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

Etienne Delannoy, Rim Zaag, Guillem Rigaill, 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
Gene co-expression using mixture models

- 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

Matrix
{ genes x log-ratios }
by stress

Gaussian Mixture Model

Mathematical Criterion to select
the cluster number (BIC)

Classification rule based on
conditional probabilities

~700 Clusters of coexpressed
genes

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

Zaag R et al., Nucleic Acids Res, 2015
From stress coexpression clusters to stress coregulation gene network

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

Coregulation network

Filters

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

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

AT5G14910 in Pulido et al. (June 2018)
Supports plastid translation under stress associated with 30S ribosome subunit

AT1G01230 in Yang et al. (Dec 2018)
Associated to ER and autophagy of FLS2
Identification of the common response to stresses

Functional validation of the communities

Phenotyping of T-DNA mutants of 12 genes with unknown function
Identification of the common response to stresses

Functional validation of the communities

chloroplast

chloroplast

chloroplast

mitochondrion

mitochondrion

ribosome

ribosome

stress

stress

stress

stress

chloroplast

chloroplast

chloroplast

mitochondrion

mitochondrion

ribosome

ribosome

mitochondrion

mitochondrion

ribosome

ribosome

stress

stress

stress

stress

mitochondrion

mitochondrion

ribosome

ribosome

stress

stress

chloroplast

chloroplast

chloroplast

mitochondrion

mitochondrion

ribosome

ribosome

stress

stress

DNA rep

transport

Cell wall

Transcription

PLASTID

MITO
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.
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 backbone of plant stress response + a shortest path analysis
The backbone of plant stress response

+ a shortest path analysis
+ GO enrichment
The backbone of plant stress response

- Stress
- Photosynthesis
- ATP synthesis
- Translation

Energy source → Conversion → Sink
The backbone of plant stress response

STRESS RESPONSES

PHOTOSYNTHESIS

ATP SYNTHESIS

TRANSLATION

Post-transcriptional?

Energy

source

conversion

sink
Conclusions

- 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
Network validation?
Perspectives

Biological Network validation?
Perspectives

Biological Network validation?

NODES validation of communities = validation of genes?
Perspectives

Biological Network validation?

NODES
EDGES?
The actors of the plant stress response are conserved

Based on Murat et al. 2017 using similarity + synteny to identify orthologs among 38 Angiosperm species + C. reinhardii + 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
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