



# **Global Analysis of CoRegulation for the identification of functional modules**

Rim Zaag, Guillem Rigaill, Jean-Philippe Tamby, Veronique Brunaud, Zakia Tariq, Sebastien Aubourg, Etienne Delannoy, Marie-Laure Martin-Magniette

Unité de Recherche en Génomique Végétale (URGV) - UMR INRA CNRS UEVE Saclay Plant Science , 2 Rue Gaston Crémieux, 91057 Evry Cedex, France.

Unité de Mathématiques et Informatique Appliquées (MIA) - UMR 518 AgroParisTech-INRA, 16 Rue Claude Bernard, 75231 Paris Cedex, France.

**Rim ZAAG**  
PhD student  
**Bioinformatics for Predictive Genomics**  
**Team**



# Context and Background

- Completely sequenced eukaryotic organisms: 20% to 40% of the predicted genes have no assigned function (Hanson et al., 2010).
- Plant model *Arabidopsis*: around of 5000 orphan genes according to the TAIR v10.



Limit: Each gene can be involved in different biological processes and can interact with different functional partners



**gene modules** represent an excellent framework to infer new functions to genes and to go toward regulatory networks

# Context and Background

- Completely sequenced eukaryotic organisms: 20% to 40% of the predicted genes have no assigned function (Hanson et al., 2010).
- Plant model *Arabidopsis*: around of 5000 orphan genes according to the TAIR v10.



Limit: Each gene can be involved in different biological processes and can interact with different functional partners



**gene modules** represent an excellent framework to infer new functions to genes and to go toward regulatory networks

## AIMS

- Identification and characterization of modules involved in stress response.
- Function inference of orphan genes.

# Dataset

GEM2Net: a new CATdb module (Cf poster 122)

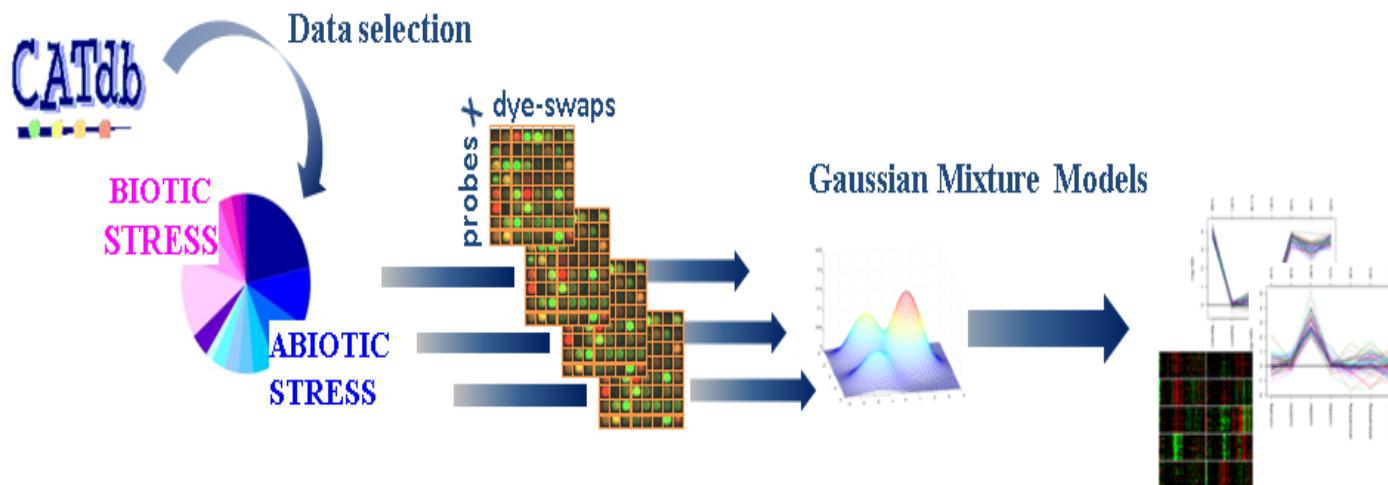


**Goal:** provide a global overview of the co-expression modules responding to stress.

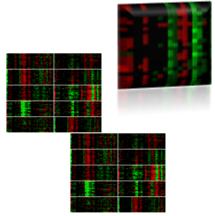
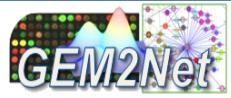
**Data:** a set of CATMA experiments of Arabidopsis transcriptome across 18 different stress categories.

**Method:** model-based clustering for each stress category

**Results:** ~700 clusters of coexpression.

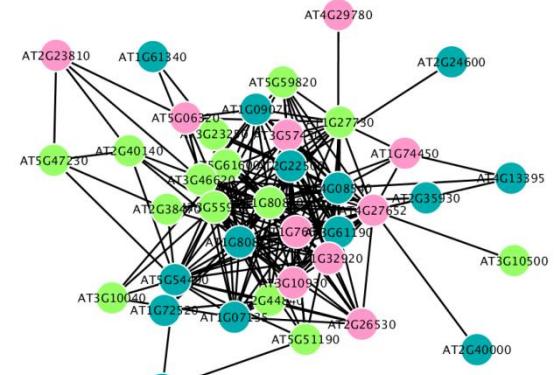


# CoRegulation



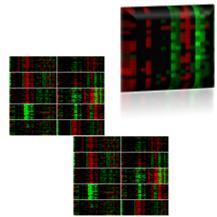
CoExpression clusters  
for each category of  
stress

Integration



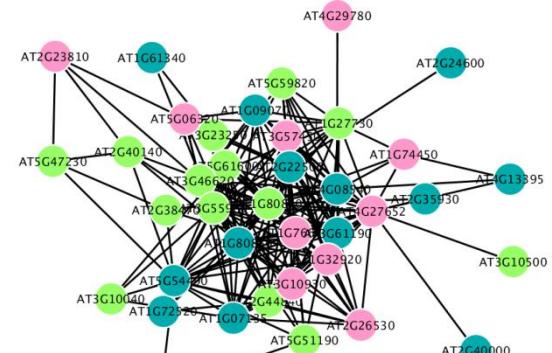
CoRegulation Network

# CoRegulation



CoExpression clusters  
for each category of  
stress

## Integration



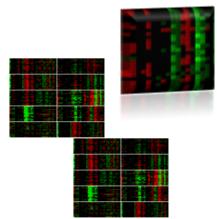
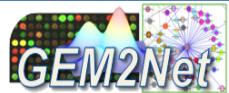
CoRegulation Network

**Identification  
of Coregulated  
genes**

**Describe groups  
of functional  
partners**

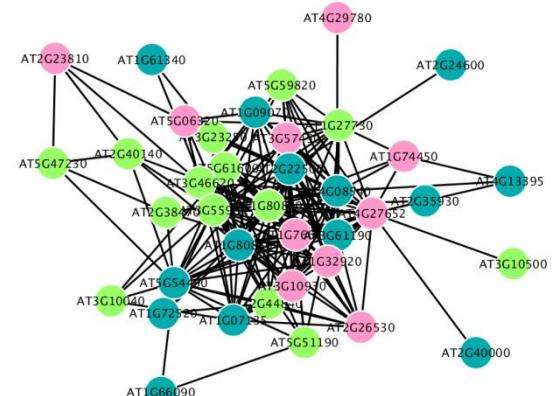
**Annotation  
of orphan  
genes**

# CoRegulation



CoExpression clusters  
for each category of  
stress

## Integration

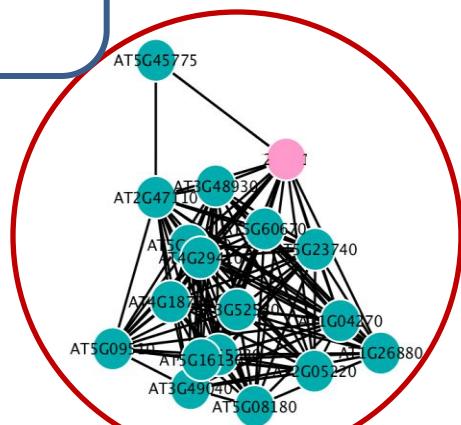


CoRegulation Network

**Identification  
of Coregulated  
genes**

**Describe groups  
of functional  
partners**

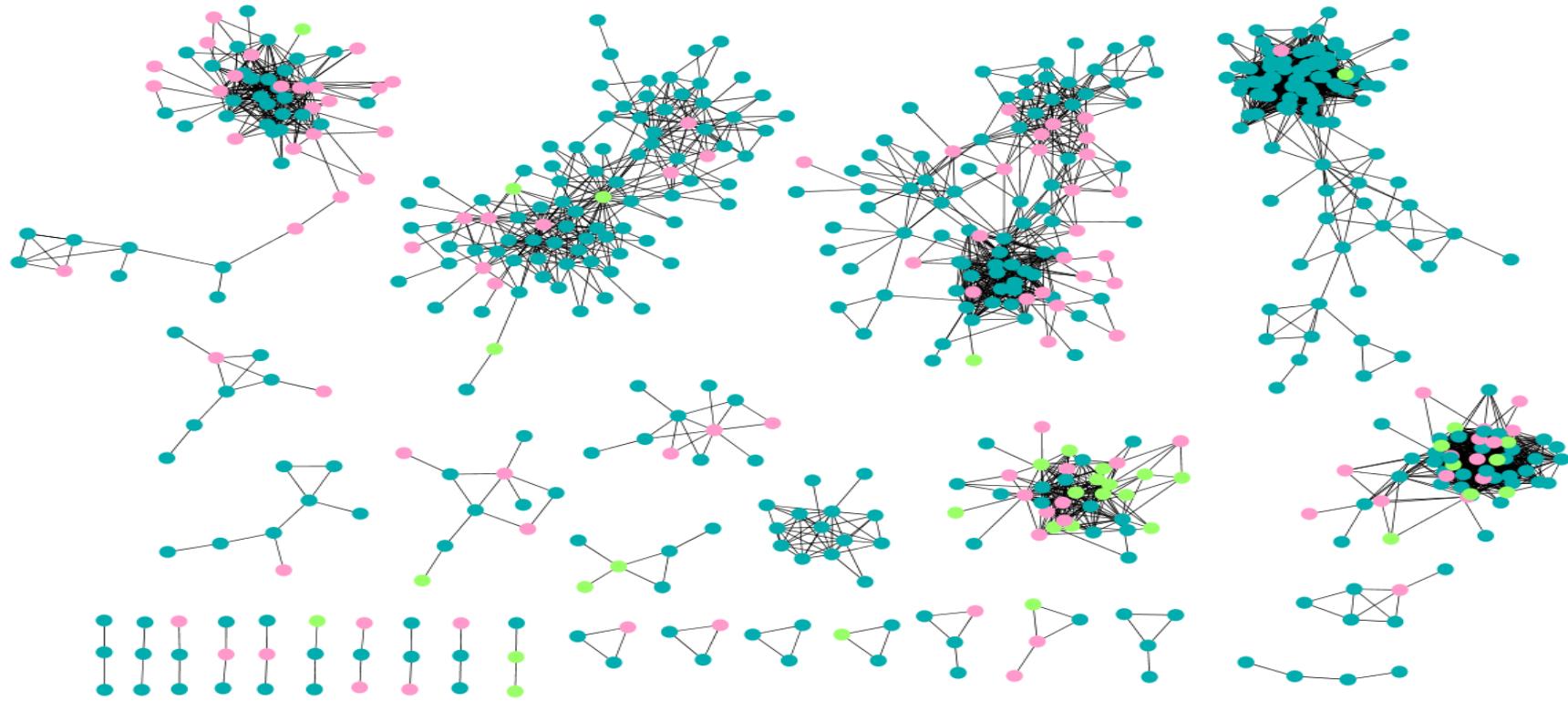
**Annotation  
of orphan  
genes**



# CoRegulation Network

POSTER n : 104

Pairs conserved in at least 7 stresses



Legend

