



Collège Sciences et technologies

## Projet de Modélisation qualitative et quantitative de la cellule :

Recherche de régulateurs clés à partir d'un  
réseau de co-expression

**Auteurs :**

Julie Blasquiz

Marc Mongy

Bordeaux  
06-11-2018  
L<sup>A</sup>T<sub>E</sub>X

## Table des matières

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Matériel et Méthodes</b>	<b>3</b>
2.1	Matériel : . . . . .	3
2.1.1	Jeu de données : . . . . .	3
2.1.2	Langage de programmation et logiciel utilisés : . . . . .	3
2.2	Méthode : . . . . .	4
2.2.1	Filtrage des données : . . . . .	4
2.2.2	Construction du réseau : . . . . .	4
2.2.3	Sélection des ontologies : . . . . .	4
<b>3</b>	<b>Résultats</b>	<b>6</b>
3.1	Filtrage des données : . . . . .	6
3.2	Réseau de co-expression : . . . . .	6
<b>4</b>	<b>Discussion</b>	<b>9</b>
	<b>Annexes :</b>	<b>10</b>
<b>A</b>	<b>Origine des données de transcriptome</b>	<b>10</b>
<b>B</b>	<b>Autres sources de données</b>	<b>11</b>
<b>C</b>	<b>Script R du filtrage des données</b>	<b>12</b>
<b>D</b>	<b>Gènes régulateurs clés de la réponse au nitrate dans les racines d'<i>Arabidopsis Thaliana</i></b>	<b>14</b>

# 1 Introduction

De tous temps, l'homme a eu recours aux plantes, que ce soit pour se nourrir, se soigner, et même se vêtir, ou encore se construire des bâtiments. Toutefois, l'agriculture a connu un développement intensif à partir des années 1960, avec un usage intensif d'engrais tel que les nitrates. Si l'impact des nitrates sur la santé humaine est relativement négligeable, l'impact de leur usage s'est fait vivement ressentir sur l'environnement, avec une rupture de l'équilibre chimique des sols, un épuisement de ceux-ci, en raison de la sur-spécialisation intensive de certaines régions (avec une réduction de la diversité des espèces cultivées) et de l'épuisement progressif de ceux-ci, mais également avec les rejets des excès d'engrais dans les mers et océans par lessivage, aboutissant régulièrement aux phénomènes dits de "marée verte" avec une prolifération anormale d'algues vertes sur les littoraux, ce phénomène ayant pour une effet une asphyxie de la faune marine. Ces algues ont également une toxicité propre, d'une part par une production potentielle de toxines (pour certaines algues), mais surtout par le dégagement de sulfure d'hydrogène lors de leur pourrissement, ce phénomène étant régulièrement responsable d'accidents, notamment sur les côtes bretonnes. Toutefois, ce phénomène n'est qu'un des aspects les plus visibles du déséquilibre environnemental qui perdure depuis l'essor des activités agricoles (même si dans le cas des marées vertes, les phosphates seraient plus à incriminer) [1]. Aujourd'hui, avec l'émergence des biocarburants, ainsi que d'une population terrestre toujours plus nombreuse qu'il faudra nourrir, le contrôle du rendement des cultures, et surtout, une utilisation raisonnée des engrais et produits phytosanitaires deviendra une pierre angulaire du développement agricole moderne, s'il ne l'est pas déjà. Le travail présenté ici a pour objectif d'identifier des gènes de régulation de la réponse d'une plante, *Arabidopsis thaliana* (plante ubiquitaire, résistante à un grand nombre de conditions difficiles, dont le génome a été complètement annoté, utilisée comme modèle pour l'étude des plantes) aux engrais azotés, en explorant un réseau PPI (Protein-Protein Interaction).

## 2 Matériel et Méthodes

### 2.1 Matériel :

#### 2.1.1 Jeu de données :

Le jeu de données utilisé dans le cadre de ce projet est un ensemble de fichiers textes dont la description de chacun d'entre eux est donnée ci-dessous :

- P\_EXPR.txt : Chaque ligne de cette table donne pour un gène d'*Arabidopsis Thaliana* (AT) les valeurs d'expression observées dans 146 mesures. Ces mesures ont été collectées à partir de 10 expériences disponibles dans différentes bases de données (cf. Annexe A). Les expériences ont toutes utilisé le microarray Affymetrix ATH1. La première colonne de la table donne le nom des locus des gènes. Les données ont déjà été nettoyées et normalisées par la méthode Robust Multiarray Analysis (RMA).
- P\_GOI.txt : cette table donne la liste des noms des locus des gènes d'AT connus pour répondre à une stimulation à l'azote (cf. Annexe B). On limitera l'étude à cette liste (GOI = Gene Of Interest). La table contient une seule colonne «id».
- P\_PPI.txt : cette table donne la liste de paires de gènes d'AT codant pour des protéines en interaction (cf. Annexe B). La table contient 3 colonnes «id1», «id2» et «type». Les colonnes «id1» et «id2» contiennent les identifiants des locus des gènes et la colonne «type» contient la valeur «PPI».
- P\_TF\_LABEL.txt : cette table donne une liste des facteurs de transcription d'AT (cf. Annexe B). La table contient 3 colonnes «id», «TF\_family» et «gen\_type». La colonne «id» contient les identifiants des locus des gènes, la colonne «TF\_family» donne le nom de la famille de facteur de transcription à laquelle appartient ce gène et la colonne «gen\_type» contient la valeur «TF».

Ces fichiers ont été préparés à partir des données aimablement mises à disposition par Tomás C. Moyano, Elena A. Vidal, Orlando Contreras-López et Rodrigo A. Gutiérrez du «Plant Systems Biology Lab» de l'université pontificale catholique du Chili. Les données proviennent du laboratoire et de différentes bases de données publiques (cf Annexe B).

#### 2.1.2 Langage de programmation et logiciel utilisés :

Les fichiers textes décrits précédemment ont été traités via un script rédigé en R. La construction du réseau de co-expression a été réalisée via le logiciel Cytoscape.

## 2.2 Méthode :

### 2.2.1 Filtrage des données :

Les données d'expression brutes (P\_EXPR.txt) comprenant l'expression de gènes courants n'ayant rien à voir avec le métabolisme de l'azote, une étape de filtrage a été nécessaire. Le filtrage des données a été effectué par le biais de la liste de gènes d'intérêt (Genes Of Interest, ou GOI) mentionnée plus haut (annexe B), afin de ne conserver que l'expression des gènes impliqués dans le métabolisme de l'azote. Les corrélations de co-expression pour chaque paire de gènes ont été ensuite calculées, et les co-expressions pour lesquelles le coefficient de corrélation était inférieur à 0.6 (valeur la plus couramment utilisée) en valeur absolue ont été considérées comme non-significatives. Une étape de suppression des doublons (paires de gènes symétriques) a été ensuite ajoutée. Ainsi, une liste de co-expressions a été obtenue. Afin de tracer un graphe PPI (Protein-Protein Interaction), la liste de paires de gènes d'AT codant pour des protéines en interaction a été chargée. La liste de co-expression, utilisée précédemment, a été utilisée comme filtre afin de retirer les interactions non pertinentes dans le contexte de notre travail. Une étape de suppression des doublons (interactions symétriques) a été également ajoutée. Plus de détails se situent dans le script R (figurant en Annexe C).

### 2.2.2 Construction du réseau :

Le réseau brut a été fragmenté en utilisant l'algorithme GLay du plugin clusterMaker. A partir de là, 6 clusters ont été obtenus. Les clusters ont été annotés par l'utilisation des plugins Cytoscape BiNGO (les réglages ont été conservés par défaut, seule l'espèce a été changée pour *Arabidopsis Thaliana*) et AutoAnnotate, qui propose pour l'annotation plusieurs algorithmes de clustering. Les 2 clusters les plus massifs, en raison de leur taille, ont été annotés en utilisant l'algorithme GLay, les autres clusters ont été annotés en utilisant l'algorithme MCL. Dans les 2 cas, les réglages (du plugin AutoAnnotate) par défaut ont été conservés. Les figures ont toutefois été finalisées sous le portail web draw.io, car si AutoAnnotate dispose effectivement d'une fonction de dessin permettant de délimiter des groupements de gènes exprimés, il ne prend absolument pas en compte la possibilité de présence de gènes dans plusieurs groupements.

### 2.2.3 Sélection des ontologies :

Afin de sélectionner les ontologies les plus pertinentes, il convient d'examiner le réseau métabolique du métabolisme de l'azote, concernant *Arabidopsis Thaliana*.

On y observe que le métabolisme des nitrates est en lien avec :

- la synthèse d'intermédiaires tels que les nitrites
- la synthèse du glutamate et de la glutamine

De fait, il est possible de dégager des annotations GO en rapport.

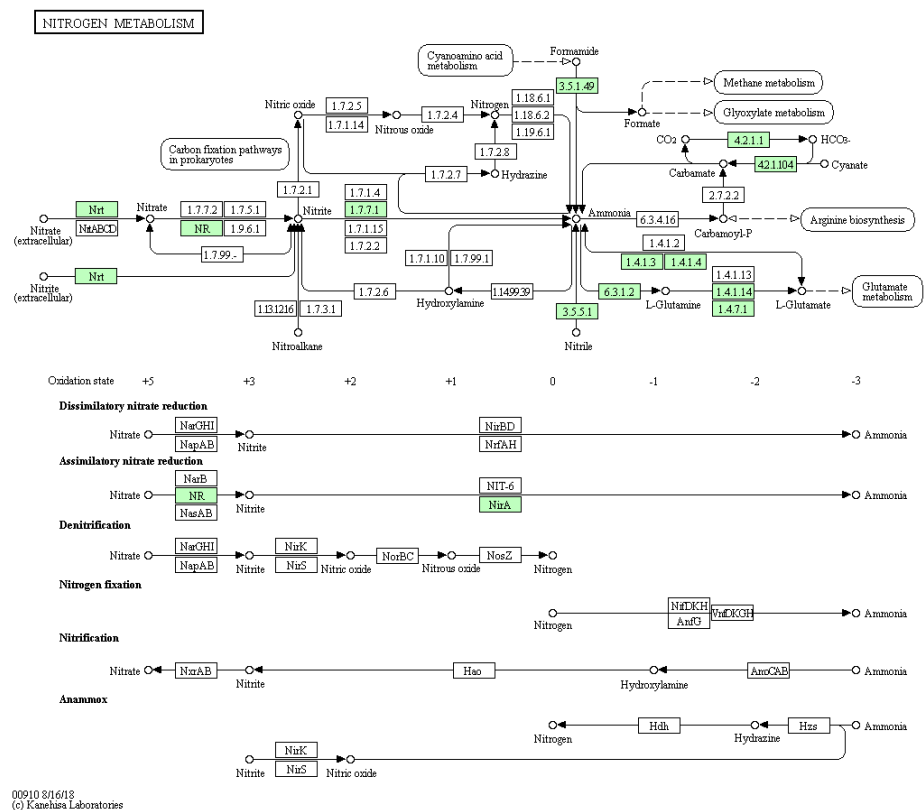


FIGURE 1 – Réseau métabolique du métabolisme de l'azote (KEGG). Les gènes surlignés en vert sont ceux exprimés chez *Arabidopsis Thaliana*.

## 3 Résultats

### 3.1 Filtrage des données :

Les trois fichiers donnés (cf Annexe B) comprennent 20 922 gènes. Après avoir effectué le filtrage sur les gènes d'intérêt, comme expliqué dans la partie "Méthode", 390 gènes et 7 934 interactions ont été retenus pour l'élaboration du réseau de co-expression.

### 3.2 Réseau de co-expression :

La liste obtenue comprend, pour chacun des gènes d'*Arabidopsis Thaliana* dont les locus des gènes sont connus pour répondre à une stimulation à l'azote (issu du fichier P\_GOI.txt), les valeurs d'expression observées dans 146 mesures (issu du fichier P\_EXPR.txt) ainsi que les paires de gènes codant pour des protéines en interaction (issues du fichier P\_PPI.txt). À partir de cette liste, un réseau de co-expression a été réalisé (cf Figure 2). Les gènes sont représentés par des carrés de couleur bleue. Les interactions entre les gènes peuvent être de deux types : corrélation pair ou PPI. Les interactions de type corrélation pair sont représentées en trait plein rouge et les interactions dites PPI sont représentées en trait sinusoïdal noir (cf Figure 2). Une liste des facteurs de transcription d'*Arabidopsis Thaliana* (provenant du fichier P\_TF\_LABEL.txt), figurant en triangle vert sur la Figure 2, a également été incorporée au réseau.

Seuls les gènes ayant des annotations GO relatives au nitrate, nitrogène, glutamate et glutamine ont été sélectionnés (le choix de la sélection de ces annotations GO a été préalablement justifié dans la sous partie "Sélection des ontologies" de la partie "Méthode"). Les trente gènes ayant ces annotations GO se trouvent dans les clusters numéro un et deux. Ces gènes ont été mis en évidence en jaune sur la Figure 2.

Afin de déterminer les annotations fonctionnelles des trois clusters majeurs (soit les clusters un, deux et trois, les trois autres contenant moins de cinq gènes chacun), des groupes d'annotations ont été réalisés. Les gènes appartenant à une figure de même couleur (ellipse ou carré) ont tous l'annotation propre à cette figure (cf Figure 2, cf Figure 3 et Figure 4). Le cluster 1, de par son nombre de gènes impliqués, est associé à l'annotation fonctionnelle "Response Temperature stimulus" (cf Figure 2, le cluster 2 également à celle de "Response Temperature stimulus" et le cluster 3 à celle de "Response multi organism" (cf Figure 4). Un récapitulatif sous forme de tableau associant les gènes régulateurs clés, leurs GO annotations et le groupe d'annotations duquel ils font partie, figure en Annexe D.

Les facteurs de transcription interagissant directement avec les gènes régulateurs clés ont été recensés dans la Table 1. 19 facteurs de transcription sont directement liés à un gène régulateur clé dans le cluster 1 et seulement 4 dans le cluster 2.

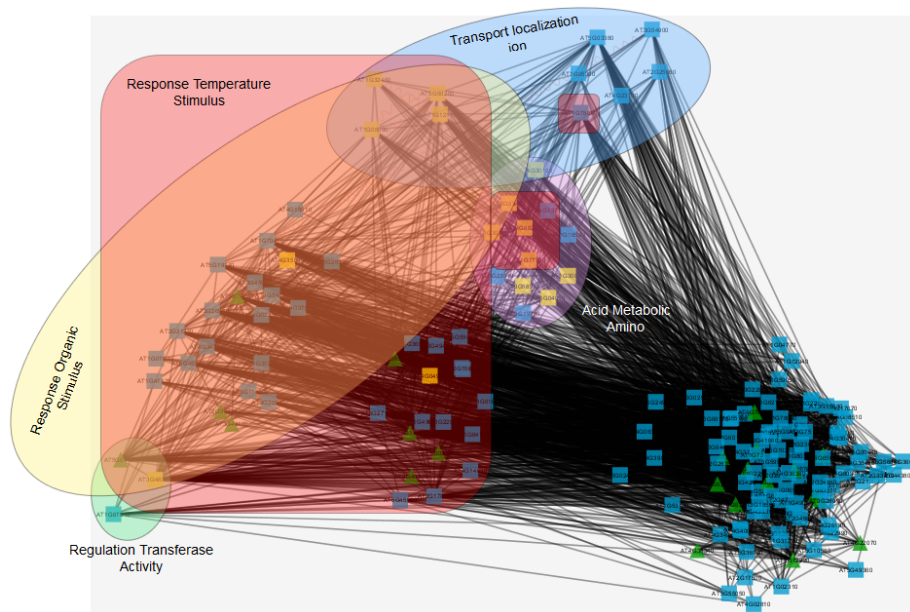


FIGURE 2 – Groupes d'annotations du cluster 1

TABLE 1: Facteurs de transcription ayant une interaction directe avec les gènes «régulateurs clés»

Cluster	Facteurs de transcription		
Cluster 1	AT5G61420, AT1G13300, AT1G74660, AT1G76350, AT5G62430, AT5G20885, AT1G61660	AT1G02340, AT2G33710, AT2G43500, AT4G39780, AT5G15830, AT4G22070,	AT5G65210, AT3G48100, AT1G68670, AT1G72220, AT1G66140, AT5G07680,
Cluster 2	AT5G52020, AT3G25790	AT5G58900,	AT5G10030,



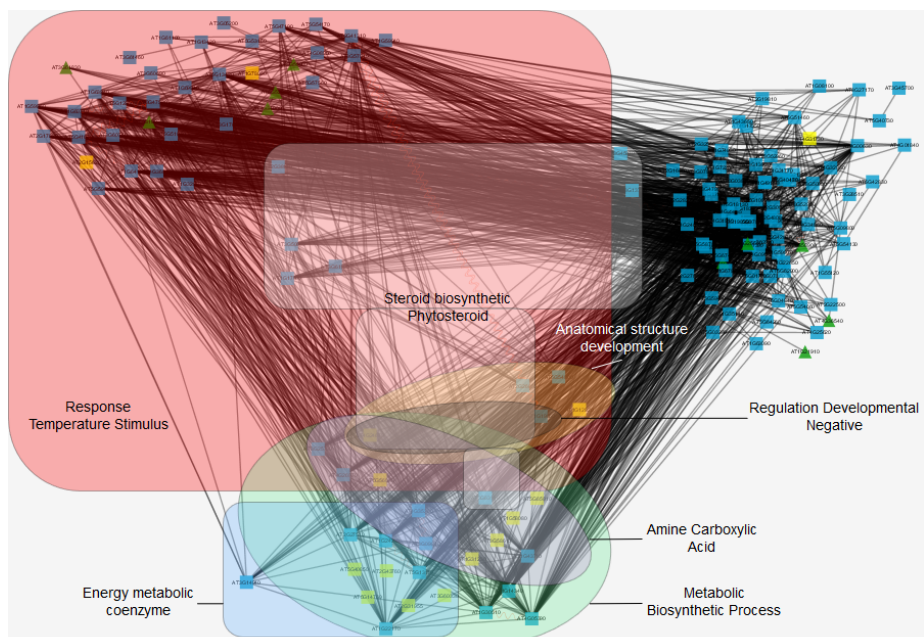


FIGURE 3 – Groupes d'annotations du cluster 2

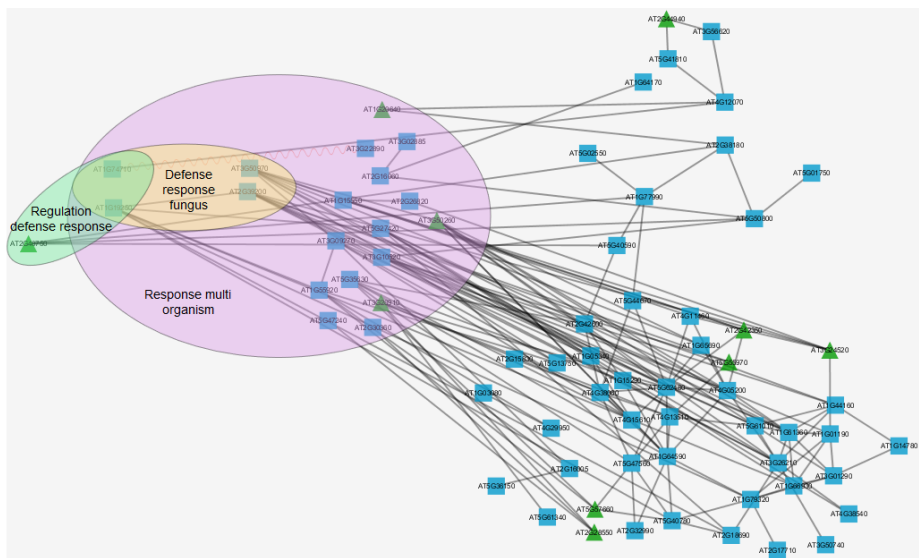


FIGURE 4 – Groupes d'annotations du cluster 3

## 4 Discussion

Les gènes régulateurs clés de la réponse au nitrate dans les racines d'*Arabidopsis Thaliana* du cluster 1 appartiennent essentiellement à l'annotation fonctionnelle de la réponse aux stimulus de température (11 sur les 15 du cluster 1). De même, les gènes «régulateurs clés» de la réponse au nitrate dans les racines d'*Arabidopsis Thaliana* du cluster 2 appartiennent essentiellement à l'annotation fonctionnelle "Metabolic Biosynthetic Process". Celle-ci étant très générale, il convient de regarder le second groupe comprenant le plus de gènes «régulateurs clés» (en jaune sur la Figure 3) à savoir la réponse aux stimulus de température (5 sur les 15 du cluster 2). Ainsi, un lien peut donc être établi entre la température et la réponse au nitrate dans les racines d'*Arabidopsis Thaliana*. Le glutamate faisant partie de la synthèse du nitrate, cette déduction est renforcée par la conclusion de l'article de Walch-Liu P[2] expliquant que dans diverses conditions environnementales, notamment la température, l'ombrage et l'apport en nitrates, une forte interaction génotype par environnement a été identifiée dans les contrôles de la réponse au glutamate.

Le cluster 3 est défini par l'annotation fonctionnelle "Response multi organism". Les gènes «régulateurs clés» de la réponse au nitrate dans les racines d'*Arabidopsis Thaliana* ne sont pas présents dans le cluster 3. Ainsi, on peut suggérer que la réponse au nitrate n'est pas corrélée à la réponse aux multi organismes.

Certains de ces gènes ont été obtenus dans d'autres études, ce qui confirme partiellement nos résultats [3]. Le réseau final obtenu repose sur une série de choix qui peut être discutée.

Le premier choix a été de conserver le coefficient de corrélation des interactions entre les gènes à 0,6. Ceci a pour conséquence de générer un réseau fortement connecté rendant la lisibilité de celui-ci plus fastidieuse mais permettant d'un autre côté de conserver un maximum d'informations.

Le second choix a été la sélection des annotations GO pour trouver les gènes «régulateurs clés» de la réponse au nitrate dans les racines d'*Arabidopsis Thaliana*. La difficulté a été de sélectionner des annotations ciblant la réponse au nitrate. Le terme "réponse" peut à la fois regrouper la réponse au niveau de la cellule, du tissu ou de la plante entière, chaque échelon de réponse dépendant du précédent. Les annotations GO choisies s'appuient sur des termes très liés au nitrate comme les termes glutamine et glutamate, responsable de la synthèse du nitrate (démonstré dans la sous partie "Sélection des ontologies" de la partie "Matériel et Méthodes"), le terme nitrogène, ancienne désignation de l'azote, et bien sûr, le terme nitrate.

Enfin, le dernier choix a porté sur les groupes d'annotations à former sur les clusters. Ces derniers ont été réalisés à partir de l'algorithme par défaut appelé MLC du plugin AutoAnnotate. Cependant, d'autres plugins (comme AllegroMCODE utilisé dans l'article de Mohammad Saberi Anvar[4]) et d'autres algorithmes proposant également la formation de groupes d'annotation sont disponibles (comme Community Cluster).

## Annexe A : Origine des données de transcriptome

Publication	Titre de l'expérience	Identifiant (base de données)
Gutierrez (2007)	Transcription profiling by array of Arabidopsis grown in nutrient solutions with various concentrations of nitrate and sucrose.	E-MEXP-828 (ArrayExpress)
Krouk (2010)	High resolution NO <sub>3</sub> <sup>-</sup> response of Arabidopsis roots	GSE20044 (GEO)
Ju (2009)	Expression data of 10-day-old wild-type and <i>chl1-5</i> plants exposed to 25 mM nitrate for 0 or 0.5 h	GSE9148 (GEO)
Ruffel (2011)	A systemic view of coordinated root responses to NO <sub>3</sub> <sup>-</sup> heterogeneous environment in Arabidopsis	GSE22966 (GEO)
Patterson (2010)	Comparison of root transcriptomes in <i>Arabidopsis thaliana</i> plants supplied with different forms of inorganic nitrogen	GSE29589 (GEO)
Vidal (2013)	Root nitrate response of Ws plants and <i>afb3-1</i> mutant plants.	GSE35544 (GEO)
Alvarez (2014)	Root nitrate response of Col-0 plants and <i>tga1/tga4</i> mutant plants	GSE43011 (GEO)
Wang (2003)	Treatment of Arabidopsis with low concentration of nitrate.	Exp479 (iPlant Collaborative)
Wang (2004)	WT vs. <i>Nitrate Reductase</i> null mutant high nitrate concentration treatment.	Exp480 (iPlant Collaborative)
Wang (2007)	Arabidopsis treated with nitrite and nitrate	Exp481 (iPlant Collaborative)

## Annexe B : Autres sources de données

La liste des gènes d'intérêt est disponible sous forme du «Supplementary Material» de la publication suivante : Canales J, Moyano TC, Villarroel E, Gutiérrez RA (2014) Systems analysis of transcriptome data provides new hypotheses about *Arabidopsis* root response to nitrate treatments [5].

Les données d'interaction protéine-protéine ont été compilées à partir des bases de données suivantes : PLAPID, ATPID, ATPIN et PAIR.

La table contenant la liste des facteurs de transcription d'*Arabidopsis Thaliana* avec leur famille a été compilée par le laboratoire Plant Systems Biology Lab de l'université pontificale catholique du Chili.

## Annexe C : Script R du filtrage des données

```
# Projet MOCELL - 11/10/08 :  
# Recherche de régulateurs clés à partir d'un réseau de co-expression :  
# Marc MONGY & Julie BLASQUIZ  
  
# Importation des fichiers :  
data <- read.table("P_EXPR.txt", header = T)  
  
# Lecture de la liste des gènes d'intérêt et chargement dans un vecteur  
focus <- read.table("P_GOI.txt", header = TRUE)  
list_focus <- as.vector(focus[, "id"])  
  
# Filtrage des données d'expression correspondant aux gènes d'intérêt  
filter_list <- data[, "id"] %in% list_focus  
data_focus <- data[filter_list,]  
  
# Labels que l'on souhaite sont dans la première colonne  
data_focus[, 1]  
rownames(data_focus) = data_focus[, 1]  
data_focus = data_focus[, -1]  
  
# Fonction t() inverse les colonnes et lignes (Nécessaire pour calculer les  
corrélations entre les gènes qui au départ correspondent aux lignes)  
data_focus = t(data_focus)  
  
# Calculer corrélation entre toutes les colonnes d'une matrice (Dataframe  
en paramètre)  
result_correlation = cor(data_focus)  
  
# Utiliser fonction melt pour convertir une matrice en liste d'interaction  
install.packages("reshape2")  
library("reshape2")  
list_correlation = melt(result_correlation)  
list_correlation = list_correlation[abs(list_correlation[, "value"]) > 0.6,]  
  
# Nécessaire de supprimer les paires redondantes (on est partie d'une ma-  
trice symétrique)  
# Création d'une nouvelle colonne Alphabétique (pour stipuler si le nome du  
gene A est ordonné selon l'aphabet par rapport au nom du gene B)  
list_correlation[, "Alphabétique"] <-  
as.character(list_correlation[, "Var1"]) < as.character(list_correlation[, "Var2"])  
  
# Supression des lignes où l'ordre de classement entre les deux noms de  
gènes ne suit par l'ordre alphabétique (permet de supprimer les doublons)  
list_correlation = list_correlation[list_correlation[, 4] == TRUE,]
```

```

# Supression de la colonne transitoire alphabétique n'est plus nécessaire
list_correlation = list_correlation[,-4]

# Supression de la colonne Value
list_correlation = list_correlation[,-3]

# Ajout de la colonne Interaction pour donner le type, ici ce sont les paires
dont l'expression est corrélée
list_correlation["type"] = "correlation_pair"

# Chargement des données PPI
data_PPI = read.table("P_PPI.txt", header = T, sep = "\t")

# Changement du nom des 2 premières colonnes pour qu'elles correspondent
dans les 2 listes
colnames(list_correlation)[1] = "id1"
colnames(list_correlation)[2] = "id2"
colnames(data_PPI)[3] = "type"

# Extraction des noms de gènes de la liste de co-expression pour les utiliser
comme filtre sur les PPI (on ne s'intéresse pas aux PPI des gènes pour lesquelles
pas de données d'expression)
id1 <- as.vector(list_correlation[, "id1"])
id2 <- as.vector(list_correlation[, "id2"])
list_net_genes <- c(id1, id2)
list_net_genes <- unique(list_net_genes)

# Filtrer PPI avec la liste de gènes d'intérêt
filtered_PPI_id1 <- data_PPI[, "id1"] %in% list_net_genes
expressed_PPI_id1 <- data_PPI[filtered_PPI_id1,]
filtered_PPI_id2 <- expressed_PPI_id1[, "id2"] %in% list_net_genes
expressed_PPI_id2 <- expressed_PPI_id1[filtered_PPI_id2,]

LIST = rbind(list_correlation_GOI.corrected.trimmed, expressed_PPI_id2)
#Ecrire fichier
write.table(LIST, "GRAPH.txt", sep="\t", row.names = FALSE, quote = F)

```

## Annexe D : Gènes régulateurs clés de la réponse au nitrate dans les racines d'*Arabidopsis Thaliana*

TABLE 2: Gènes «régulateurs clés» de la réponse au nitrate dans  
les racines d'*Arabidopsis Thaliana*

Gènes	GO annotations	Groupes d'annotation	cluster
AT5G53460	Glutamate metabolic process Glutamate biosynthetic process Glutamine family amino acid biosynthetic process Nitrate metabolic process Nitrate assimilation	Response Temperature stimulus, Response organic stimulus, Acid metabolic amino	1
AT3G30775	Glutamate metabolic process Glutamine family amino acid biosynthetic process Cellular nitrogen compound biosynthetic process	Response organic stimulus, Acid metabolic amino	1
AT4G04955	Cellular response to nitrogen starvation Cellular response to nitrogen levels	Response Temperature stimulus	1
AT4G35090	Cellular response to nitrogen starvation	Response Temperature stimulus, Response organic stimulus	1
AT1G12110	Response to nitrate Nitrate transport	Response Temperature stimu- lus, Response organic stimulus, Transport localization ion	1
AT5G50200	Response to nitrate Nitrate transport	Response Temperature stimu- lus, Response organic stimulus, Transport localization ion	1
AT1G32450	Response to nitrate Nitrate transport	Response Temperature stimu- lus, Response organic stimulus, Transport localization ion	1
AT1G08090	Response to nitrate Nitrate transport	Response Temperature stimu- lus, Response organic stimulus , Transport localization ion	1
AT3G48360	Response to nitrate	Regulation transferase activity, Response temperature stimulus, Respond organic stimulus	1
AT1G77760	Nitrate metabolic process Nitrate assimilation		

TABLE 2: Gènes «régulateurs clés» de la réponse au nitrate dans les racines d'*Arabidopsis Thaliana* (suite).

Gènes	GO annotations	Groupe annotation	cluster
	Cellular nitrogen compound biosynthetic process	Response Temperature stimulus, Acid metabolic amino	1
AT1G66200	Nitrate metabolic process Nitrate assimilation	Acid metabolic amino	1
AT4G04955	Cellular response to nitrogen levels	Response temperature stimulus	1
AT1G30820	Cellular nitrogen compound biosynthetic process	Acid metabolic amino	1
AT5G04950	Cellular nitrogen compound biosynthetic process	Acid metabolic amino	1
AT3G58990	Cellular nitrogen compound biosynthetic process	Acid metabolic amino	1
AT4G31730	Glutamine transport Glutamine secretion		2
AT2G15620	Response to nitrate	Response temperature stimulus	2
AT1G12820	Response to nitrate	Response temperature stimulus, Anatomical structure development	2
AT1G78050	Response to nitrate	Response temperature stimulus	2
AT3G58610	Cellular nitrogen compound biosynthetic process	Metabolic biosynthetic process, Amine carboxylic acid	2
AT3G60330	Cellular nitrogen compound biosynthetic process	Metabolic biosynthetic process, Energy metabolic coenzyme	2
AT5G14760	Cellular nitrogen compound biosynthetic process	Metabolic biosynthetic process, Energy metabolic coenzyme	2
AT1G31230	Cellular nitrogen compound biosynthetic process	Metabolic biosynthetic process, Amine carboxylic acid	2
AT5G40850	Cellular nitrogen compound biosynthetic process	Metabolic biosynthetic process, Energy metabolic coenzyme	2
AT2G31955	Cellular nitrogen compound biosynthetic process	Metabolic biosynthetic process, Energy metabolic coenzyme	2
AT5G56080	Cellular nitrogen compound biosynthetic process	Response temperature stimulus, Metabolic biosynthetic process	2
AT5G65010	Cellular nitrogen compound biosynthetic process	Metabolic biosynthetic process, Amine carboxylic acid	2
AT1G58080	Cellular nitrogen compound biosynthetic process	Metabolic biosynthetic process, Amine carboxylic acid	2
AT2G43760	Cellular nitrogen compound biosynthetic process	Metabolic biosynthetic process, Energy metabolic coenzyme	2



TABLE 2: Gènes «régulateurs clés» de la réponse au nitrate dans les racines d'*Arabidopsis Thaliana* (suite).

Gènes	GO annotations	Groupe annotation	cluster
AT4G24670	Cellular nitrogen compound biosynthetic process	Response temperature stimulus, Amine carboxylic acid, Anatomical structure development, Regulation developmental negative, Metabolic biosynthetic process, Steroid biosynthetic phytosteroid	2

## Références

- [1] G. Barroin, “Gestion des risques. santé et environnement : le cas des nitrates phosphore, azote et prolifération des végétaux aquatiques.” Assises internationales envirobio 13-14/11/2000 Paris, 2000.
- [2] A. T. F. B. Walch-Liu P, Meyer RC, “Qtl analysis of the developmental response to l-glutamate in arabidopsis roots and its genotype-by-environment interactions,” *Journal of Experimental Botany*, vol. 68, no. 11, pp. 2919–2931, 2017.
- [3] S. Das, R. Pathak, D. Choudhury, and N. Raghuram, “Genomewide computational analysis of nitrate response elements in rice and arabidopsis,” *Molecular Genetics and Genomics*, vol. 278, pp. 519–525, 12 2007.
- [4] M. S. Anvar, M. S. Zarrin Minuchehr, and S. Kheitan, “Gastric cancer biomarkers; a systems biology approach,” *Biochemistry and Biophysics Reports*, no. 13, pp. 141–146, 2018.
- [5] C. J, M. TC, V. E, and G. RA, “Systems analysis of transcriptome data provides new hypotheses about arabidopsis root response to nitrate treatments,” *Front Plant Sci.*, vol. 5, pp. 1–14, 02 2014.