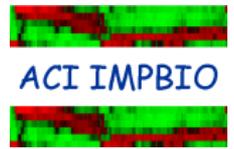


# DRUGNETWORKS

(Projet de Recherche 2004)

## RESEAUX GENETIQUES, GROUPES DE GENES CO-REGULES ET EVOLUTION DE LA RESISTANCE AUX DROGUES

<http://www.biologie.ens.fr/drugnetworks>



### CONTEXTE

L'action de l'**environnement** sur le **fonctionnement des génomes** est un processus complexe qui implique, sur plusieurs échelles de temps, la mise en place de **casades de réseaux** de régulations géniques. Cette **réponse cognitive** de la cellule à la nature du milieu extérieur engage très rapidement un changement dans le programme d'expression du génome qui conduira à des mécanismes d'adaptation et de résistance.

La plupart des études ayant actuellement abordé cette question ont délaissé l'aspect **dynamique** de cette réponse, de plus, elle ne permettent pas de distinguer les relations directes des relations indirectes entre les phénomènes décrits.

Les **éléments précoces** de cette réponse cognitive du génome à l'environnement déterminent certainement le devenir de la cellule et son phénotype ultérieur. Il n'y a aujourd'hui que peu d'informations sur la chronologie de mise en place des réseaux de régulation géniques et sur la relation entre ce phénomène et la nature des signaux environnementaux.

### OBJECTIFS

Ce projet vise à développer des outils biologiques et bioinformatiques pour décrire, au niveau du fonctionnement du génome, la **réponse précoce** de la cellule à des stress environnementaux.

Le modèle choisi sera la levure ***S. cerevisiae***, **eucaryote simple**, et sa réaction à différents stress chimiques.

Outre la volonté de restreindre l'analyse de la **variation du transcriptome** aux premiers instants de mise en présence de la cellule avec la drogue, cette étude s'attachera à définir avec une **haute densité** (espaces de temps courts) les variations de la quantité des transcrits de chacun des gènes. Ainsi une **étude cinétique** typique conduite sur 40 minutes contiendra 6 temps d'analyse qui répétés trois fois dans 5 contextes génétiques différents généreront 500 000 données dont il faudra extraire les informations importantes.

Une étude cinétique pourra, après mises au point, être réalisée avec plusieurs types de stress chimiques (bénomyl, fluphénazine, H<sub>2</sub>O<sub>2</sub>, sélénium,...) et sur des espèces de levures différentes (*S. cerevisiae*, *C. albicans*,...); le point important étant que l'ensemble de ces expériences seront conduites d'une façon équivalente et donc comparables.

Sur le plan informatique, le grand nombre de données homogènes et leur caractère cinétique impliquent des approches plus exigeantes et efficaces que le simple « guilt by association » basé sur la similarité de profils d'expression. En particulier :

- Les données cinétiques seront exploitées grâce à une méthodologie en cours de développement fondée sur une classification de type « **model-driven clustering** » c'est à dire guidée par les « formes » possibles des profils cinétiques des gènes.

- Pour l'étude de l'expression différentielle des gènes en fonction des drogues étudiées, une des approches envisagées repose sur le « **biclustering des profils d'expression** » c'est à dire la détermination de groupes de gènes qui ont un profil d'expression semblable pour un sous-ensemble de drogues.

- Les résultats obtenus dans les deux levures *S. cerevisiae* et *C. albicans* seront analysés en utilisant une approche en cours de développement et fondée sur des procédures de « **multidimensional scaling** » ou MDS

### PROJET

De ces études nous espérons une meilleure caractérisation des réseaux de régulation et de leurs interactions. Nous devrions, en particulier, être capables de décrire les relations de hiérarchie dans l'établissement de ces réseaux. De plus, l'analyse de l'évolution de ces groupes de gènes régulés au cours de 150 millions d'années présente à la fois un intérêt théorique mais aussi pratique puisque l'on peut espérer définir des spécificités de la levure pathogène *C. albicans*. En outre, les outils bioinformatiques mis en œuvre seront applicables à d'autres problématiques.

### PUBLICATIONS PRINCIPALES

Lelandais G, Lucau-Danila A, Kozovska Z, Tanty V, Delaveau T, Devaux F, Jacq C.

**Early expression of yeast genes affected by chemical stress.**  
Mol Cell Biol. 2005 Mar;25(5):1860-8.

Lelandais G, Vincens P, Badel-Chagnon A, Vialette S, Jacq C, Hazout S

**Comparing Gene Expression Networks in a 3D Space to Extract Similarities and Differences between Organisms**

Soumris

Kieffer Y, Lelandais G, Vincens P, Hazout S, Badel-Chagnon A.

**BiSIAN: Biclustering in gene expression data by Simulated Annealing**

En préparation

### LABORATOIRES IMPLIQUES

- Laboratoire de Génétique Moléculaire CNRS UMR 8541, ENS
- Equipe de Bioinformatique Génomique et Moléculaire, INSERM U726, Université P7

### MOTS CLES

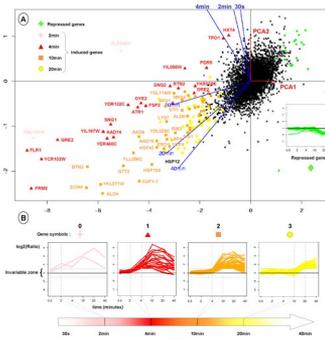
Transcriptome, cinétique, réseaux de régulation, classification, « multidimensional scaling »

### RESULTATS

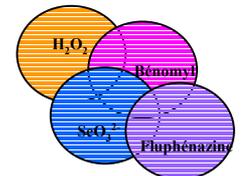
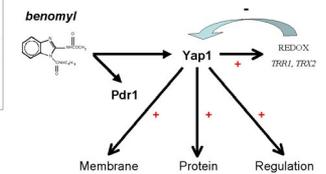
#### • Sur le plan biologique :

Le choix de la première drogue étudiée s'est porté sur le **bénomyl** (composé antimitotique qui inhibe la synthèse des microtubules).

- Mise en évidence de 4 classes de gènes qui diffèrent par leur délais d'induction (119 gènes au total).
- Mise en évidence d'une réponse **rapide** et **transitoire**.



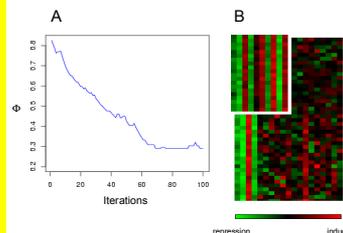
➢ Les événements transcriptionnels précoces sont sous le contrôle des facteurs de transcription **Yap1** et **Pdr1**



Des études similaires sont actuellement menées avec **des stress de nature diverse** afin de caractériser les éléments **communs** et **spécifiques** des différentes réponses.

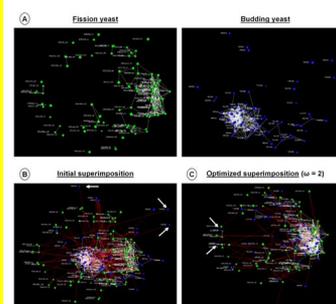
#### • Sur le plan Bioinformatique :

- **Biclustering par recuit simulé appliqué à des profils d'expression**



La méthode de biclustering par recuit simulé est basée sur la minimisation d'un indice d'hétérogénéité  $\Phi$  (A) en réalisant une série de transformations de la matrice initiale. L'algorithme utilisé permet d'identifier des groupes de gènes présentant des profils d'expression similaires pour un nombre limité d'expériences (B).

- **Extraction de similitudes et de différences dans les expressions de gènes entre deux organismes différents**



Les gènes sont représentés par des points dans un espace à 2 ou 3 dimensions, de telle sorte que les distances euclidiennes entre les points reflètent au mieux les similitudes mesurées entre les profils d'expression des gènes. Afin de comparer la structure des nuages obtenus, une superposition dans un unique espace est réalisée de manière à rapprocher les gènes orthologues tout en conservant la structure initiale de chacun des nuages.