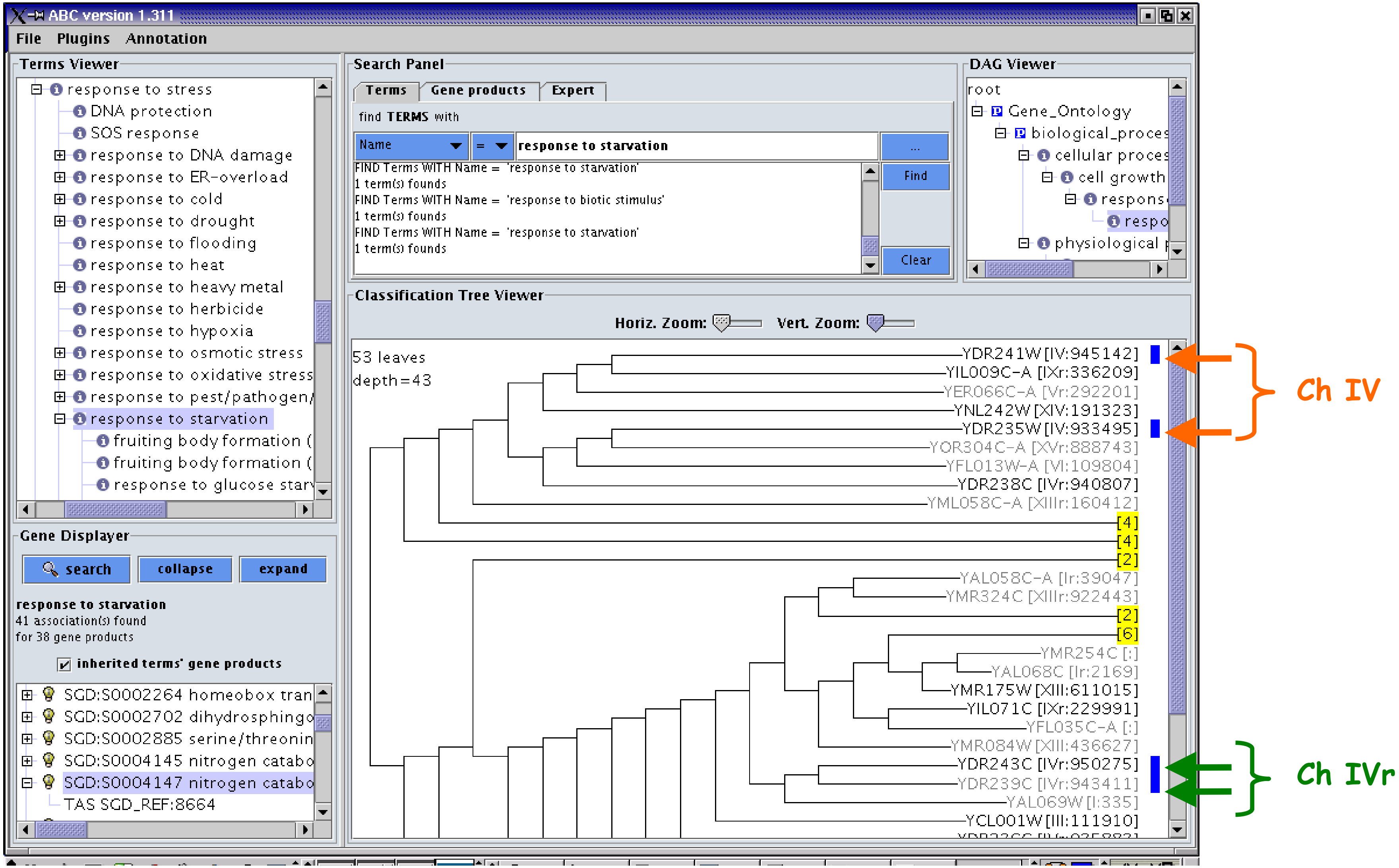
6221 leaves
depth=0

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

Détail des gènes co-localisés sur le même brin



Name = response to starvation

```
FIND Terms WITH Name =  'response to starvation'
1 term(s) founds
FIND Terms WITH Name =  'response to biotic stimulus'
1 term(s) founds
FIND Terms WITH Name =  'response to starvation'
1 term(s) founds
```

53 leaves
depth=43

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]
[4]
[4]
[2]
YAL058C-A [Ir:39047]
YMR324C [XIIIr:922443]
[2]
[6]
YMR254C [:]
YAL068C [Ir:2169]
YMR175W [XIII:611015]
YIL071C [IXr:229991]
YFL035C-A [:]
YMR084W [XIII:436627]
YDR243C [IVr:950275]
YDR239C [IVr:943411]
YAL069W [I:335]
YCL001W [III:111910]

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

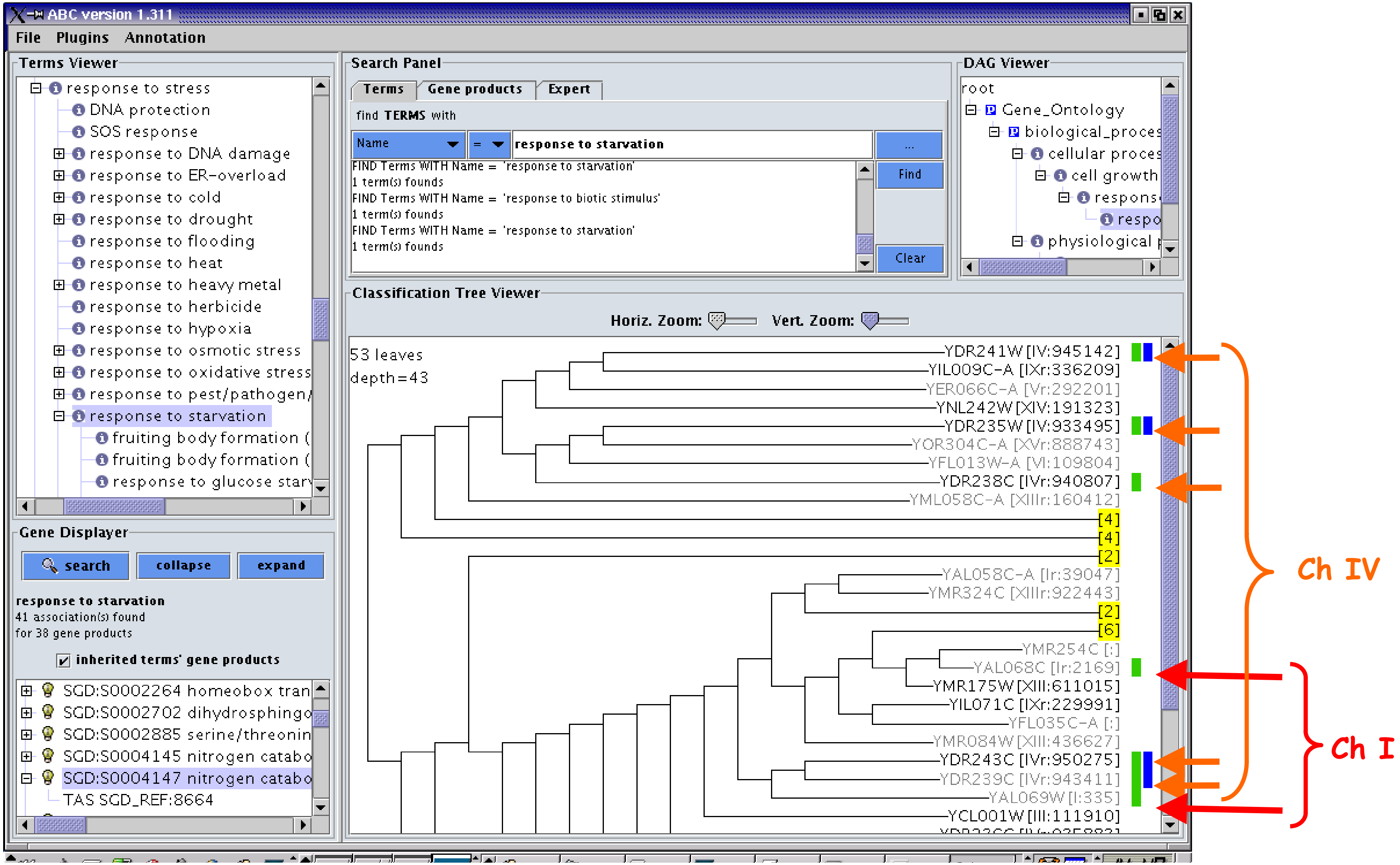
☑ inherited terms' gene products

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

Ch IV

Ch IVr

Détail des gènes co-localisés sur le même brin

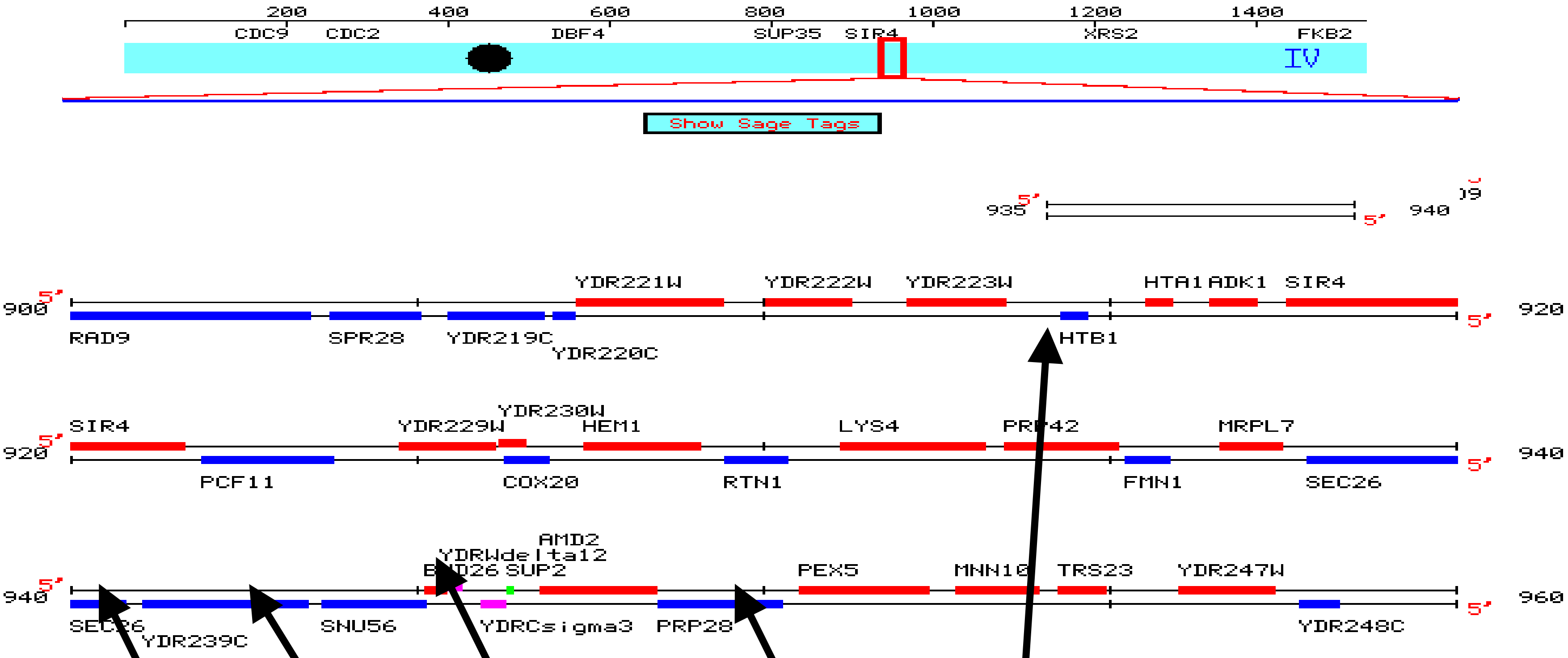


```
FIND Terms WITH Name = 'response to starvation'
1 term(s) founds
FIND Terms WITH Name = 'response to biotic stimulus'
1 term(s) founds
FIND Terms WITH Name = 'response to starvation'
1 term(s) founds
```

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

Localisation des gènes sur le chromosome IV

Coordonnées 920000 à 960000



YDR235W (PRP42)

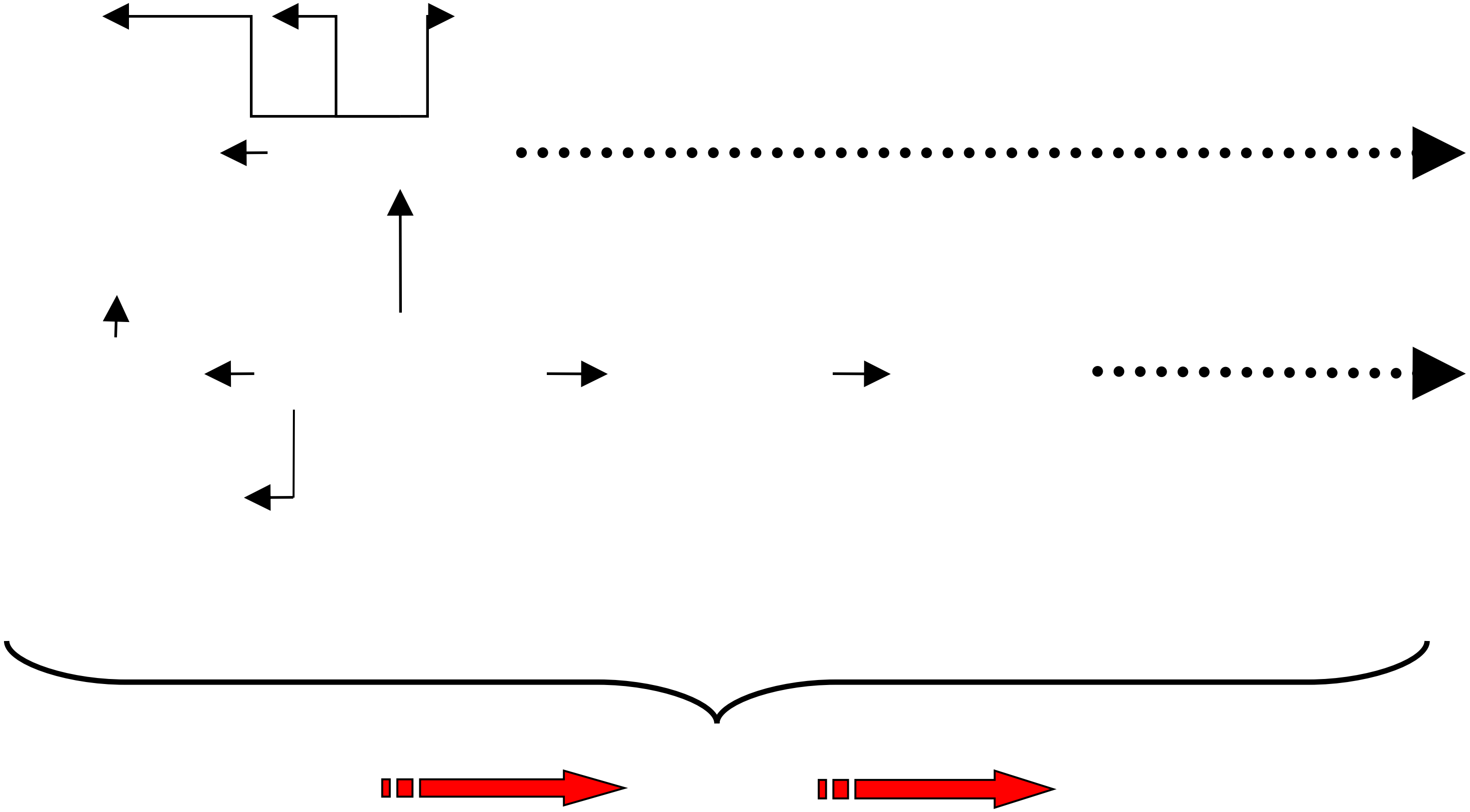
YDR241W (BUD26)

YDR238C (SEC26)

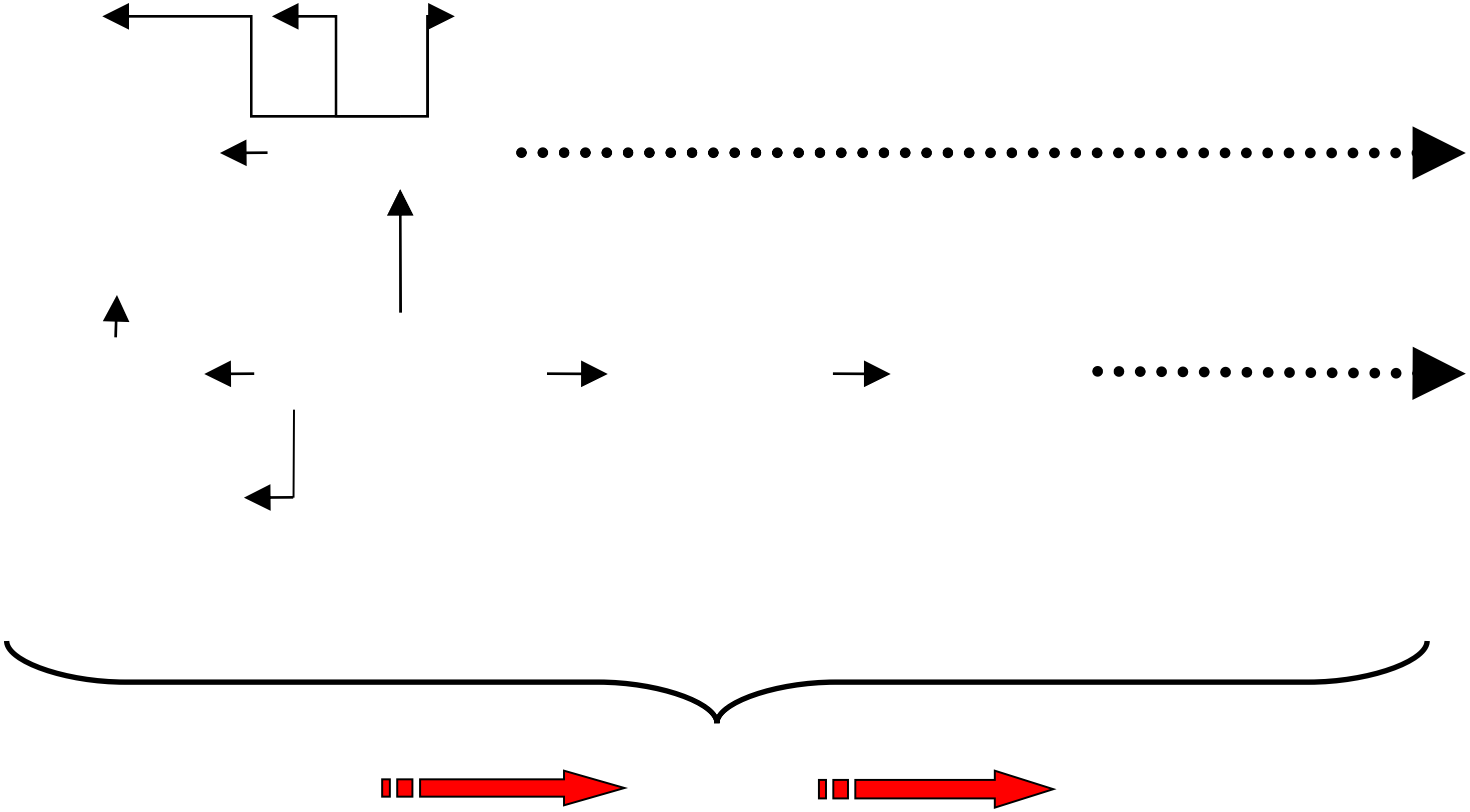
YDR243C (PRP28)

YDR239C

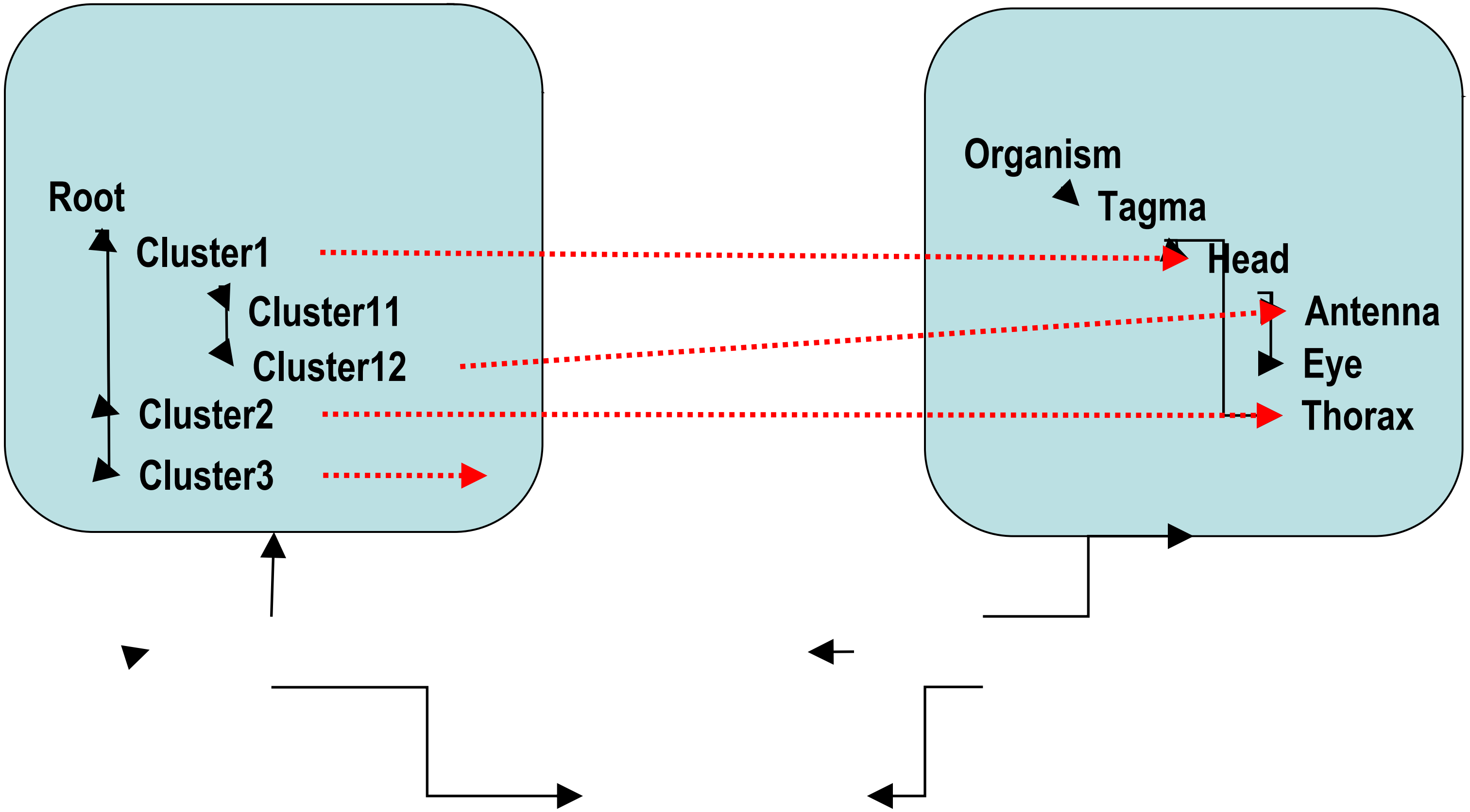
Schéma actuel



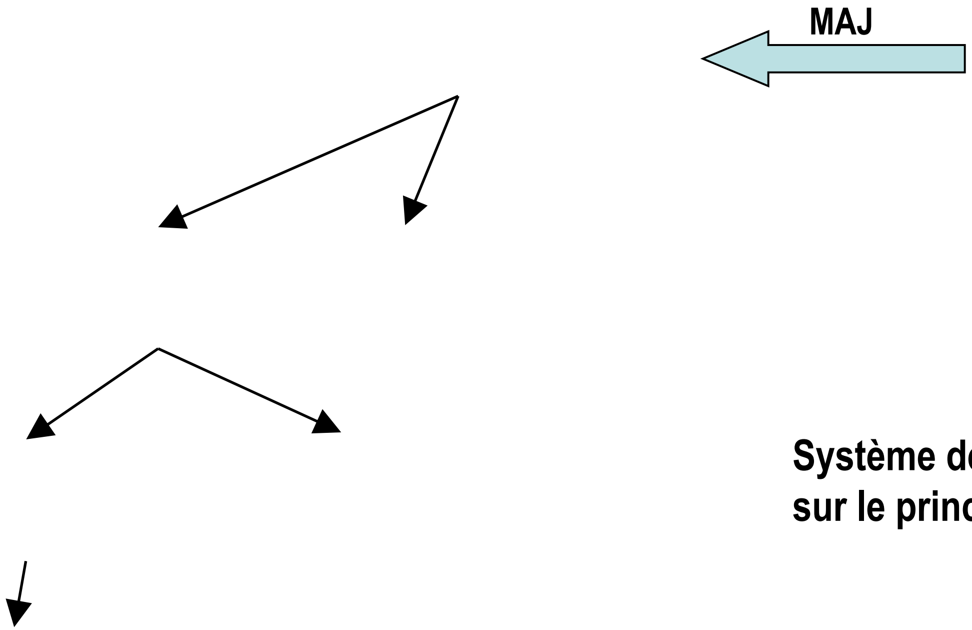
Evolutions futures



Réutilisation des annotations



Arborescence de vues



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

Cadre du projet

CNRS
CENTRE NATIONAL
DE LA RECHERCHE
SCIENTIFIQUE

Université
Nice SOPHIA ANTIPOLIS

UMR 6543 Institut de signalisation, biologie du développement et cancer
Université de Nice - Sophia Antipolis
CENTRE DE BIOCHIMIE
Parc Valrose
06108 NICE CEDEX 2