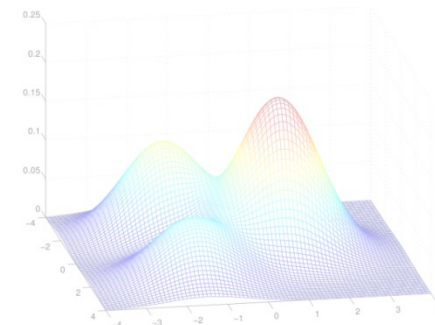
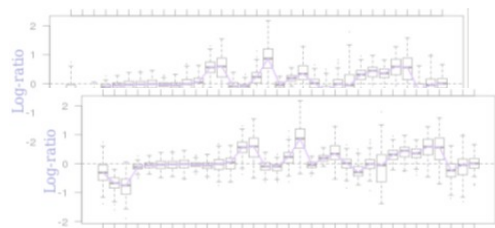
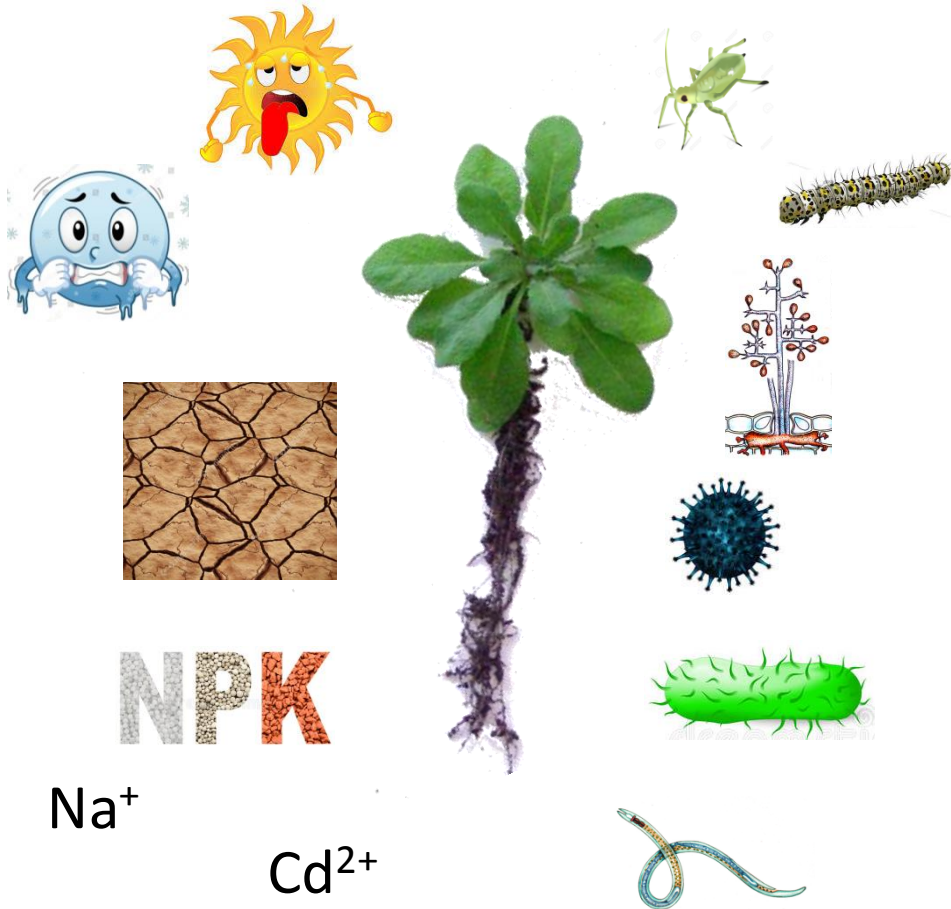


# A transcriptome meta-analysis identifies the response of plant to stresses

Etienne Delannoy, Rim Zaag, Guillem Rigail, Marie-Laure Martin-Magniette



# Biological context



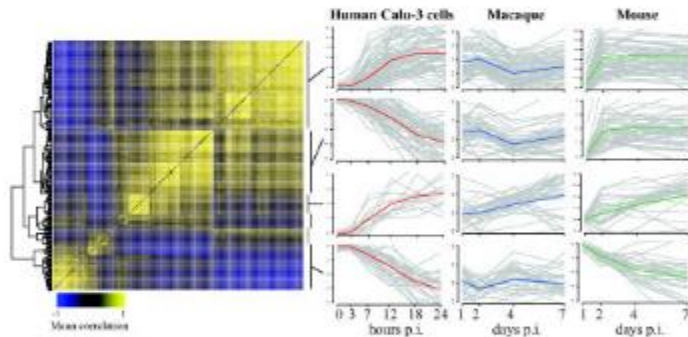
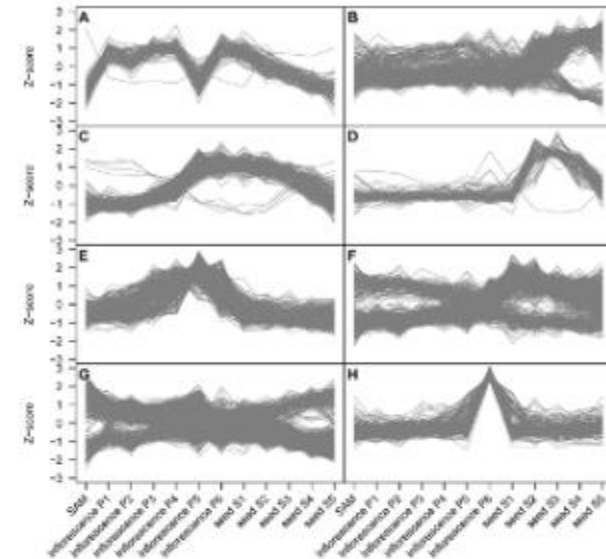
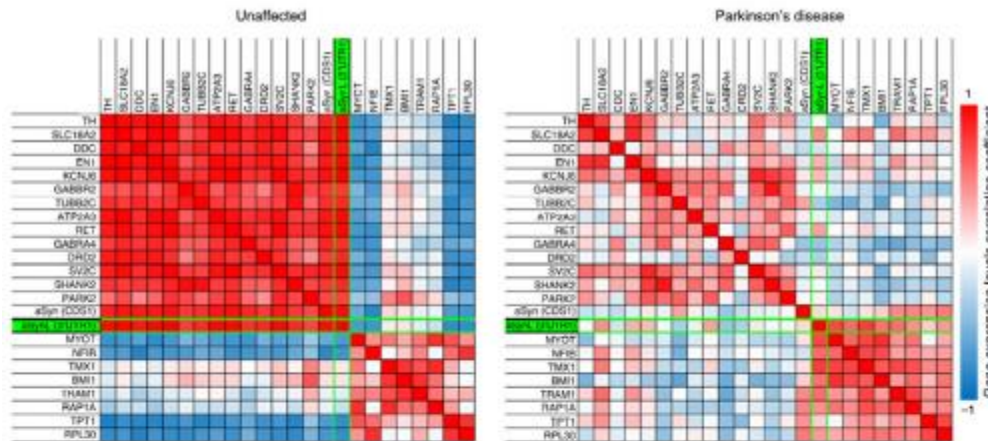
Multiple biotic and abiotic stresses impacting plant growth

Coordinated response to stresses in general ?

Numerous single stress transcriptomic data sets available

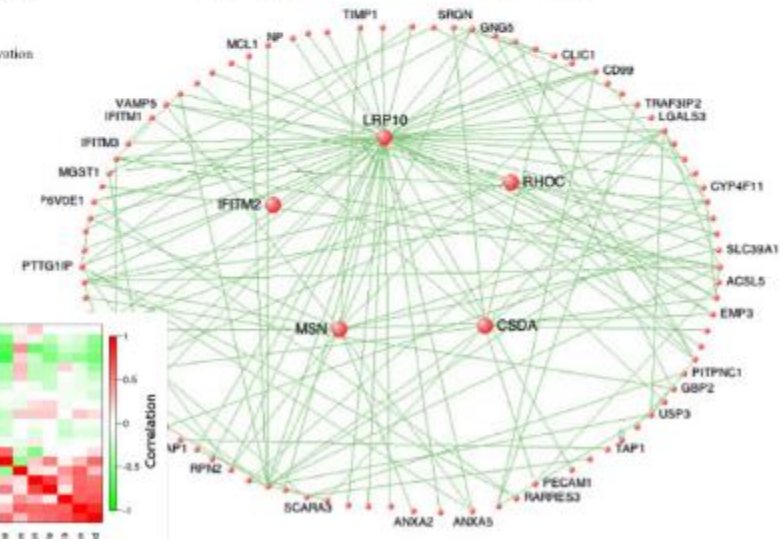
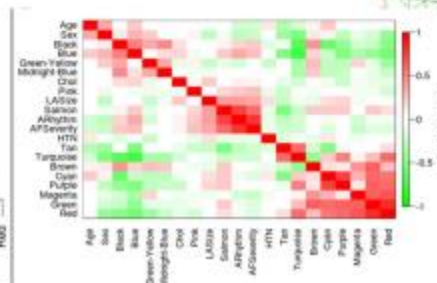
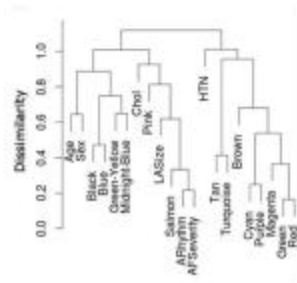
# Gene co-expression

Study patterns of relative gene expression across several conditions



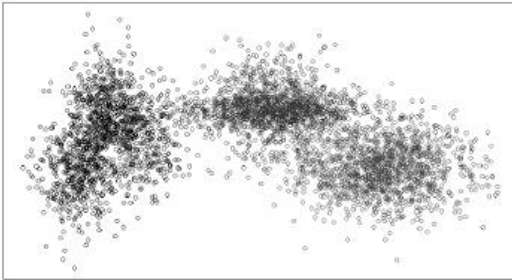
Enriched functions

- response to virus
- inflammatory response
- lymphocyte activation
- immune response
- macrophage activation
- mitosis/microtub
- cell adhesion



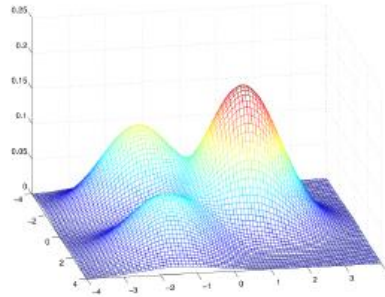
# Gene co-expression using mixture models

what we observe

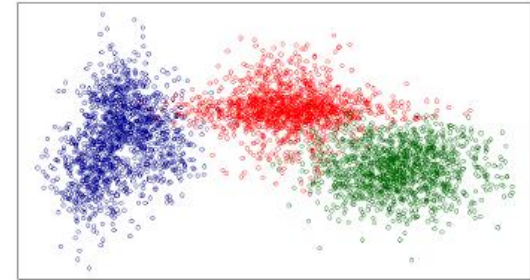


$Z = ?$

the model



the expected results



$Z : 1 = \bullet, 2 = \bullet, 3 = \bullet$

- Global modelling of the entire gene population
- Genes are assumed to come from different subpopulations
- Rigorous framework for parameters estimation and the choice of the number of subpopulations
- Each gene is assigned to a subpopulation with a probability.

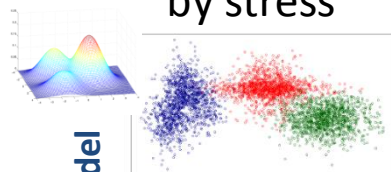
# Coexpression analyses of 18 stress responses

**CATdb**: 387 comparisons in 18 stress categories: 9 biotic and 9 abiotic

<http://tools.ips2.u-psud.fr/CATdb>

Matrix

{ genes x log-ratios }  
by stress



Gaussian Mixture Model

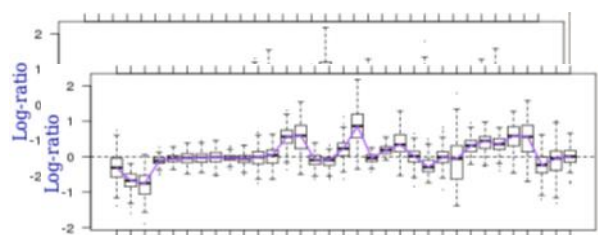


<http://tools.ips2.u-psud.fr//GEM2NET/>

Mathematical Criterion to select  
the cluster number (BIC)

Classification rule based on  
conditional probabilities

**~700 Clusters of coexpressed  
genes**

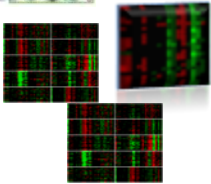


Stress category	Sample_nb	Gene_nb	Cluster_nb
Drought	17	8167	34
Gamma ray	25	5419	32
Heavy metals	45	10533	57
Nitrogen	46	13807	60
Oxidative stress	16	10027	52
Salt	15	5786	30
Temperature	45	11199	34
UV	7	7903	37
Other abiotic	8	3944	24
Fungi	21	9705	51
Biotrophic bacteria	40	11817	56
Necrotrophic bacteria	26	11030	50
Nematodes	10	7487	29
Oomycetes	14	5591	31
Rhodococcus	7	1965	13
Stifenia	6	1565	17
Virus	33	11685	54
Other biotic	6	3803	20



# From stress coexpression clusters to stress coregulation gene network

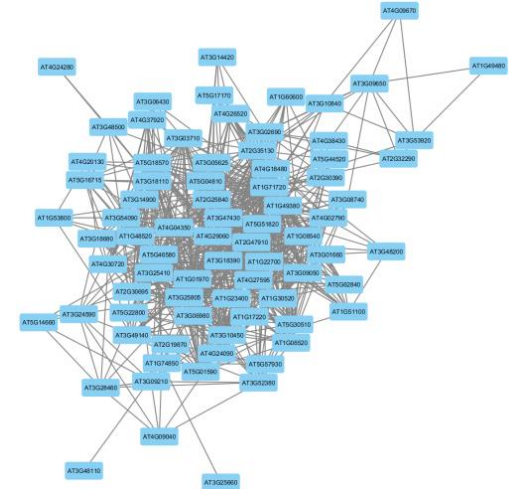
GEM2Net



Coexpression clusters for each category of stress

## Integration

Occurrence of pairs of coexpressed genes conserved in several stresses among the 18 considered stress categories



Coexpression network

1) Compared with random networks, only edges providing a  $FDR < 1\%$  were kept

**Filters**

2) Only genes involved in triangles were considered as co-regulated



Coregulation network

# Arabidopsis stress co-regulation network

4475 genes and 56487 co-regulation links

86% of the co-regulation links are supported by both types of stress

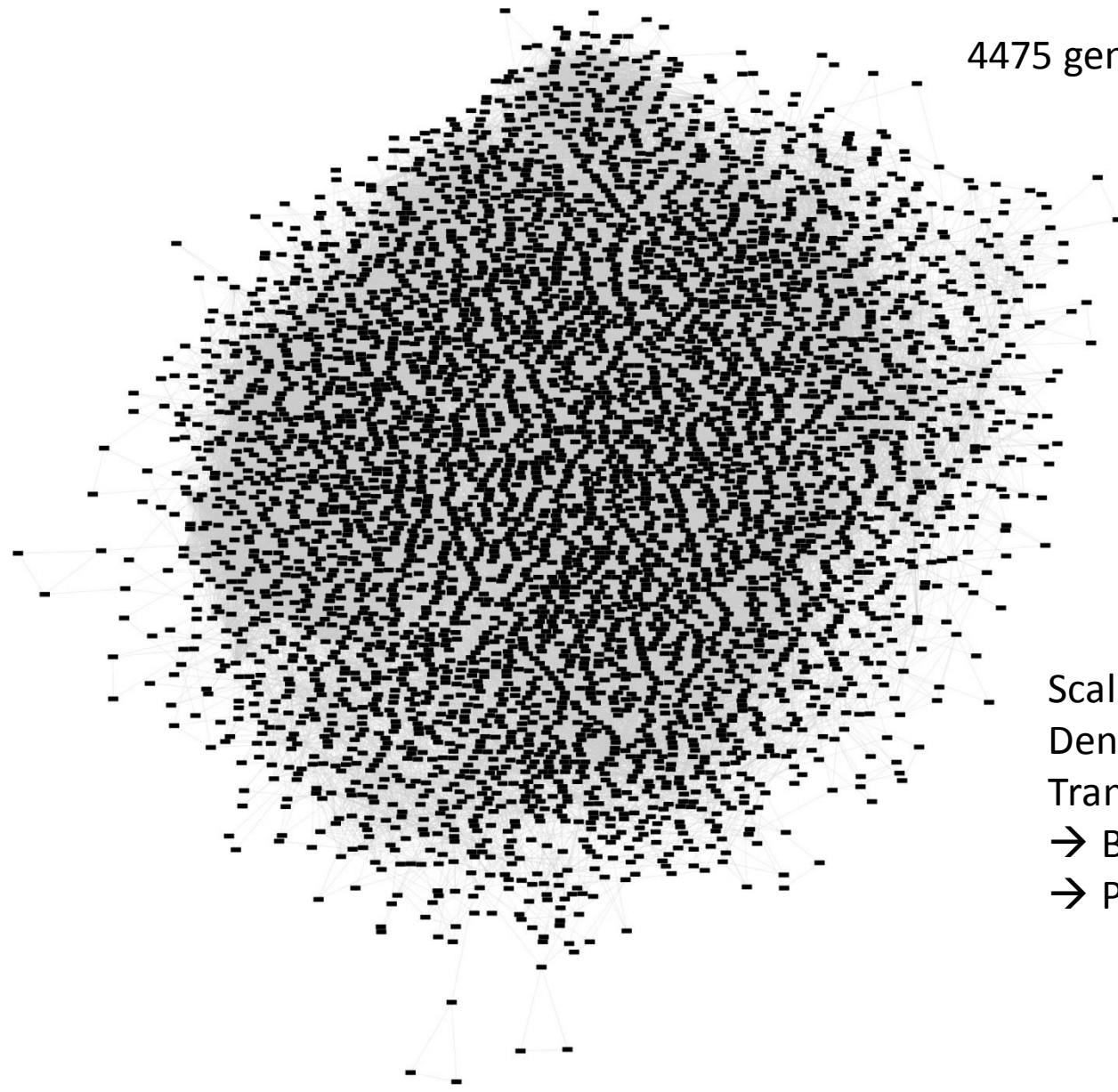
Scale-free network

Density = 0.006

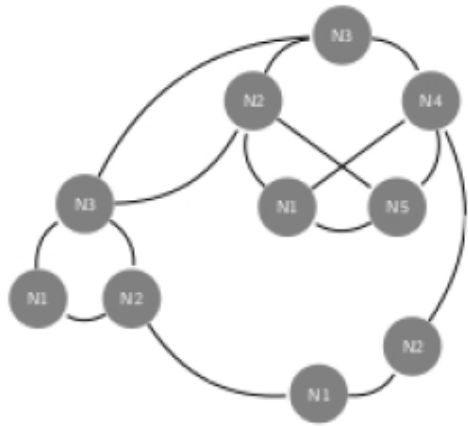
Transitivity=0.54

→ Biological network

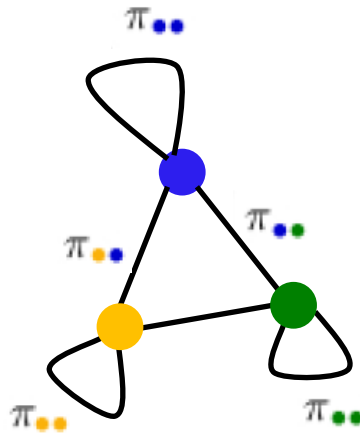
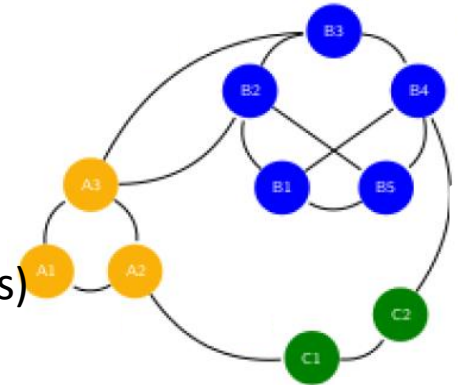
→ Presence of gene clusters



# Identification of communities within the network

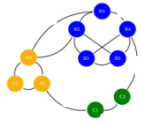


Communities are defined  
by their probabilities of interaction  
(within-community and between communities)



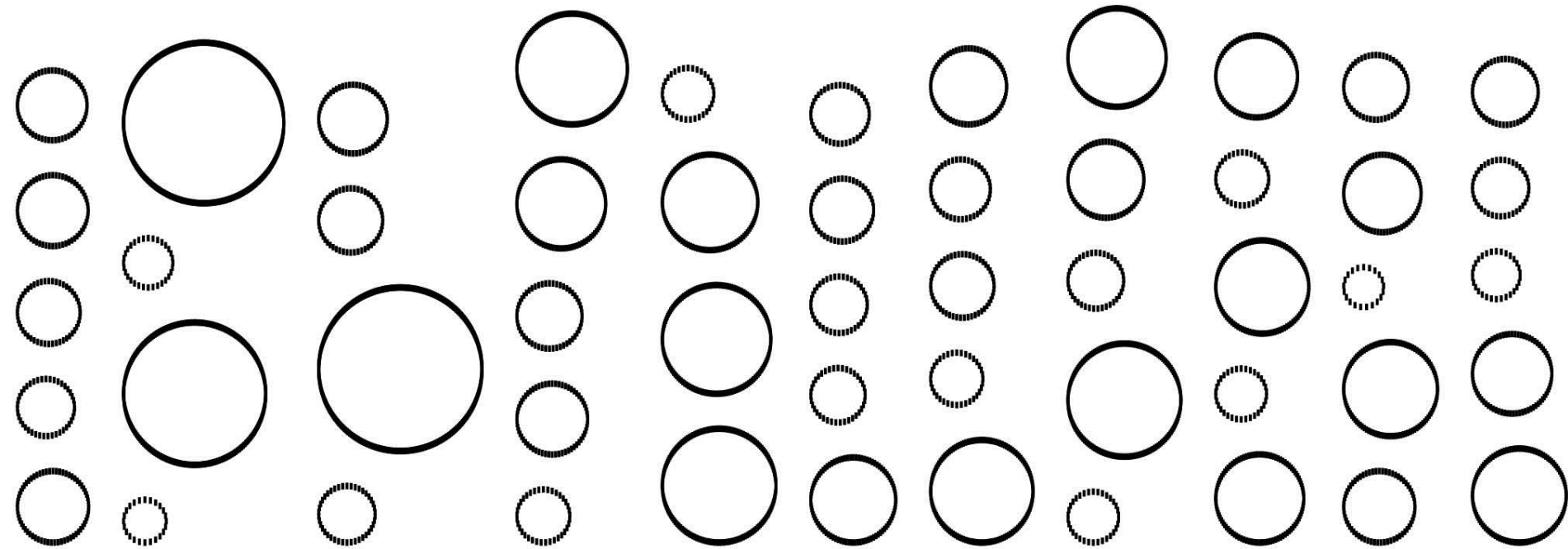
Synthetic representation:  
a node is a community  
An edge is a probability of interaction  
between two communities



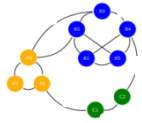


## Identification of gene communities within the network

52 communities of 21 to 351 genes

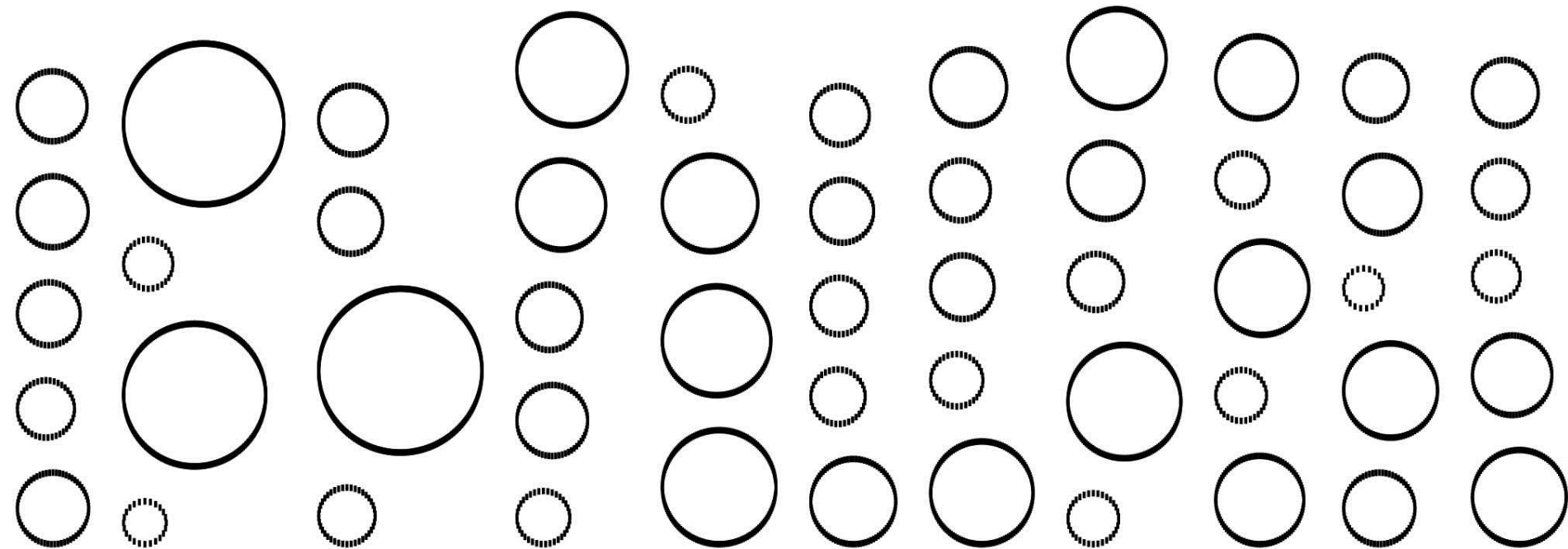


Stability of the communities?



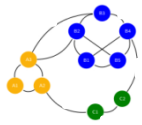
## Identification of gene communities within the network

52 communities of 21 to 351 genes



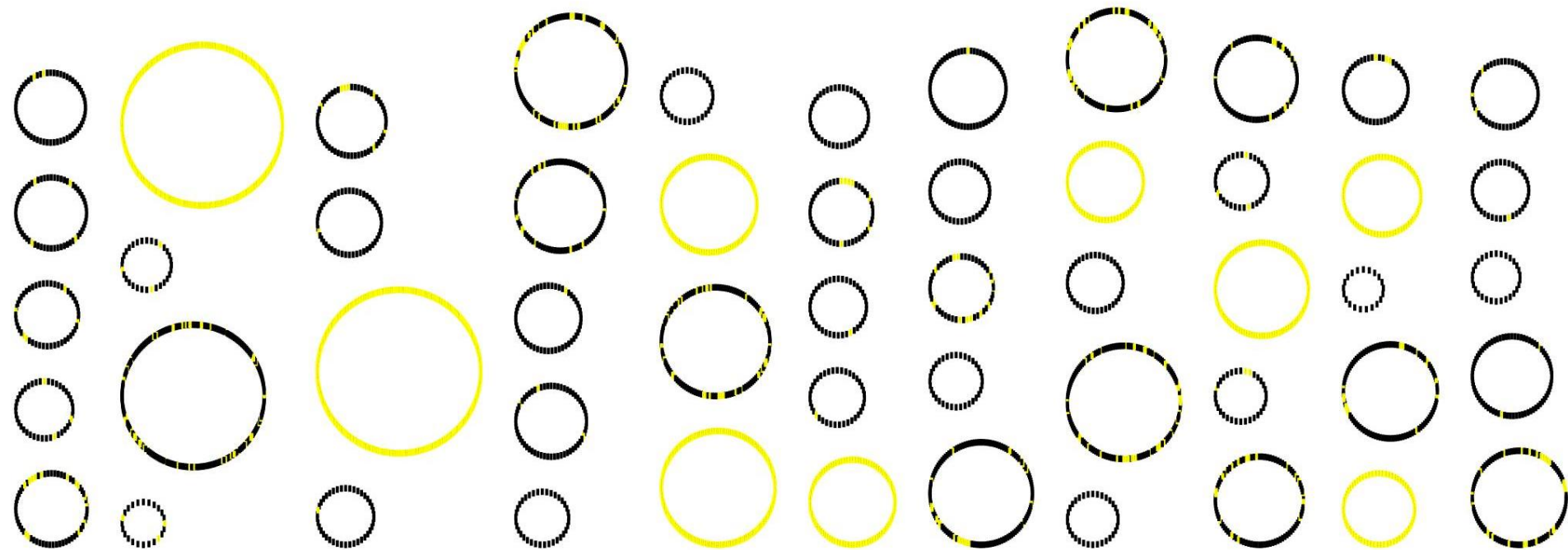
Stability measured with a **cross-validation procedure**

- For each stress category, create a network from the 17 others
- Find communities using mixture of graphs
- Comparison of these 18 results with the network built from all the categories



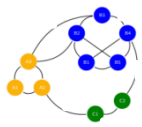
## Identification of the common response to stresses

2674 genes in 43 communities describe the common response to stresses



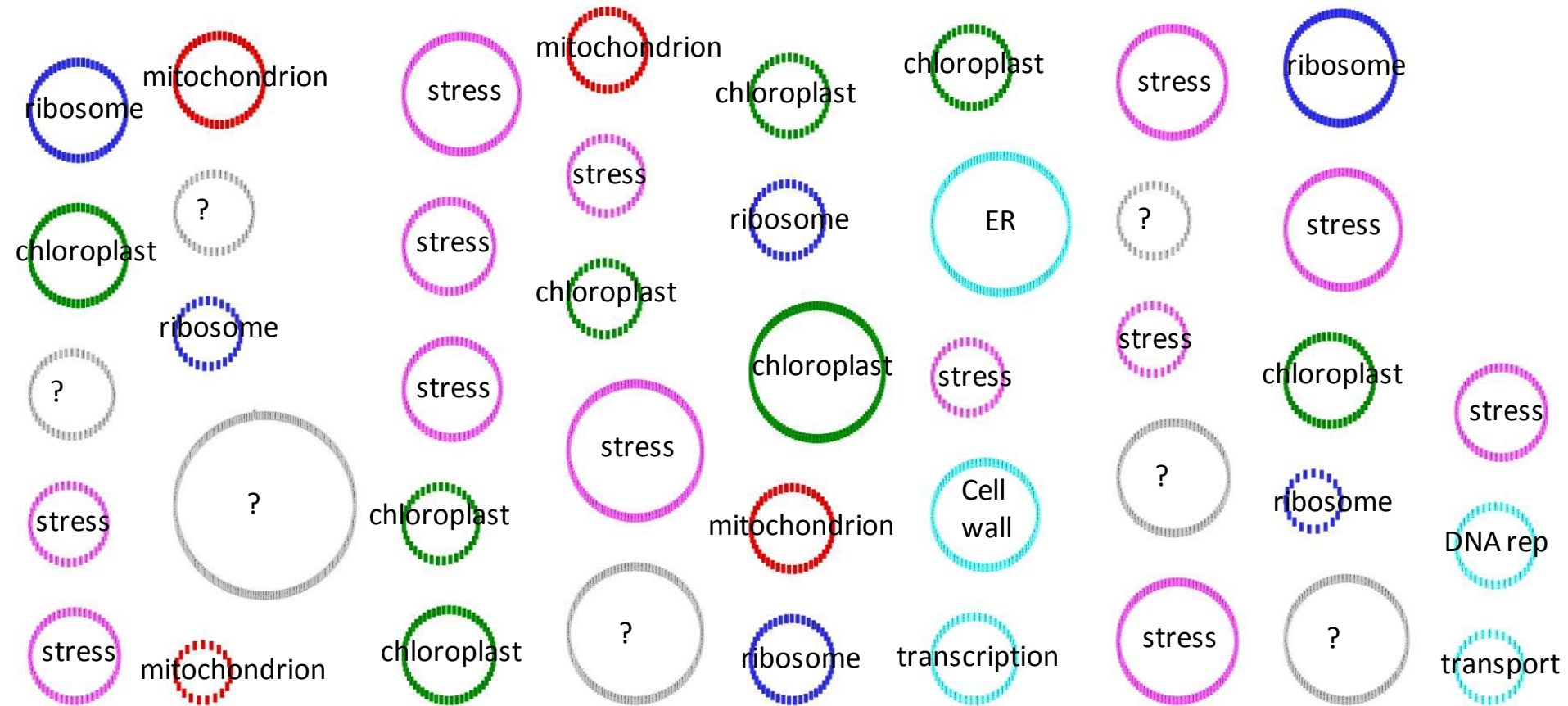
### Cross-validation procedure

- For each stress category, create a network from the 17 others
- Find communities using mixture of graphs
- Comparison of these 18 results with the network built from all the categories



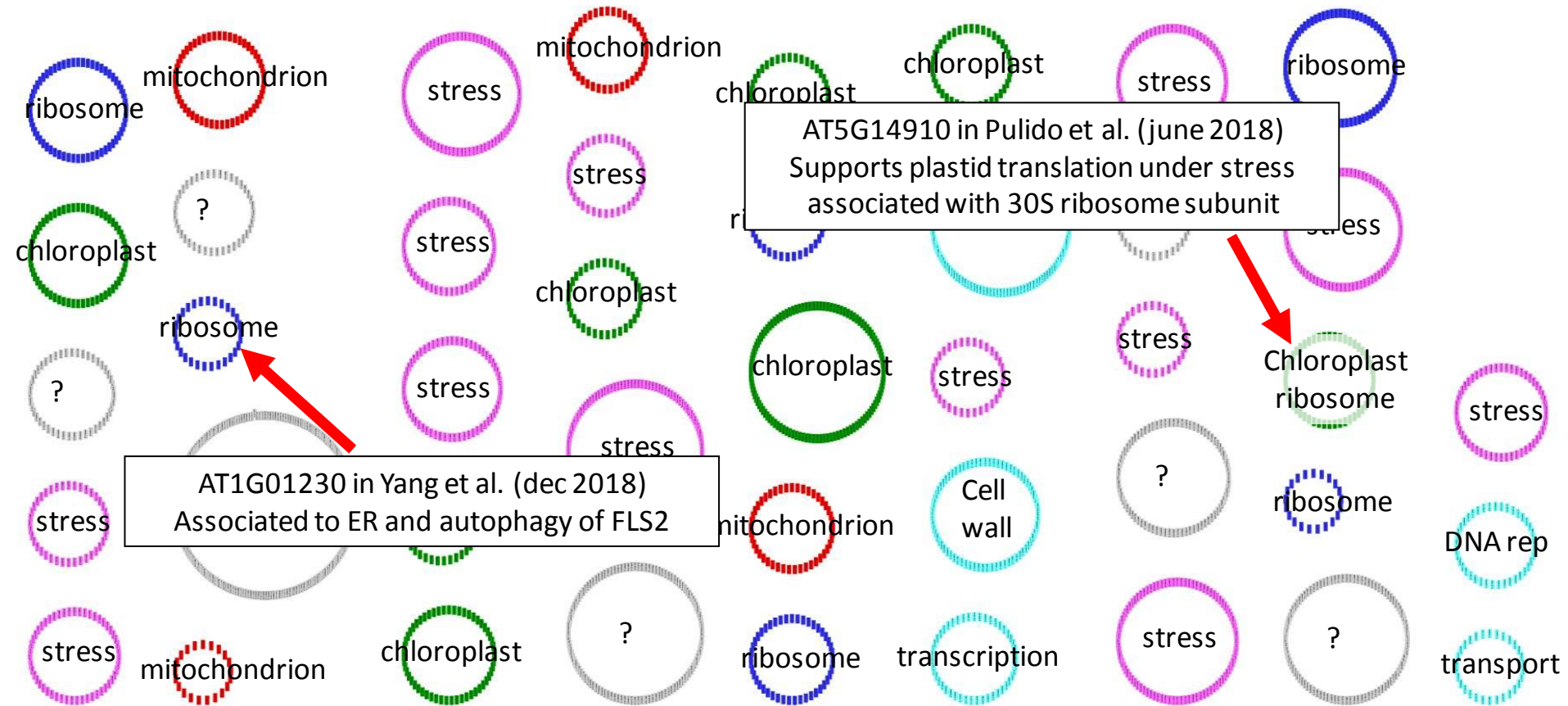
# Identification of the common response to stresses

## Most communities with GO enrichments



# Identification of the common response to stresses

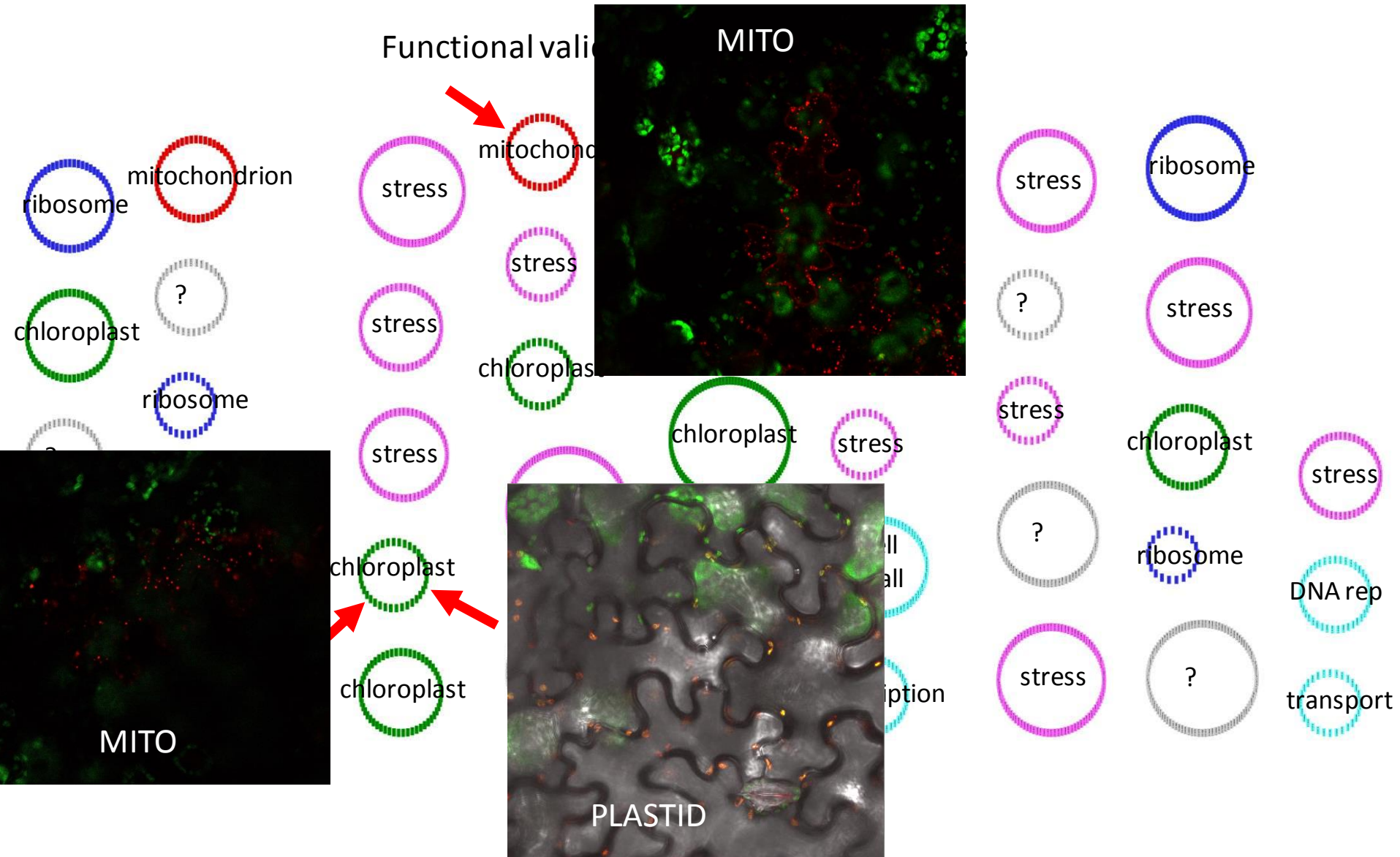
## Functional validation of the communities



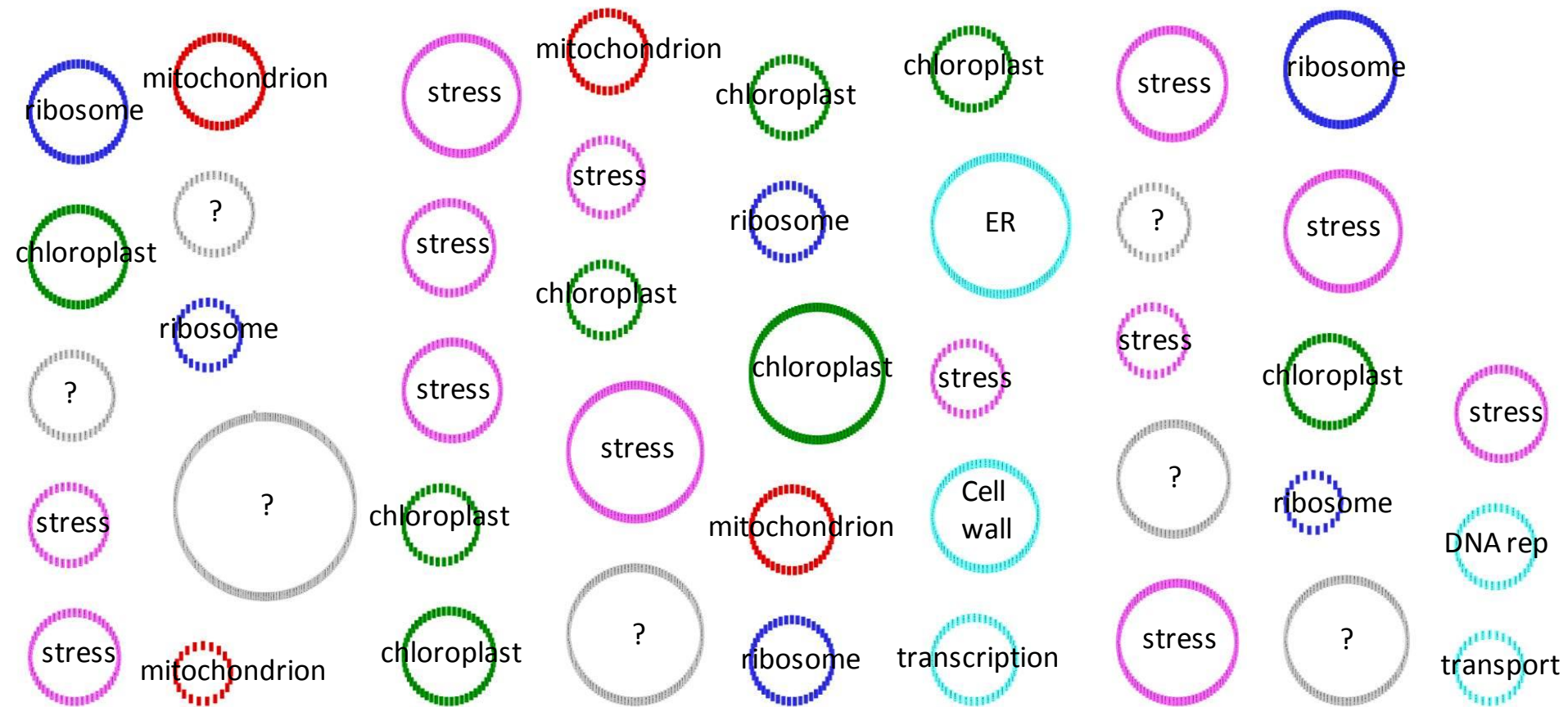




# Identification of the common response to stresses

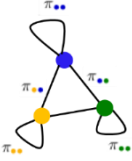


# Stresses impact the energy metabolism

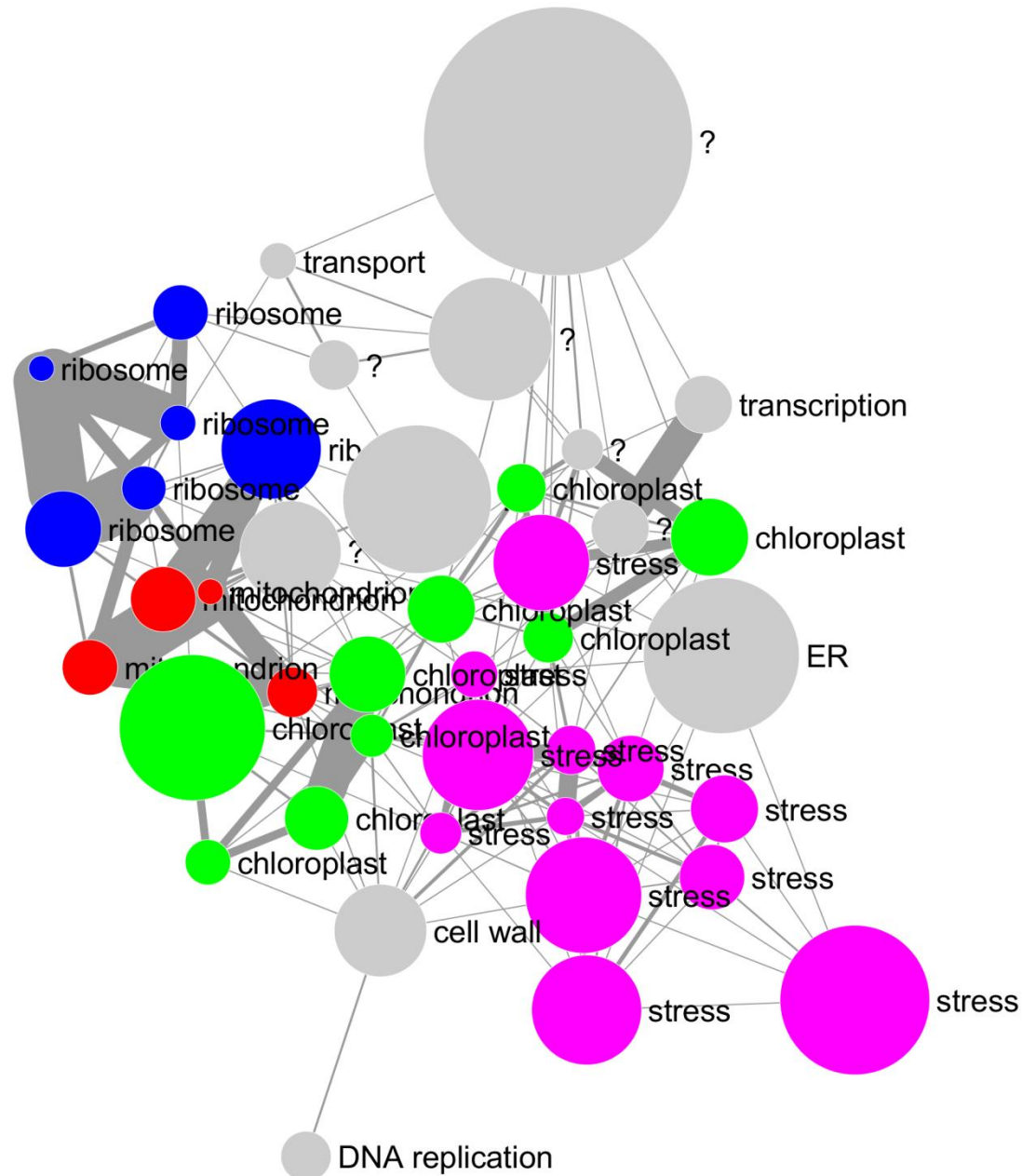


In addition to stress, predominance of chloroplast, mitochondrion and ribosome clusters

A common point to stresses is their impact on the energy metabolism

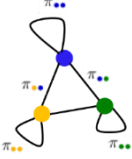


# The organisation of the stress response



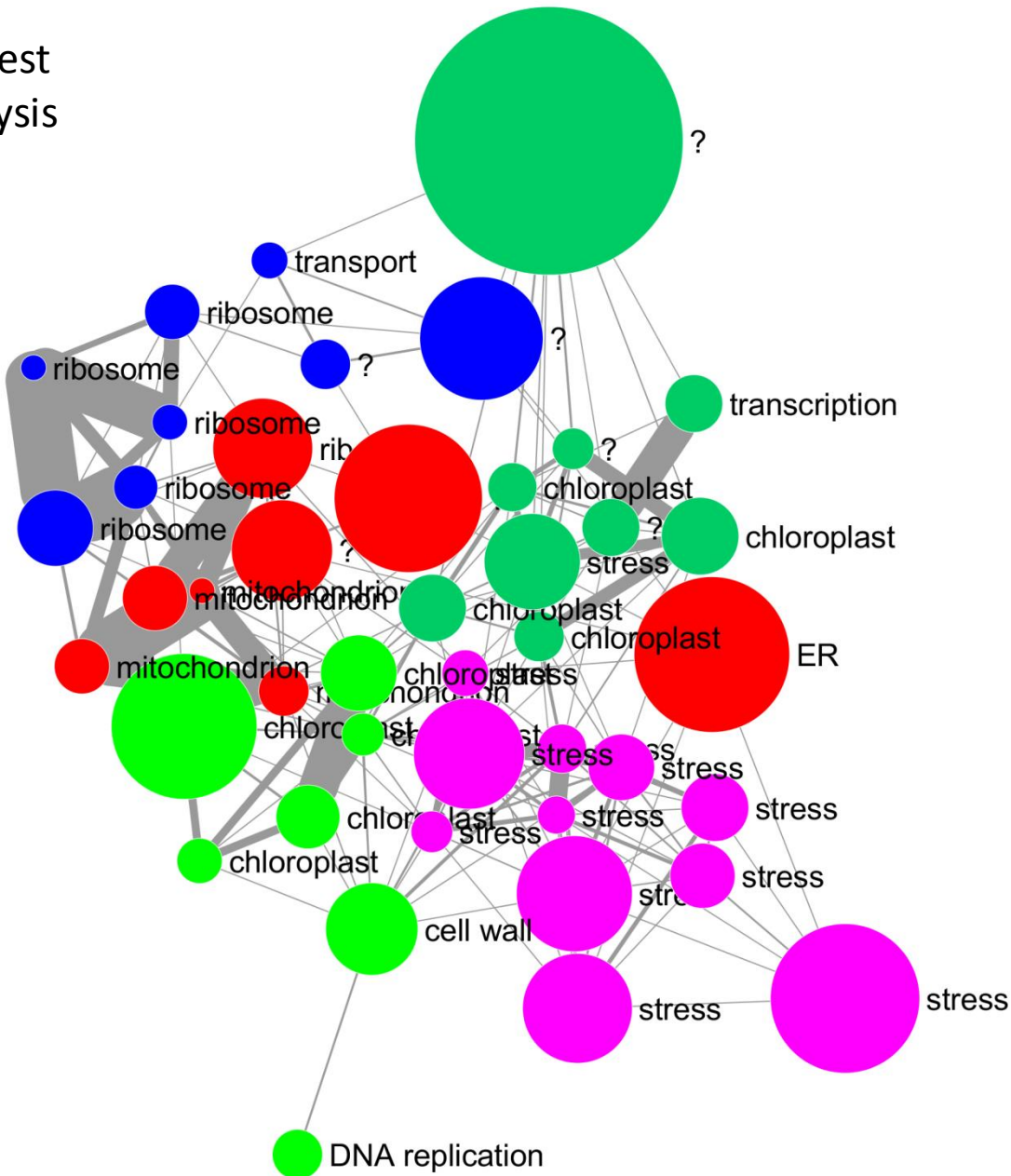
The functional modules are not randomly located and interacting in the network





+ a shortest path analysis

# The organisation of the stress response



The functional modules are not randomly located and interacting in the network

4 biological functions in 5 groups

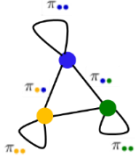




The diagram illustrates a complex network of biological processes. The nodes are categorized by color and size, representing different functional groups and their relative importance or frequency. The network is highly interconnected, with many edges (grey lines) connecting the nodes. The main clusters include:

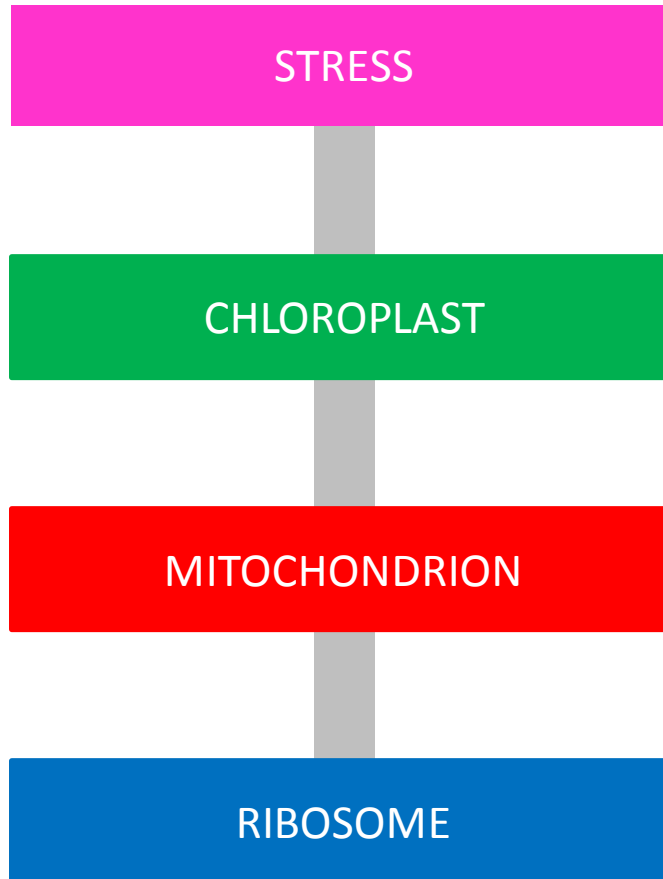
- Stress (Magenta):** A large cluster of nodes labeled "stress" on the left side of the diagram.
- Chloroplast (Green):** Multiple nodes labeled "chloroplast" are distributed across the top, bottom, and right-center of the network.
- ER/Mitochondrion (Red):** A cluster of nodes labeled "ER" and "mitochondrion" is located in the center-right area.
- Ribosome (Blue):** A cluster of nodes labeled "ribosome" is located on the far right side of the diagram.
- Other Processes:** Nodes for "transcription", "cell wall", "DNA replication", and "transport" are also present, often acting as bridges between the main clusters.

Many nodes are connected by thick grey lines, indicating strong or frequent interactions. Some nodes are marked with a question mark, suggesting areas of uncertainty or further investigation in the network analysis.

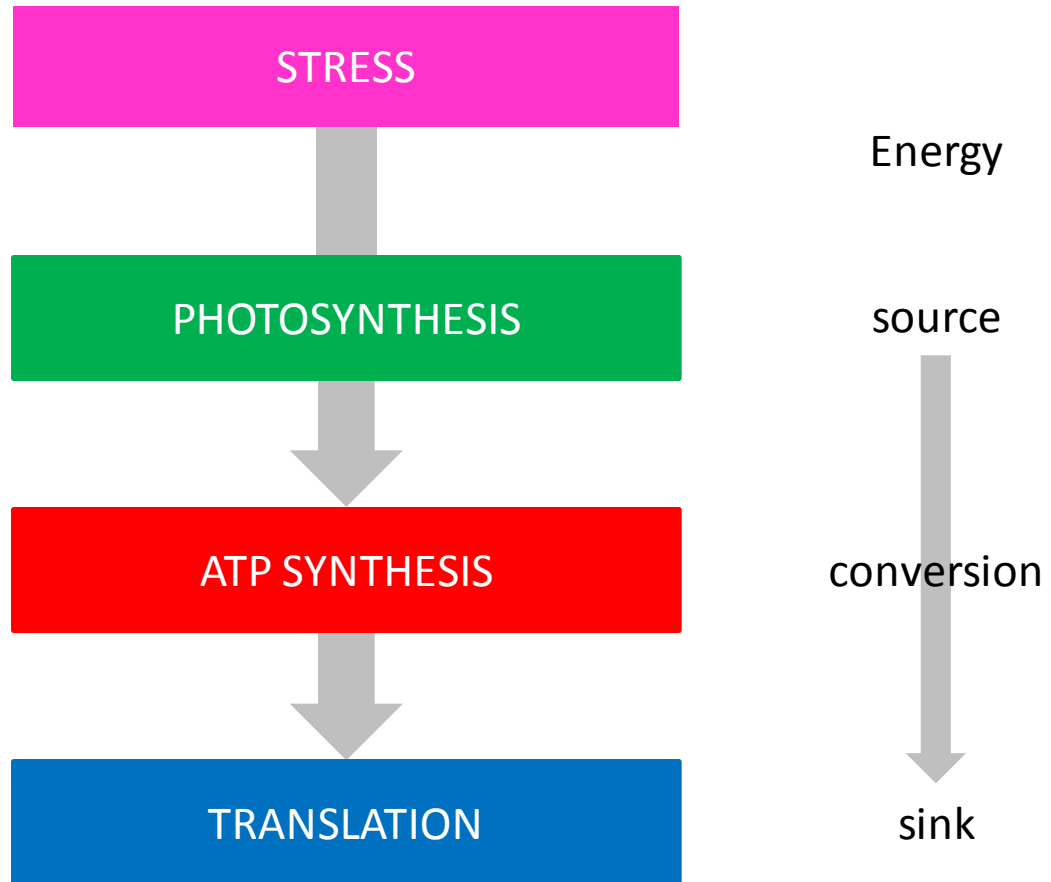


# The backbone of plant stress response

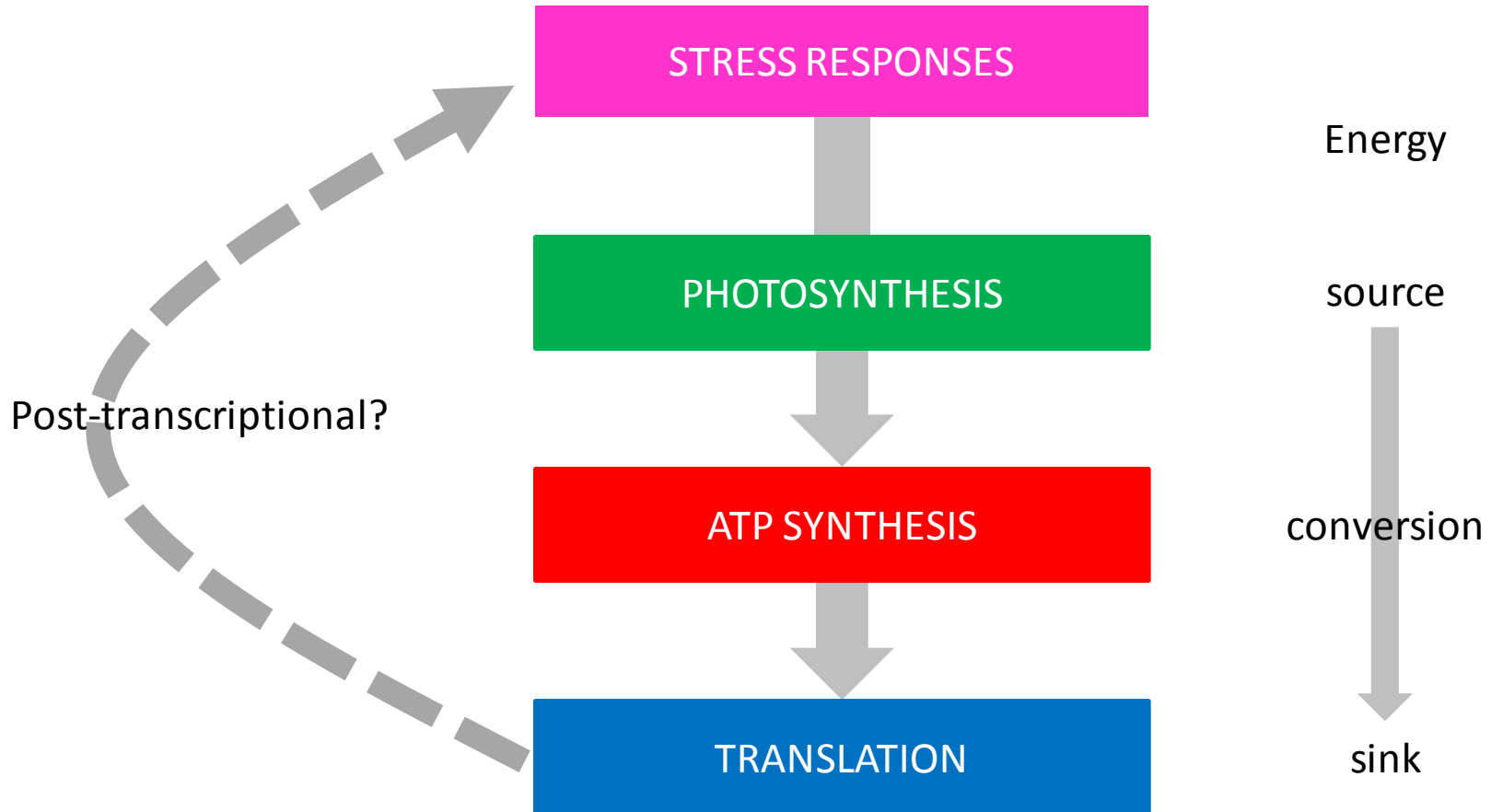
+ a shortest  
path analysis  
+ GO  
enrichment



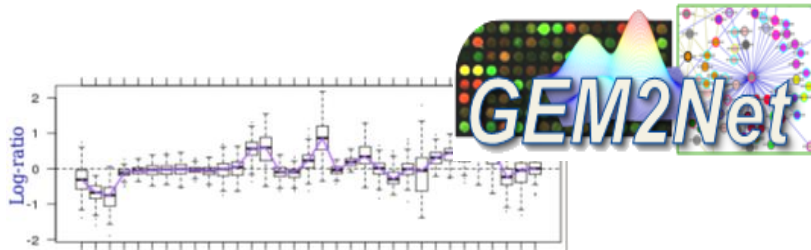
# The backbone of plant stress response



# The backbone of plant stress response



# Conclusions



Transcriptome  
meta-analysis



STRESS

PHOTOSYNTHESIS

ATP SYNTHESIS

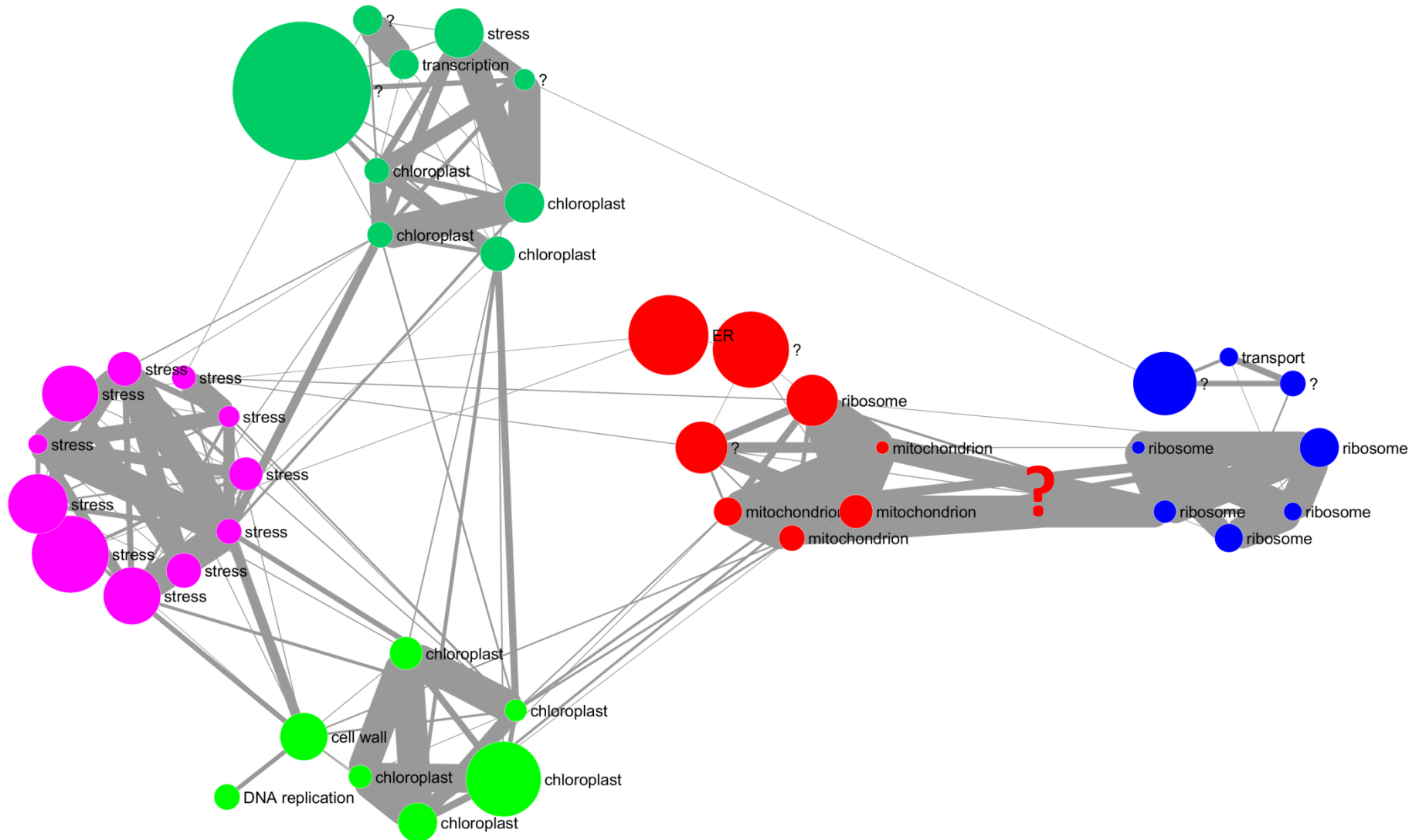
TRANSLATION

- Identification of a common response to stress
- Involvement of the energy organelles and translation
- Hierarchical organization along the energy gradient
- Statistical modelling: a relevant approach to formulate biological hypotheses



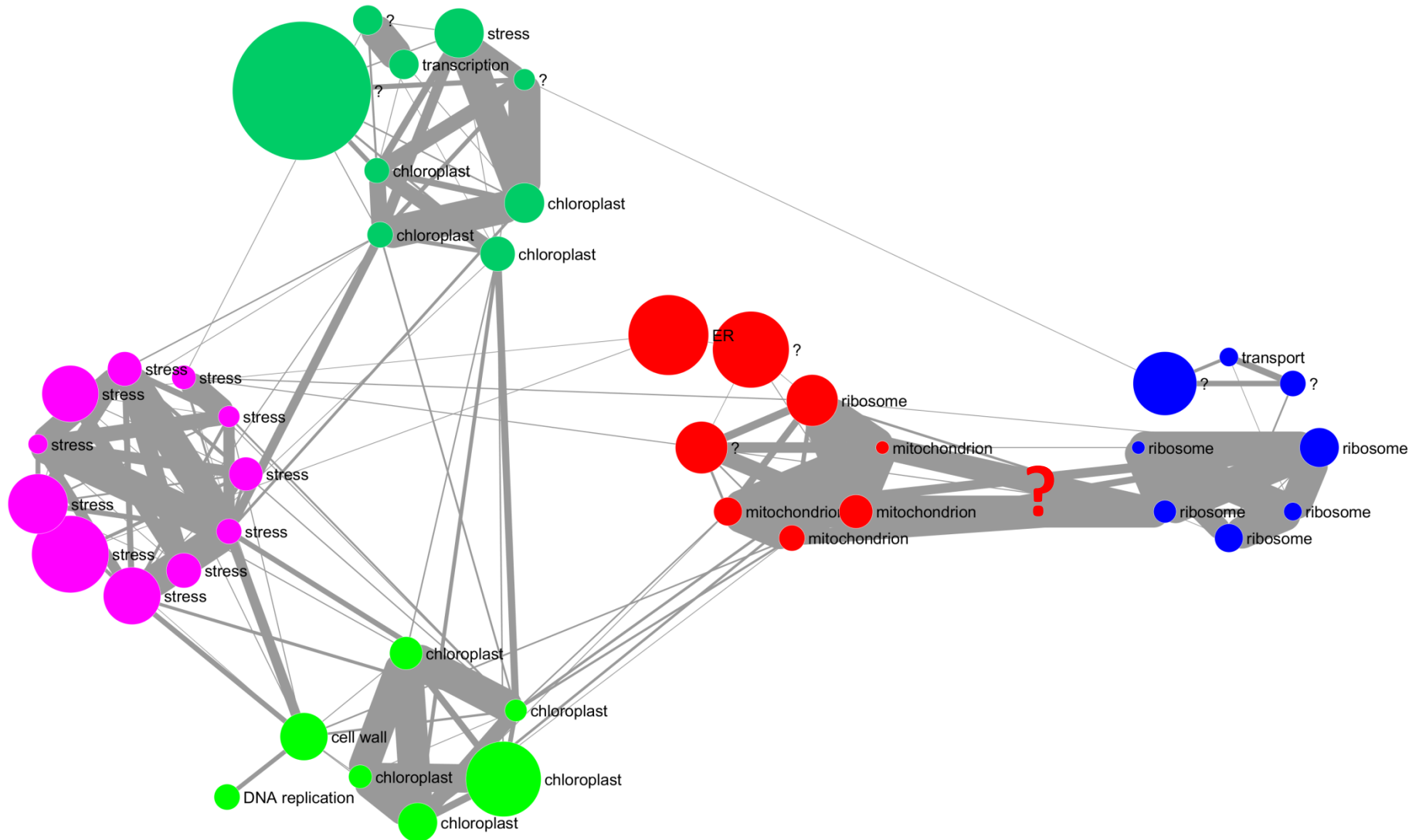
# Perspectives

## Network validation?



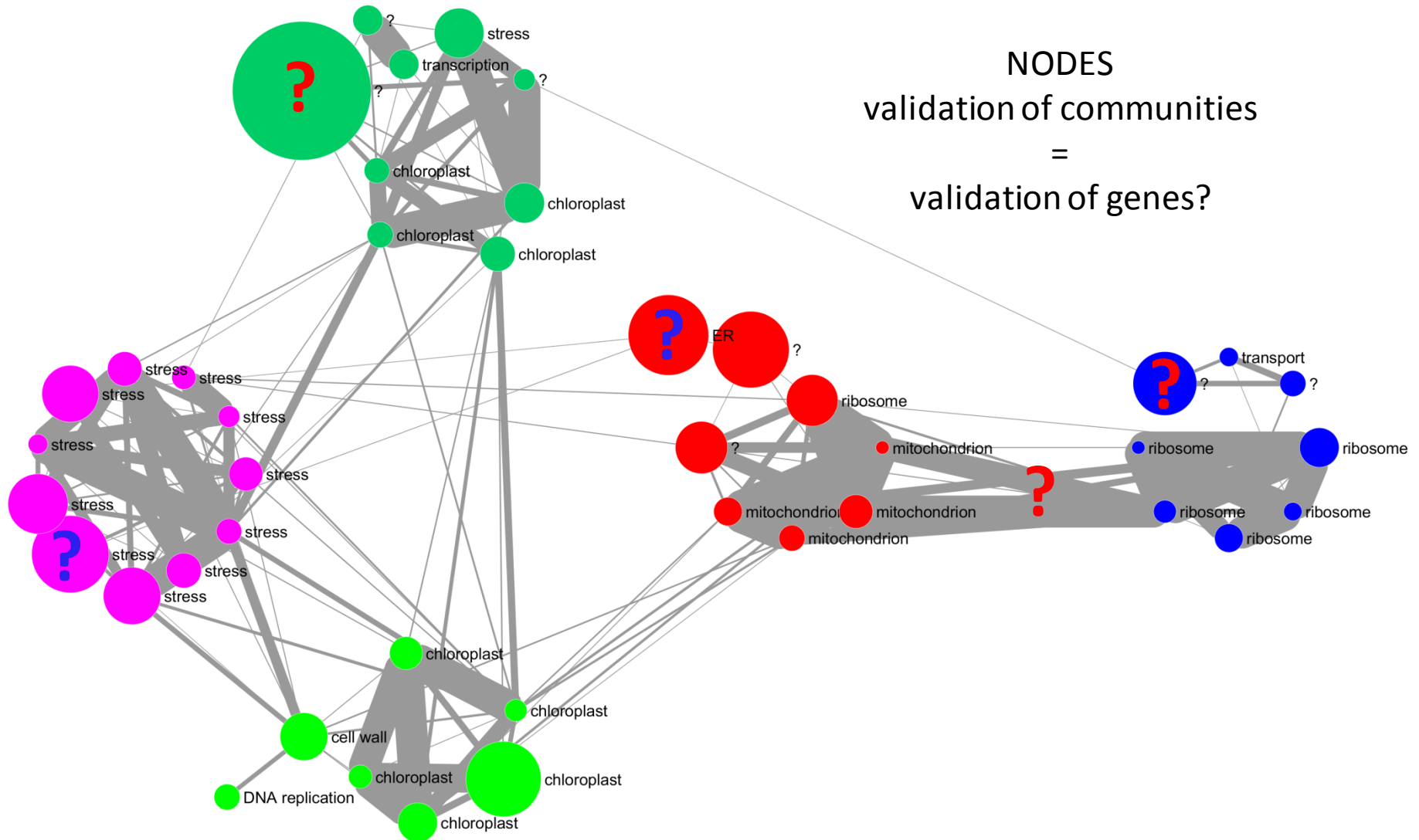
# Perspectives

## Biological Network validation?



## Perspectives

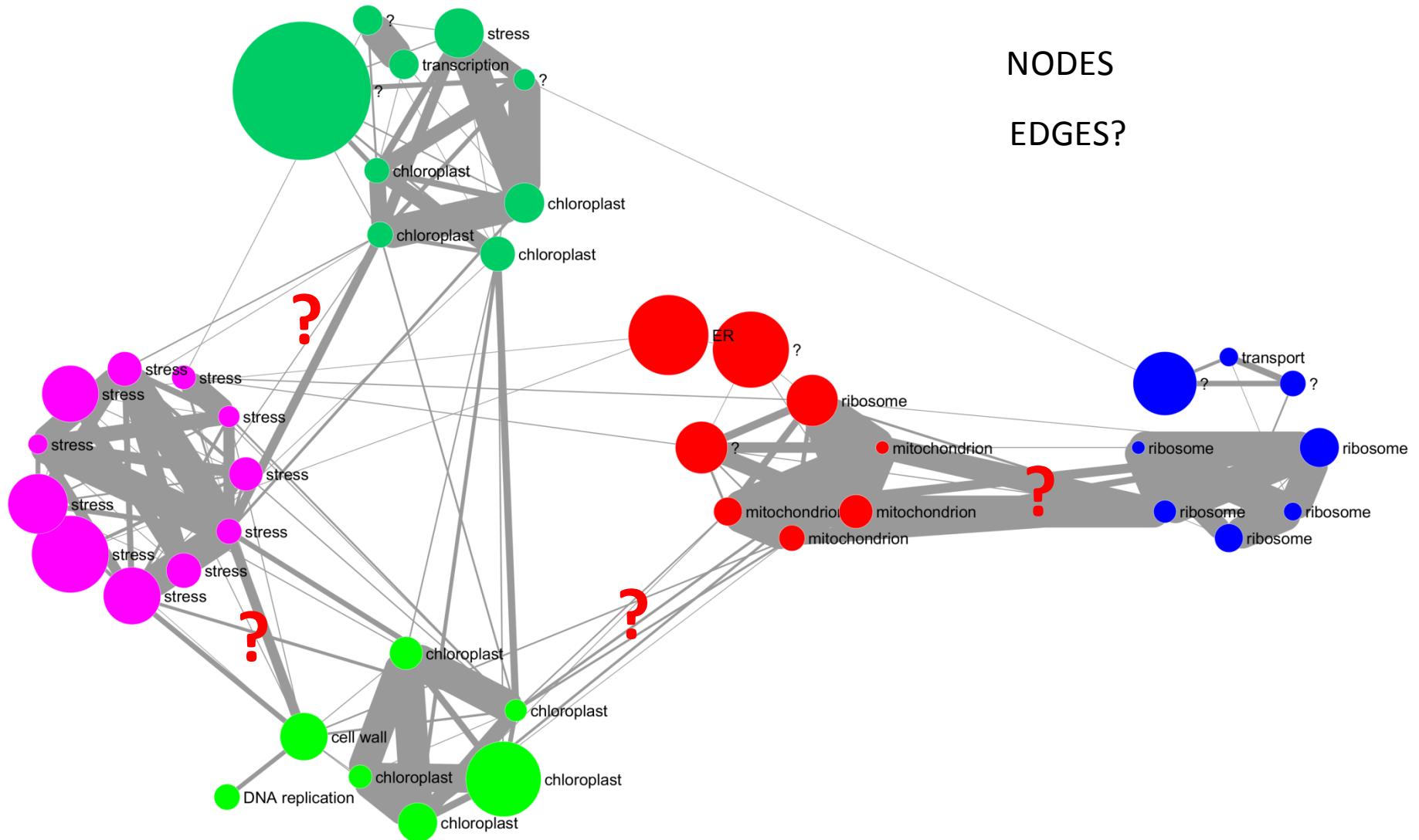
# Biological Network validation?



NODES  
validation of communities  
=  
validation of genes?

# Perspectives

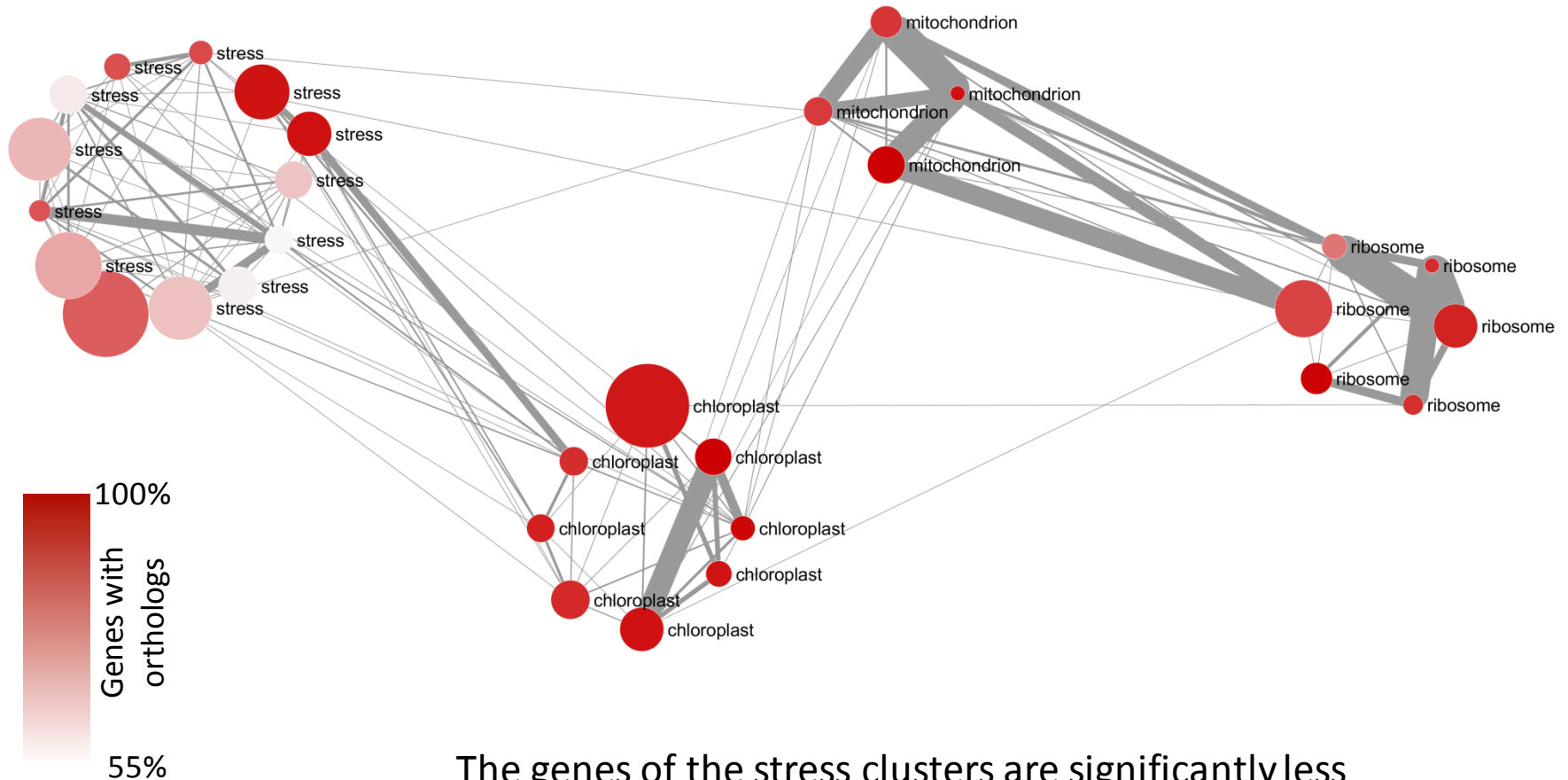
## Biological Network validation?





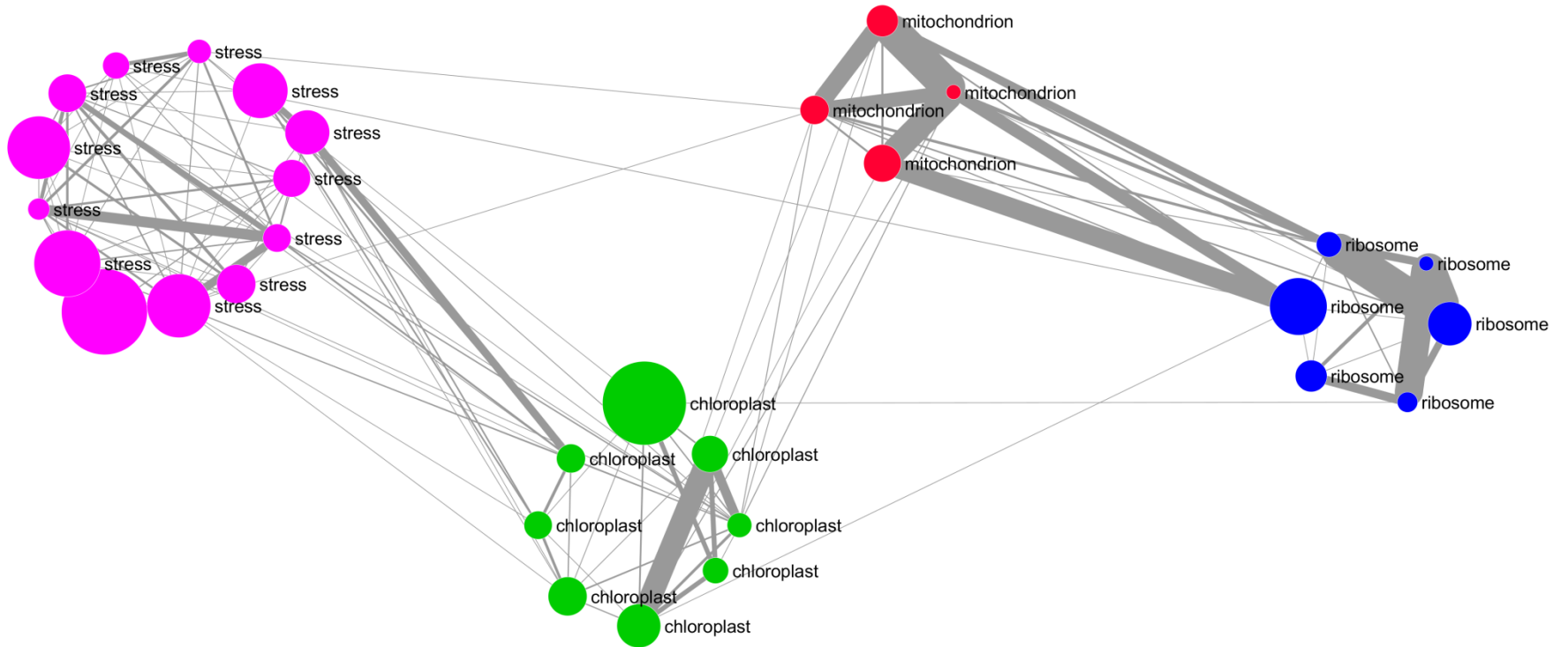
# The actors of the plant stress response are conserved

Based on Murat et al. 2017 using similarity + syteny to identify orthologs among 38 Angiosperm species + *C. reinhardtii* + *P. patens*



The genes of the stress clusters are significantly less conserved than the other genes

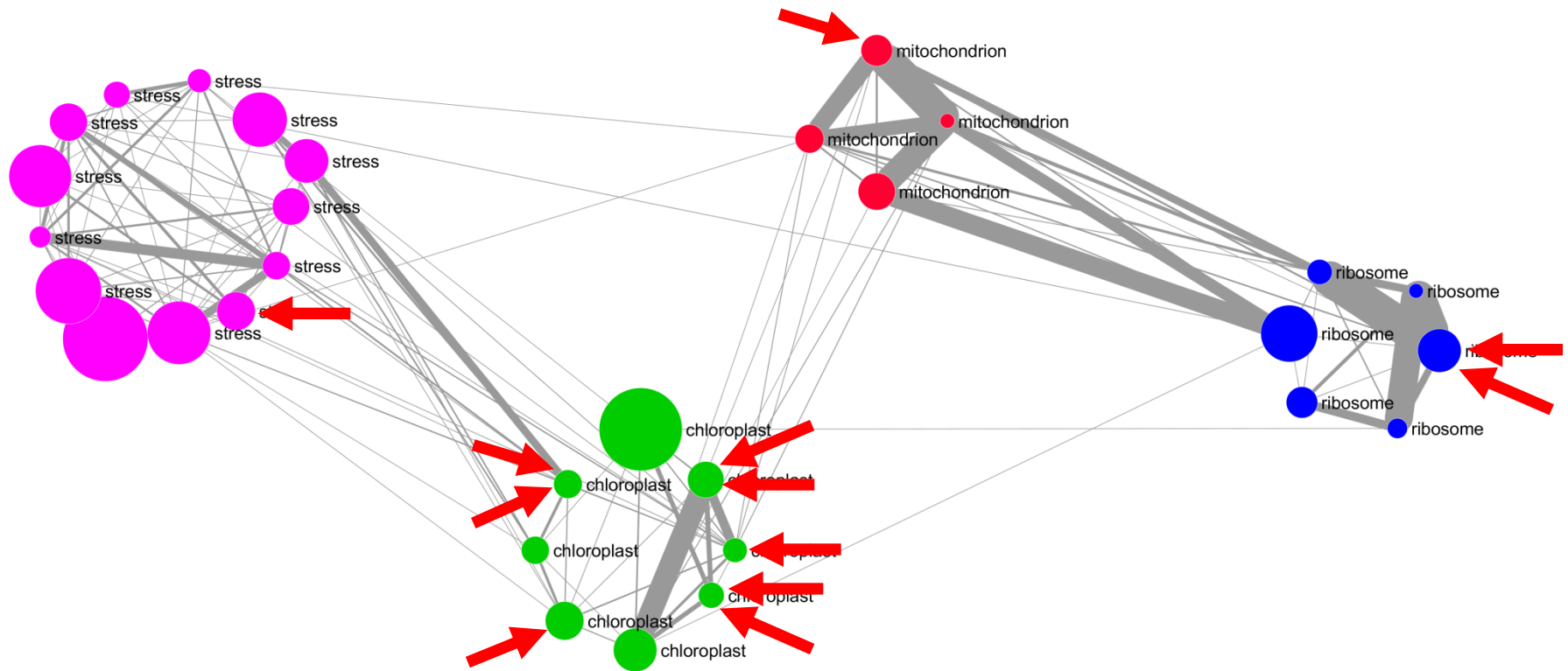
# Functional validation





# Functional validation

## T-DNA mutants for 12 genes of unknown function

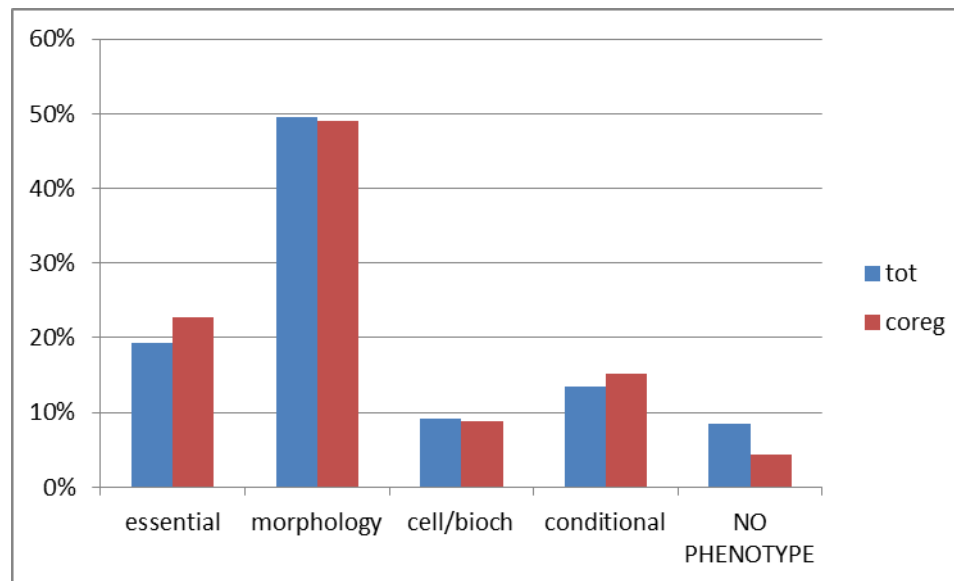


## phenotypes

Sur la base des Phenotypes du papier Ahmed et al. Nature comm 2018.

4344 AGI with curated phenotypes

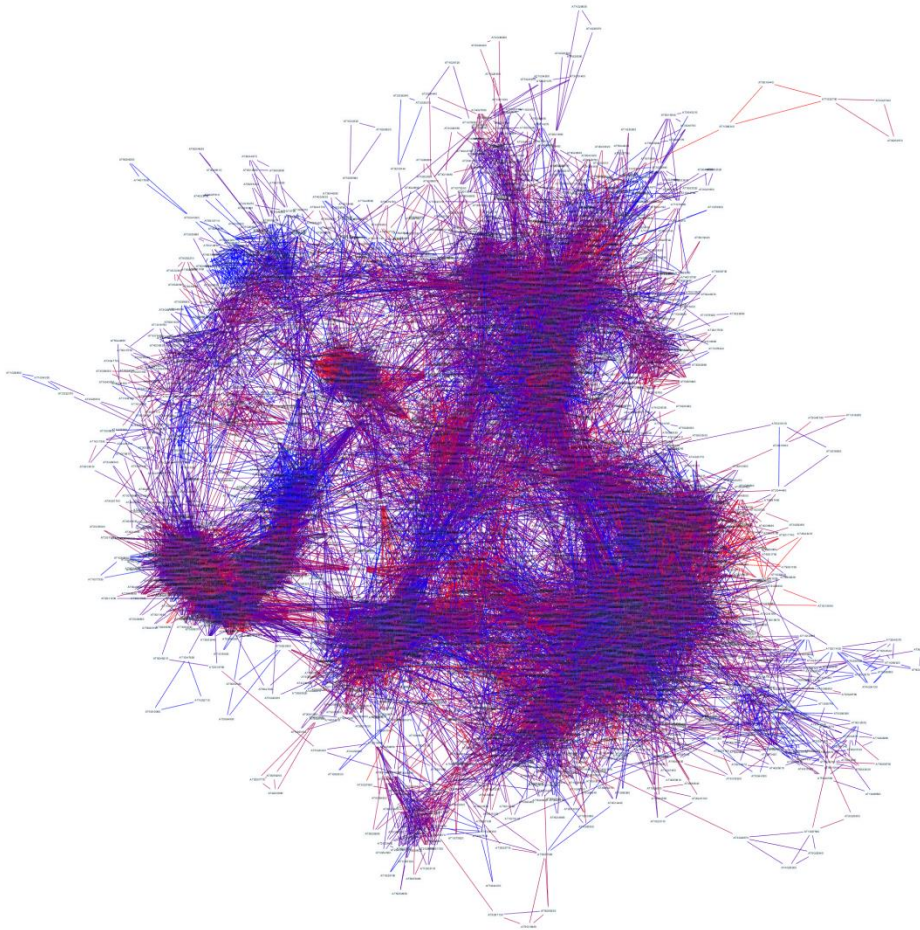
648 AGI dans le réseau après CV



Léger enrichissement en AGI essential et conditional

C'est difficile de faire la même analyse par communauté car seules 10 communautés ont >20 gènes avec phenotype

# Plant stress response



- 86% of the edges are supported by both types of stresses
- Homogenous distribution of edges

→ Large overlap between biotic and abiotic stress responses

Analysis of the gene modules identified by the mixture model