



الجمهورية الجزائرية الديمقراطية الشعبية
RÉPUBLIQUE ALGÉRIENNE DÉMOCRATIQUE ET POPULAIRE

وزارة التعليم العالي والبحث العلمي
MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE SCIENTIFIQUE



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

جامعة الاخوة منتوري قسنطينة
كلية علوم الطبيعة والحياة

Département : Biologie et Ecologie végétale

قسم : بيولوجيا و علم البيئة النباتية

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

Domaine : Sciences de la Nature et de la Vie

Filière : Sciences Biologiques

Spécialité : Biologie et Génomique Végétale

Intitulé :

**Recherche de gènes impliqués dans la réponse aux stress
hydrique et salin chez le modèle biologique *Medicago
truncatula* par analyse *in silico* de données de transcription.**

Le : 20/06/2017

Présenté et soutenu par : Mr *BELOUCIF Mohamed Fath Eddine.*

M^{lle} *SAIDI SIEF Roumeissa.*

Jury d'évaluation :

Président du jury : Mr. HAMIDECHI Mohamed Abdelhafid. Prof. Université Constantine1.

Rapporteur : Mr. TEMAGOULT Mahmoud. M.A.A. Université Constantine1.

Examineurs : Mme. MAOUGAL Rim Tinhinen M.C.B. Université Constantine1.

Année universitaire 2016 - 2017

Dédicaces

*Avant de dédier ce travaille on tient d'abord à remercier 'ALLAH', le
tout puissant qui nous a permis de mener à bien ce modeste travaille,*

A la mémoire de nos parents,

Nos frères et sœurs :

Samy, Gamra et Sid ali, Oumnia,

A tous nos amis pour leur soutien et leurs encouragements,

Une sincère pensée pour ceux qui nous ont quittés

« Père et Grands-parents »

"A Dieu nous appartenons et à lui nous retournons"

*A nos professeurs de Biotechnologie et Génomique Végétale en
espérant qu'ils trouveront dans ce travail la fierté d'un savoir bien acquis.*

A nos familles respectives.

« BELOUCIF, SAIDI SIEF »

Remerciement

On dit souvent que le trajet est aussi important que la destination .Ses années ont permis de bien comprendre la signification de cette phrase toute simple.

*On tient à la fin de ce travail à remercier **Allah** le tout puissant et soyons reconnaissants pour les faveurs qu'il Nous octroyé.*
*En premier lieu, on tiens à remercier « **Mr TEMAGOULT Mahmoud** » qui a encadré ce travail il s'est toujours montré à l'écoute et très disponible tout au long de la réalisation de ce mémoire, ainsi pour l'inspiration, l'aide et le temps qu'il a bien voulu nous consacrer et la grande patience dont il a su faire preuve et pour La confiance que l'on a reçue de sa part nous a permis de progresser.*

*On tiens également à remercier les membres de jurys, 'Mr **HAMIDECHI Mohamed Abdelhafid**' Professeur à l'université Constantine 1 non seulement pour qu'il soit président mais aussi pour tous les efforts qu'il fait pour le bien des étudiants et 'Mme **MAOUGAL Rim Tinhinen**' maître de conférence à l'Université Constantine 1 qui ont accepté de juger ce travail.*

Nos remerciement vont également à nos parents de tous les sacrifices qu'ils ont consentis pour suivre nos études dans les meilleurs conditions possibles et n'avoir jamais cessé de nous encourager tout au long de nos années d'étude.

Enfin on remercie tous les professeurs de la spécialité de Biologie et Génomique Végétale de l'université de Constantine 1.

Liste des abréviations

ADN :	Acide désoxyribonucléique.
ABA :	Acide abscissique.
ARNm :	Acide ribonucléique messenger.
BLAST:	Basic Local Alignment Search Tool.
cm :	Centimètre.
CRK :	Cysteine Rich receptor Kinase.
G :	Gramme.
C° :	Degré Celsius.
CO2 :	Dioxyde de carbone.
DREB :	Dehydration Responsive Element Binding factor.
ERF :	Ethylene Respons Factor.
EST :	Expressed sequence tag.
IMGAG:	International <i>Medicago</i> Genome Annotation Group.
Legoo:	Bioinformatics gateway towards integrative legume biology.
EMBL:	European Molecular Biology Laboratory.
LEA :	Late embryogenesis abundant protein.
LIS:	Legume Information System.
mM:	Millie mole.
MtGEA:	<i>Medicago truncatula</i> Gene Expression Atlas.
Mpb:	Mega paire de bases.
µm :	Micromètre.
Na Cl :	Chlorure de sodiumm.
NADPH :	Nicotinamide adénine dinucléotide phosphate
NCED:	Nine-cis-epoxycarotenoid dioxygenase.
NCBI :	National Center for Biotechnology Information.
NTF2 :	Nuclear Transport Factor 2.
PCR :	Amplification en chaîne par polymérase.
RefSeq :	Référence de séquence.
SLG :	Spotted Leaf Gens.
TF :	Transcription Factor.
TIGR:	The Institute for Genomic Research.
2wk:	Deux semaines.

Liste des figures

Figure 1	Aspects morphologiques de plusieurs organes de <i>Medicago truncatula</i> .	2
Figure 2	Aire de répartition de la luzerne annuelle <i>Medicago truncatula</i> .	3
Figure 3	Stades de développement de <i>Medicago truncatula</i> .	5
Figure 4	Schématisation de l'approche <i>in silico</i> de la biologie.	7
Figure 5	Les différents stress environnementaux.	10
Figure 6	La tolérance à la salinité chez quelques espèces végétales.	12
Figure 7	Effet du stress salin sur les plantes.	13
Figure 8	Effets du stress hydrique sur les plantes.	14
Figure 9	Structure de l'ABA.	16
Figure 10	Biosynthèse de l'ABA, d'après Wasilewska et al, 2008.	16
Figure 11	La fonction <i>Nicknames</i> sur <i>Legoo</i> .	19
Figure 12	Interface du moteur de recherche <i>MtGEA</i> .	20
Figure 13	Le profil d'expression du gène NCED3 sur <i>MtGEA</i> .	21
Figure 14	L'interface de R en quatre quadrants.	24
Figure 15	Clustering de la totalité des gènes relatifs au stress hydrique.	26
Figure 16	Position du clade numéro 16 sur le dendrogramme.	27
Figure 17	Clustering de la totalité des gènes relatifs au stress salin.	28
Figure 18	Position du clade numéro 6 sur le dendrogramme.	29

Liste des tableaux

Tableau 1	Quelques bases de données disponibles pour <i>M. truncatula</i> .	9
Tableau 2	Représentation des différents identifiants du gène <i>NCED3</i> .	18
Tableau 3	Représentation d'identifiant <i>probset ID</i> du gène <i>NCED3</i> de <i>Medicago truncatula</i> .	20
Tableau 4	Les données relatives au stress salin.	21
Tableau 5	Les données relatives au stress hydrique.	22
Tableau 6	Conditions sélectionnées pour le stress hydrique.	22
Tableau 7	Les données des gènes relatives au stress hydrique.	22
Tableau 8	Conditions sélectionnées pour le stress salin.	23
Tableau 9	Les données des gènes relatives au stress salin.	23
Tableau 10	Les formules des commandes utilisées dans R.	25
Tableau 11	Le nombre de gènes dans chaque clade.	26
Tableau 12	Les gènes du 16 ^{ème} clade.	27
Tableau 13	Le nombre de gènes dans chaque clade.	28
Tableau 14	Les gènes du 6 ^{ème} clade.	29

Résumé

Recherche de gènes impliqués dans la réponse aux stress hydrique et salin chez le modèle biologique *Medicago truncatula* par analyse *in silico* de données de transcription.

L'objectif du présent travail est la recherche de gènes impliqués dans la réponse au stress hydrique et au stress salin dans la racine du modèle biologique *Medicago truncatula*, par une analyse *in silico* de données de transcription issues de puces *Affymetrix*, et en utilisant un gène clé dans la biosynthèse de l'ABA comme gène de référence (*NCED3*).

Le Clustering des gènes étudiés à un pourcentage de dissimilarité de 3 classe le *NCED3* dans un clade de 8 gènes pour le stress hydrique et dans un clade de 10 gènes pour le stress salin.

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, 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 hydrique, Stress salin, ABA, *Medicago truncatula*.

Abstract

Search for genes involved in the response to water and salt stress in the biological model *Medicago truncatula* by in silico analysis of transcription data.

The aim of this work is to search for genes involved in the response to drought stress and salt stress in the root of the biological model *Medicago truncatula*, by an in silico analysis of transcription data from Affymetrix chips, and by using a key gene in the biosynthesis of ABA as a reference gene (NCED3).

Clustering of the genes studied to a dissimilarity percentage of 3 classifies NCED3 in a clade of 8 genes for drought stress and in a clade of 10 genes for salt stress.

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.

Keywords: Drought stress, salt stress, ABA, *Medicago truncatula*.

Sommaire

	<u>Page</u>
Introduction.....	1
 <i>Chapitre I : Revue bibliographique</i> 	
I.1 Présentation du modèle biologique <i>Medicago truncatula</i>	2
I.1.1 Botanique	2
I.1.2 Systématique	2
I.1.3 Origine et distribution géographique	3
I.1.4 Biologie	4
I.1.5 Intérêt	5
I.1.5.1 Intérêt économique	5
I.1.5.2 Intérêt scientifique	6
I.2 La Bioinformatique.....	7
I.2.1 Définition de la Bioinformatique	7
I.2.2 Les bases et les banques de données	8
I.2.3 Les domaines de la Bioinformatique	8
I.2.3.1 Phylogénie	8
I.2.3.2 Génomique structurale et fonctionnelle	9
I.2.3.3 Biotechnologie	9
I.2.3.4 Analyse du transcriptome	9
I.2.4 <i>Medicago truncatula</i> et bioinformatique	9
I.3 Le stress.....	10
I.3.1 Définition du mot stress	10
I.3.2 Les stress abiotiques	11
I.3.3 Le stress salin	11
I.3.3.1 Définition	11
I.3.3.2 Effets du stress salin	11
I.3.3.3 Conséquences de la contrainte osmotique	12

Sommaire

I.3.3.4	Conséquences de la contrainte ionique	12
I.3.4	Le stress hydrique	13
I.3.4.1	Définition	13
I.3.4.2	Effets du stress hydrique	13
I.3.5	Stratégies développées par les plantes	14
I.3.5.1	L'esquive	14
I.3.5.2	La tolérance	15
I.3.5.3	L'adaptation	15
I.3.6	Aspects moléculaire de la réponse des plantes aux stress	15

Chapitre II : Matériels et méthodes

II.1	Matériel	18
II.2	Méthodes.....	18
II.2.1	Recherche du <i>Probsset ID</i> correspondant au gène NCED3	18
II.2.2	Collecte des données d'expression du gène NCED3 en conditions de stress hydrique et salin.....	20
II.2.3	Collecte des données d'expression de gènes relatives au stress hydrique et stress salin.....	22
II.2.4	Etude statistique par R.....	23

Chapitre III : Résultats et discussion

III.1	Résultats.....	26
III.1.1	Partie stress hydrique	26
III.1.2	Partie stress salin.....	28
III.2	Discussion.....	30
	Conclusion.....	33
	Références bibliographique.....	34

Introduction

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 comme des conditions extrêmes de sécheresse, de température sévère, de salinité accentuée...

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 rechercher les gènes impliqués dans la réponse au stress hydrique et salin dans la racine du modèle biologique *Medicago truncatula*, par une analyse *in silico* de 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 Présentation du modèle biologique *Medicago truncatula*

I.1.1 Botanique :

Medicago truncatula, appelée communément la luzerne tronquée ou encore la luzerne sauvage, fait partie des luzernes surnommées «Medics» (Doyle et al, 2003). 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'

C'est une plante herbacée, de 35 à 65 cm, velue, ramifiée à partir de la base, à port variable, souvent prostré, parfois sérigé. Elle porte des feuilles trifoliolées Le pédoncule est plus court que le pétiole et l'inflorescence porte de une à cinq fleurs jaunes groupées en racèmes, la fleur hermaphrodite a une longueur comprise entre 5 et 8 mm, contient 10 ovules, des étamines diadelphes (réunies par leurs filets en deux faisceaux égaux). Les gousses sont cylindriques, en forme de vrilles, glabres, très dures, comprenant 2 à 8 spires jointives et serrées, aux épines recourbées, souvent perpendiculaires au plan de spires. Elles contiennent de 3 à 12 graines Le poids de 1 000 graines varie de 3.3 à 6 g (Huguet T et al, 1995), (Lesins et Lesins, 1979).

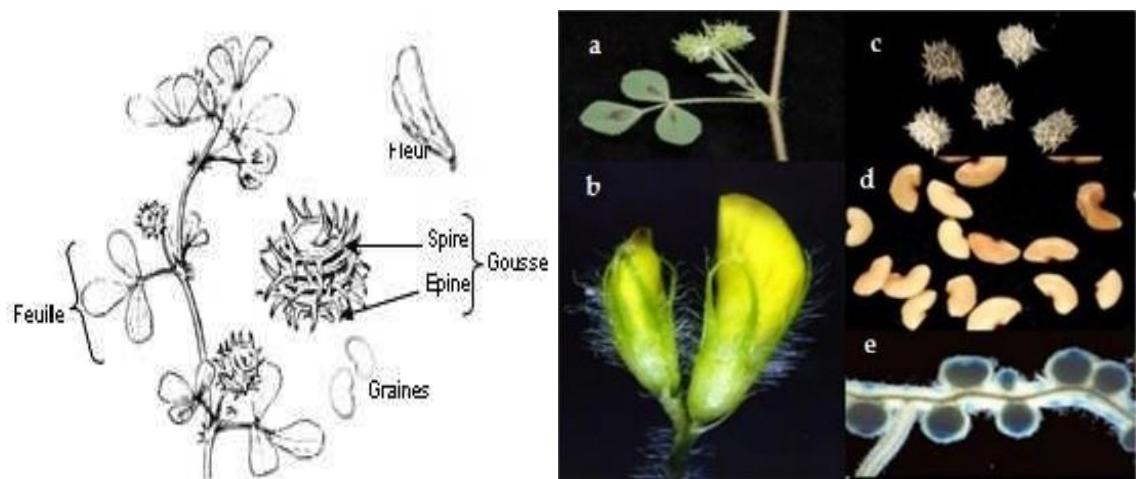


Figure 1 : Aspects morphologiques de plusieurs organes de *Medicago truncatula*

I.1.2 Systématique :

Medicago truncatula appartient à l'ordre des *Fabales* 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.1.3 Origine et distribution géographique :

C'est une espèce localisée principalement dans les zones méditerranéennes chaudes et de basse altitude. D'après **Lesins et Lesins 1979** in (**Delalande et al, 2007**), les formes les plus anciennes des espèces de *Medicago* auraient été pérennes probablement ligneuse et allogames, Il est reconnu que les aires d'origine de toutes les espèces du genre *Medicago* sont «le croissant fertile», recouvrant les pays ou régions actuelles de Turquie, Iran, Irak, sud du Caucase et le pourtour méditerranéen. Ces espèces ont ensuite conquis l'ensemble de la zone méditerranéenne et les steppes avoisinantes au cours du XIXème siècle, elles ont ensuite envahi d'autres parties du monde, en particulier les continents américains et australiens à l'occasion des différents courants de la colonisation humaine

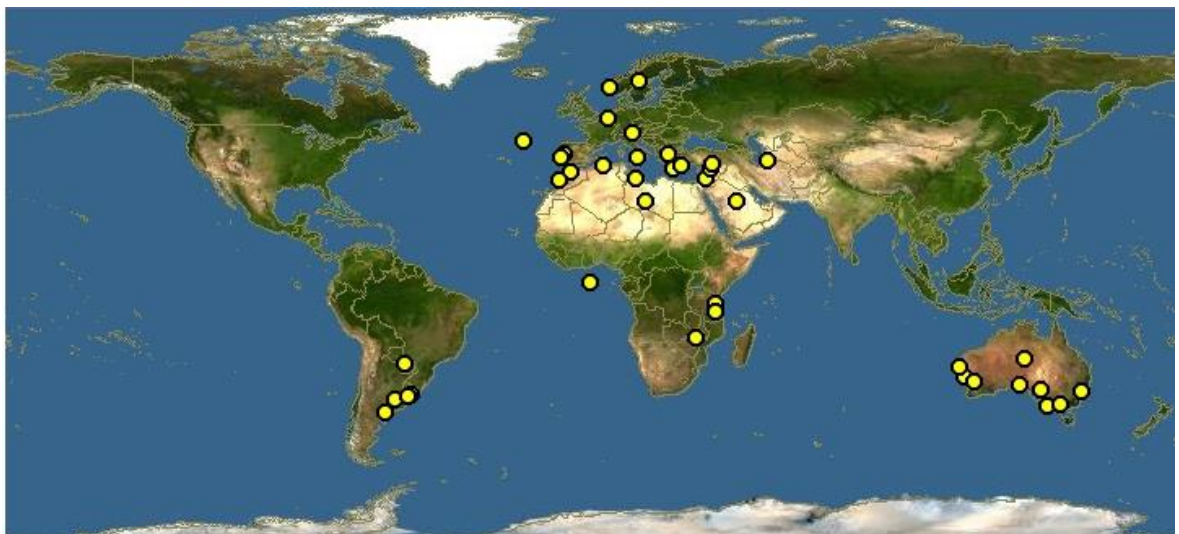


Figure 2 : Aire de répartition de la luzerne annuelle *Medicago truncatula* ([Pickering et Smith, 2006](#))

I.1.4 Biologie :

La mise en place d'un cadre d'analyse du développement de *Medicago truncatula* avec une terminologie précise d'identification des différents organes et d'une échelle de notations des stades de développement est très récente. (**Guedon et al, 2005**).

Medicago truncatula a un mode de reproduction principalement autogame avec un temps de génération de 10 à 12 semaines et les plantes peuvent produire jusqu'à trois cycles par an de graines à graines dans de bonnes conditions de culture (**Young et al, 2005**).

Elle présente une adaptabilité importante (**Mathison et al, 1980**) et peut croître sur une large gamme de milieux incluant les sols lourds, marneux ou argileux (**Lesins et Lesins 1979**).

En conditions naturelles, la croissance végétative a lieu principalement en automne, hiver et au début du printemps, c'est-à-dire sous des températures fraîches et des photopériodes courtes. (**Chabaud et al, 2006**). La plante produit tout d'abord deux cotylédons puis la première feuille unifoliolée appelée feuille 0. Viennent ensuite les feuilles trifoliolées (1 à 5 suivant l'ordre d'apparition) sur l'axe principal, organisées en rosette selon une phyllotaxie qui peut être dextre ou sénestre. A partir de la feuille 6, la disposition devient alterne sur l'axe principal. Le phyllochrone qui représente l'intervalle de temps entre l'apparition de deux feuilles étalées sur un axe, est de 2.5 à 3 jours en croissance rapide de la plante à une température moyenne de 15.3°C, soit 37 à 45 °C.J pour une feuille (**Chabaud et al, 2006**).

Les ramifications primaires sont nommées d'après leur position sur l'axe principal. Généralement des ramifications secondaires et tertiaires apparaissent à la base axillaire des feuilles de ramifications primaires et secondaires respectivement. La croissance indéterminée de *M. truncatula* se caractérise par l'aptitude de l'apex à rester végétatif. La progression séquentielle du développement reproductif s'effectue le long des axes depuis le premier nœud jusqu'à l'apex. La position du premier nœud reproducteur varie selon l'axe mais il n'existe pas de modèle qui explique son emplacement exact. Un nœud reproducteur porte de deux à trois fleurs ou inflorescences, souvent en plus d'un axe d'ordre supérieur.

La floraison se produit au printemps sous des photopériodes encore courtes et les graines mûres à la fin du printemps ou au début de l'été.

Les stades végétatifs sont définis par le nombre de feuilles trifoliolées déployées par axe ainsi que par le nombre d'entre-nœuds. Les stades reproducteurs sont quant à eux déterminés par l'identification de la position du premier nœud reproducteur pour chaque axe et le nombre de nœuds reproducteurs par axe. Cette méthode de notation a l'avantage de permettre une caractérisation phénotypique plus complète qu'une simple description à l'échelle d'un organe individuel ou de la nature des organes présents sur la plante.

C'est une plante qui produit des graines en abondance qui ont une durée de vie importante supérieure à 40 ans, et une dormance qui peut être levée facilement (**Huguet et al, 1995**).



Figure 3 : Stades de développement de *Medicago truncatula*.

I.1.5 Intérêt :

I.1.5.1 Intérêt économique :

Les populations spontanées de *Medicago truncatula* sont fréquemment exploitées dans les zones de parcours, souvent par des ovins au pâturage. La culture de cette espèce est très marginale dans le pourtour méditerranéen. A l'opposé, dans les systèmes prairiaux australiens *M. truncatula* occupe une place prépondérante avec plus de 4.5 millions d'hectares cultivés et représente ainsi l'espèce la plus cultivée parmi les médics. La principale particularité des cultivars sélectionnés en Australie réside dans leurs précocités et leurs adaptations aux conditions pédoclimatiques. (**Nair et al, 2006**).

D'autre part, *M. truncatula* limite les pertes de nitrate par lessivage, améliore la structure des sols et protège contre l'érosion (**Le Gall, 1993**). Elle est très riche en acides aminés (éléments de base des protéines végétales), mais aussi en minéraux dont le calcium, le fer, le phosphore, le zinc ou le cuivre.

I.1.5.2 Intérêt scientifique : *Medicago truncatula* comme modèle biologique pour l'étude des légumineuses

Plusieurs caractéristiques de la luzerne annuelle *Medicago truncatula* ont amené la communauté scientifique internationale qui travaille sur les légumineuses à la proposer comme plante modèle, non seulement pour l'analyse de nombreux processus biologiques et physiologiques relatifs tant au développement de la plante, mais aussi pour l'étude des effets des stress abiotiques et des symbioses ainsi que l'analyse des interactions légumineuses-parasites. **(Cuire D.R, 1999).**

Medicago truncatula offre de nombreux avantages permettant et facilitant les analyses génétiques et moléculaires. Elle est diploïde ($2n=16$), autogame et annuelle. Elle possède un génome de faible taille, 550 Mpb soit environ 4 fois celui d'*Arabidopsis thaliana* **(Cuire D.R, 1999)**. Elle présente une forte variabilité biologique pour le port, la périodicité des plantes, le nombre des graines par gousses. Par ailleurs, le temps de générations de graine à graine est relativement court (10 à 12 semaines en conditions contrôlées) et la production est importante chaque plante produit 500 à 1 000 graines, ce qui assure une disponibilité rapide du matériel. Elle répond au greffage, à l'embryogénèse somatique, à la transformation par *Agrobacterium tumefaciens* et *Agrobacterium rhizogenes*. De plus, *Medicago truncatula* forme des interactions symbiotiques avec l'une des bactéries du sol les plus étudiées qui est *Sinorhizobium liloti* ainsi qu'avec des champignons pour former des symbioses endomycorhiziennes **(Julier et al. 2006)**.

De plus, *Medicago truncatula* est caractérisée par une forte variabilité morphologique et génétique intra- et inter-populations et par une importante homozygotie au niveau individuel. Vu la diversification des biotopes qu'elle occupe, elle regroupe de nombreux génotypes, les plus utilisés en laboratoire sont JemalongA17 et R-108 **(Tivoli et al, 2006)**.

I.2 La Bioinformatique

I.2.1 Définition de la Bioinformatique :

La bioinformatique 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**).

La bioinformatique est donc une branche théorique de la Biologie est c'est l'analyse " *in silico* " de l'information biologique, elle analyse et interprète des données biologiques. (**Bonsai, 2014**). Elle permet aussi de mettre l'outil informatique au service des biologistes c'est à dire stocker, extraire, organiser, analyser, interpréter et utiliser les données biologiques. (**Khuri, 2010**). Elle analyse les gènes les génomes et les protéines à l'aide des algorithmes et des banques de données. Elle s'applique à tous types de données biologiques (séquences d'ADN, structure d'ARN, les puces à ADN, les arbres phylogéniques...) Il s'agit dans tous les cas d'un champ de recherche multidisciplinaire qui associe informaticiens, mathématiciens, physiciens et biologistes. (**Khuri, 2010**).

Ce développement de la bioinformatique a été rendu possible par les énormes progrès réalisés au niveau des capacités de calcul et de stockage des ordinateurs. Sans ces progrès, il n'eût pas été envisageable de construire des banques capables de manipuler l'intégralité des séquences biologiques publiées ou de développer des logiciels susceptibles d'effectuer des traitements sur de très larges sous-ensembles de ces banques. (**Perrière, 2000**).

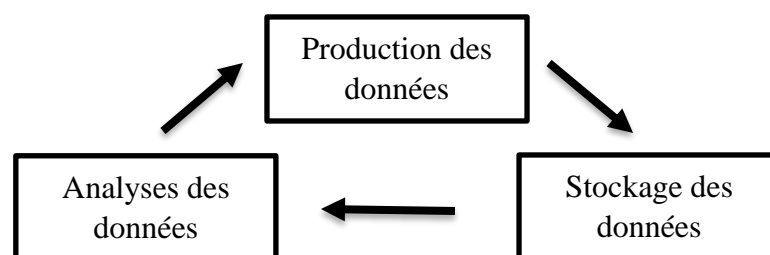


Figure 4 : Schématisation de l'approche *in silico* de la biologie (**Bonsai, 2014**)

I.2.2 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 : European Molecular Biology Laboratory* (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**)

- Moteur de recherche :

Un moteur de recherche est un site indexant tous les autres sites Internet et permet de poser une question ou taper des mots pour faire une recherche. Le moteur va ensuite retourner les résultats les plus pertinents.

I.2.3 Les domaines de la Bioinformatique :

I.2.3.1 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 facile grâce aux logiciels et les outils informatiques. (**Chollier, 2006**)

I.2.3.2 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 passé 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.2.3.3 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 bioinformatique et ces applications et ces logiciels.

I.2.3.4 Analyse du transcriptome :

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 utilisée pour comprendre les mécanismes de régulation transcriptionnelle, ainsi que pour certaines applications médicales.

I.2.4 *Medicago truncatula* et bioinformatique :

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)	<i>MTGD: The Medicago truncatula Genome Database</i> (http://www.MedicagoGenome.org)
EST	<i>TIGR :The Institute for Genomic Research</i> (http://www.jcvi.org)
Alignements nucléotidique	Université d'Oklahoma (https://www.ou.edu/)
EST, ORF, cartes génétiques	<i>LIS : Legume Information System</i> (https://legumeinfo.org/)

I.3 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.3.1 Définition du mot stress :

Le mot 'Stress' est un mot Anglais qui est apparu autour de l'année 1940, il est employé en mécanique, en physique ou en biologie, et qui a pour signification une force ou une influence qui empêche un système normal de fonctionner.

D'un point de vu 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).

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).

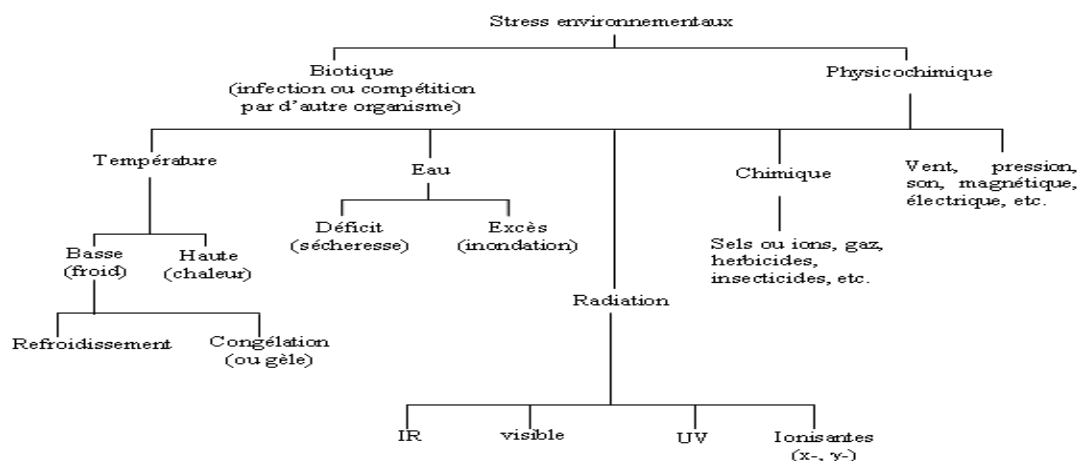


Figure 5 : Les différents stress environnementaux (Levitt, 1980).

I.3.2 Les stress abiotiques :

Selon Tal (1984), c'est le type de stress dont les plantes cultivées souffrent depuis leur semis jusqu'à leur récolte, il est dû principalement à des facteurs environnementaux qui sont susceptibles de provoquer chez une espèce végétale une réduction de la croissance des individus ou une augmentation du taux de mortalité au sein de la population, ces conditions peuvent être assimilées à des conditions stressantes, comme les éléments du sol (salinité, lessivage des éléments de la rhizosphère...), les rayons solaires, l'humidité relative, la température, le gèle, la sécheresse, les pluies non saisonnières, la pollution de l'air, les produits oxydés formés à partir des réactions de pesticides et d'autres propriétés physicochimiques du sol capables d'engendrer un stress à la plante telle qu'une structure compacte ou un sol asphyxiant.

Parmi les différents stress abiotiques, notre étude portera sur le stress salin causé par une augmentation du taux de NaCl et le stress hydrique causé par une alimentation hydrique sub-optimal.

I.3.2.1 Le stress salin :

I.3.2.1.1 Définition :

Le stress salin ou la salinité est défini comme étant une contrainte du milieu causée par une concentration excessive en sel. Le terme stress salin s'applique surtout à un excès des ions, en particulier Na^+ et Cl^- (**Hopkins, 2003**). Selon Krys (2000), une condition stressante en NaCl correspond à une concentration en sel supérieur à 100 mM dans le milieu extérieur.

I.3.2.1.2 Effets du stress salin :

La salinité affecte les plantes à différents niveaux en engendrant une sécheresse physiologique dû à une pression osmotique élevée (stress osmotique) qui perturbe la capacité des racines à extraire de l'eau, et des concentrations ioniques élevées provoquent une toxicité ionique entraînant une inhibition de nombreux processus physiologiques et biochimiques tels que l'absorption des nutriments et de l'assimilation. L'action conjointe de ces deux composantes aura pour effet une diminution de la croissance, le développement et la survie de la plante (**Toshio et Eduardo, 2005**).

Une concentration de NaCl entre 100mM et 200mM peut être considérer comme un stress modéré. Par contre, une concentration supérieure à 300 mM de NaCl peut causer un stress sévère aux plantes.

Selon la tolérance au sel, on peut définir deux groupes des végétaux, Les halophytes supportent les concentrations en sels entre 200 et 500 mM et Les glycophytes représentent la majorité des

espèces végétales dont leur croissance est ralentie dès que la concentration des milieux externes dépasse 100 mM et devient létale à partir de 300 mM. (GREENWAY *et al*, 1980).

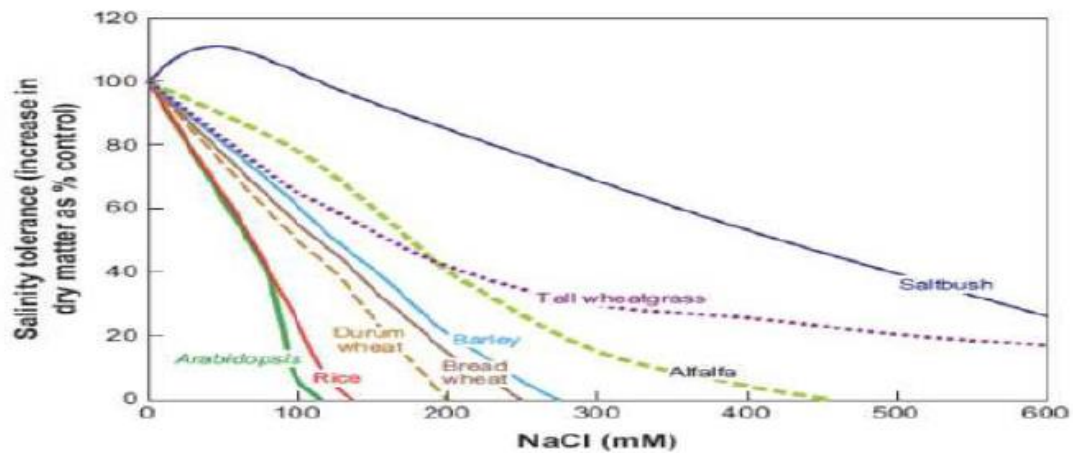


Figure 6 : La tolérance à la salinité chez quelques espèces végétales (Mahajan, 2005).

I.3.2.1.3 Conséquences de la contrainte osmotique :

L'effet osmotique entraîne un accroissement des dépenses d'énergie pour tirer l'eau du sol. Ainsi, la présence d'une forte concentration de sels solubles dans le sol crée une pression osmotique élevée dans l'environnement racinaire qui réduit la disponibilité de l'eau du sol pour la plante conduisant à une sécheresse physiologique qui aura comme conséquence une réduction de l'approvisionnement en CO₂ à cause de la fermeture des stomates (Parida et Das, 2005) (Maricle *et al*, 2007).

I.3.2.1.4 Conséquences de la contrainte ionique :

L'entrée du sel dans la plante provoque généralement un déséquilibre ionique, qui se traduit par des carences ou des excès en certains éléments. Selon la composition ionique de la solution saline, la toxicité ionique ou les déficiences nutritionnelles peuvent survenir à cause de la prédominance d'un ion spécifique ou à cause des effets compétitifs entre cations et anions (Bernstein, 1974).

Dans le cas d'un stress salin causé par une forte concentration en NaCl, l'accumulation des ions Na⁺ dans la plante limite l'absorption des cations indispensables tels que K⁺ et Ca⁺⁺ à cause de la compétition pour les mêmes sites de fixation apoplasmique.

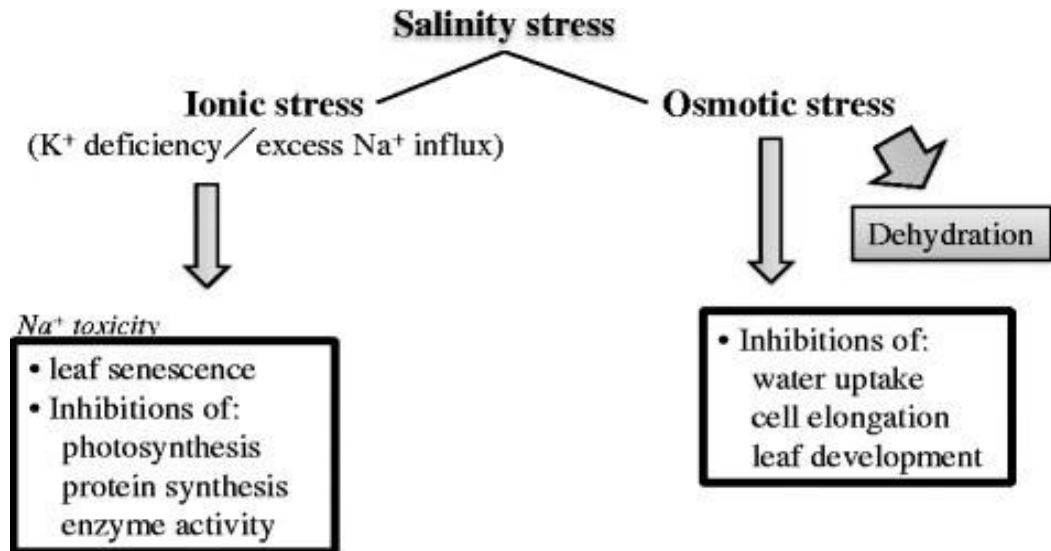


Figure 7 : Effet du stress salin sur les plantes. (Horie *et al.*, 2012)

I.3.2.2 Le stress hydrique :

I.3.2.2.1 Définition :

Le stress hydrique ou la sécheresse est défini comme étant un manque d'eau au cours d'une période prolongée qui concerne exclusivement des plantes localisées en plein champ. C'est un phénomène climatique défini comme D'un point de vue écologique, une sécheresse interfèrera défavorablement avec le fonctionnement des écosystèmes. (HOPKINS, 2003).

I.3.2.2.2 Effets du stress hydrique :

Le stress hydrique affecte les plantes à différents niveaux suivant sa durée et son intensité (YOKOTA *et al.* 2006). Selon la durée de l'exposition, on peut classer les effets du manque d'eau en deux catégories :

- **Effets à court terme :**

Lors d'un déficit hydrique, l'activité physiologique de la feuille, et plus particulièrement la photosynthèse est affectée (Lowlor, 2002 ; Lowlor et Cornic, 2002). Cette diminution est due à la fois à la fermeture des stomates, et à la diffusion du CO₂ qui cause une limitation biochimique du chloroplaste à fixer le CO₂ (Tardieu et Simoneau, 1998 ; Escolana *et al.*, 1999 Flexas et Medrano, 2002.)

- **Effet à long terme :**

Des modifications morphologiques, anatomiques, physiologiques et développementales de la plante sont observées dès qu'une sécheresse s'installe (Boyer, 1970 ; LAMAZE *et al.*, 1994). Elles comprennent principalement une réduction de la surface foliaire qui peut provenir d'une

diminution de l'expansion foliaire et/ou d'une sénescence accélérée de la feuille (**Kramer et Boyer, 1995**).

Lors d'un stress hydrique de faible intensité, le potentiel hydrique de la plante peut être maintenu par un ajustement osmotique. Cet ajustement implique l'accumulation, au niveau cellulaire, de sucres, d'acides aminés, d'ions ou d'autres solutés compatibles c'est-à-dire non toxiques. L'accumulation d'osmolites permet de créer un influx d'eau dans la cellule ou tout du moins d'éviter un efflux, en augmentant la force de rétention des molécules d'eau. Le maintien de cette quantité d'eau permet ainsi de conserver la turgescence nécessaire à la croissance des cellules. (**Crowe et al, 1992**).

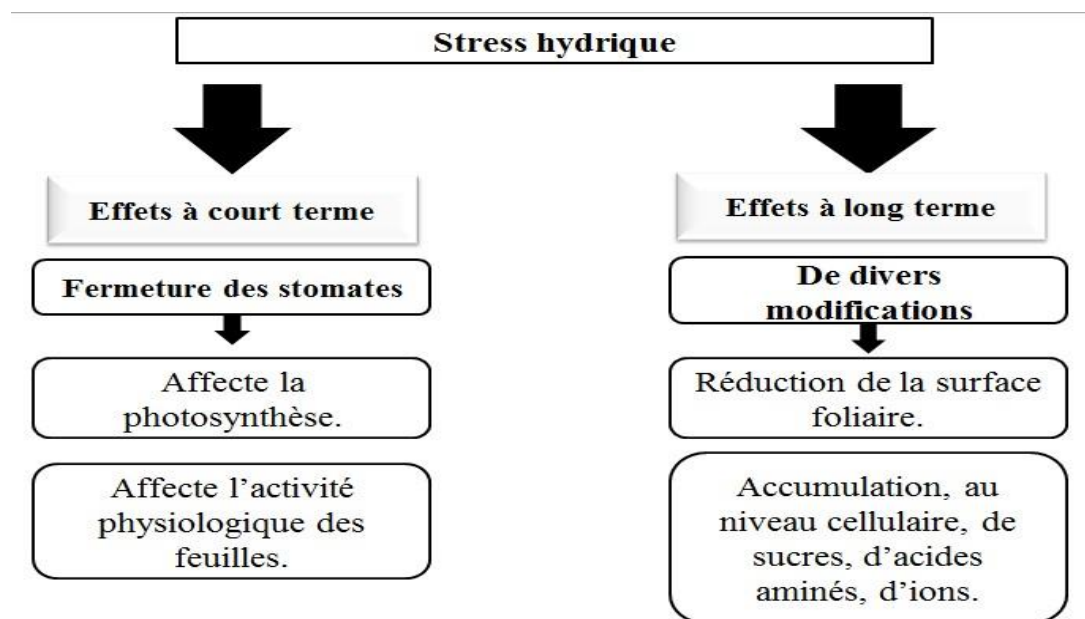


Figure 8 : Effets du stress hydrique sur les plantes.

I.3.3 Stratégies développées par les plantes :

Lorsque l'alimentation hydrique est interrompue, les végétaux sont caractérisés par une grande capacité de résistance à des variations importantes de la teneur en eau du sol. Les stratégies développées par les plantes pour faire face au manque d'eau peuvent être classées en trois catégories : l'évitement, la tolérance et l'adaptation. (**JONES, 1992**).

I.3.3.1 L'esquive :

La première façon d'éviter la sécheresse est l'esquive, qui est un changement dans la longévité du cycle phénologique. Elle permet à la plante de réduire ou d'annuler les effets du stress hydrique en évitant que le stress ne se produise au cours d'une phase sensible ou critique. Le développement phénologique rapide avec une floraison précoce, permet à la plante d'éviter les périodes sèches. Cette stratégie appliquée aux espèces cultivées a amené à décaler la date de semis et/ou à sélectionner des variétés plus précoces permettant d'éviter les déficits hydriques de fin de cycle (ATTIA, 2006).

I.3.3.2 La tolérance :

La tolérance est la stratégie qui permet à la plante d'assurer ses fonctions physiologiques malgré une dégradation de son état hydrique. Le maintien de la turgescence lors d'un déficit hydrique permet de retarder la fermeture des stomates, de maintenir le volume chloroplastique et de réduire le flétrissement foliaire (BLUM et EBERCON, 1981 ; MEDRANO et al, 2003). En outre, le maintien de la turgescence cellulaire est à la base de la préservation de plusieurs fonctions physiologiques comme la photosynthèse, la transpiration, l'assimilation du carbone et l'élongation cellulaire. Cette aptitude confère à la plante une meilleure tolérance au déficit hydrique (ATTIA, 2006).

I.3.3.3 L'adaptation :

La sécheresse se manifeste par la combinaison d'une part, de la restriction de la disponibilité en eau du sol et, d'autre part, de l'augmentation de la demande évaporatrice. Dans les zones arides, les plantes ont développé des mécanismes de régulation assurant leur survie, en général aux dépens de la productivité. D'un point de vue agronomique, l'adaptation à la sécheresse est la capacité de la plante à croître et donner des rendements satisfaisants dans des zones sujettes à des déficits hydriques épisodiques (CHAVES et OLIVEIRA, 2002).

I.3.4 Aspects moléculaire de la réponse des plantes aux stress :

Les conditions stressantes de l'environnement provoquent un certain nombre de changements moléculaires, biochimiques et physiologiques au niveau de la cellule et de la plante entière. L'un des principaux changements biochimiques en réponse au stress est une élévation du taux de l'acide abscissique (ABA), il aide les graines à surpasser les conditions de stress, de ne germer que dans des conditions qui favorisent la germination et la croissance, il empêche aussi la germination précoce des embryons prématurés, et il limite les pertes d'eau intracellulaire par la fermeture des stomates. On le nomme aussi hormone de stress. De ce fait, l'hormone végétale ABA a été identifiée comme un régulateur clé impliqué dans la réponse des plantes à de multiples contraintes. (Dana et al, 2016).

L'acide abscissique, un sesquiterpène (C₁₅H₂₀O₄) (Figure 9) présent de manière ubiquitaire dans plusieurs règnes a été découvert dans les années 1960 et initialement nommé « dormine » ou « abscissine ».

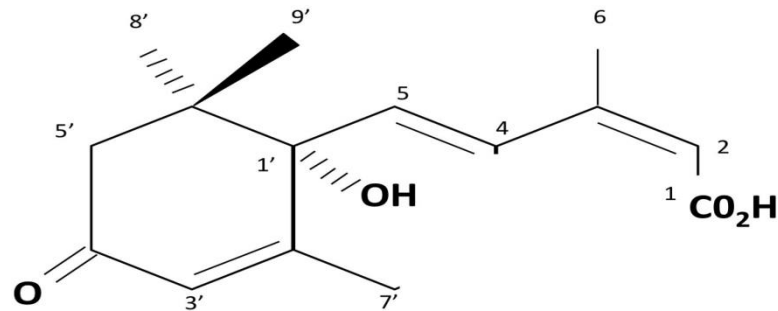


Figure 9 : Structure de l'ABA (Dana et al, 2016).

La sécheresse et la salinité exercent sur les plantes un stress osmotique au niveau cellulaire conduisant à une élévation du niveau l'ABA endogène. Cette accumulation mène à l'activation des facteurs de transcription (comme AREB1 / ABF2, AREB2 / ABF4, ABF3 et MYB41) associés à l'expression des gènes de réponse au stress dites « gènes ABA-sensibles », et l'expression de ces gènes codant des protéines qui gèrent la situation de stress pour atteindre un équilibre du bilan hydrique de la plante par la régulation des cellules de garde et la tolérance des cellules à la déshydratation. (Yamaguchi et al, 2002).

La biosynthèse de l'ABA se résume en un clivage d'un précurseur caroténoïde C40, suivi d'une conversion en deux étapes de la xanthoxine intermédiaire en ABA via ABA-aldéhyde.

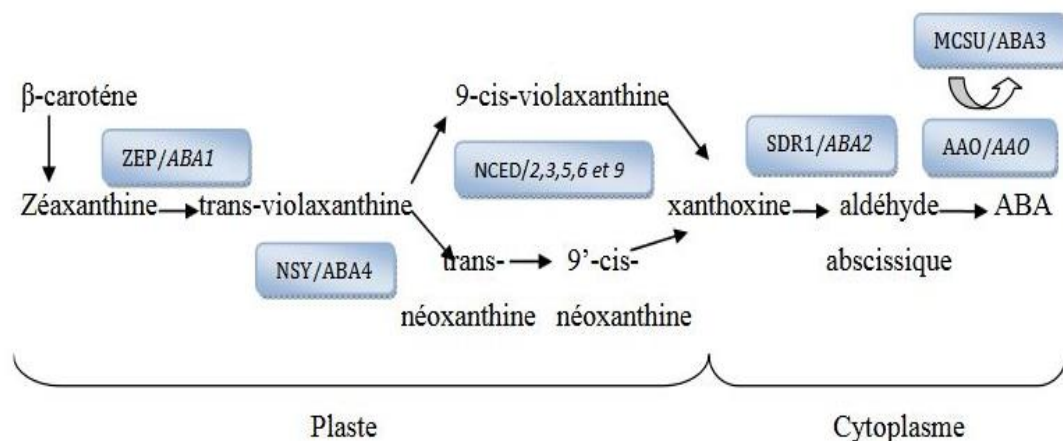


Figure 10 : Biosynthèse de l'ABA, d'après Wasilewska et al, 2008.

Le passage du caroténoïde à la xanthoxine implique la formation d'un précurseur 9- cis – époxycaroténoïde catalysée par l'enzyme NCED3. (Ian et *al*, 2000)

Chapitre II :

Matériel et méthodes

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*¹.

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 du gène de référence *NCED3*. Ensuite, la collecte et le classement des données d'expression de gènes relatives au stress hydrique et stress salin. Et se termine par une analyse statistique par **R**.

II.2.1 Recherche du *Probset ID* correspondant au gène *NCED3* :

La séquence nucléotidique de base utilisée est représenté par l'orthologue (Deux séquences homologues de deux espèces différentes) du gène *NCED3* sur le génome d'*Arabidopsis thaliana* et qui a été récupérée sur la banque de données *Genbank* à partir du moteur de recherche *TAIR*² : sous l'identifiant 'AT3G14440.1'

On allant de cette séquence de base, un blast 'N' a été réalisé sur la banque de données afin de récupérer la séquence correspondante sur le génome de *Medicago truncatula*.

¹ *Medicago truncatula* Gene Expression Atlas: <https://mtgea.noble.org>.

² The Arabidopsis Information Resource : <https://www.arabidopsis.org/>

Tableau 2 : Représentation des différents identifiants du gène *NCED3*

<i>NCED3</i>	<i>TAIR</i>	<i>Genbank</i>	<i>Craig Venter</i>
Organisme	<i>A. thaliana</i>	<i>A. thaliana</i>	<i>M. truncatula</i>
Identifiant	AT3G14440.1	NM_112304.3	Medtr2g070460.1

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*".³

A partir de l'identifiant obtenu sur *CVI*, la *ProbSet ID* correspondante a été retrouvé sur le moteur de recherche *Legoo*⁴.

Legoo est un moteur de recherche qui offre de multiples fonctionnalités aux chercheurs qui travaillent sur les légumineuses, essentiellement sur *Medicago truncatula* :

- Une fonction nommée '*Nicknames*' qui permet de trouver rapidement la correspondance entre plusieurs identifiants d'un même gène.
- La fonction '*LegumeProteomes*' qui est dédiée à l'annotation fonctionnelle de *M.truncatula*, *Glycine max* et *Lotus japonicus*.
- Une fonction pour l'étude *in silico* des données nommée '*Workflows*'.
- Une fonction '*Knowledgebase*' qui représente les résultats des expérimentations des articles sous forme de model.

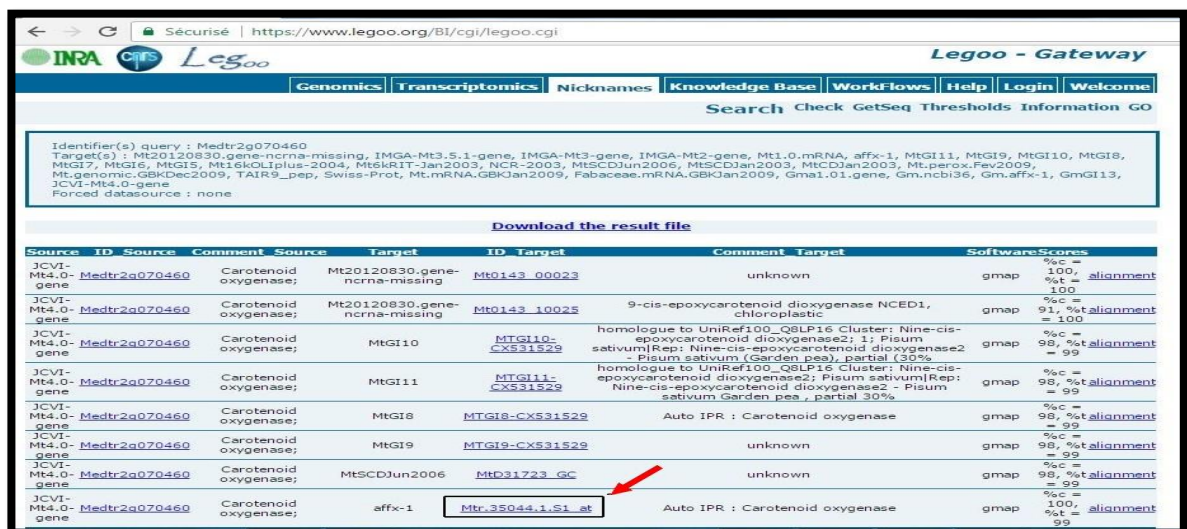


Figure 11 : La fonction *Nicknames* sur *Legoo*.

³ *J. Craig Venter* : <http://www.jcvi.org/cms/home/>

⁴ *Legoo* : <https://www.legoo.org/>

Tableau3 :Représentation d'identifiant *probset ID* du gène *NCED3* de *Medicago truncatula*

<i>NCED3</i>	<i>CVI</i>	<i>Legoo</i>
Organisme	<i>M. truncatula</i>	<i>M.truncatula</i>
Identifiant	Medtr2g070460.1	Mtr.35044.1.S1_at

II.2.2 Collecte des données d'expression du gène *NCED3* en conditions de stress hydrique et salin :

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.

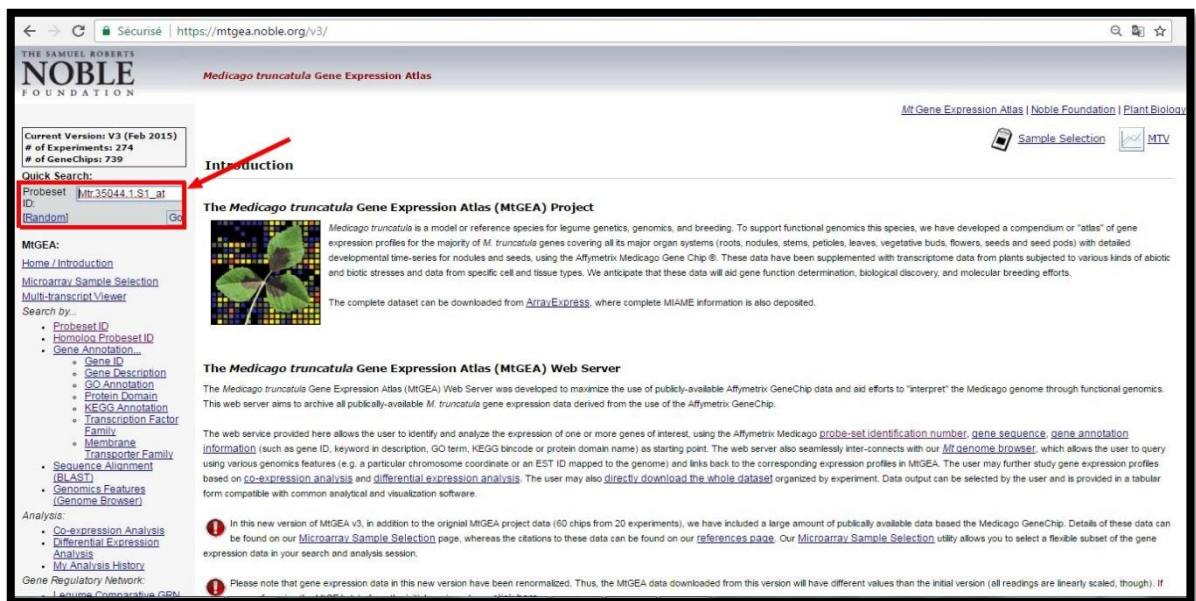


Figure 12 : Interface du moteur de recherche *MtGEA*.

Parmi les fonctionnalités qu'offre le moteur de recherche *MtGEA* est la présentation du profil d'expression sous forme de graphe et de tableau. Chaque profil d'expression est relié à une *probset ID* correspondant au gène étudié.

Ainsi, L'utilisation du *probset ID* correspondant au gène *NCED3* ([Mtr.35044.1.S1_at](#)) nous permis d'obtenir le profil d'expression suivant :

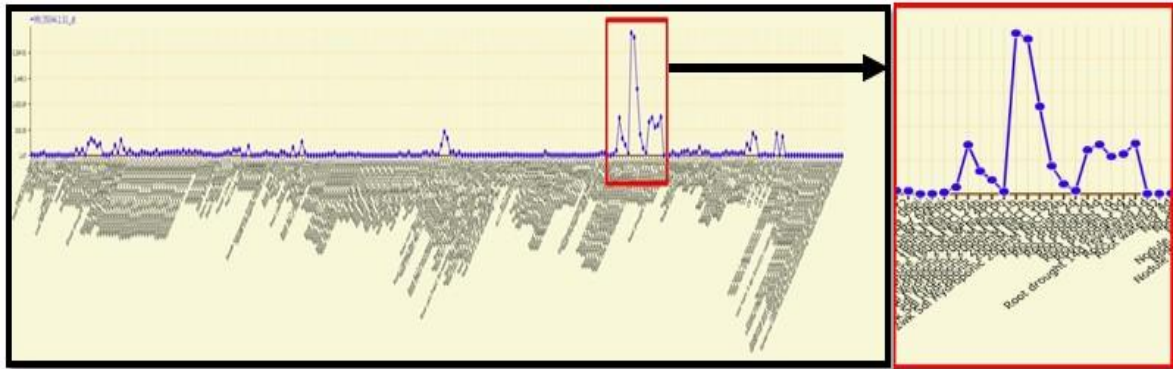














Figure 13 : Le profil d'expression du gène *NCED3* sur *MtGEA*

Les valeurs d'expression correspondants à chaque traitement (stress hydrique et stress salin) ont été récupéré sous forme de tableaux. Ces valeurs vont être ensuite utilisées afin de rechercher d'autres profils d'expression proche de celui du *NCED3*.

Tableau 4 : Les données relatives au stress hydrique.

Root drought 2d	Root drought 3d	Root drought 4d	Root drought 7d	Root drought 10d	Root drought 14d	Root drought 14d rewatered 1d
86.6	1069.4	1199.1	910.0	969.9	1224.8	12.7
Root drought 2d	Root drought 3d	Root drought 4d	Root drought 7d	Root drought 10d	Root drought 14d	Root drought 14d rewatered 1d

Tableau 5 : Les données relatives au stress salin.

Root 2wk SdI Hydroponic 200mM NaCl 0h 	Root 2wk SdI Hydroponic 200mM NaCl 1h 	Root 2wk SdI Hydroponic 200mM NaCl 2h 	Root 2wk SdI Hydroponic 200mM NaCl 5h 	Root 2wk SdI Hydroponic 200mM NaCl 10h 	Root 2wk SdI Hydroponic 200mM NaCl 24h 
63.8	3886.2	3746.5	2118.0	680.9	247.5
Root 2wk SdI Hydroponic 200mM NaCl 0h 	Root 2wk SdI Hydroponic 200mM NaCl 1h 	Root 2wk SdI Hydroponic 200mM NaCl 2h 	Root 2wk SdI Hydroponic 200mM NaCl 5h 	Root 2wk SdI Hydroponic 200mM NaCl 10h 	Root 2wk SdI Hydroponic 200mM NaCl 24h 

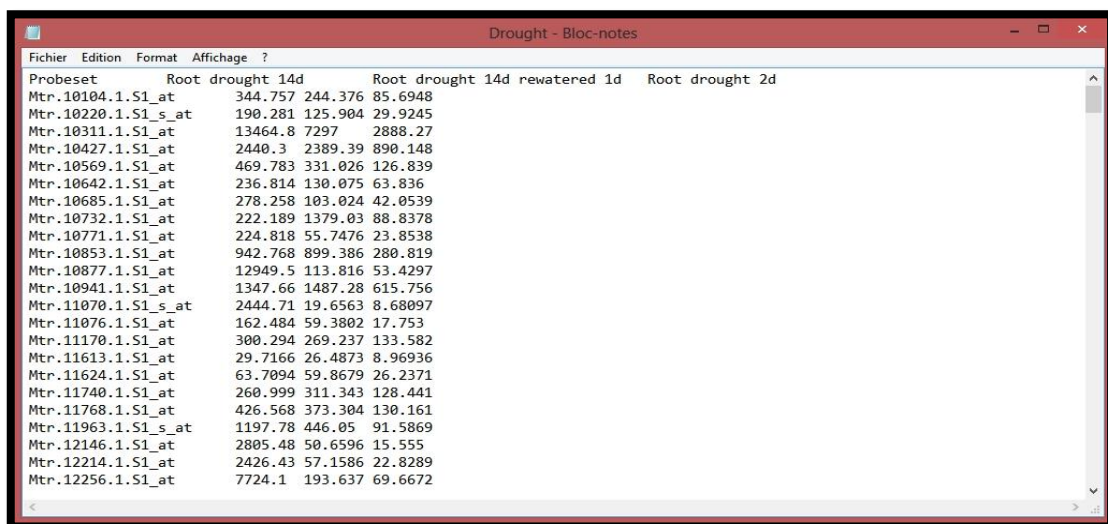
II.2.3 Collecte des données d'expression de gènes relatives au stress hydrique et stress salin :

Par l'utilisation de la fonctionnalité *Differential Expression Analysis*, le moteur de recherche *MtGEA* donne le choix de sélectionner une ou plusieurs conditions par rapport à une condition cible. Les données d'expression sont fournies en format texte (.txt).

Tableau 6 : Conditions sélectionnées pour le stress hydrique.

	Condition cible	Condition stressante	
Stress Hydrique	<i>Root drought 2d</i>	<i>Root drought 14d</i>	<i>Root drought 14d rewatered 1d</i>

Tableau 7 : Les données des gènes relatives au stress hydrique.



Probeset	Root drought 14d	Root drought 14d rewatered 1d	Root drought 2d
Mtr.10104.1.S1_at	344.757	244.376	85.6948
Mtr.10220.1.S1_s_at	190.281	125.904	29.9245
Mtr.10311.1.S1_at	13464.8	7297	2888.27
Mtr.10427.1.S1_at	2440.3	2389.39	890.148
Mtr.10569.1.S1_at	469.783	331.026	126.839
Mtr.10642.1.S1_at	236.814	130.075	63.836
Mtr.10685.1.S1_at	278.258	103.024	42.0539
Mtr.10732.1.S1_at	222.189	1379.03	88.8378
Mtr.10771.1.S1_at	224.818	55.7476	23.8538
Mtr.10853.1.S1_at	942.768	899.386	280.819
Mtr.10877.1.S1_at	12949.5	113.816	53.4297
Mtr.10941.1.S1_at	1347.66	1487.28	615.756
Mtr.11070.1.S1_s_at	2444.71	19.6563	8.68097
Mtr.11076.1.S1_at	162.484	59.3802	17.753
Mtr.11170.1.S1_at	300.294	269.237	133.582
Mtr.11613.1.S1_at	29.7166	26.4873	8.96936
Mtr.11624.1.S1_at	63.7094	59.8679	26.2371
Mtr.11740.1.S1_at	260.999	311.343	128.441
Mtr.11768.1.S1_at	426.568	373.304	130.161
Mtr.11963.1.S1_s_at	1197.78	446.05	91.5869
Mtr.12146.1.S1_at	2805.48	50.6596	15.555
Mtr.12214.1.S1_at	2426.43	57.1586	22.8289
Mtr.12256.1.S1_at	7724.1	193.637	69.6672

Tableau 8 : Conditions sélectionnées pour le stress salin.

	Condition cible	Condition stressante	
Stress salin	<p><i>Root 2wk Sdl</i></p> <p><i>hydroponic 200mM</i></p> <p><i>Nacl 0h</i></p>	<p><i>Root 2wk Sdl</i></p> <p><i>hydroponic 200mM</i></p> <p><i>Nacl 2h</i></p>	<p><i>Root 2wk Sdl</i></p> <p><i>hydroponic 200mM</i></p> <p><i>Nacl 24h</i></p>

Tableau 9 : Les données des gènes relatives au stress salin.

Probeset	Root 2wk Sdl Hydroponic 200mM NaCl 0h	Root 2wk Sdl Hydroponic 200mM NaCl 24h	Root 2wk Sdl Hydroponic 200mM NaCl 24h
Mtr.10045.1.S1_at	526.58	2169.61	4185.25
Mtr.10112.1.S1_at	52.8094	576.751	120.874
Mtr.10121.1.S1_at	108.808	615.819	399.152
Mtr.10127.1.S1_s_at	50.2042	212.938	136.673
Mtr.10159.1.S1_at	76.3664	241.641	175.57
Mtr.10205.1.S1_at	87.5416	191.36	191.008
Mtr.10215.1.S1_at	123.734	323.246	249.614
Mtr.10235.1.S1_at	61.895	150.852	147.096
Mtr.10265.1.S1_at	43.6388	124.074	155.693
Mtr.10265.1.S1_s_at	276.107	639.903	860.598
Mtr.10338.1.S1_s_at	820.252	3150.78	5814.81
Mtr.10349.1.S1_s_at	2918.82	7470.73	9997.1
Mtr.10439.1.S1_at	906.773	2320.67	2182.89
Mtr.10488.1.S1_at	2107.83	6669.2	4942.04
Mtr.10489.1.S1_s_at	1656.63	5988.62	4259.12
Mtr.10502.1.S1_at	953.002	3777.68	2782.55
Mtr.10502.1.S1_s_at	2031.82	7947.23	7326.35
Mtr.10558.1.S1_at	93.4048	289.925	286.82
Mtr.10597.1.S1_at	48.8489	136.23	304.339
Mtr.10601.1.S1_at	108.617	881.242	246.282
Mtr.10642.1.S1_at	431.948	1939.85	1483.97
Mtr.10714.1.S1_at	451.395	908.62	1068.67
Mtr.10780.1.S1_at	17.8013	75.7653	184.752

Enfin, Pour l'étude statistique, on a classé les données obtenues (Tableau 7 et 9) et on a pris les 200 premiers.

II.2.4 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 sauves a 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 afficher 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 inter-classe afin d'obtenir des classes les plus homogènes possibles.

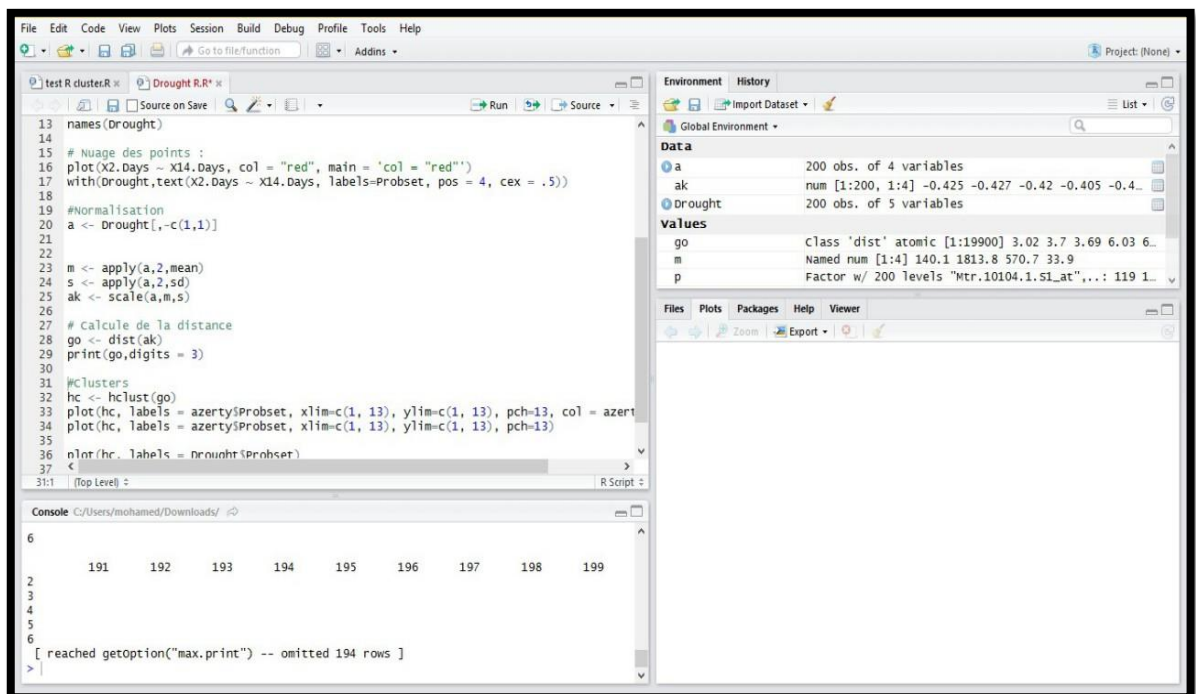


Figure 14 : L'interface de R en quatre quadrants

Tableau 10 : 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.
setdata	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.
Lenght	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.

Chapitre III :

Résultats et discussion

III Résultats et discussion :

III.1 Résultats :

III.1.1 Partie stress hydrique :

Après l'analyse statistique menée sur la totalité des données relatives au stress hydrique, l'application nous a permis de répartir les 200 gènes utilisées pour l'étude en plusieurs clades (figure 15), l'obtention des résultats sous forme de phylogramme.

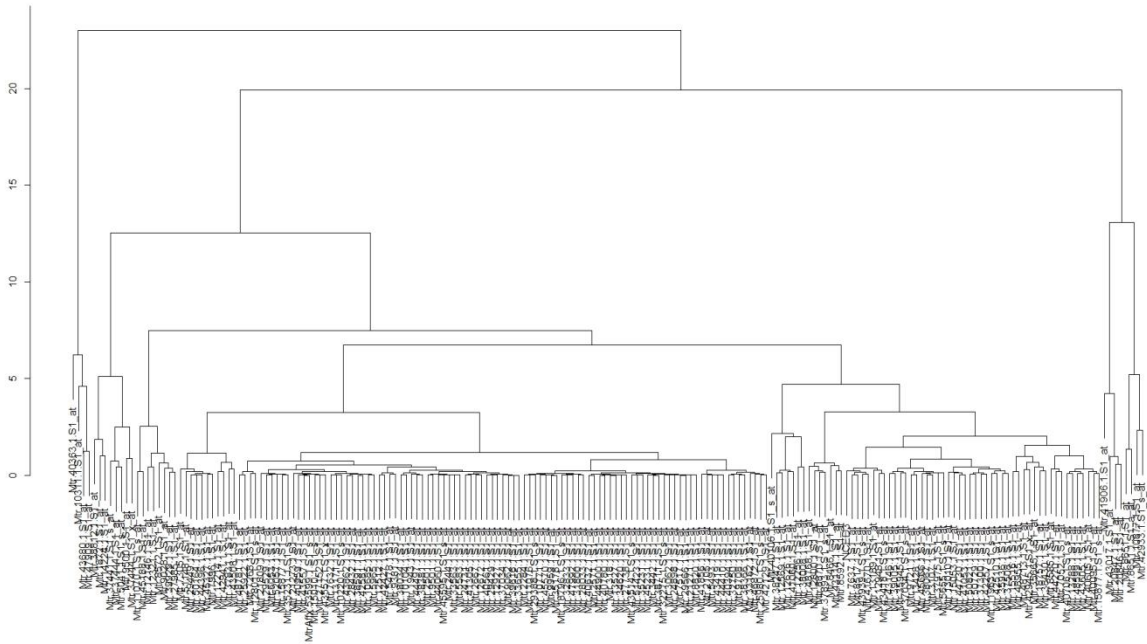


Figure 15 : Clustering de la totalité des gènes relatifs au stress hydrique.

Selon un pourcentage de dissimilarité inférieur ou égale à 3, le phylogramme est divisé en 24 clades, chaque clade regroupe un nombre de gènes. (tableau 11)

Tableau 11 : le nombre de gènes dans chaque clade.

Clades	Nombre de gènes	Clade	Nombre de gènes
Clade 1	1	Clade 13	4
Clade 2	1	Clade 14	2
Clade 3	2	Clade 15	7
Clade 4	1	Clade 16	8
Clade 5	2	Clade 17	23
Clade 6	3	Clade 18	7
Clade 7	2	Clade 19	10
Clade 8	2	Clade 20	1

Clade 9	6	Clade 21	3
Clade 10	11	Clade 22	2
Clade 11	98	Clade 23	1
Clade 12	1	Clade 24	1

Le gène *NCED3* appartient au 16^{ème} clade avec un groupe de 7gènes (figure 16).



Figure 16 : Position du clade numéro 16 sur le phylogramme.

La recherche des fonctions de différents gènes du 16eme clade a donné les Résultats suivants :

Tableau 12 : Les identifiants et les fonctions des gènes du 16^{ème} clade

<i>Probset ID</i>	Identifiant	Fonction
Mtr.7637.1.S1_at	Medtr4g062450.1	<i>Heavy-metal-associated domain protein</i>
Mtr.8517.1.S1_at	Medtr5g031250.2	<i>Cinnamyl alcohol dehydrogenase</i>
Mtr.43939.1.S1_s_at	Medtr4g082880.1	<i>Late embryogenesis abundant protein (LEA) family protein</i>
Mtr.41670.1.S1_at	Medtr1g029500.1	<i>F-box SKP2A protein</i>
Mtr.13189.1.S1_at	Medtr1g090630.1	<i>Nuclear transport factor 2 (NTF2) family protein</i>
Mtr.12948.1.S1_at	Medtr5g083340.2	<i>Dehydration responsive element-binding protein</i>
Mtr.43418.1.S1_s_at	Medtr1g104890.3	<i>Cysteine-rich receptor-kinase</i>

III.2 Partie stress salin :

Après l'analyse statistique menée sur la totalité des données relatives au stress salin, l'application nous a permis de repartir les 200 gènes utilisées pour l'étude en plusieurs clades (figure 17), l'obtention des résultats sous forme de phylogramme.

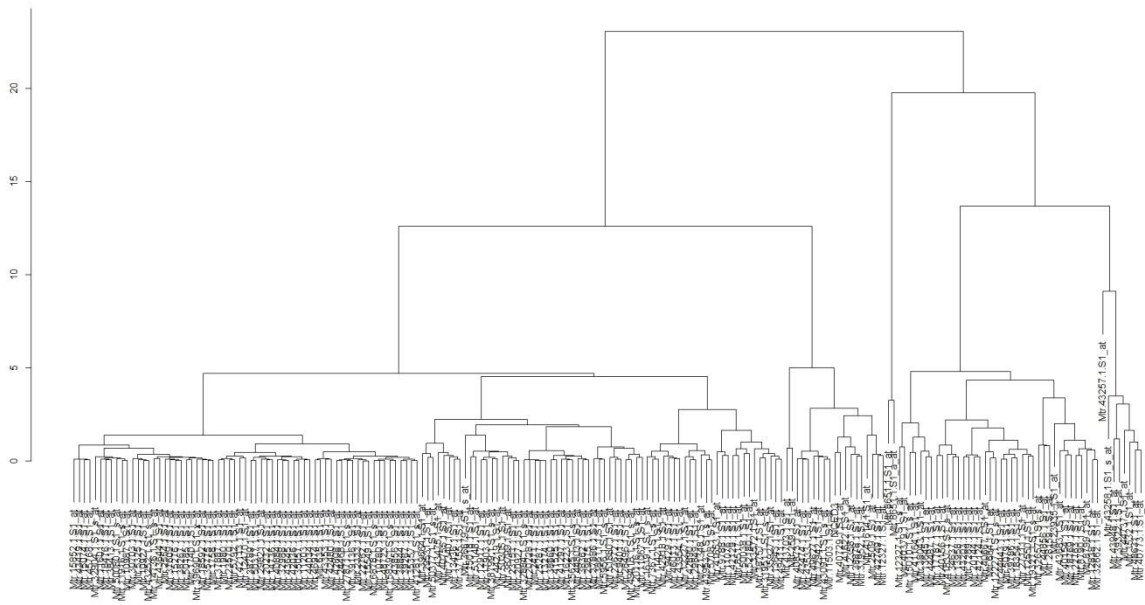


Figure 17 : Clustering de la totalité des gènes relatifs au stress salin.

Selon un pourcentage de dissimilarité inférieur ou égale à 3 , le phylogramme est divisé en 16 clades, chaque clade regroupe un nombre de gènes. (Tableau 13).

Tableau 13 : le nombre de gènes dans chaque clade.

Clades	Nombre de gènes	Clade	Nombre de gènes
Clade 1	65	Clade 9	7
Clade 2	42	Clade 10	19
Clade 3	26	Clade 11	3
Clade 4	2	Clade 12	9
Clade 5	7	Clade 13	1
Clade 6	10	Clade 14	1
Clade 7	1	Clade 15	2
Clade 8	1	Clade 16	4

L'observation du phylogramme nous a permis de voir que le gène *NCED3* appartient au 6^{ème} clade avec un groupe de 9 gènes.



Figure 18 : Position du clade numéro 6 sur le phylogramme.

L'utilisation des *probset ID* des différents gènes qui appartiennent au même clade que le *NCED3* nous a permis d'identifier les différents gènes correspondants.

Tableau 14 : Les identifiants et les fonctions des gènes du 6^{ème} clade

<i>Probset ID</i>	Identifiant	Fonction
Mtr.40729.1.S1_at	Medtr6g088475.1	<i>Lipid transporter</i>
Mtr.41622.1.S1_at	Medtr3g067720.1	<i>Cold regulated protein</i>
Mtr.12358.1.S1_at	Medtr7g093160.1	<i>Seed maturation protein</i>
Mtr.38878.1.S1_at	Medtr1g101600.1	<i>Dehydration responsive element binding protein</i>
Mtr.31952.1.S1_at	Medtr5g071560.1	<i>Ethylene respons factor</i>
Mtr.41614.1.S1_at	Medtr4g107010.1	<i>Spotted leaf protein</i>
Mtr.2509.1.S1_at	/	/
Mtr.42749.1.S1_at	/	<i>Late embryogenesis abundant protein</i>
Mtr.12327.1.S1_at	Medtr4g011250.1	<i>Late embryogenesis abundant protein</i>

Globalement, on considérant les deux stress au même temps (stress hydrique et salin), on peut distinguer, selon la fonction, 5 grand groupes : des gènes qui appartiennent au groupe *LEA*, des facteurs de transcription, des gènes qui codent pour des protéines réceptrices, des gènes codants pour des protéines de transport et des gènes qui ont des fonctions différentes.

III.2 Discussion :

Les plantes soumises à des conditions environnementales extrêmes telles que le manque d'eau et la salinité 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 la recherche des gènes impliqués dans la réponse des plantes aux stress hydrique et salin en utilisant la voie ABA comme référence, nous avons pu définir, selon la fonction, plusieurs groupes de gènes :

- Les gènes codant pour des protéines de types *LEA* (*Late embryogenesis abundant protein*) : sont bien associées à la tolérance au stress hydrique et stress salin, ces protéines s'accumulent lors de la déshydratation des semences au dernier stade de l'embryogenèse, elles s'accumulent également en réponse à la limitation de l'eau dans les tissus végétatifs. En plus du rôle qu'elles jouent dans la tolérance à la déshydratation cellulaire, elles ont un rôle dans la stabilisation des protéines, la protection de l'activité enzymatique, l'association et la stabilisation des membranes, la fonction antioxydante, la liaison métal-ion ou la protection de l'ADN, etc. **(Carrillo et al, 2011) (Battaglia et Covarrubias, 2013).**

-Les gènes codant pour des facteurs de transcription (*TF*) : sont connus pour leur rôle dans la régulation de l'expression des gènes et sont des acteurs clés de l'immunité innée des plantes. Parmi les différents types de facteurs de transcription, les familles présentées dans les résultats de notre étude sont les *ERF*, les *DREB* et *SLG*.

Les facteurs de réponse à l'éthylène ou bien *ERF* sont connus pour être impliqués dans la réponse moléculaire à l'attaque des agents pathogènes. Cependant, des études récentes suggèrent que plusieurs *ERF* se lient également à des éléments sensibles à la déshydratation et constituent un centre régulateur clé dans les réponses des plantes aux contraintes abiotiques. **(Müller et Munné-Bosch S, 2015) (Huang et al, 2016)**

Les facteurs de transcription appelés les *DREB* (*Dehydration Responsive Element Binding factor*) sont des facteurs de transcription qui régulent l'expression de nombreux gènes inductibles par le stress et jouent un rôle essentiel dans l'amélioration de la tolérance au stress abiotique des plantes. Ils se lient spécifiquement à la séquence d'ADN 5' - [AG] CCGAC-3'. La liaison à l'élément *C-repeat / DRE* favorise une transcription inductible à haute salinité et à la déshydratation. **(Sakuma et al, 2002)**

Les facteurs de transcriptions *SLG* sont des facteurs de transcription codant pour *spotted leaf protein*, qui sont des protéines liées à la défense qui régule négativement la mort cellulaire programmée (la sénescence cellulaire). **(Zeng et al, 2004)**

Un troisième groupe qui code pour des protéines réceptrices comme la *F-box SKP2A protein* et les *CRK*.

La *F-box SKP2A protein* (une ubiquitine) est une protéine multifonctionnelle qui agit comme récepteur d'auxine et joue un rôle important dans le contrôle de plusieurs processus de développement, les réponses aux changements environnementaux, le réglage de la stabilité des facteurs de transcription et aussi la division cellulaire (Confère une augmentation de la tolérance au stress osmotique en favorisant la division cellulaire). **(del Pozo et al. 2002) (Jurado et al, 2008)**

Les *CRK* (*cysteine rich receptor kinase*) sont des régulateurs impliqués dans la signalisation d'acide abscisique (ABA) et les mouvements des stomates. La surexpression des *CRK* améliore la tolérance à la sécheresse des plantes sans affecter la croissance et la productivité des plantes. **(Burdiak et al, 2015) (Kai Lu et al, 2016)**

Un quatrième groupe qui code pour des protéines de transport comme *lipid 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.

Par ailleurs, on distingue un cinquième groupe composé de différents gènes dont les rôles connus jusqu'à l'heure ne sont pas en relation directe avec les stress hydrique et salin : Les gènes codant pour '*Cold regulated proteins*' qui augmentent la tolérance des chloroplastes et des protoplastes au gel, et qui ont un rôle dans la stabilisation des membranes contre les dommages induits par la congélation. **(Thomashow et al, 1997)**

Les gènes qui codent pour *Cinnamyl alcohol dehydrogenases* sont des catalyseurs de l'étape finale de la biosynthèse du monolignol, la conversion des aldéhydes cinnamyls en alcools, en utilisant le NADPH comme cofacteur. Ils conduisent à la biosynthèse de la lignine dans les plantes. **(Bukh et al, 2012)**

En réponse aux stress hydrique et salin, les plantes développent une voie de signalisation hormonale contrôlée par plusieurs groupes de gènes comme les kinases, les facteurs de transcription et les canaux ioniques membranaires.

Le NCED3 est un gène clé dans la biosynthèse de l'acide abscisique qui est une hormone clé dans la régulation de la réponse des plantes car elle joue un rôle important dans l'expression des gènes sensibles au stress. Conduisant toutes aux même stratégies de protection de la cellule, tel que le rétablissement de l'équilibre osmotique, la protection des membranes et des protéines et la détoxification du milieu.

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 réponse au stress hydrique et salin chez le modèle biologique *Medicago truncatula*, nous avons réalisé une analyse *in silico* de données de transcriptomique issus de puces de type *Affymetrix*.

Au terme de notre travail, les résultats obtenus montrent que ces gènes peuvent être répartis, suivant leurs rôles qu'ils jouent, en plusieurs groupes : les gènes codant pour des protéines de types *LEA*, des gènes codant pour des facteurs de transcription (*TF*), des gènes qui codent pour des protéines réceptrices, des protéines de transport 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 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, tel que le rétablissement de l'équilibre osmotique, la protection des membranes et des protéines et la détoxification du milieu.

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 interférents (*RNAi*).

Références bibliographique

Références bibliographique

1. **Abbé Coste. 2011.** *Medicago truncatula* Gaertn. Tela Botanica [En Ligne], <https://tel.archives-ouvertes.fr/tel-00460435/document> [consulté le 03/04/2017]
2. **Asish Kumar Paridaa and Anath Bandhu Dasa. 2005.** Salt tolerance and salinity effects on plants: a review. 324–349
3. **Attia F., 2007 :** Effet du stress hydrique sur le comportement ecophysologique et la maturité phénolique de la vigne *Vitis vinifera* L : Etude de cinq cepages autochtones de midi-pyrenees. Université de Toulouse. Formation doctorale : Science Agronomique. P: 13-47.
4. **B. TIVOLI A. BARANGER K. SIVASITHAMPARAM M. J. BARBETTI. 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. 98 (6): 1117-1128
5. **Bernadette Julier, Thierry Huguet, Fabien Chardon, Radia Ayadi, Jean-Baptiste Pierre, Jean-Marie Prospero, Philippe Barre and Christian Huyghe. 2006.** Identification of quantitative trait loci influencing aerial morphogenesis in the model legume *Medicago truncatula*. 114:1391–1406
6. **Blum A, Ebercon A., 1981:** Cell membrane stability as a measure of drought and heat tolerance in wheat. Crop Sci. 21. P: 43 - 47.
7. **Bonsai. (2014).** Bioinformatique et données biologiques [en ligne], http://crystal.univ-lille.fr/~noe/enseignement/m1-genpro/Cours/bioinfo_bio1-2x3.pdf [consulté le 07/04/2017]

8. **Boukadida Jawer, Denis Julien, 2004. Te Bioinformatique [En ligne]**
<http://deptinfo.unice.fr/twiki/pub/Linfo/PlanningDesSoutenances20032004/Boukadida-Denis.pdf> [Consulté le 15/03/2017].
9. **Bukh C, Nord-Larsen PH and Rasmussen SK. 2012.** Phylogeny and structure of the cinnamyl alcohol dehydrogenase gene family in *Brachypodium distachyon*. *J Exp Bot.* 63(17):6223-36.
10. **C. Beroud. (2011).** Bases de données et outils bio-informatiques utiles en génétique [En ligne],
<http://campus.cerimes.fr/genetiquemedicale/enseignement/genetique28/site/html/cours.pdf> [consulté le 07/04/2017]
11. **Chaves M M, Oliveira M M., 2002:** Mechanisms underlying plant resilience to water deficits: prospects for water-saving agriculture. *Journal of Experimental Botany* 55. P : 2365-2384.
12. **Cuire D.R. 1999.** *Medicago truncatula* - un modèle dans la fabrication ! [En Ligne],
<http://www1.montpellier.inra.fr/BRCMTR/accueil.php?menu=medicago&page=menu11> [consulté le 02/04/2017]
13. **Del Pozo J.C., Boniotti M.B. and Gutierrez C. 2002.** Arabidopsis E2Fc functions in cell division and is degraded by the ubiquitin-SCF (AtSKP2) pathway in response to light. *Plant Cell* 14:3057-3071
14. **Delalande Magalie et al. 2007.** Wild accessions / populations [En Ligne],
<https://www.noble.org/globalassets/docs/medicago-handbook/wild-accessions-populations.pdf> [consulté le 30/03/2017]
15. **Djemel N, Guedon D, Lechevalier A, Salon C, Miquel M, Prosperi JM, Rochat C, Boutin JP. 2005.** Development and composition of the seeds of nine genotypes of the

- Medicago truncatula species complex. [En Ligne],
<https://www.ncbi.nlm.nih.gov/pubmed/15967671> [consulté le 03/04/2017]
16. **Georges Gardarin. 2006.** Base de données. [En ligne]
http://georges.gardarin.free.fr/Livre_BD_Contentu/XX-TotalBD.pdf [Consulté le 08/03/2017]
17. **GREENWAY H. and MUNNS R. 1980.** Mechanisms of salt tolerance in non-halophytes. *Ann. Rev. Plant Physiol* **31**, p. 149–190.
18. **Guy PERRIÈRE.2000.** Bases de données et outils d'analyse pour la génomique bactérienne [En ligne] ftp://biomol.univ-lyon1.fr/pub/theses/HDR_GP.pdf [Consulté le 19/05/2017].
19. **Hemmat. Ibrahim Khattab. 2005.** Responses of Eucalypt trees to insect feeding (gall-forming psyllid).
20. **Hopkins. 2003.** Physiologie végétale.
21. **Horie T, Karahara I, Katsuhara M. 2012.** Salinity tolerance mechanisms in glycophytes: An overview with the central focus on rice plants. *The Rice Journal* **5** (1) 1-18.
22. **Huang PY, Catinot J and Zimmerli L. 2016.** Ethylene response factors in Arabidopsis immunity. *J Exp Bot.* 1231-41
23. **Huguet T., Prosperi J.M.. 1995.** *Medicago truncatula*: a legume model-plant. 171-175
24. **Ian B. Taylor Alan Burbidge Andrew J. Thompson, 2000** Contrôle de la synthèse de l'acide abscissique. *J Exp Bot* **51** (350): 1563-1574.
25. **Jean-Michel Claverie. (2013).** La Bioinformatique [EN ligne],
<http://campus.cerimes.fr/genetiquemedicale/enseignement/genetique28/site/html/cours.pdf> . [Consulté le 07/04/2017].

26. **John Klingler, Robert Creasy, Lingling Gao, Ramakrishnan M. Nair, Alonso Suazo Calix, Helen Spafford Jacob, Owain R. Edwards, and Karam B. Singh . 2005.** Aphid Resistance in *Medicago truncatula* Involves Antixenosis and Phloem-Specific, Inducible Antibiosis, and Maps to a Single Locus Flanked by NBS-LRR Resistance Gene Analogs. 1174–1181.
- John Pickering and Kate Smith, 2006.** [en ligne] <http://www.discoverlife.org/20/q?kind=Quercus,Tuxedo> [consulté le 07/06/2017]
27. **Jones HG., 1992:** Plant and Microclimates (Ed): A quantitative approach to environmental plant physiology, Cambridge University Press, London.
28. **Jurado S, Triviño SD, Abraham Z, Manzano C, Gutierrez C and Del Pozo C. 2008.** SKP2A protein, an F-box that regulates cell division, is degraded via the ubiquitin pathway. *Plant Signal Behav.* 810-2.
29. **Kai Lu, Shan Liang, Zhen Wu, Chao Bi, Yong-Tao Yu, Xiao-Fang Wang and Da-Peng Zhang. 2016.** Overexpression of an Arabidopsis cysteine-rich receptor-like protein kinase, CRK5, enhances abscisic acid sensitivity and confers drought tolerance. *J Exp Bot* 67 (17): 5009-5027
30. **Le GALL. 1993.** Les grandes légumineuses : situation actuelle, atouts et perspectives dans le nouveau paysage fourrager français. 134, 121-144
31. **Leon Bernstein. 1963.** Osmotic Adjustment of Plants to Saline Media. II. Dynamic Phase. Vol. 50, No. 4, pp. 360-370
32. **Marina Battaglia and Alejandra A. Covarrubias. 2013.** Late Embryogenesis Abundant (LEA) proteins in legumes *Front Plant Sci.* 190
33. **Mathison et al. 1980.** Annual medics and subterranean clover in Southern Australia. [En Ligne], <https://www.cabdirect.org/cabdirect/abstract/19800704699> [consulté le 03/04/2017]

34. **Medrano H, Escalona JM, Cifre J, Bota J, Flexas J., 2003:** A ten year study on the physiology of two Spanish grapevine cultivars under field conditions: effects of water availability from leaf photosynthesis to grape yield and quality. *Functional Plant Biology* 30. P: 607-619.
35. **Michel Caboche .2006.** L'essor de la génomique végétale. [En ligne] http://www-urgv.versailles.inra.fr/pub/24-26Biof265_Caboche.pdf [Consulté le 03/05/2017].
36. **Mireille Chabaud, Judith Lichtenzweig, Simon Ellwood, Theo Pfaff and Etienne-Pascal Journet. 2005.** Vernalization, crossings and testing for pollen viability
37. **Morgane Thomas—Chollier. 2006.** Introduction à la phylogénie [En ligne], http://www.biologie.ens.fr/~mthomas/L3/phylogenie/ThomasChollier_phylogenie.pdf [Consulté le 12/05/2017].
38. **Müller M and Munné-Bosch S. 2015.** Ethylene Response Factors: A Key Regulatory Hub in Hormone and Stress Signaling..*Plant Physiol.* 32-41
39. **Nevin D. Young, Steven B. Cannon, Shusei Sato, Dongjin Kim, Douglas R. Cook, Chris D. Town, Bruce A. Roe, and Satoshi Tabata. 2005.** Sequencing the Genespaces of *Medicago truncatula* and *Lotus japonicas*. 1445–1455.
40. **Paweł Burdiak, Anna Rusaczonok, Damian Witoń, Dawid Głów and Stanisław Karpiński. 2015.** Cysteine-rich receptor-like kinase CRK5 as a regulator of growth, development, and ultraviolet radiation responses in *Arabidopsis thaliana*. *J Exp Bot* 66 (11): 3325-3337
41. **Sakuma Y., Liu Q., Dubouzet J.G., Abe H., Shinozaki K. and Yamaguchi-Shinozaki K. 2002.** DNA-binding specificity of the ERF/AP2 domain of *Arabidopsis* DREBs, transcription factors involved in dehydration- and cold-inducible gene expression. *Biochem. Biophys. Res. Commun.* 290:998-1009

42. **Sami Khuri. (2010).** Bioinformatique [En ligne], <http://www.cs.sjsu.edu/~khuri/Yverdon2010/French/YverdonBioinformaticsFrench6parpage.pdf> [consulté le 07/04/2017]
43. **Sami Khuri. (2010).** Bioinformatique [En ligne], <http://deptinfo.unice.fr/twiki/pub/Linfo/SuiviDesTE/rapport-imbsd-sayedham.pdf> [consulté le 07/04/2017]
44. **Schulze E-D. et al. 2005.** Plant Ecology P 7
45. **Thomashow, M.F., Stockinger, E.J., Jaglo-Ottosen, K.R. et al. 1997.** Function and regulation of Arabidopsis thaliana COR (cold-regulated) genes Acta Physiol Plant 19: 497.
46. **Toshio Yamaguchi and Eduardo Blumwald. 2005.** Developing salt-tolerant crop plants: challenges and opportunities. Vol.10 No.12
47. **V. C. Dilukshi Fernando and Dana F. Schroeder. 2016.** Role of ABA in Arabidopsis Salt, Drought, and Desiccation Tolerance. Chapitre 22. DOI: 10.5772/61957
48. **Victor Jongeneel. 2000.** Biologie, la Bioinformatique [En ligne] <http://www.universalis.fr/encyclopedie/biologie-la-bio-informatique/1-definition-et-enjeux-de-la-discipline/> [Consulté le 26/04/2017].
49. **Yadira Olvera-Carrillo, José Luis Reyes, and Alejandra A Covarrubias. 2011.** Late embryogenesis abundant proteins versatile players in the plant adaptation to water limiting environments. Plant Signal Behav. 586–589.
50. **Yamaguchi-Shinozaki, K., M. Kasuga, Q. Liu, K. Nakashima, Y. Sakuma, H. Abe, Z.K. Shinwari, M. Seki, K. Shinozaki. 2002.** Biological mechanisms of drought stress response. JIRCAS Working Report, 1-8.

51. **Zeng L.-R., Qu S., Bordeos A., Yang C., Baraoidan M., Yan H., Xie Q., Nahm B.H., Leung H. and Wang G.-L. 2004.** Spotted leaf11, a negative regulator of plant cell death and defense, encodes a U-Box/Armadillo repeat protein endowed with E3 Ubiquitin ligase activity. *Plant Cell* 16:2795-2808.

Année universitaire : 2016/2017

Présenté par : BELOUCIF Mohamed Fath Eddine.

SAIDI SIEF Roumeissa.

Recherche de gènes impliqués dans la réponse aux stress hydrique et salin chez le modèle biologique *Medicago truncatula* par analyse *in silico* de données de transcription.

Mémoire de fin de cycle pour l'obtention du diplôme de Master en Biologie et Génomique Végétale.

L'objectif du présent travail est la recherche de gènes impliqués dans la réponse au stress hydrique et au stress salin dans la racine du modèle biologique *Medicago truncatula*, par une analyse *in silico* de données de transcription issues de puces *Affymetrix*, et en utilisant un gène clé dans la biosynthèse de l'ABA comme gène de référence (*NCED3*).

Le Clustering des gènes étudiés à un pourcentage de dissimilarité de 3 classe le *NCED3* dans un clade de 7 gènes pour le stress hydrique et dans un clade de 9 gènes pour le stress salin.

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*, 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 hydrique, Stress salin, ABA, *Medicago truncatula*.

Jury d'évaluation :

Président du jury :	Mr. HAMIDECHI Mohamed Abdelhafid.	Prof. Université Constantine1.
Rapporteur :	Mr. TEMAGOULT Mahmoud.	M.A.A. Université Constantine1.
Examineur :	Mme. MAOUGAL Rim Tinhinen	M.C.B. Université Constantine1.

Date de soutenance : 20/06/2017