



Université Frères Mentouri Constantine 1
Faculté des Sciences de la Nature et de la Vie

جامعة الإخوة منتوري قسنطينة 1
كلية علوم الطبيعة والحياة

Département : Biologie appliquée

قسم : البيولوجيا التطبيقية

Mémoire présenté en vue de l'obtention du Diplôme de Master

Domaine : Sciences de la Nature et de la Vie

Filière : Sciences Biologiques

Spécialité : Bioinformatique

Intitulé :

**Analyse *in silico* des données de transcription et
identification des gènes impliqués dans la réponse au
stress abiotique (stress d'azote) chez le modèle
biologique *Medicago truncatula***

Présenté et soutenu par : M^{lle} BENMENIA RayenRebiha.

Le : 07/10/2020

M^{lle} CHELOUCHE Roufaïda.

Jury d'évaluation :

Président du jury : Mr. BELKHIRI Loutfi

Prof. Université Constantine 1.

Rapporteurs : Mr. TEMAGOULT Mahmoud
Mr. KELOU Kamel

M.A.A. Université Constantine 1
M.A.A. Université Constantine 1

Examineur : Mr. DAAS Mouhamed Skander

M.C.B. Université Constantine 1

Année universitaire 2019 - 2020

Remerciement :

Avant tout, nous remercions "Allah" le tout puissant de nous avoir donné la santé, la force, le courage, la patience, la persistance et nous a permis d'exploiter les moyens disponibles à fin d'accomplir ce modeste travail. Merci de nous avoir éclairé le chemin de la réussite.

*Nous adressons nos plus sincères remerciements à notre encadreur Monsieur Dr. **TAMAGOULT Mahmoud** et Dr. **KELLOU kamal** qui nous ont dirigés dans ce travail avec une grande rigueur scientifique, leur patience, leurs conseils, leur grande disponibilité tout au long de la réalisation de ce mémoire, ainsi pour le temps qu'ils ont bien voulu nous consacrer et sans eux ce mémoire n'aurait jamais vu le jour.*

*Nous tenons aussi à remercier profondément les membres du jury : Monsieur Pr **BELKHIRI** et Monsieur Dr **DAAS** qui Nous ont fait l'honneur d'examiner ce travail.*

Nous souhaiterons également remercier nos professeurs de la faculté des Sciences de la Nature et de la Vie pendant les cinq années de notre parcours.

Dédicaces

*Avant de dédier ce travail je tiens d'abord à remercier "ALLAH",
Le tout puissant qui nous a permis de mener à bien ce modeste travail,*

A mes chers parents: Yamina et Mohammed Zin Eddine

A ma sœur : Esma et mon frère : Abed El Djallil

A mon neveu : Racim et ma petite nièce : Tamara

A mon responsable de travail : M^{ELLE} LAABOUN Hanane

pour leur soutien et leur encouragement

A mes chers amis: Louiza ,Imen, Hamza ,Rahil ,et Nahla

A toute ma familles respective.

RAYEN

Dédicace :

Je dédie ce modeste travail en signe de respect, de reconnaissance et de gratitude

A dieu de tout puissant de m'avoir donné le courage, la santé, et m'a accordé son soutien durant les périodes les plus difficiles.

A mes très chers parents, que j'admire beaucoup, qui m'ont toujours aidés dans ma vie et qui n'ont cessés de m'encourager et de me soutenir tout au long de mes études, que dieu les garde en bonne santé.

A mes frères Amani Amin Salim pour leurs tendresse et leurs permanentes présence à mes cotés.

Je le dédie aussi à tous les enseignants de notre faculté qui ont toujours guidé tout au long de mon parcours éducatif

A mes amis Djalil Karim Sara Khawla Ines Safia Rania Yassmine Hiba Rania qui ont cru en moi et qui ont toujours encouragé, et avec qui j'ai passé des années inoubliables.

Liste des abréviations

%	: Pourcent
2wk	: Deux semaines.
ATP	: Adénosine Tri-Phosphate
ARN	: Acide ribonucléique
ARNm	: Acide Ribonucléique messenger
BACs	: Bankers' Automated Clearing System
BLAST	: Basic Local AlignmentSearchTool.
°C	: Degrés Celsius
CP	: Cystéine Protéases
EMBL	: Européen Moléculaire Biologie Laboratoire
ESTs	: Expressed Sequence Tag
FNs	: Facteurs de Nodulation
Lb	: Leghémoglobine
LCOs	: Lipo-Chito-Oligosaccharides
LEA	: Late Embryogenesis Abundant Protein
LIS	: Legume Information System
Mpb	: Mega paire de bases.
MtGEA	: <i>Medicago truncatula</i> Gene Expression Atlas
Mt	: Tonnes métriques
MTGD	: The <i>Medicago truncatula</i> Genome Database
NO₃⁻	: Nitrates
Nm	: Nanomolaire
NCBI	: National Center for Biotechnology Information
ORF	: Open Reading Frame
PCR	: Amplification en chaîne par polymérase
PH	: Potentiel hydrogène
QPCR	: Amplification en chaîne par polymérase quantitative
RNAi	: Acide Ribonucléique Interférence
SRL	: Symbiose Rhizobium Légumineuse
TF	: Transcription Factor.
TIGR	: The Institute for GenomicResearch.

Liste des figures

Figure 1: Phylogénie des légumineuses.....	3
Figure 2: Un rameau de <i>Medicago truncatula</i> ; fleur; gousse; graines. (Lahrech, 2011).....	5
Figure 3: Répartition géographique du genre <i>Medicago</i>	6
Figure 4: Structure du complexe enzymatique de la nitrogénase.....	10
Figure 5: Représentation schématique des principales étapes de la symbiose <i>Rhizobium</i> - légumineuses.....	11
Figure 6: Schéma du cycle de l'azote terrestre.....	12
Figure 7 : Schématisation de l'approche <i>in silico</i> de la biologie (Bonsai, 2014)	13
Figure 8 : Interface du moteur de recherche MtGEA.....	18
Figure 9 : l'option Batch Download by Experiment sur le moteur de recherche MtGEA.....	18
Figure 10: Interface du Batch Download by Experiment.....	19
Figure 11 : L'interface de R en quatre quadrants.....	22
Figure12 : Cluster ING de la totalité des gènes sur logiciel R.....	25
Figure 13: présente le premier sous- clade dans la phylogramme.....	26
Figure14 : présente le deuxième clade dans la phylogramme.....	26
Figure 15: présente le troisième sous- clade dans la phylogramme.....	27
Figure 16: présente le quatrième sous- clade dans la phylogramme.....	28
Figure 17: présente le cinquième sous- clade dans la phylogramme.....	28
Figure 18: présente le sixième sous- clade dans la phylogramme.....	29
Figure 19 : présente le septième clade dans la phylogramme.....	30

Liste des tableaux

Tableau 1 : Quelques bases de données bioinformatiques disponibles pour <i>M. truncatula</i>	15
Tableau 2 : Conditions sélectionnées pour la Symbiose et la Senescence.....	19
Tableau 3 : Les données des gènes relatives aux 08 conditions.....	19
Tableau 4 : Les données des gènes classé sur la feuille de calcul Excel.....	20
Tableau 5 : Les formules des commandes utilisées dans R (Barnier et al, 2017).....	22
Tableau 6 : présente les données de l'expression des gènes sur le logiciel.....	23
Tableau 7: le nombre de gènes dans chaque clade.....	25
Tableau 8: présente l'identifiant et la description du premier sous-clade.....	26
Tableau 9: présente l'identifiant et la description de le deuxième sous- clade.....	26
Tableau 10 : présente l'identifiant et la description du troisième sous-clade.....	27
Tableau 11: présente l'identifiant et la description de le quatrième sous- clade.....	28
Tableau 12: présente l'identifiant et la description du cinquième sous-clade.....	29
Tableau 13 : présente l'identifiant et la description du sixième sous- clade.....	29
Tableau 14: présente l'identifiant et la description de le septième clade.....	30

Résumé

Analyse *in silico* des données de transcription et identification des gènes impliqués dans la réponse au stress abiotique (stress d'azote) chez le modèle biologique *Medicago truncatula*

L'objectif du présent travail est la recherche des gènes impliqués dans la réponse au stress abiotique fixation d'azote dans la racine du modèle biologique *Medicago truncatula*, par une analyse *in silico* des données de transcription issues de puces Affymetrix, et en utilisant le profil d'expression de 20 gènes différents aux huit différentes conditions. L'arbre phylogénétique tracé par le logiciel statistique R classe les 20 gènes en 2 clades; le premier clade dispos 4 gènes et le deuxième clade dispose 16 gènes. Les résultats obtenus montrent que les gènes impliqués dans cette réponse sont très diversifiés et appartiennent à plusieurs groupes comme les gènes codant pour les protéines de types LEA, Les facteurs NOD, différents facteurs de transcription (TF), des protéines réceptrices et de transport ainsi que d'autres gènes qui codent pour d'autres fonctions.

Mots clés : Stress azote ; *Medicago truncatula* ; L'arbre phylogénétique.

Abstract

In silico* analysis of transcription data and identification of genes involved in the abiotic response (nitrogen stress) in the biological model *Medicago truncatula

The objective of this work is the search for genes involved in the response to nitrogen stress in the root of the biological model *Medicago truncatula*, by an *in silico* analysis of transcription data from Affymetrix chips, and by using the expression profile of 20 vas deferens in eight deferent conditions

The phylogenetic tree traced by the statistical software R classifies the 20 genes into 2 clades; the first clade has 4 genes and the second clade has 16 genes. The results obtained show that the genes involved in this response are very diverse and belong to several groups such as genes encoding LEA type proteins, various transcription factors (TF), receptor and transport proteins as well as other genes. Which code for other functions.

Key words: Nitrogen stress ; *Medicago truncatula* ; phylogenetic tree.

المخلص

تحليل السيليكو لبيانات النسخ وتحديد الجينات المشاركة في الاستجابة
اللاأحيائية (إجهاد النيتروجين) في النموذج البيولوجي *Medicago truncatula*

الهدف من هذا العمل هو البحث عن الجينات المشاركة في الاستجابة لضغط النيتروجين في جذر النموذج البيولوجي *Medicago truncatula* ، عن طريق التحليل السليكي لبيانات النسخ من رقائق Affymetrix ، وباستخدام ملف تعريف التعبير لـ 20 أسهر في ثمانية ظروف مختلفة. تصنف شجرة النشوء والتطور التي تم تتبعها بواسطة البرنامج الإحصائي R الجينات العشرين في كتلتين ؛ يحتوي الكليد الأول على 4 جينات والفرع الثاني يحتوي على 16 جيناً. تظهر النتائج التي تم الحصول عليها أن الجينات المشاركة في هذه الاستجابة متنوعة للغاية وتنتمي إلى عدة مجموعات مثل الجينات التي تشفر بروتينات نوع LEA وعوامل النسخ المختلفة (TF) والمستقبلات وبروتينات النقل بالإضافة إلى الجينات الأخرى. أي رمز لوظائف أخرى.

Sommaire

Introduction.....	1
Chapitre I : Revue bibliographique	
I.1 Les Légumineuses.....	3
I.1.1 Importance et rôle des légumineuses.....	3
I.1.2 Medicago truncatula (modèle d'étude).....	4
I.1.2.1 Systématique.....	5
I.1.2.2 Origine et distribution géographique.....	6
I.1.2.3 Caractéristiques et intérêts génétiques et génomiques.....	6
I.1.2.4 Intérêts agronomiques.....	7
I.1.2.5 La biologie et Medicago-truncatula.....	7
I.2 Le stress.....	7
I.2.1 Définition du mot stress.....	7
I.2.2 Le stress biotiques et stress abiotique.....	8
I.2.3 Le stress d'azote.....	8
I.3 La symbiose fixatrice d'azote.....	8
I.3.1 La symbiose légumineuse-Rhizobium.....	8
I.3.2 Les partenaires symbiotiques.....	8
a. Les légumineuses	
b. Les rhizobia	
I.3.3 Les différentes étapes de la formation de la nodosité.....	9
I.3.3.1 Dialogue moléculaire entre les deux partenaires.....	9
I.3.3.2 Infection et développement de la nodosité.....	9
I.3.3.3 Le fonctionnement de la nodosité.....	10
I.3.3.4 La sénescence du processus symbiotique.....	10
I.3.4 Le cycle d'azote.....	12
a. La fixation de l'azote moléculaire N₂	
b. La nitrification	
c. La dénitrification	
I.4 La Bioinformatique.....	13
I.4.1 Définition de la Bioinformatique.....	13
I.4.2 Les bases et les banques de données.....	13
I.4.3 Les domaines de la Bioinformatique.....	14
I.4.3.1 Phylogénie.....	14
I.4.3.2 Génomique structurale et fonctionnelle.....	14
I.4.3.3 Biotechnologie.....	14
I.4.3.4 Analyse de la transcription.....	14
I.4.4 Medicago truncatula et bioinformatique.....	15

Chapitre II : Matériels et méthodes

II.1	Matériel	17
II.2	Méthodes.....	17
II.2.1	Recherche du <u>Batch Download by Experiment</u> dans la base de données MtGEA.....	17
II.2.2	Collecte des données d'expression des gènes sur la page d'accueil Batch Download by Experiment.....	19
II.3	Etude statistique par R.....	21

Chapitre III : Résultats et discussion

III.1	Résultats.....	25
III.2	Discussion.....	31
	Conclusion.....	34
	Références bibliographique.....	36

Introduction

Les plantes dans leur milieu sont exposées à différentes contraintes environnementales qui peuvent limiter leur croissance et leur développement. De telles conditions environnementales sont dites conditions stressantes ou stress. Selon la nature de stress, on peut distinguer les stress biotiques et stress abiotiques. Les stress abiotiques sont un ensemble de pression causées par le milieu soit physique ou chimique de la plante.

L'amélioration de la résistance des plantes aux stress abiotiques a été depuis longtemps un des principaux objectifs de l'amélioration végétale. De nombreuses approches ont été conduites pour mieux comprendre les mécanismes impliqués et identifier les caractères qui peuvent être utilisés comme critères de sélection de la tolérance, la plus récente de ces approches est l'approche moléculaire couplée à l'utilisation de l'outil informatique. De cette combinaison interdisciplinaire est née la bioinformatique qui trouve son champ d'application dans l'analyse et l'interprétation des données biologiques, au moyen de méthodes informatiques en vue de la minimisation du temps et du coût des projets de recherche.

Dans cet ordre d'idée, le présent travail a pour objectif recherché les gènes impliqués dans la réponse au stress azoté du modèle biologique *Medicago truncatula*, par une analyse *in silico* des données de transcription issues de puces Affymetrix.

Ainsi, ce mémoire est structuré en trois grands chapitres :

- Un premier chapitre qui représente qui est une synthèse bibliographique sur la thématique du projet.
- Un deuxième chapitre pour la description du matériel et des différentes méthodes utilisées.
- Le troisième chapitre sera consacré au traitement des résultats obtenus et à leur discussion.
- Finalement, une conclusion avec le dégagement de quelques perspectives.

Chapitre I :

Revue bibliographique

I.1. Les Légumineuses

La famille des légumineuses, estimée à environ 20 000 espèces, se compose essentiellement de 3 sous familles, à savoir les *Mimosoideae*, les *Caesalpinioideae*, et les *Papilionoideae* (Doyle, 2003). Ce qui en fait la troisième plus large famille des plantes supérieures (Gepts, 2005) (figure.1).

Les légumineuses, largement définies par leur structure florale inhabituelle, leurs fruits et leur capacité à former des nodules avec des *Rhizobia*, viennent en deuxième place parmi les plantes cultivées après les graminées (Graham et Vance, 2003).

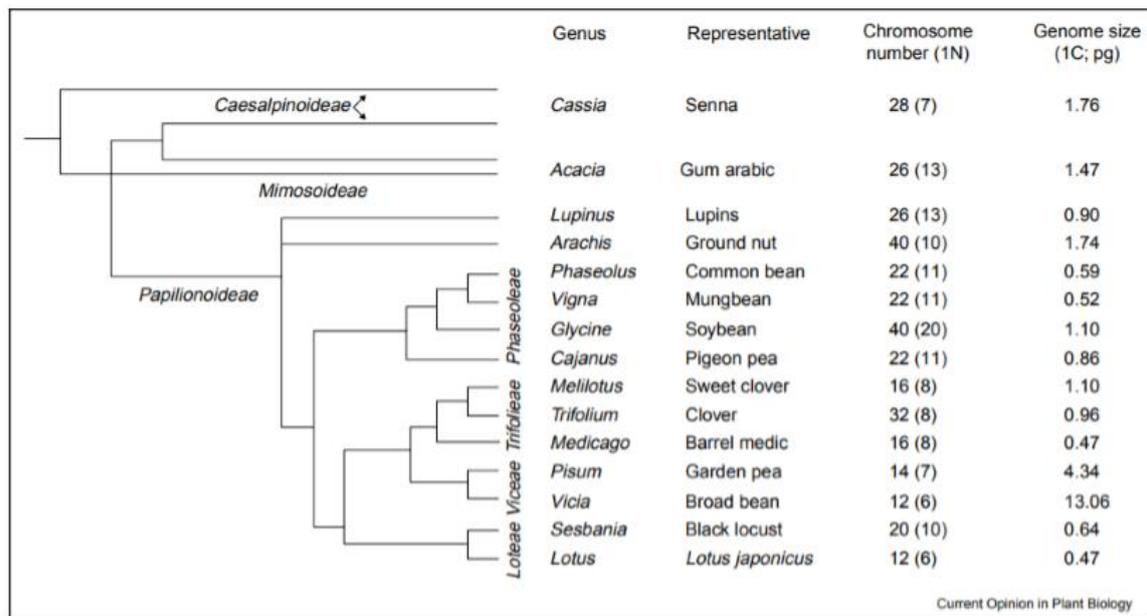


Figure 1 : Phylogénie des légumineuses.

Les trois grandes sous-familles des légumineuses sont représentées (*les Mimosoïdées*, *les Caesalpinioïdées* et *les Papilionoïdées*), et des estimations du nombre de chromosomes et de la taille du génome sont mentionnées. (Young et al., 2003).

I.1.1. Importance et rôle des légumineuses

L'importance des légumineuses découle principalement de leur pouvoir à établir une association symbiotique fixatrice de l'azote atmosphérique (N₂) par association avec des bactéries rhizosphériques, appelées rhizobiums au sein de structures racinaires appelées nodules ou nodosités (Graham et Vance, 2003).

Ceci les rend indépendantes de source d'azote sous forme de nitrates ou d'ammoniaque présents dans le sol, ou apportés par des engrais. On estime la quantité de N₂ fixé par les légumineuses à intérêts agronomiques à environ 40 à 60 millions de tonnes métriques (Mt)

chaque année, avec 3 à 5 millions de tonnes supplémentaires fixées par des légumineuses dans des écosystèmes naturels (Smil, 1999).

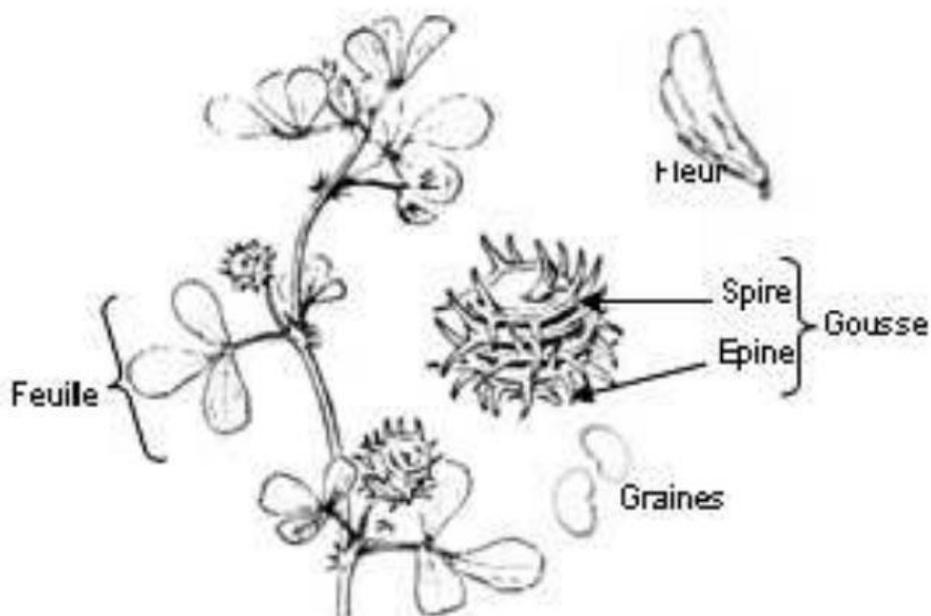
Au cours de la première phase de la croissance, les légumineuses doivent cependant s'appuyer principalement sur l'absorption de N du sol similaire aux espèces non légumineuses, avant de pouvoir initier la nodulation et de développer des nodules efficaces infectés par les Rhizobiums. En outre, les légumineuses à graines peuvent compter aussi sur la réserve d'azote de semences pendant un temps important, en raison des grandes graines avec une teneur en azote relativement élevée (Herdina, 1990).

I.1.2. *Medicago truncatula* en tant que plante modèle

Le terme *Medicago* vient du mot latin "Medica" (herbes de médic), il fut modifié par Dalechamps en 1587, et devient *Medicago*. Le nom américain donné à la luzerne "alfalfa" proviendrait de l'Arabe 'البرسيم' 'الفصفاة'.

La luzerne tronquée (*Medicago truncatula*) est une espèce végétale herbacée des régions méditerranéennes. Elle appartient à la famille botanique des Fabacées, plus couramment nommées Légumineuses ou Papilionacées. (Doyle and Luckow, 2003).

Chaque plante peut produire jusqu'à 500 à 1000 graines- et ayant un cycle de vie relativement court qui varie entre 2 et 3 mois. La plante fait généralement entre 15 et 80 cm de hauteur. Ses feuilles sont trifoliées (Figure 2) et ses fleurs sont jaunes. L'autofécondation donne des gousses cylindriques dures contenant chacune 3 à 12 graines avec une durée de vie supérieures à 40 ans. (Lesins et Lesins, 1979).



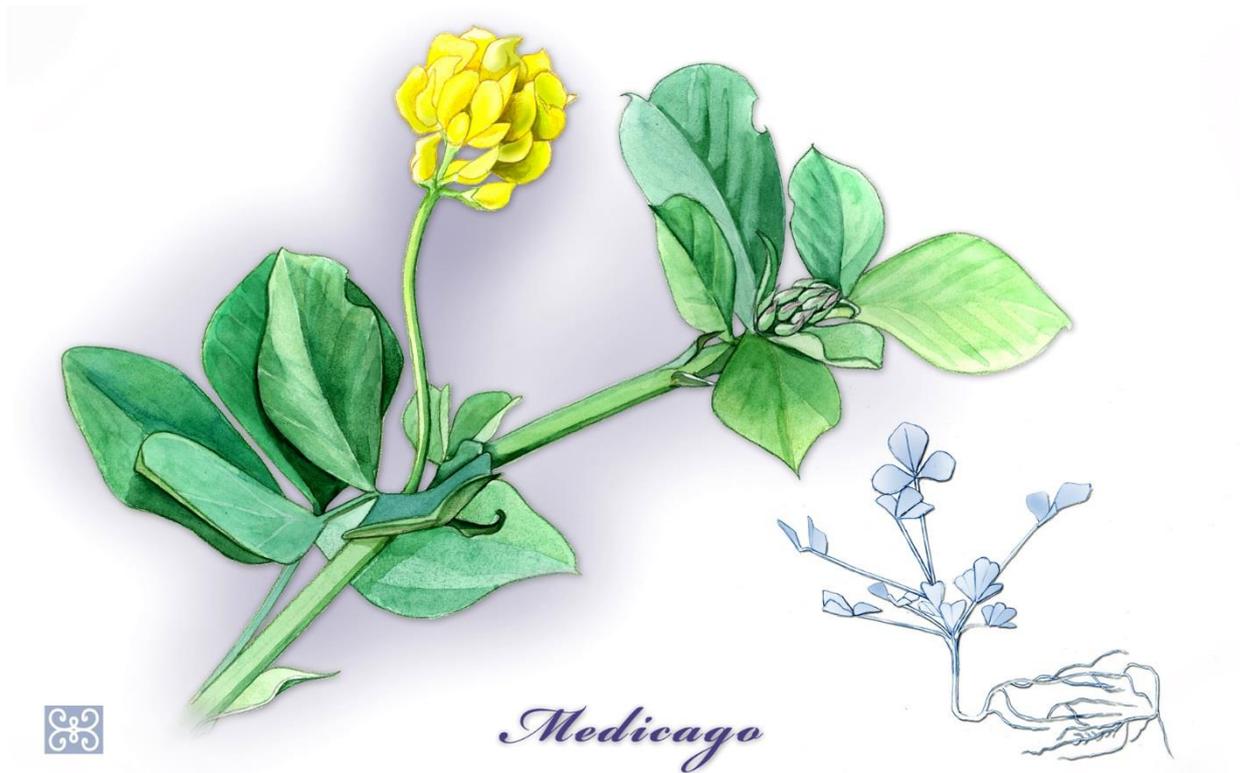


Figure 2. : Un rameau de *Medicago truncatula*; fleur; gousse; graines. (Lahrech, 2011).

I.1.2.1. Systématique

Medicago truncatula appartient à l'ordre des Fables et à la famille des Fabaceae (ex légumineuses), qui comprend plus de 14 000 espèces réparties en trois sous-familles : les *Césalpiniées*, les *Mimosacées* et les *Papilionacées*. Les *Papilionacées* renferment deux clades : les *Phaseolides* d'origine tropicale (ex. le soja *Glycine max*) et les Galégoïdes d'origine tempérée de l'hémisphère nord comme les pois et la luzerne (Doyle et al, 2003).

D'où la classification de *Medicago truncatula* :

Règne : *Plantae*

Division : *Magnoliophyta*

Classe : *Magnoliopsida*

Ordre : *Fabales*

Famille : *Fabaceae*

Sous-famille : *Papilionoideae*

Tribu : *Trifolieae*

Genre : *Medicago*

Espèce : *Medicago truncatula*

I.1.2.2. Origine et distribution géographique

La présence de *M. truncatula* sur tout le pourtour méditerranéen (**Figure 3**) représente une source de biodiversité, à travers la présence de plusieurs variétés qui montrent des adaptations différentes aux facteurs climatiques et biotiques de leur milieu. Cette caractéristique est donc importante pour l'étude de la génétique et de la génomique.

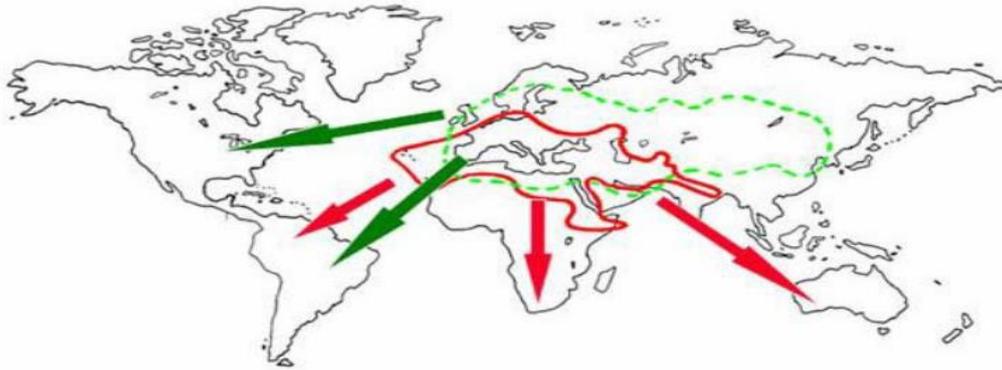


Figure 3 : Répartition géographique du genre *Medicago*.

Les régions encadrées correspondent aux régions d'origine de l'espèce pérenne *M. sativa* (en rouge) et l'espèce annuelle *M. truncatula* (en vert). Les flèches indiquent l'expansion vers de nouvelles régions. (Delalande et al., 2007).

I.1.2.3. Caractéristiques et intérêts génétiques et génomiques

Elle est diploïde ($2n=16$), autogame et annuelle. Elle possède un génome de faible taille, 550 Mpb soit environ 4 fois celui d'*A. Thaliana* (Blondon et al., 1994).

Elle présente une forte variabilité biologique pour le port, la périodicité des plantes, le nombre des graines par gousses, l'efficacité de nodulation et de fixation de l'azote. Il existe une forte synergie entre le génome de *M. truncatula* et de plusieurs autres légumineuses, notamment de *M. sativa* et du pois (Zhu et al., 2005 ; Aubert et al., 2006).

Des outils génétiques (cartes génétiques) ont été générés par plusieurs laboratoires (Thoquet et al., 2002 ; Choi et al., 2004 ; Julier et al., 2007 ; Kamphuis et al., 2008).

truncatula est facilement transformée par *Agrobacterium rhizogenes* (Boisson-Dernier et al., 2001).

Par ailleurs, de nombreux outils et ressources génétiques et génomiques, tels que le séquençage en cours de son génome, des cartes génétiques, des banques de BACs et d'ESTs, des puces de microarray, des mutants... sont maintenant disponibles (Sato et al., 2007 ; Ané et al., 2008) et ont permis d'aborder l'analyse de nombreux processus biologiques et physiologiques relatifs tant au développement de la plante (Julier et al., 2007 ; Pierre et al., 2008) qu'à l'analyse des interactions biotiques ou abiotiques (pour revue : Rose, 2008).

Ces constats ont amené la communauté scientifique internationale qui travaille sur les Légumineuses à proposer *Medicago truncatula* comme plante modèle, non seulement pour l'étude des effets des stress abiotiques et des symbioses, mais aussi pour l'analyse des interactions Légumineuses-parasites (Cook, 1999 ; Young et al., 2003).

I.1.2.4. Intérêts agronomiques

Venant à l'origine du pourtour méditerranéen, *M. truncatula* est présente sur des sols lourds, marneux ou argileux et est adaptée aux conditions semi-arides (Lazrek ben Friha, 2008).

L'intérêt agronomique pour la production fourragère et la diminution de l'épandage d'engrais azotés en zone sèche s'avère donc très important. Bien qu'elle soit annuelle, *M. truncatula* pourrait être assimilée à une espèce pérenne, grâce à sa capacité à se ressemer naturellement d'une année à l'autre, ce qui aboutit à la diminution du phénomène d'érosion des sols très présent dans le bassin méditerranéen. La plupart des variétés commerciales de *M. truncatula* sont essentiellement cultivées en rotation avec des céréales.

I.1.2.5. La biologie de *Medicago truncatula*

M. truncatula a déjà fait l'objet de nombreux travaux de biologie moléculaire, qui ont permis d'étudier plusieurs dizaines de gènes impliqués dans l'interaction symbiotique avec *Sinorhizobium* dans la symbiose endomycorhizienne.

I.2. Le stress

Les plantes sont souvent soumises à des conditions environnementales défavorables qui affectent leur croissance, leur développement et leur productivité et qu'on peut regrouper sous le nom 'stress'. Tous les stress impliquent des réactions de signalisation capables de déclencher une réponse chez la plante.

I.2.1. Définition du mot stress en biologie

D'un point de vue biologique, un stress est toute condition externe capable de provoquer des changements ou des dégâts chez l'être vivant. (Hopkins, 2003).

Si la contrainte est suffisamment sévère, l'effet sur l'organisme vivant peut être permanent, dommage ou mort. Aussi comme pour les corps physiques, les organismes vivants présentent une particularité vis-à-vis de chaque stress (Levitt, 1980).

I.2.2. Le stress biotiques et le stress abiotique

Selon l'agent stressant, on peut distinguer les stress biotiques résultant d'une interaction avec d'autres organismes vivant, par exemple, une infection ou un parasitisme, et les stress abiotiques causés par des conditions défavorables de l'environnement physique de l'être vivant comme la température, l'humidité, l'intensité lumineuse, l'approvisionnement en eau ou en nutriments (Schulze et al, 2005).

I.2.3. Le stress d'azote

La capacité à pouvoir fixer l'azote atmosphérique via la mise en place d'une symbiose avec *Rhizobium* méliot au niveau racinaire fait que cette culture ne nécessite pas d'engrais azote (Spaink, 2000 ; Perret et al. 2000).

Sa composition en acides aminés est en outre bien équilibrée. Les acides aminés essentiels pour l'alimentation animale sont tous présents dans la matière azotée de cette plante. Enfin, et de manière plus récente, elle est utilisée pour la production de molécules d'intérêt pharmacologique lui conférant un attrait industriel non négligeable.

I.3. la symbiose fixatrice d'azote

I.3.1 La symbiose légumineuse-Rhizobium

L'importance des légumineuses est largement attribuée à leur capacité à établir au niveau de leurs racines une symbiose fixatrice de l'azote atmosphérique (N₂) avec les rhizobiums. Ces bactéries sont hébergées dans des structures appelées nodules ou nodosités dans lesquelles elles trouvent les conditions idéales pour réaliser les réactions de réduction d'azote.

I.3.2 Les partenaires symbiotiques

a. Les légumineuses

Les Légumineuses possèdent l'aptitude de former des symbioses fixatrices d'azote avec des bactéries du sol (De Faria et al., 1989).

L'association entre les plantes de la famille des légumineuses et les bactéries rhizobia est l'interaction symbiotique la plus efficace en termes de fixation azotée, (Ferguson et al., 2010). Sur un plan scientifique, la Légumineuse modèle *Medicago truncatula*, est l'une des plus utilisées.

b. Les rhizobia

Le terme *Rhizobium* vient du grec *Rhiza*=racine et *bios*=vie et provient de la première appellation au XIX^{ème} siècle du premier genre bactérien vivant dans le sol et pouvant nodules les légumineuses.

Les rhizobiums sont des bactéries de la famille des Rhizobiacées, ils sont Gram négatifs, non sporulant, et ils apparaissent en forme de bâtonnets et sont mobiles. Ce sont des

microorganismes aérobies de petite taille dont la température de croissance optimale est de 28°C et le pH optimal est de 6 à 7.

I.3.3. Les différentes étapes de la formation de la nodosité

La mise en place de la symbiose Légumineuse-Rhizobium se caractérise par 3 étapes principales (1) une étape de reconnaissance entre la plante et la bactérie par l'intermédiaire d'un dialogue moléculaire, (2) l'infection de la racine par la bactérie et le développement du primordiaux nodulaire, et (3) la formation d'un nouvel organe racinaire, la nodosité au sein duquel se déroule la réduction et la fixation de l'azote atmosphérique.

I.3.3.1. Dialogue moléculaire entre les deux partenaires

L'établissement de la symbiose Rhizobium/légumineuse (SRL) est possible grâce à un dialogue moléculaire complexe entre la plante hôte et la bactérie (**Gough et Cullimore, 2011; Oldroyd, 2013**).

En condition de carence azotée, les racines des légumineuses secrètent un cocktail de composés phénoliques, les flavonoïdes, qui va être reconnu par certaines espèces de Rhizobia (**Redmond et al., 1986**).

La reconnaissance des flavonoïdes a deux effets sur les bactéries : un effet d'attraction en direction de la plante et l'activation de l'expression des gènes nod (nodulation). L'expression de ces gènes va permettre la production et la sécrétion de lipo-chito-oligosaccharides (LCOs) particuliers, appelés communément Facteurs de Nodulation (ou facteurs Nod, FNs). Chaque souche de *Rhizobia* produit des FNs qui lui sont propres et qui permettent une reconnaissance spécifique entre les deux partenaires de la symbiose (**Caetano-Anolles et Gresshoff, 1991; Dénarié, Debellé et Promé, 1996; Spaink, 2000**).

I.3.3.2. Infection et développement de la nodosité

Les facteurs Nod sont perçus par la plante jusqu'à de très faibles concentrations (10⁻⁴ nM) grâce à la présence de récepteurs membranaires LysM-RKs (**Gough & Cullimore, 2011**).

Ces récepteurs sont composés d'un domaine LysM extracellulaire (pour la liaison au facteur Nod) et d'un domaine kinase cytosolique (**Arrighi et al., 2006**).

La perception des facteurs Nod induit chez la plante de multiples réponses, et permet simultanément l'invasion bactérienne et le développement des divisions cellulaires, ce qui aboutit à la formation du primordiaux nodulaire.

I.3.3.3. Le fonctionnement de la nodosité

Un des changements clefs de la différenciation des bactéries en bactéroïdes est l'induction des gènes codant les différentes sous-unités de la nitrogénase (**figure 4**), l'enzyme responsable de la fixation de l'azote atmosphérique.

La réduction d'une molécule d'azote en deux molécules d'ammoniaque par cette enzyme consomme 16 ATP. L'énergie nécessaire au maintien de l'activité nitrogénase est assurée par la plante qui fournit aux bactéroïdes des composés carbonés issus de la photosynthèse (**Zirilli et al., 1988**).

En contrepartie, l'ammoniaque produit par le bactéroïdes est assimilé dans le cytoplasme de la cellule végétale.

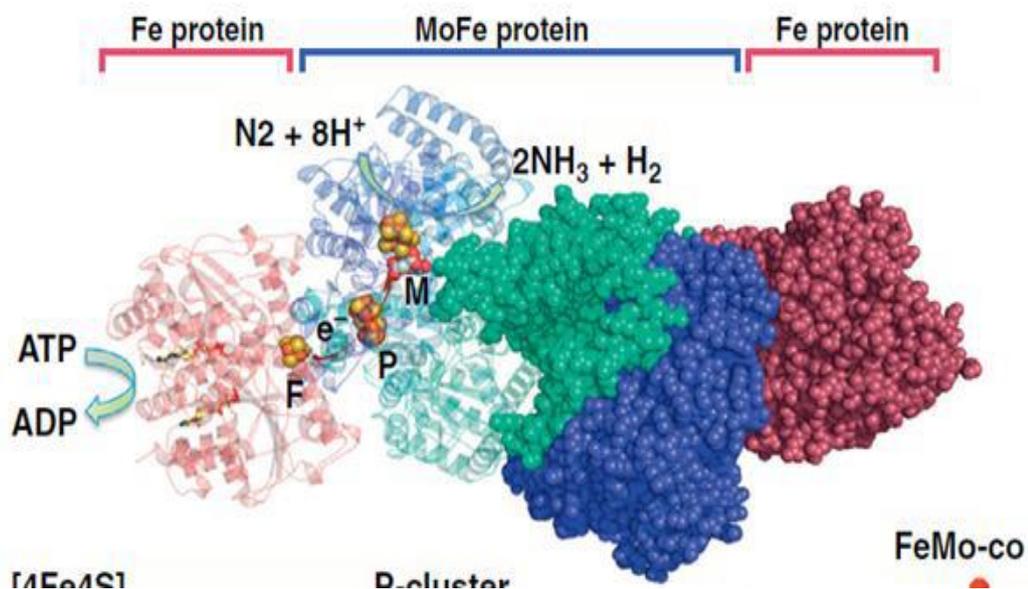


Figure 4 : Structure du complexe enzymatique de la nitrogénase.

I.3.3.4. La sénescence du processus symbiotique

La fixation de l'azote est optimale pendant les premières semaines post-infection. Au-delà de cette période, l'efficacité de fixation diminue et conduit à la mise en place progressive d'un processus de sénescence (**Puppo et al., 2005**).

La sénescence nodositaire est caractérisée par un changement de couleur de la zone de fixation qui passe du rose au vert ce changement est en partie provoqué par la dégradation de leghémoglobine. Le développement de la sénescence peut être divisé en deux étapes.

Une première étape, au cours de laquelle les bactéroïdes sont dégradés et les symbiosomes fusionnent. Cette fusion forme des compartiments de lyse qui permettent la remobilisation des nutriments (**Vasse et al., 1990; Limpens et al., 2009**).

La deuxième étape correspond à la dégradation des cellules végétales hôtes (Vasse et al., 1990; Van de Velde et al., 2006).

Le phénomène de dégradation cellulaire, s'accompagne de la mise en place d'une forte activité protéolytique via l'activation de gènes codant pour des protéases telles que les cystéines protéases (CP), des nucléases, des lipases et différents types d'hydrolases (Van de Velde et al., 2006).

La diminution de l'expression des gènes codants pour les leghémoglobines, ainsi que la diminution de la teneur en Lb, sont corrélées à la mise en place du processus de sénescence (Puppo et al., 2005).

Il est possible de déclencher artificiellement la sénescence nodositaire en appliquant soit un stress carboné en privant les plantes de lumière, soit en provoquant un stress nitrate en augmentant les quantités de nitrate dans le milieu (Matamoros et al., 1999).

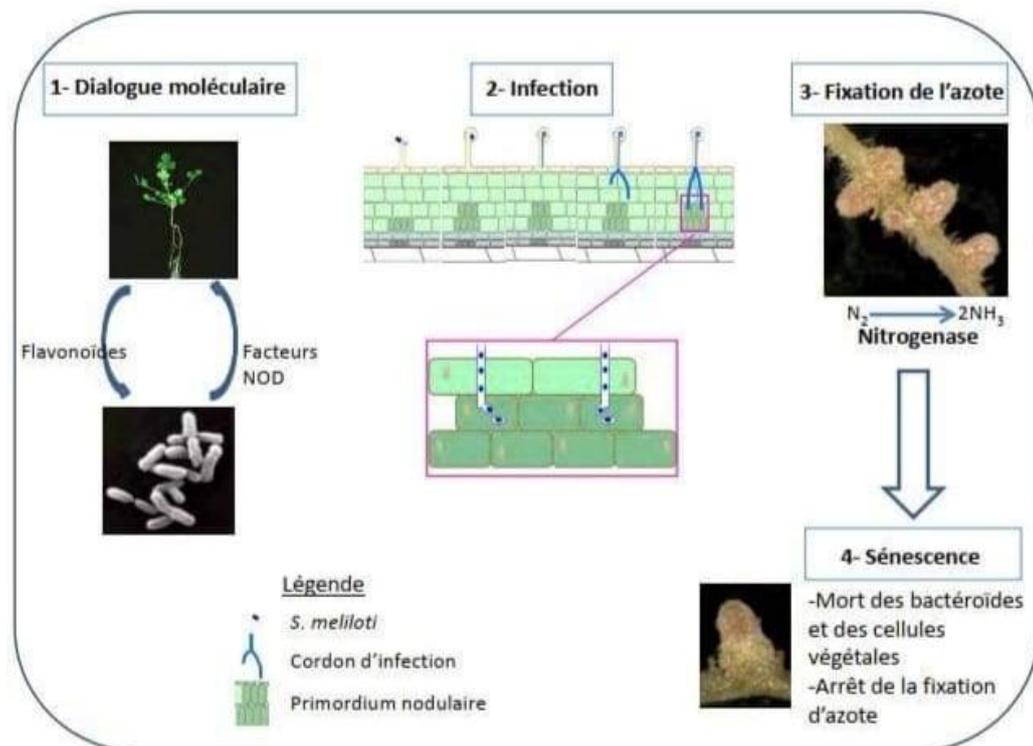


Figure 5 : Représentation schématique des principales étapes de la symbiose *Rhizobium*-légumineuses.

I.3.4. Le cycle d'azote

Il se compose principalement de trois processus de base (figure 6) :

(a).La fixation de l'azote moléculaire N_2

Correspond à la conversion de l'azote atmosphérique en azote utilisable par les plantes et les animaux.

L'ammonium (NH_4^+) se transforme en ammoniac (NH_3) gazeux, dans les sols où le pH est élevé

(b) La nitrification :

correspond à la transformation des produits de la fixation (NH_3 et NH_4^+) par l'action des bactéries (Costa et al., 2006) **nitrosantes** (Nitrosomonas, Nitrosococcus) en nitrites (NO_2^-) les nitrites sont transformés en nitrates (NO_3^-) par l'action de bactéries nitrifiantes (Nitrobacter, Nitrococcus).

(c) La dénitrification :

Il s'agit d'une réaction de réduction des nitrates (NO_3^-), par l'intermédiaire de bactéries anaérobies (Pseudomonas) qui transforme la matière organique, en une forme gazeuse N_2 qui retourne dans l'atmosphère et ainsi le cycle recommence.

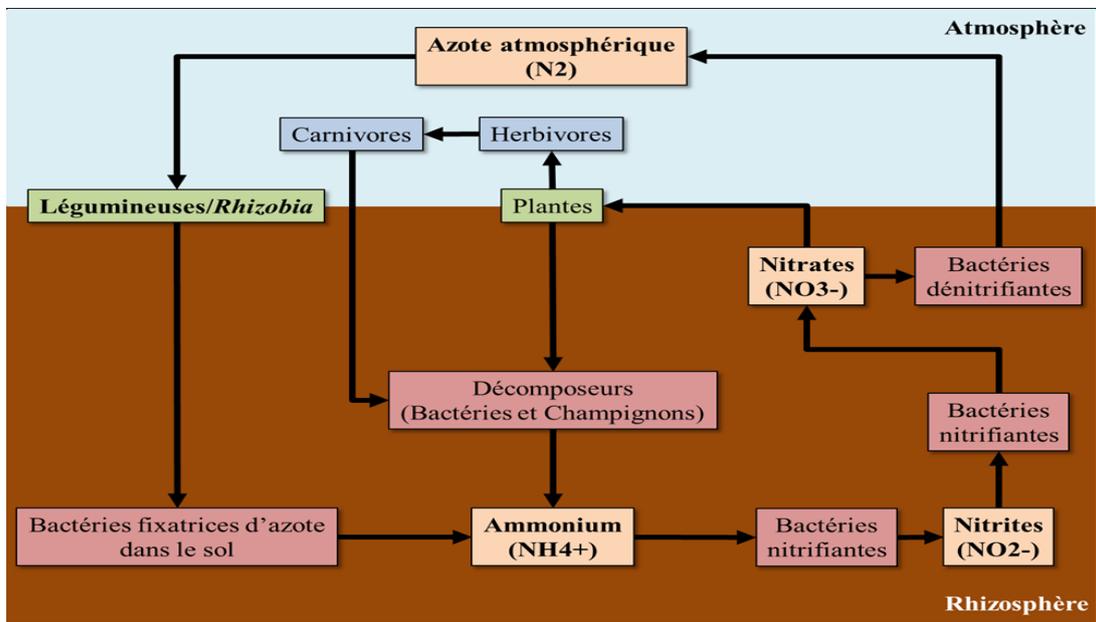


Figure 6 : Schéma du cycle de l'azote terrestre.

I.4. La Bio-informatique

I.4.1. Définition de la Bio-informatique

La bio-informatique correspond à l'utilisation de l'informatique pour stocker et analyser les données de la biologie moléculaire (Beroud, 2011).

On peut également la définir comme étant la discipline émergente de la recherche qui se place à l'interface de la biologie et de l'informatique (Jongennel, 2000).

Elle est constituée par l'ensemble des concepts et des techniques nécessaires à l'interprétation de l'information génétique et structurale donc c'est le décryptage de la 'bio-information' et donc en anglais 'Computational biology' (Claverie, 2013).

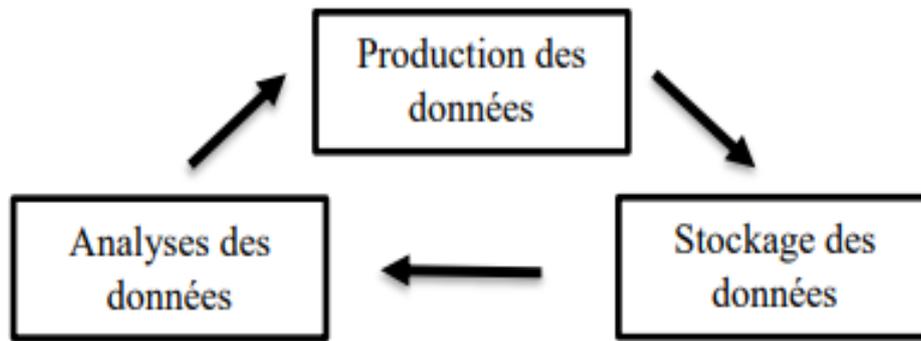


Figure 7 : Schématisation de l'approche in silico de la biologie (Bonsai, 2014).

I.4.2 Les bases et les banques de données

Les techniques récentes de biologie moléculaire génèrent une quantité massive de données qui ne se sont pas gérables par les techniques de publication traditionnelle (Chiapello et al, 2000).

Pour cela, Il existe deux types de structure de stockage :

- ✓ Celles qui offrent des informations plutôt hétérogènes.
- ✓ Celles qui correspondent à des données plus homogènes d'espèces précises.

Il est fréquent d'appeler les premières 'banques de données' et les secondes 'bases de données', mais cette distinction n'est pas très connue par les non-biologistes (Boukadida et al, 2004).

- ❖ Les banques de données ou banques généralistes : Ensemble de données relatives à un domaine d'étude et organisées pour être offertes à la consultation des utilisateurs. Par exemple, NCBI : National Center for Biotechnology Information (banque américaine créée en 1988) et EMBL : Européen Moléculaire Biologie Laboratoire (banque européenne qui existe depuis 1980) sont les grandes banques de séquences généralistes. Leur mission est de rendre publiques les séquences qui ont été déterminées. On trouve également une expertise biologique directement liées aux séquences traitées (Gardaring, 2003).
- ❖ Les bases de données : Ensemble de données relatives à un domaine d'étude spécifique. Elles ont pour but de recenser des familles de séquences ayant des caractéristiques biologiques communes comme les gènes identiques issus d'espèces différentes. Elles peuvent aussi regrouper des classes spécifiques de séquences comme les vecteurs de clonage ou la séquence complète d'un même génome (Gardaring, 2003).

I.4.3. Les domaines de la Bioinformatique

I.4.3.1. Phylogénie

La phylogénie c'est l'étude des relations de parentés entre différents êtres vivants en vue de comprendre l'évolution des organismes vivants. Il existe plusieurs méthodes pour la construction des arbres phylogéniques tels que la parcimonie. Par le biais de la bioinformatique ces méthodes sont devenues faciles grâce aux logiciels et les outils informatiques (Chollier, 2006).

I.4.3.2. Génomique structurale et fonctionnelle

La génomique est une branche de la biologie qui porte sur l'étude des génomes, qui constituent le support moléculaire des caractères héréditaires des êtres vivants. Grâce à la bioinformatique et au séquençage, l'analyse en génomique est passée d'un gène au génome complet (Caboche, 2006).

L'annotation puis l'intégration de ces données vise à donner un sens à l'information stockée. C'est le passage du génome séquencé au génome annoté et aux bases de données secondaires tel que 'GENE' (NCBI). Pour une recherche de gène candidat.

I.4.3.3. Biotechnologie

Les évolutions des biotechnologies dans les domaines de l'environnement, de la santé, de la recherche et de l'industrie ont besoin de diagnostics précis, d'analyses fiables et de productions innovantes, conduites avec un souci d'assurance qualité et de sécurité biologique à l'aide de la bio-informatique et ces applications et ces logiciels.

I.4.3.4. Analyse de la transcription

La transcription des gènes est précisément régulée, chaque gène est exprimé à un niveau spécifique en fonction du type cellulaire, du tissu, du temps, des conditions intra- et extracellulaires. Le transcriptome est défini comme l'ensemble de toutes les molécules d'ARN transcrites à partir d'un génome il est utilisé pour comprendre les mécanismes de régulation transcriptionnelle, ainsi que pour certaines applications médicales.

I.4.4. *Medicago truncatula* et bio-informatique

A l'heure actuelle, de nombreux outils bioinformatiques sont disponibles chez *M. truncatula*. Ils concernent les prédictions et annotations de gènes, d'EST (Expressed Sequence Tag) et de protéines, mais permettent aussi la réalisation d'alignements de séquences sur les bases de données spécifiques de *M. truncatula* ou des légumineuses.

Tableau 1 : Quelques bases de données bioinformatiques disponibles pour *M. truncatula*.

Données	Nom
Gènes (prédictions et annotations)	MTGD: The <i>Medicago truncatula</i> Genome Database (http://www.MedicagoGenome.org)
EST	<i>TIGR</i> :TheInstituteforGenomicResearch(http://www.jcvi.org)
Alignements nucléotidique	Université d'Oklahoma (https://www.ou.edu/)
EST, ORF, cartes Génétiques	<i>LIS</i> : <i>Legume Information System</i> (https://legumeinfo.org/)

Chapitre II :

Matériel et méthodes

II.1 Matériel

Le matériel d'étude sur lequel nous avons travaillé est constitué d'une collection de données d'expression de gènes issues de puces de type Affymetrix disposée online sur la base de données MtGEA.

L'utilisation de puces à oligonucléotides de type Affymetrix permet de quantifier en théorie l'abondance absolue de chaque ARNm transcrit. Le principe de fabrication de ces puces est la synthèse in situ de centaines de milliers de séquences distinctes d'oligonucléotides sur une matrice de verre, fixées par photolithographie et chimie combinatoire. Les ARNm de l'échantillon à analyser sont amplifiés, fragmentés et marqués par un système de couplage biotines-treptavidine pour l'hybridation sur la puce. Chaque gène est représenté sur la puce par une quinzaine de sondes constituées d'oligonucléotides courts de 20-25 bases couvrant différentes portions spécifiques du gène. Chaque ensemble d'oligonucléotides est appelé ProbSet auquel un identifiant est associé (ProbSet ID).

Une estimation directe du niveau d'expression de chaque gène est obtenue en calculant le signal moyen sur l'ensemble des sondes représentant le gène.

II.2 Méthodes

La méthodologie de travail débute par la recherche de ProbSet ID et la collecte des données d'expression des gènes relative à la symbiose et la sénescence.

Ensuite, la collecte et le classement des données d'expression de ces gènes dans 8 conditions différentes. Et se termine par une analyse statistique par **R**.

II.2.1 Recherche du Batch Download by Experiment dans la base de données MtGEA

Le projet de séquençage du génome de *Medicago truncatula* a été lancé en 2003 par la Samuel Roberts Noble Foundation et l'université d'Oklahoma. Les données sont rassemblées et mise à jour de façon continue par l'institut "**J. Craig Venter**".

Le serveur Web *Medicago truncatula* Gene Expression Atlas (MtGEA) est une plate-forme centralisée pour l'analyse du transcriptome de *M. truncatula*.

Actuellement, il héberge les données d'expression génique de 156 GeneChip du génome de *M. truncatula* dans 64 expériences différentes, couvrant un large éventail de conditions environnementales et de stades de développements. Le serveur permet des analyses multiples de données de transcription et fournir une série d'informations sur les gènes, y compris les différents types d'annotations et des liens vers la séquence de génome, qui aident les utilisateurs à formuler des hypothèses sur la fonction des gènes.

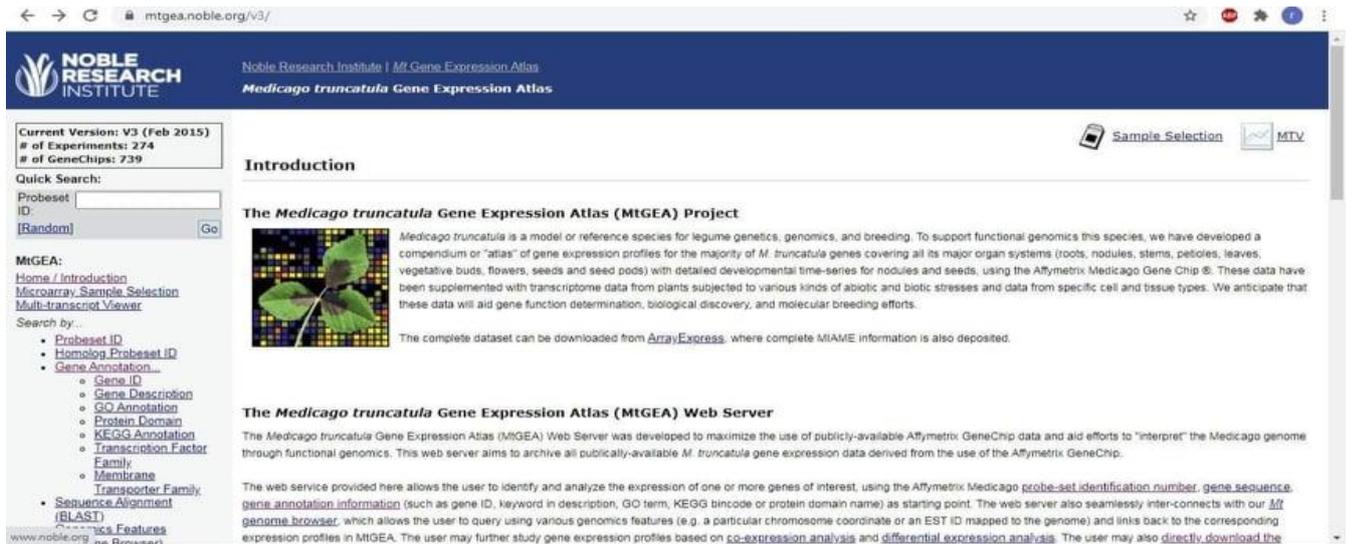


Figure 8 : Interface du moteur de recherche MtGEA.

Parmi les fonctionnalités qu'offre le moteur de recherche *MtGEA* est la page d'accueil 'Batch Download by Experiment'. Cette page répertorie les données d'expression par expérience (appelées «puce», «diapositive» ailleurs) et le téléchargement de toutes les expressions.

Les expériences sont regroupées selon leurs sources de données (auteurs).

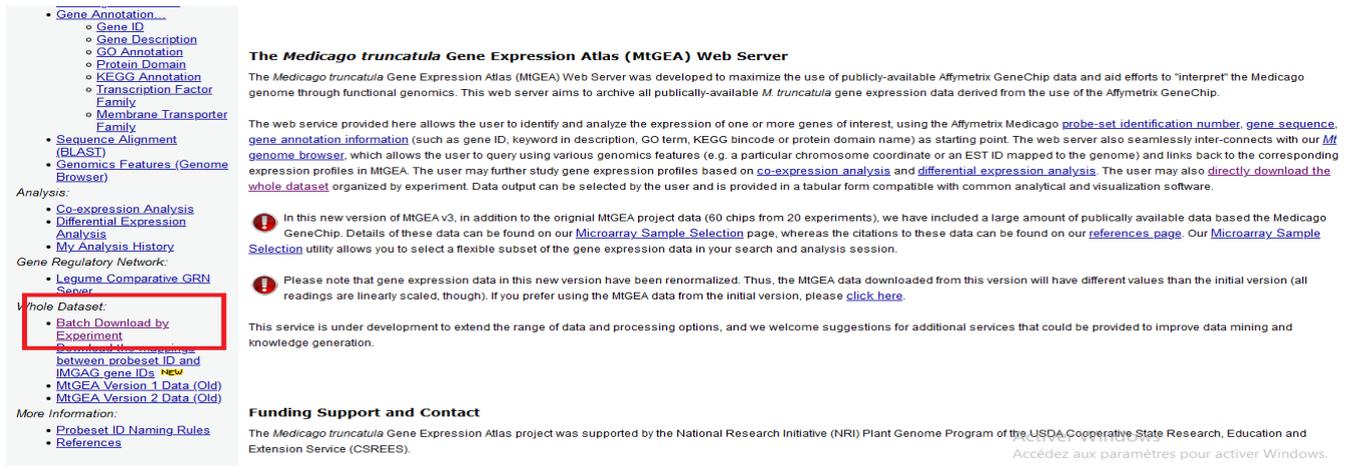


Figure 9 : l'option Batch Download by Experimentsur le moteur de recherche MtGEA.

The screenshot shows the 'Batch Download by Experiment' interface. At the top, it displays 'Current Version: V3 (Feb 2015)', '# of Experiments: 274', and '# of GeneChips: 739'. Below this is a 'Quick Search' section with a 'Probeset ID' input field and a '[Random]' button. A navigation menu on the left includes links like 'Home / Introduction', 'Microarray Sample Selection', and 'Multi-transcript Viewer'. The main content area is titled 'Batch Download by Experiment' and contains a table of experiments. Each row represents an experiment with columns for 'Organ', 'Treatment Type', 'Treatment', and 'Download' options (checkboxes for individual gene chips and a 'Mean' button).

Organ	Treatment Type	Treatment	Download Options
Hypocotyl A17 10C 35C-day1	physical treatments	A17 10 P1	<input type="checkbox"/> Hyptl_A17_10C_35C_day1_1 <input type="checkbox"/> Mean (Hyptl_A17_10C_35C_day1)
Hypocotyl A17 10C 35C-day2	physical treatments	A17 10 P1	<input type="checkbox"/> Hyptl_A17_10C_35C_day2_1 <input type="checkbox"/> Mean (Hyptl_A17_10C_35C_day2)
Hypocotyl F83 10C 35C-day1	physical treatments	F8310 P1	<input type="checkbox"/> Hyptl_F83_10C_35C_day1_1 <input type="checkbox"/> Mean (Hyptl_F83_10C_35C_day1)
Hypocotyl F83 10C 35C-day2	physical treatments	F8310 P1	<input type="checkbox"/> Hyptl_F83_10C_35C_day2_1 <input type="checkbox"/> Mean (Hyptl_F83_10C_35C_day2)
Hypocotyl A17 10C 50C-day1	physical treatments	A17 10 P2	

Figure 10: Interface du Batch Download by Experiment.

II.2.2. Collecte des données d'expression des gènes sur la page d'accueil Batch Download by Experiment

Par l'utilisation de la fonctionnalité batch downloadexperiment on a sélectionné 08 conditions (KNO3, Sinorhizobium ,Mycorisation 6h ,Mycorization 24 h ,Nod4 day ,Nod 14 day , RootMyc 3wk infection , Nodule 20) différents relient au deux phénomène la symbiose et la senescence.

Tableau 2 : Conditions sélectionnées pour la Symbiose et la Senescence.

Conditions	Leurs traitements
KNO3	Traitement chimique et biologique: nitrate de potassium 10mM
SinorhizobiumMeliloti	Traitement biologique(racine du plante avec bactérie sinorhizobium)
Nod 4Day/14Day	Traitement biologique après l'inoculation rhizobienne (grumeaux)
Mycorisation 6h /24h	Traitement chimique et biologique (racine du plante de type sauvage traitées avec des sMyco-LCO Symbiose)
Root-Myc 3wek infection	Traitement biologique : interaction plante micro-organisme nodulation
Nodule 20 Day	Traitement chimique: plante âgé de 20 jours

Tableau 3 : Les données des gènes relatives aux 08 conditions.

Probeset	RT_Myc_3wks_infection (Root Myc 3wk infection)	RT_J5_10mM_KNO3 (Root J5 10mM KNO3)	Nod_4dpi (Nod 4dpi)	Nod_14dpi (Nod 14dpi)
RT_WT_sMyc_LCOs_6h (Root WT sMyc-LCOs 6h)	RT_WT_sMyc_LCOs_24h (Root WT sMyc-LCOs 24h)	Root_A17_5d_seedling_1dpi_Sinorhizobium (Root A17 5d seedling - 1dpi)		
Sinorhizobium Nodule_20d (Nodule (20d))				
Mtr.1.1.S1_at	25.7997	73.6634	88.8242	131.567
Mtr.1.1.S1_s_at	125.486	629.244	617.26	700.63
Mtr.10.1.S1_at	30.6789	10.3547	8.6724	9.76679
Mtr.100.1.S1_at	9.63896	11.0095	10.8219	8.97696
Mtr.1000.1.S1_at	15.5059	14.5822	14.5549	16.8103
Mtr.10000.1.S1_at	67.1122	341.988	397.59	428.954
Mtr.10001.1.S1_at	1570.61	192.211	100.241	67.6953
Mtr.10001.1.S1_s_at	38.4944	517.394	405.706	319.143
Mtr.10002.1.S1_at	102.791	24.9683	18.5877	19.4479
Mtr.10003.1.S1_at	572.05	6.09666	8.24435	2230.37
Mtr.10004.1.S1_at	41.9729	9.92198	8.34574	9.42182
Mtr.10005.1.S1_at	15.5544	122.461	296.636	158.489
Mtr.10006.1.S1_at	86.8717	55.7217	41.9869	38.9101
Mtr.10007.1.S1_at	46.1087	300.506	130.467	82.503
Mtr.10008.1.S1_at	22.456	26.1703	29.6854	31.0054
Mtr.10009.1.S1_at	509.807	3069.69	870.51	321.331
Mtr.1001.1.S1_at	15.1641	7.11582	6.48684	6.4408
Mtr.10010.1.S1_at	377.704	217.982	246.177	218.688
Mtr.10011.1.S1_at	22.6541	7.34594	6.62672	6.9997
Mtr.10012.1.S1_at	24.2779	8.08657	7.71566	9.97764
Mtr.10013.1.S1_at	61.5426	1579.33	2551.39	1742.04
Mtr.10014.1.S1_at	36.2216	8.68847	8.71966	7.46403
Mtr.10015.1.S1_at	394.359	37.958	41.5199	40.8783
Mtr.10016.1.S1_at	42.1076	462.82	332.736	262.116
Mtr.10017.1.S1_at	15.1382	22.4816	20.0242	22.2097
Mtr.10018.1.S1_at	48.2144	418.81	159.076	248.1
Mtr.10019.1.S1_at	1052.54	100.107	199.505	95.882
Mtr.1002.1.S1_at	24.2779	8.08657	7.71566	9.97764
Mtr.10020.1.S1_at	124.437	33.8964	34.4276	31.0893
Mtr.10021.1.S1_at	12.9038	15.0332	9.46152	9.2926
Mtr.10022.1.S1_at	741.187	197.053	198.027	178.708
Mtr.10023.1.S1_at	316.611	133.932	635.018	160.227
Mtr.10024.1.S1_at	11.4302	8.78698	7.46299	7.40095
Mtr.10025.1.S1_at	16.8251	28.8297	31.8249	15.1478
Mtr.10026.1.S1_at	114.799	35.2195	39.7007	62.8167
Mtr.10026.1.S1_s_at	746.249	221.647	168.423	249.271
Mtr.10027.1.S1_at	10.856	10.6781	10.6237	11.6923
Mtr.10028.1.S1_at	1187.38	131.801	140.371	133.375

Les collectes des données sont classe sur l'Excel sous former de tableaux.Ces valeurs vont être ensuite triées du plus grand au plus petit.

Tableau 4 : Les données des gènes classé sur la feuille de calcul Excel.

Probeset	RT_J5_10mM_KNO3	RT_WT_sMyc_LCOs_6h	RT_WT_sMyc_LCOs_24h	RT_Myc_3wks_infection	Nodule_20d
Mtr.2238.1.S1_s_at	21195.7	17685.7	27781.1	28024.1	6465.88
Mtr.39601.1.S1_s_at	23867.5	20073.2	19344.4	13374.2	19474.5
Mtr.27459.1.S1_s_at	20909.1	19480.2	16706.9	16883.4	13390
Mtr.8445.1.S1_s_at	20501	22314.3	21055.3	15204.1	23559.4
Mtr.42967.1.S1_at	19651	26040.6	4103.31	1798.47	18876.5
Mtr.36255.1.S1_at	19458.8	16156.2	13335.4	12210.1	17254.2
Mtr.40144.1.S1_at	19248.7	16438.8	11877.8	11981.9	14632.4
Mtr.24829.1.S1_at	18531.7	20312.8	10601.8	8833.03	17876.3
Mtr.42926.1.S1_s_at	18291.2	13417.9	11989.2	13912.5	18876.5
Mtr.10392.1.S1_at	18184.1	16057.6	14600.5	12631.9	11787.8
Mtr.5687.1.S1_s_at	17992.6	14537.8	13871.7	9844.93	16710.5
Mtr.49523.1.S1_at	17841.7	17013.8	14976.9	13060.4	15254.8
Mtr.10361.1.S1_at	17567.3	16057.6	10432.6	3289.6	12974.2
Mtr.8505.1.S1_at	17528	19360	12725.1	11416.2	7882.51
Mtr.12276.1.S1_at	17436.7	15147.1	18021.6	12009.3	13369.7
Mtr.34487.1.S1_s_at	17191.3	15992.1	17566.5	12514.4	14235.4
Mtr.10297.1.S1_at	17090.8	19526	17050.6	14790.7	16025.3
Mtr.23143.1.S1_at	17051	14711.9	11236.2	7900.1	10181.5
Mtr.12333.1.S1_a_at	16809.8	16834.5	14273.5	8608.92	14885.7
Mtr.40065.1.S1_at	16585.7	13596.3	13963.8	11600.3	11728.2

Enfin, Pour l'étude statistique, on a classé les données obtenues dans l'Excel et on a pris les 20 premiers.

II.3 Etude statistique par R

R est un système d'analyse statistique et graphique créé par Ross Ihaka et Robert Gentleman. C'est à la fois un logiciel et un langage qualifié de dialecte du langage S créé par AT&T Bell Labs. Comporte de nombreuses fonctions pour les analyses statistiques et les graphiques ; ceux-ci sont visualisés immédiatement dans une fenêtre propre et peuvent être exportés sous divers formats. Les résultats des analyses statistiques sont affichés à l'écran, certains résultats partiels peuvent être sauvegardés à part, exportés dans un fichier ou utilisés dans des analyses ultérieures.

-L'utilisation de **R** présente plusieurs avantages :

- C'est un logiciel multiplateforme, qui fonctionne aussi bien sur des systèmes Linux, Mac OS X ou Windows.
- C'est un logiciel libre, développé par ses utilisateurs et modifiable, gratuit.
- **R** est de plus utilisé dans tous les secteurs scientifiques, y compris dans le domaine des analyses d'enquêtes et, plus généralement, des sciences sociales.

-L'interface de **R** est divisée en quatre quadrants :

- Le quadrant supérieur gauche est dédié aux différents fichiers de travail.
- Le quadrant inférieur gauche correspond à ce que l'on appelle la console.
- Le quadrant supérieur droit permet de connaître la liste des objets en mémoire ou environnement de travail, ainsi que l'historique des commandes saisies dans la console.
- Le quadrant inférieur droit affiche la liste des fichiers du répertoire de travail, les graphiques réalisés et la liste des extensions disponibles, l'aide en ligne et un Viewer utilisé pour visualiser certains types de graphiques au format web.

De nombreuses solutions existent (distance maximum, moyenne, Ward...). Chacune d'elle produira un dendrogramme différent. Cependant, à l'usage, on privilégiera le plus souvent la méthode de Ward. Cette méthode se distingue de toutes les autres en ce sens qu'elle utilise une analyse de la variance approchée afin d'évaluer les distances entre groupes. En résumé, cette méthode cherche à minimiser l'inertie intra-classe et à maximiser l'inertie interclasse afin d'obtenir des classes les plus homogènes possibles.

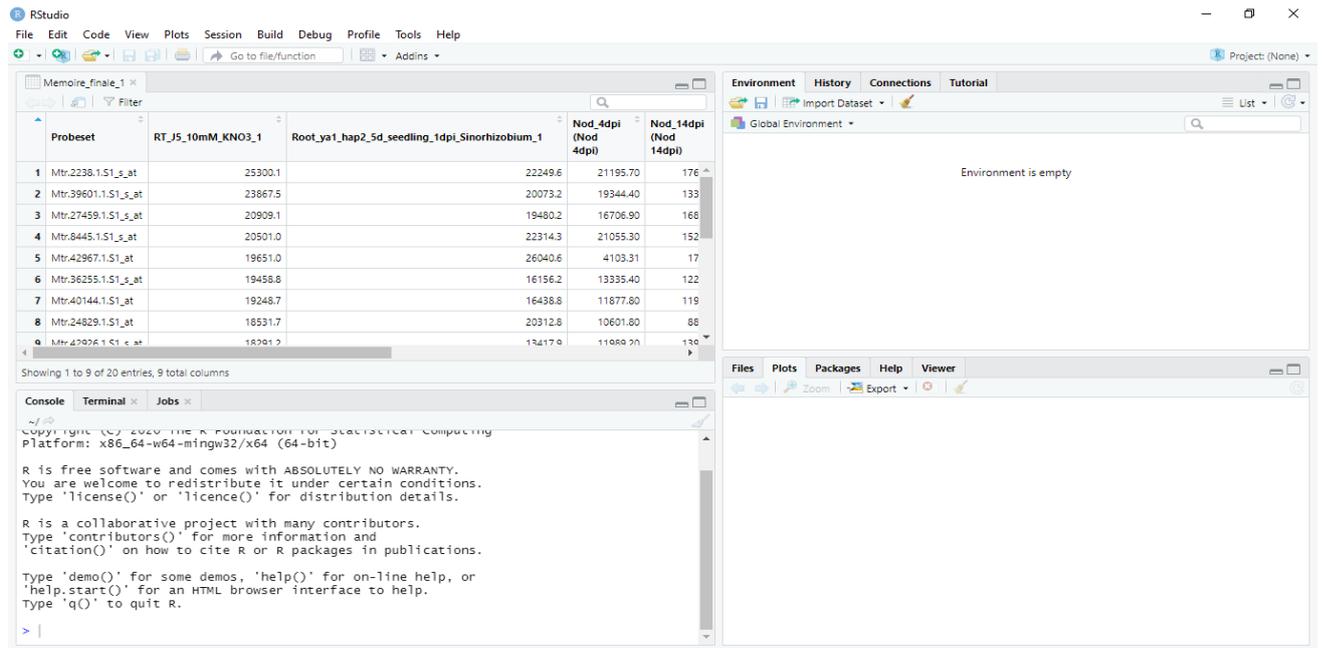


Figure 11 : L’interface de R en quatre quadrants.

Tableau 5 : Les formules des commandes utilisées dans R (Barnier et al, 2017).

<-	Opérateur d’assignation. Il prend une valeur quelconque à droite et la place dans l’objet indiqué à gauche.
Set data	Le répertoire des documents.
Setwd	Le répertoire de travail peut être modifié par cette fonction.
Read The Data	Lire le document/fichier.
Nuage des points /plot	Représentations graphiques, notamment grâce à des extensions dédiées. Pour l’heure contentons-nous d’un premier essai à l’aide de la fonction générique, afficher le dendrogramme obtenus.
Normalisation	S’accorder sur une procédure, sur une règle, qui sera utilisée par le plus grand nombre.
Apply	La fonction est liée à la notion de sous-population.
Scale	La fonction est centrée.
~	Le symbole correspond au répertoire utilisateur système, dont l’emplacement dépend du système d’exploitation.
Dist	Calcule de la distance.
hc / hclust	permet de réduire significativement le temps de calcul.
Length	Nous renvoie le nombre d’éléments du vecteur.
Mean	Nous donne la moyenne des éléments du vecteur.
Var	Nous donne la variance des éléments.
Labels	Permet de modifier les noms dans le facteur résultat.

Tableau 6 : présente les données de l'expression des gènes sur le logiciel

	Probeset	RT_J5_10mM_KNO3_1	Root_ya1_hap2_5d_seedling_1dpi_Sinorhizobium_1	Nod_4dpi (Nod 4dpi)	Nod_14dpi (Nod 14dpi)	RT_WT_sMyc_LCOs_6h (Root WT sMyc-LCOs 6h)	RT_WT_sMyc_LCOs_24h (Root WT sMyc-LCOs 24h)	RT_Myc_3wks_infection (Root Myc 3wk infection)	Nodule_20d (Nodule (20d))
1	Mtr.2238.1.S1_s_at	25300.1	22249.6	21195.70	17685.70	27781.10	28024.1	6465.860	20014.60
2	Mtr.39601.1.S1_s_at	23867.5	20073.2	19344.40	13374.20	16951.60	22240.4	19474.500	11437.50
3	Mtr.27459.1.S1_s_at	20909.1	19480.2	16706.90	16883.40	20248.10	20424.9	13390.000	18206.40
4	Mtr.8445.1.S1_s_at	20501.0	22314.3	21055.30	15204.10	22202.40	23559.4	9153.480	15179.60
5	Mtr.42967.1.S1_s_at	19651.0	26040.6	4103.31	1798.47	18876.50	17566.9	483.918	1908.52
6	Mtr.36255.1.S1_s_at	19458.8	16156.2	13335.40	12210.10	13349.80	17254.2	4710.750	12922.00
7	Mtr.40144.1.S1_s_at	19248.7	16438.8	11877.80	11981.90	14632.40	16159.9	14999.400	10882.60
8	Mtr.24829.1.S1_s_at	18531.7	20312.8	10601.80	8833.03	17876.30	17511.0	10295.400	7619.22
9	Mtr.42926.1.S1_s_at	18291.2	13417.9	11989.20	13912.50	18876.50	17566.9	9954.220	10397.40
10	Mtr.10392.1.S1_s_at	18184.1	16057.6	14600.50	12631.90	11787.80	18424.8	3317.150	13676.90
11	Mtr.5687.1.S1_s_at	17992.6	14537.8	13871.70	9844.93	16710.50	15937.0	1972.360	7718.30
12	Mtr.49523.1.S1_s_at	17841.7	17013.8	14976.90	13060.40	15254.80	15027.4	1792.620	14411.10
13	Mtr.10361.1.S1_s_at	17567.3	16057.6	10432.60	3289.60	12974.20	14541.1	9057.840	2459.71
14	Mtr.8505.1.S1_s_at	17528.0	19360.0	12725.10	11416.20	7882.51	11912.2	6904.800	11933.20
15	Mtr.12276.1.S1_s_at	17436.7	15147.1	18021.60	12009.30	13369.70	14073.2	7807.880	10043.20
16	Mtr.34487.1.S1_s_at	17191.3	15992.1	17566.50	12514.40	14235.40	15176.1	4390.070	11699.40
17	Mtr.10297.1.S1_s_at	17090.8	19526.0	17050.60	14790.70	16025.30	15187.8	6228.830	14938.40
18	Mtr.23143.1.S1_s_at	17051.0	14711.9	11236.20	7900.10	10181.50	12549.6	170.088	6429.72
19	Mtr.12333.1.S1_a_at	16809.8	16834.5	14273.50	8608.92	14885.70	14720.9	5872.970	11862.20
20	Mtr.40065.1.S1_s_at	16585.7	13596.3	13963.80	11600.30	11728.20	11769.3	4115.120	9849.56

Chapitre III :
Résultats et discussion

III. Résultats et discussion

III.1. Résultats

Après l'analyse statistique menée sur la totalité des données relatives aux 08 conditions pour la symbiose et senescence, l'application nous a permis de répartir les 20 gènes utilisés pour l'étude en plusieurs clades, l'obtention des résultats sous forme de phylogramme.

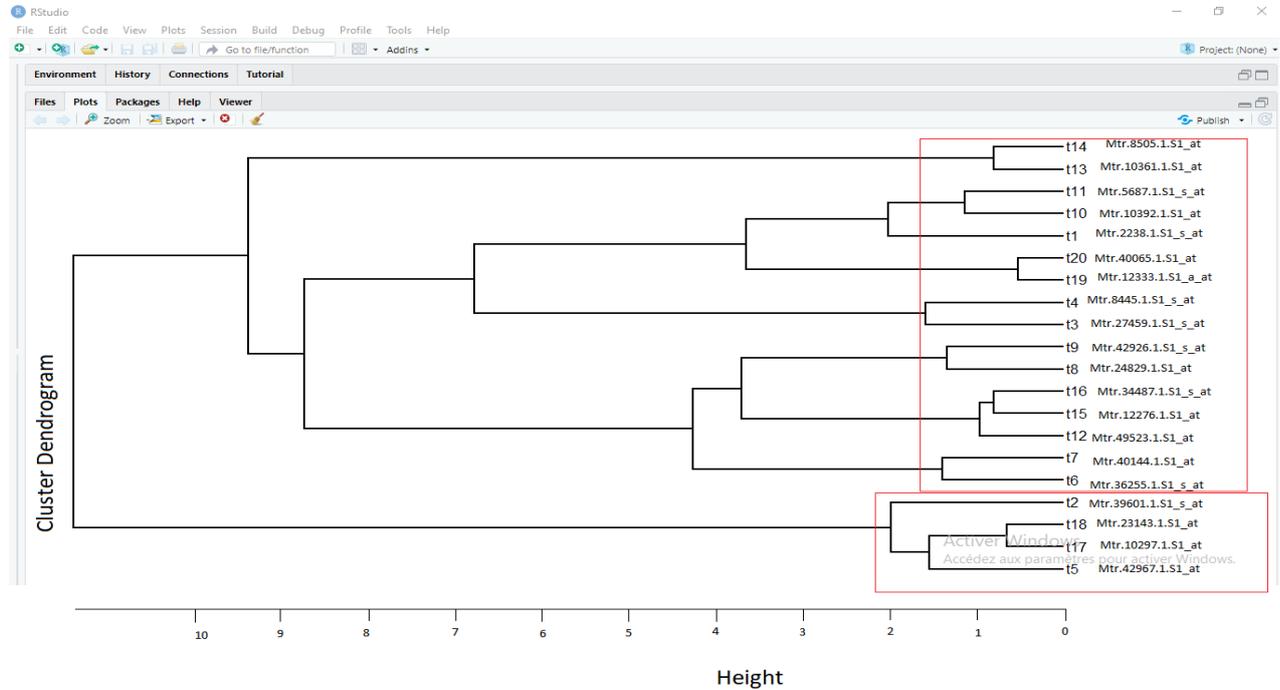


Figure 12 : Cluster ING de la totalité des gènes sur logiciel R.

Selon un pourcentage de dissimilarité inférieur ou égale à 9,5, le phylogramme est divisé en 2 clades, chaque clade regroupe un nombre de gènes (Tableau).

Tableau 7: le nombre de gènes dans chaque clade.

Clades	Nombres de gène
Calade 1	16
Calade 2	4

L'observation du phylogramme nous a permis de voir chaque clade est divisé en plusieurs sous-clades:

Le 1^{er} clade

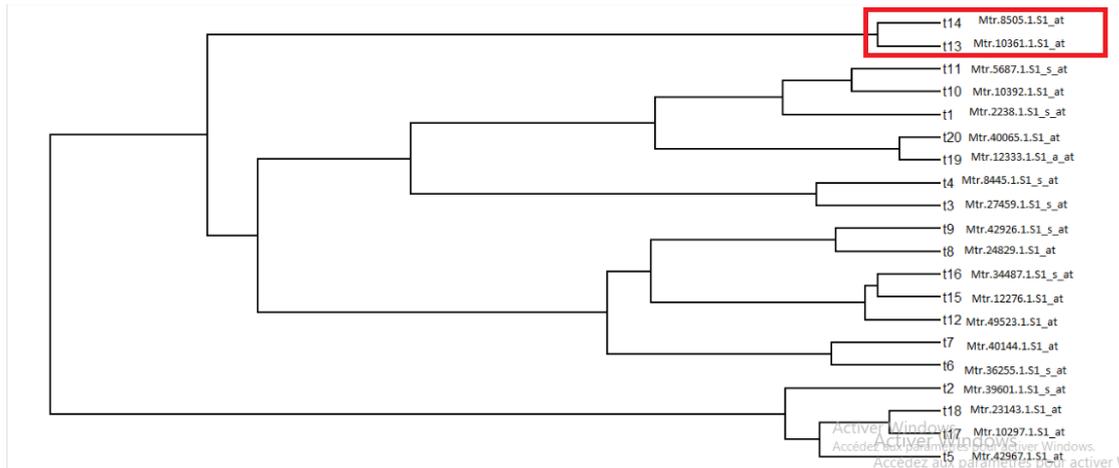


Figure 13 : présente le premier sous- clade dans la phylogramme.

Tableau 8: présente l'identifiant et la description du premier sous-clade.

Probset ID	Identifiant	Description
Mtr.8505.1.S1_at	Medtr2g081770.1	embryonicabundant-likeprotein
Mtr.10361.1.S1_at	Medtr8g045490.1	pathogenesis-relatedproteinbet V I familyprotein

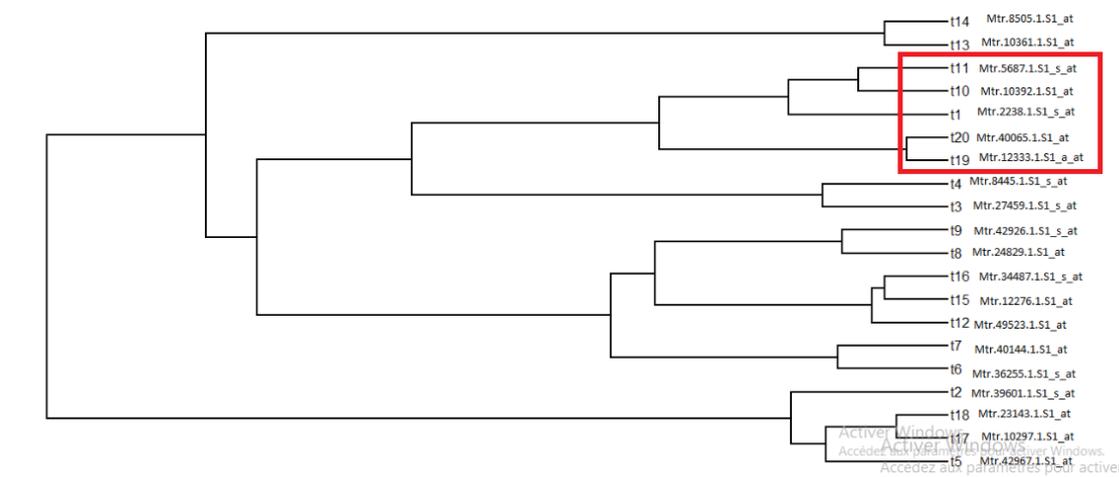


Figure 14 : présente le deuxième clade dans la phylogramme

Tableau 9: présente l'identifiant et la description de le deuxième sous- clade

Probset ID	Identifiant	Description
Mtr.5687.1.S1_s_at	Medtr4g101280.1	Lipide Transfer protéine

Mtr.10392.1.S1_at	Medtr2g094270.2 Medtr2g094270.1	major intrinsicprotein (MIP) familytransporter major intrinsicprotein (MIP) familytransporter
Mtr.2238.1.S1_s_at	Medtr4g029600.1	extensin- likerepeatprotein
Mtr.40065.1.S1_at	Medtr3g085490.1 Medtr3g085490.2	60S ribosomal protein L17A 60S ribosomal protein L17A
Mtr.12333.1.S1_a_at	Medtr3g114190.1 Medtr3g113780.3 Medtr3g113780.1 Medtr3g113780.2 Medtr3g114180.1	transmembraneprotein, putative hypotheticalprotein hypotheticalprotein hypotheticalprotein

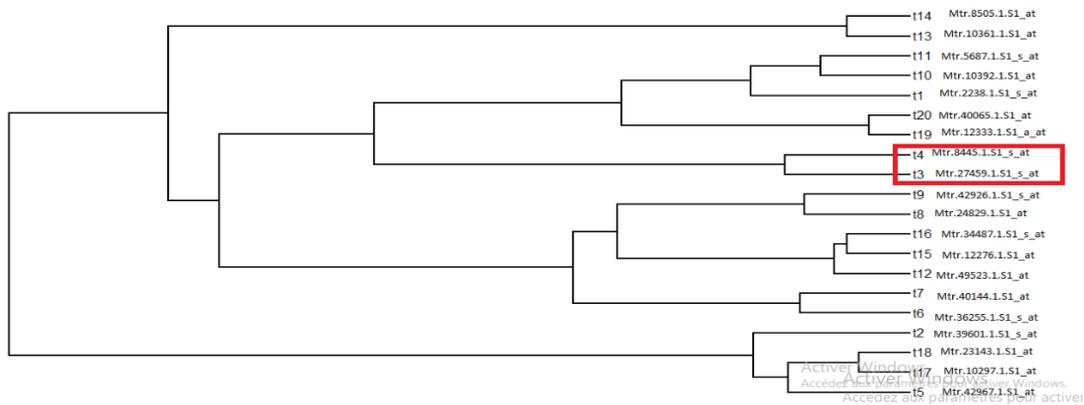


Figure 15: présente le troisième sous- clade dans la phylogramme.

Tableau 10 : présente l'identifiant et la description du troisième sous-clade.

Probst ID	Identifiant	Description
Mtr.8445.1.S1_s_at	Medtr1g116947.1	pollen proteinOle E I- likeprotein
Mtr.27459.1.S1_s_at	Medtr2g436620.1	translationally- controlledtumor- likeprotein

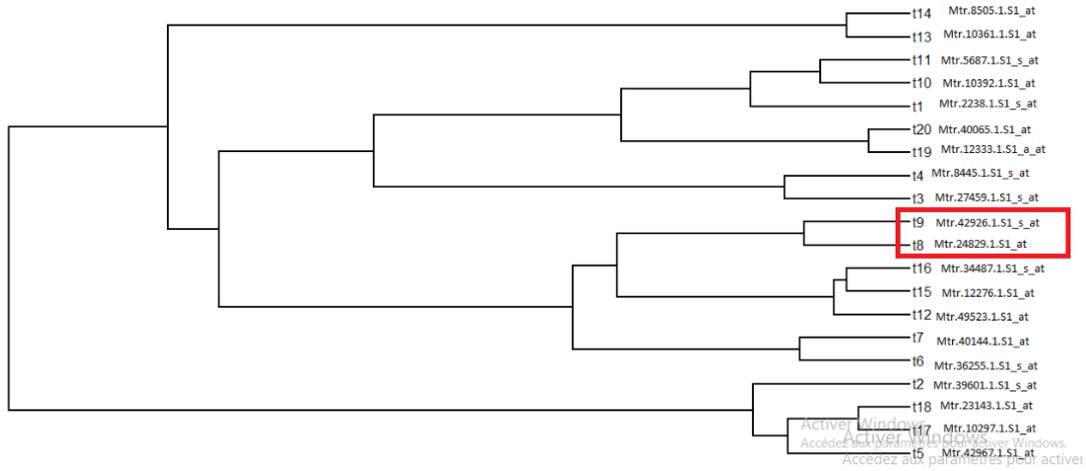


Figure 16: présente le quatrième sous- clade dans la phylogramme.

Tableau 11: présente l'identifiant et la description de le quatrième sous- clade.

Probset ID	Identifiant	Description
Mtr.42926.1.S1_s_at	Medtr7g112860.2 Medtr7g112860.1	dormancy/auxinassociatedprotein dormancy/auxinassociatedprotein
Mtr.24829.1.S1_at	No information availab.	No information available.

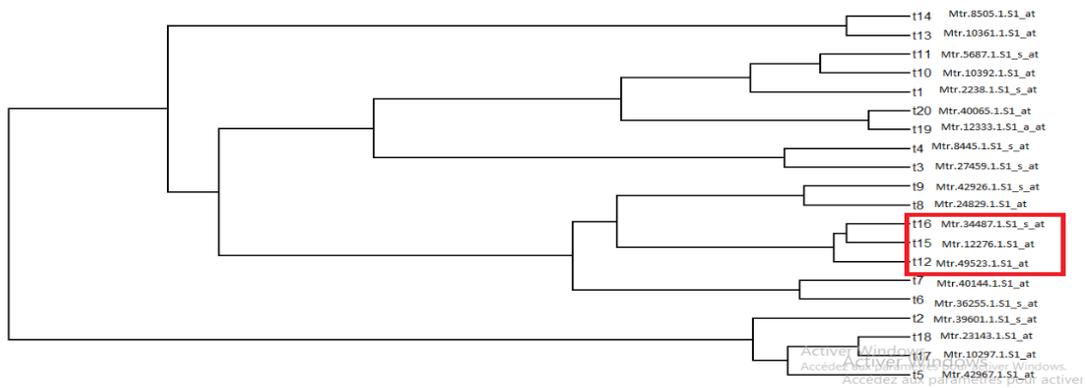


Figure 17: présente le cinquième sous- clade dans la phylogramme.

Tableau 12 : présente l'identifiant et la description du cinquième sous-clade.

Probset ID	Identifiant	Description
Mtr.34487.1.S1_s_at	Medtr5g018940.2 Medtr5g018940.1 Medtr8g052030.1	40S ribosomal S4-like protein 40S ribosomal S4-like protein 40S ribosomal S4-like protein
Mtr.12276.1.S1_at	Medtr4g132110.1 Medtr4g132110.2	cationicperoxidase cationicperoxidase
Mtr.49523.1.S1_at	Medtr1g098220.1	ribosomal protein S13P/S18e

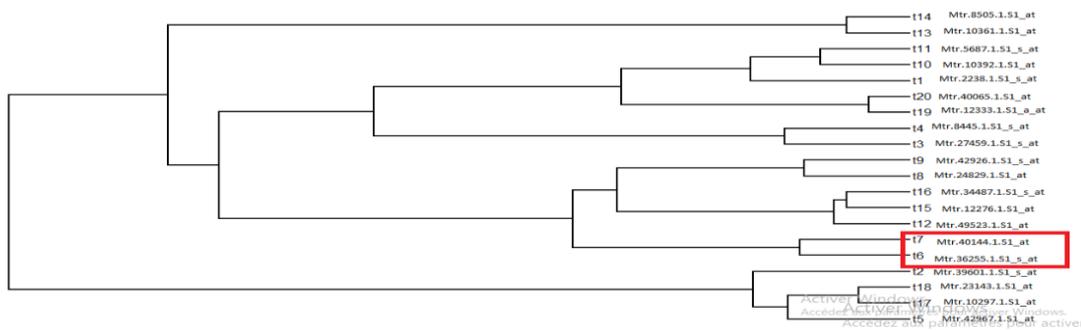


Figure 18: présente le sixième sous-clade dans la phylogramme.

Tableau 13 : présente l'identifiant et la description du sixième sous-clade.

Probset ID	Identifiant	Description
Mtr.40144.1.S1_at	Medtr2g025780.1	hypotheticalprotein
Mtr.36255.1.S1_s_at	Medtr3g070210.1 Medtr3g070210.2 Medtr3g070210.4 Medtr3g070210.3	major intrinsicprotein (MIP) familytransporter major intrinsicprotein (MIP) familytransporter major intrinsicprotein (MIP) familytransporter major intrinsicprotein (MIP) familytransporter

• Le 2^{eme} clade

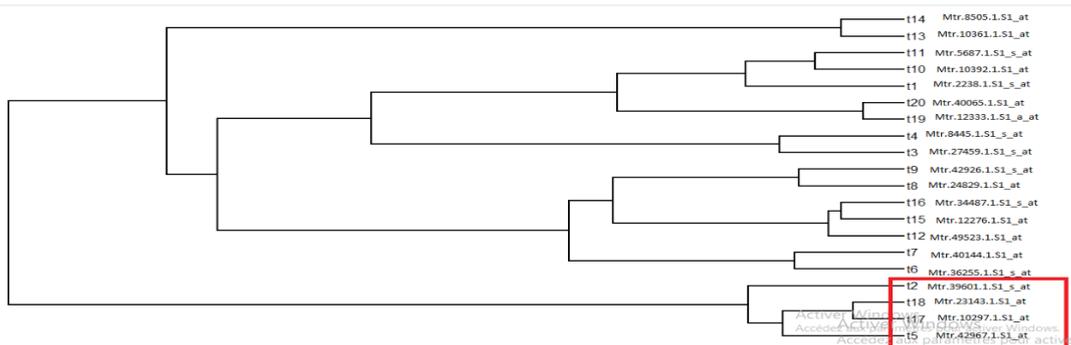


Figure 19 : présente le septième clade dans la phylogramme.

Tableau 14: présente l'identifiant et la description de le septième clade.

Probsset ID	Identifiant	Description
Mtr.39601.1.S1_s_at	Medtr7g103030.1	tonoplastintrinsicprotein
Mtr.23143.1.S1_at	Medtr3g067430.1	albumin I
Mtr.10297.1.S1_at	Medtr4g075290.1	peptidyl-prolyl cis-transisomerase

III.2 Discussion

Les plantes soumises à des conditions environnementales extrêmes développent des réactions qui sont contrôlées à la base par les voies de signalisation qui à leur tour sont conditionnées par les modifications au niveau de l'expression des gènes.

Ainsi, lors de la présente étude qui a porté sur l'analyse *in silico* des gènes en relation avec la nutrition d'azote chez le modèle biologique *Medicago truncatula*, nous avons pu définir, selon la fonction, plusieurs groupes de gènes :

- Les gènes codant pour des protéines de types LEA (Late embryogenesis abundant) protein : détectés dans les poils des racines, qui sont des organes fragiles mais essentiels non seulement pour l'absorption d'eau mais aussi pour établir une symbiose entre les légumineuses et les rhizobiums nécessaires à la fixation de l'azote. Les nodules contiennent des protéines LEA; cependant, tous les groupes de protéines LEA ne semblent pas être présents dans les différents types de nodules. Une analyse plus détaillée est nécessaire, où les niveaux de transcription de la protéine LEA de différents groupes sont déterminés en tenant compte de différentes étapes au cours du développement des nodules. (**Olvera-Carrillo et al. 2010**).

-Les gènes codant pour des protéines de types pathogenesis-related protein VI family protein

La croissance et la survie des plantes sont toujours influencées par plusieurs facteurs, dont les stress abiotiques et biotiques. Les plantes répondent à ces facteurs en induisant leur mécanisme de défense qui comprend l'expression de plusieurs effecteurs, récepteurs, molécules de signalisation et de protection.

L'une des protéines les plus couramment induites au cours du mécanisme de défense des plantes est la protéine liée à la pathogénèse (PR). L'accumulation de protéines PR fait partie intégrante des réponses immunitaires innées des plantes lors d'attaques d'agents pathogènes ou dans des conditions de stress abiotique. Les protéines PR s'accumulent non seulement localement dans la feuille infectée, mais sont également associées au développement d'une réponse hypersensible (HR) ou d'une résistance systémique acquise (SAR) contre l'infection par des champignons, des bactéries et des virus (**Guérin, 2019**).

-Un troisième groupe qui code pour des protéines de transport comme lipide transporter, Heavy metal-associated domain protein et les NTF2 ou les Nuclear transport factor 2 qui remplissent une multitude de fonctions. À l'origine identifiés comme des transporteurs impliqués dans des processus de désintoxication, ils ont été démontrés plus tard pour la croissance des organes, la nutrition des plantes, le développement des plantes, la réponse aux contraintes abiotiques, la

résistance aux agents pathogènes et l'interaction de la plante avec son environnement **(Thomashow et al, 1997)**.

-Les bactéries de la famille des *rhizobiacées* peuvent infecter les racines des légumineuses entraînant la formation de structures appelés nodosités ou nodules. Par ces nodules, la plante hôte (la légumineuse) offre un micro habitat exceptionnellement favorable à la bactérie tout en lui procurant des substrats carbonés provenant de la photosynthèse. Le processus de la fixation, lui-même, consiste en la réduction de l'azote atmosphérique N₂ sous forme ammoniacale. Cette réaction est catalysée par un complexe enzymatique appelé Nitrogénase d'origine bactérienne **(Downie 2005)**. Cette association à bénéfice réciproque entre la légumineuse et les bactéries est appelée symbiose fixatrice de l'azote atmosphérique.

-Le site de fixation symbiotique est le nodule, le seul organe localisé sur la racine qui fonctionne pour la survie des bactéries et l'activité de la Nitrogénase. La formation des nodules est le résultat d'un dialogue moléculaire entre le microsymbiont et la plante hôte **(Foucher et Kondorosi 2000; Limpens et Bisseling 2003)**. L'interaction commence avec la colonisation de jeunes poils absorbants par les rhizobia et un échange de molécules-signaux. Les bactéries reconnaissent des flavonoïdes qui sont sécrétées par la plante hôte. Ces molécules induisent la production de facteurs NOD par les rhizobia **(Oldroyd 2001)**.

Les facteurs NOD sont des lipo-chitoooligosaccharides (LCO) **(Dénarié et al. 1996)** émis par la bactérie, à l'origine de la reconnaissance spécifique entre les deux symbiotes et du déclenchement du programme d'organogenèse nodulaire chez le végétal par une cascade d'expression de gènes spécifiques. Les facteurs NOD agissent essentiellement sur deux types de cellules au niveau de la racine: Les cellules épidermiques et corticales. Au niveau des cellules épidermiques, les facteurs NOD induisent une dépolarisation de la membrane plasmique, une oscillation du flux de Ca²⁺, une induction de l'expression de gènes spécifiques et une modification de la croissance polaire des poils absorbants formant une structure dite en «crosse de berger» qui enferme les *rhizobia* **(Esseling et al. 2003)**.

Conclusion et perspectives

Conclusion et perspectives

La réaction des plantes face à un état de stress implique plusieurs mécanismes physiologiques accompagnés par des modifications dans le profil d'expression des gènes. Dans le but de rechercher des gènes impliqués dans la nutrition d'azote chez le modèle biologique *Medicago truncatula*, nous avons réalisé une analyse in silico de données de transcriptomique issues de puces de type Affymetrix.

Au terme de notre travail, les résultats obtenus montre que ces gènes peuvent être réparti, suivant leurs rôles qu'ils jouent, en plusieurs groupes :

les gènes codant pour des protéines de types LEA, des protéines de pathogène, des gènes codant pour des facteurs de transcription (TF), des gènes qui codent pour des protéines réceptrices, des protéines de transport, les facteurs NOD ainsi que d'autres gènes qui codent pour différentes fonctions. La consultation de la bibliographie relative à ces différents gènes montre qu'il s'agit de gènes dont l'implication dans la réponse au stress d'azote a été démontrée dans plusieurs organismes. Cette réponse met en place des mécanismes de tolérance qui impliquent parfois les mêmes voies de signalisations, conduisant toutes aux mêmes stratégies de protection de la cellule.

Les résultats obtenus laissent entrevoir de nombreuses perspectives afin de les confirmer comme par exemple l'utilisation de la PCR en temps réel (QPCR) et l'utilisation des techniques de génétique inversée comme les ARN interférant (RNAi).

Références bibliographiques

***Ané JM, Zhu H, Frugoli J (2008)** Recent Advances in *Medicago truncatula* Genomics. *Int J Plant Genomics* 2008, 256597.

***Aubert G, Morin J, Jacquin F, Loridon K, Quillet MC, Petit A, Rameau C, Lejeune-Hénaut I, Huguet T, Burstin J (2006)** Functional mapping in pea, as an aid to the candidate gene selection and for investigating synteny with the model legume *Medicago truncatula*. *TheorAppl Genet*, 112: 1024-1041.

•**Béroud, S., Yon, K., Dessen, M., Gantois, M., Guillaume, C., & Kesselman, D. (2011).** La loi du 20 août 2008 et ses implications sur les pratiques syndicales en entreprise: sociologie des appropriations pratiques d'un nouveau dispositif juridique .

* **Boisson-Dernier, A., Chabaud, M., Garcia, G., Bécard, G., Rosenberg, C. et Barker, D.G. (2001).** *Agrobacterium rhizogenes*-transformed roots of *Medicago truncatula* for the study of nitrogen-fixing and endomycorrhizal symbiotic associations. *Mol. Plant-Microbe Interact.* 14 (6) : 695-700.

***Bonsai.(2014).** **Bioinformatique et données biologiques [en ligne],**
http://cristal.univlille.fr/~noe/enseignement/m1-genpro/Cours/bioinfo_bio1-2x3.pdf.

***Bonsai. (2014).** **Bioinformatique et données biologiques [en ligne],**
http://cristal.univlille.fr/~noe/enseignement/m1-genpro/Cours/bioinfo_bio1-2x3.pdf.

***Boisson-Dernier, A., Chabaud, M., Garcia, G., Bécard, G., Rosenberg, C. et Barker, D.G. (2001).** *Agrobacterium rhizogenes*-transformed roots of *Medicago truncatula* for the study of nitrogen-fixing and endomycorrhizal symbiotic associations. *Mol. Plant- Microbe Interact.* 14 (6) : 695-700.

***BoukadidaJawer, Denis Julien, 2004.** Te Bioinformatique [En ligne]
http://deptinfo.unice.fr/twiki/pub/Linfo/PlanningDesSoutenances20032004/Boukadida_Denis.pdf.

- ***Bordeleau, L.M. et Prévost, D. (1994).**Nodulation and nitrogen fixation in extreme environments. *Plant Soil* 161:115-125.
- ***Caetano-Anolles, G. and Gresshoff, P. M. (1991)** ‘Plant Genetic Control of Nodulation’, *Annual Review of Microbiology*, 45(1), pp. 345–382. doi: 10.1146/annurev.mi.45.100191.002021.
- * **Choi HK, Kim D, Uhm T, Limpens E, Lim H, Mun JH, Kalo P, Penmetsa RV, Seres A, Kulikova O, Roe BA, Bisseling T, Kiss GB, Cook DR (2004)** A Sequence- Based Genetic Map of *Medicago truncatula* and Comparison of Marker Colinearity with *M. sativa*. *Genetics*, 166: 1463-1502.
- * **Cook DR, Kim DJ, Zhu HY, Uribe P (2000)** Plant-pathogen interactions in *Medicago truncatula*. *Grain Legumes*, 28: 20.
- Dénarié, J., Debelle, F. and Promé, J.-C. (1996)** ‘Rhizobium Lipo-Chitooligosaccharide Nodulation Factors: Signaling Molecules Mediating Recognition and Morphogenesis’, *Annual Review of Biochemistry*, 65(1), pp. 503–535. doi: 10.1146/annurev.bi.65.070196.002443.
- ***Doyle, J.J. (2003).**The Rest of the Iceberg.Legume Diversity and Evolution in a Phylogenetic Context.*Plant Physiol.* 131 (3) : 900-910.
- ***Doyle, J.J. (2003).**The Rest of the Iceberg.Legume Diversity and Evolution in a Phylogenetic Context.*Plant Physiol.* 131 (3) : 900-910.
- * **Doyle JJ, Luckow MA (2003)**The rest of the iceberg. Legume diversity and evolution in a phylogenetic context. *Plant Physiol* 131: 900-910.
- * **Doyle JJ, Luckow MA (2003)**The rest of the iceberg. Legume diversity and evolution in a phylogenetic context.*Plant Physiol* 131: 900-910.
- * **Esseling, J. J., Lhuissier, F. G., &Emons, A. M. C. (2003).**Nod factor-induced root hair curling: continuous polar growth towards the point of nod factor application. *Plant Physiology*, 132(4), 1982-1988.

- * **Foucher, F., & Kondorosi, E. (2000).** Cell cycle regulation in the course of nodule organogenesis in *Medicago*. In *The Plant Cell Cycle* (pp. 229-242). Springer, Dordrecht.
- * **Gepts, P. (2005).** Legumes as a Model Plant Family. *Genomics for Food and Feed Report of the Cross-Legume Advances through Genomics Conference*. *Plant Physiol.* 137 (4) : 1228-35.
- * **Graham, P.H. et Vance, C.P. (2003).** Legumes : Importance and Constraints to Greater Use. *Plant Physiology* 131 (3) : 872-877.
- * **Guérin, C. (2019).** Analyse des facteurs de transcription de la famille NAC chez le blé tendre (*Triticum aestivum L.*) et leur implication dans la réponse à des stress abiotiques (Doctoral dissertation).
- * **Julier B, Huguet T, Chardon F, Ayadi R, Pierre J, Prosperi JM, Barre P, Huyghe C (2007)** Identification of quantitative trait loci influencing aerial morphogenesis in the model legume *Medicago truncatula*. *TheorAppl Genet* 114: 1391-1406.
- * **Kamphuis LG, Lichtenzveig J, Oliver RP, Ellwood SR (2008)** Two alternative recessive quantitative trait loci influence resistance to spring black stem and leaf spot in *Medicago truncatula*. *BMC Plant Biology* 8: 30.
- * **Kaló, P., Gleason, C., Edwards, A., Marsh, J., Mitra, R. M., Hirsch, S., ...& Kiss, G. B. (2005).** Nodulation signaling in legumes requires NSP2, a member of the GRAS family of transcriptional regulators. *Science*, 308(5729), 1786-1789.
- * **Lahrech, N.E.H. (2011).** Etude de la variabilité chez quelques lignées *Medicago truncatula* soumises au stress salin. 214p. thèse de magistère.
- * **Lazrek-Ben Friha, F. (2008).** Analyse de la diversité génétique et symbiotique des populations naturelles tunisiennes de *Medicago truncatula* et recherche de QTL liés au stress salin. Thèse de doctorat. pp 250.
- * **Lesins, K.A. et Lesins, I. (1979).** *Genus Medicago (Leguminosae)*. Dordrecht : Springer Netherlands.

- ***Limpens, E., Franken, C., Smit, P., Willemse, J., Bisseling, T., & Geurts, R. (2003).** LysM domain receptor kinases regulating rhizobial Nod factor-induced infection. *Science*, 302(5645), 630-633.
- * **Oldroyd, G. E., Engstrom, E. M., & Long, S. R. (2001).** Ethylene inhibits the Nod factor signal transduction pathway of *Medicago truncatula*. *The Plant Cell*, 13(8), 1835-1849.
- * **Pierre JB (2008)** Recherche de déterminants génétiques de la date de floraison chez la Légumineuse modèle, *Medicago truncatula*. Thèse. 143 pages.
- ***Puppo A, Groten K, Bastian F, Carzaniga R, Soussi M, Lucas MM, De Felipe MR, 935 Harrison J, Vanacker H, Foyer CH. 2005.** Legume nodule senescence: Roles for redox and 936 hormone signalling in the orchestration of the natural aging process. *New Phytologist* 165,937 683–701.
- * **Rose RJ (2008)** *Medicago truncatula* as a model for understanding plant interactions with other organisms, plant development and stress biology: past, present and future. *Functional Plant Biol* 35: 253-264.
- ***Tivoli B, Baranger A, Sivasithamparam K, and Barbetti MJ (2006)** Annual *Medicago*: From a model crop challenged by a spectrum of necrotrophic pathogens to a model plant to explore the Nature of disease resistance. *Ann Botany* 98: 1117–1128.
- ***Thoquet P, Gherardi M, Journet EP, Kereszt A, Ane JM, Prosperi JM, and Huguet T (2002)** The molecular genetic linkage map of the model legume *Medicago truncatula*: an essential tool for the comparative legume genomics and the isolation of agronomically important genes. *Plant Biol* 2: 1-13.
- ***Van de Velde W, Guerra JCP, De Keyser A, De Rycke R, Rombauts S, Maunoury N, 1010 Mergaert P, Kondorosi E, Holsters M, Goormachtig S. 2006.** Aging in legume symbiosis. 1011 A molecular view on nodule senescence in *Medicago truncatula*. *Plant physiology* 141, 711–1012 20.

***Young, N.D., Mudge, J. et Ellis, T.H.N. (2003).** Legume genomes: more than peas in a pod. *Curr.Opin. Plant Biol.* 6 : 199-204.

* **Young, N.D., Mudge, J. et Ellis, T.H.N. (2003).** Legume genomes: more than peas in a pod. *Curr.Opin. Plant Biol.* 6 : 199-204.

* **Young, N.D., Debelle, F., Oldroyd, G.E.D., Geurts, R., Cannon, S.B., Udvardi, M.K., Bedito, V.A. et al. (2011).** The Medicago genome provides insight into the evolution of rhizobial symbioses. *Nature.*

***Zhang, Y., Fernando, W. G., Kievit, T. R. D., Berry, C., Daayf, F., &Paulitz, T. C. (2006).**Detection of antibiotic-related genes from bacterial biocontrol agents with polymerase chain reaction.*Canadian journal of microbiology*, 52(5), 476-481.

Analyse *in silico* des données de transcription et identification des gènes impliqués dans la réponse au stress abiotique (stress d'azote) chez le modèle biologique *Medicago truncatula*

Mémoire de fin de cycle pour l'obtention du diplôme de Master en Sciences Biologiques

L'objectif du présent travail est la recherche des gènes impliqués dans la réponse au stress abiotique fixation d'azote dans la racine du modèle biologique *Medicago truncatula*, par une analyse *in silico* des données de transcription issues de puces Affymetrix, et en utilisant le profil d'expression de 20 gènes différents aux huit différentes conditions. L'arbre phylogénétique tracé par le logiciel statistique R classe les 20 gènes en 2 clades; le premier clade dispose 4 gènes et le deuxième clade dispose 16 gènes. Les résultats obtenus montrent que les gènes impliqués dans cette réponse sont très diversifiés et appartiennent à plusieurs groupes comme les gènes codant pour les protéines de types LEA, Les facteurs NOD, différents facteurs de transcription (TF), des protéines réceptrices et de transport ainsi que d'autres gènes qui codent pour d'autres fonctions.

Mots clés : Stress azote , *Medicago truncatula* , L'arbre phylogénétique.

Jury d'évaluation :

Président du jury : Mr. BELKHIRI Loutfi

Prof. Université Constantine1.

Rapporteurs : Mr. TEMAGOULT Mahmoud
Mr. KELOU Kamel

M.A.A. Université Constantine1
M.A.A. Université Constantine1

Examineur : Mr. DAAS Mouhamed Skander

M.C.B. Université Constantine1

Date de soutenance : Le : 07/10/2020