

**République Algérienne Démocratique et Populaire**  
**Ministère de l'Enseignement Supérieur et de la Recherche Scientifique**



**Université de Saad Dahleb Blida 1**

**Faculté des Sciences de la Nature et de la Vie**  
**Département de Biologie des populations et des organismes**

## **MEMOIRE DE MASTER**

**En vue de l'obtention du diplôme de Master en Biologie**

**Option : Génomique et Biotechnologie Végétal**

### **Thème**

**Les Méthodes de reconstruction phylogénétique:  
Modèles, Exploitation des Logiciels et Applications à  
*Medicago L. (Fabaceae.)***

**Présenté par : Mme Benmouhoub Djamilia**

**Date de soutenance : 29-09-2016**

**Devant de jury :**

- |                            |     |       |              |
|----------------------------|-----|-------|--------------|
| • Melle ZATRA Y.           | MAA | UB1   | Président    |
| • Melle ABDUL-HUSSAIN.A.S. | MCB | UB1   | Examinatrice |
| • Melle DEROUICHE L.       | MCB | EPSNV | Promotrice   |

## Les Méthodes de reconstruction phylogénétique : Modèles, Exploitation des Logiciels et Applications à *Vicia Medicago L.* (Fabaceae).

**Résumé :** Dans ce mémoire, nous avons tout d'abord exposé la méthodologie de la reconstruction phylogénétique : Nous avons abordé les problèmes de l'échantillonnage, le choix des caractères donc, nous avons récupéré les séquences nucléotidiques de la base de données *GenBank*, on a trouvé que *M.truncatula* et *M.sativa* sont les deux espèces les plus étudiées dans le genre *Medicago*, et le gène qui est le plus utilisé pour la phylogénie et le gène chloroplastique, qui s'est avéré le plus utile pour les analyses phylogénétiques. Les séquences chloroplastiques obtenues seront traitées par le logiciel MEGA 5 en passant par un alignement multiple en utilisant le programme *ClustalW*, puis la méthode (UPGMA) pour inférer les phylogénies du genre *Medicago*. Les groupements taxonomiques générés par les phylogénies obtenues pour 43 espèces sur une combinaison des séquences chloroplastiques sont congruents entre eux mais pas avec ceux d'autres auteurs sur d'autres données et avec la classification traditionnelle excepté pour *Medicago Sect. Intertextae*. Nos résultats démontrent aussi la monophylie du genre *Medicago* vis-à-vis du genre proche *Trigonella*.

**Mots clés :** Modèles, Phylogénie, *Vicia L.*, *Hyoseris L.*, *Medicago L.*, Taxonomie, Hérité paternelle.

### طرق الاستدلال الفيلوجيني: النماذج، استغلال برامج الإعلام الآلي و تطبيقاتها على النوع *Medicago*

**الملخص:** في هذه المذكرة، قمنا أولاً بعرض منهجية الاستدلال الفيلوجيني. تطرقنا إلى إشكاليات جمع العينات، اختيار المميزات البيولوجية و تشفيرها و استدلال أشجار الفيلوجينية. بعد ذلك، سحبنا استرداد تسلسل النوكليوتيدات من قاعدة بيانات بنك الجينات، و وقد وجدت *M.truncatula* و *M.sativa* هما أكثر الأنواع درسا في جنس *Medicago*، والجينات الأكثر استخداما للنسالة هي الجينات بلاستيديات الخضراء، التي أثبتت أكثر فائدة للتحليل النشوء والتطور. و سيتم تجهيز تسلسل بلاستيديات الخضراء التي حصل عليها البرنامج MEGA 5 عن طريق محاذاة متعددة باستخدام برنامج *ClustalW*، ثم طريقة (UPGMA) لاستنتاج شجرة الأنساب من *Medicago*. المجموعات التصنيفية المتولدة شجرة الأنساب التي تم الحصول عليها عن 43 نوعا على مجموعة من متواليات بلاستيديات الخضراء هي المنسجمة مع بعضها البعض ولكن ليس مع تلك التي تحصل عليها الباحثون مسبقا على أساس معطيات أخرى و مع التصنيف التقليدي ما عدا اقسام *Intertextae*: نتائجا تبين كذلك وحدانية الأصل للنوع *Medicago* مقارنة مع النوع القريب له *Trigonella*.

الكلمات الرئيسية: فيلوجينيات، *Medicago L.*، التصنيف، وراثية أبوية.

### The methods of phylogenetic reconstruction: Models, Software Exploitations and Application to *Medicago L.* (Fabaceae).

**Abstract:** In this paper, we first exposed the methodology of phylogenetic reconstruction: We have addressed the problems of sampling, choice of characters so we retrieve the nucleotide sequences of the *GenBank* database, we have found that *M.truncatula* and *M.sativa* are the two most studied species in the genus *Medicago*, and the gene that is most used for phylogeny and the chloroplast gene, which proved most useful for the phylogenetic analyzes. Chloroplast sequences obtained will be processed by the MEGA software 5 via a multiple alignment using the *ClustalW* program, then the method (UPGMA) to infer phylogenies of *Medicago*. Taxonomic groupings generated phylogenies obtained for 43 species on a combination of chloroplast sequences are congruent with each other but not with those of other authors on other data and with the traditional classification except for *Medicago Sect. Intertextae*. Our results also demonstrate the monophyly of *Medicago* vis-à-vis the closest kind *Trigonella*

**Key words:** Models, Phylogeny, *Medicago L.*, Taxonomy, Paternal heredity

# Remerciement

*D'abord les plus forts de mes remerciements sont pour mon Dieu «ALLAH» le tout puissant, de m'avoir guidé et m'avoir accordé la force et la patience pour la volonté, la santé pour réaliser ce travail, par la manière que lui admette de nous, "Amine".*

*Je tiens a exprimé mes vifs remerciements et ma profonde gratitude à ma promotrice **DEROUICHE Louiza**, qui a bien voulu me consacrer tout son temps pour me transmettre leur savoir avec une bonne volonté , leur soutien leur aide , leur conseils et leur bienveillance durant l'élaboration de ce mémoire  
Merci pour votre disponibilité et de la confiance que vous m'aviez accordé.*

*J'exprime mes remerciements à **Melle ZATRA Yamina** pour m'avoir faire l'honneur de présider le jury de soutenance.*

*Je remercie vivement **Melle ABDULHUSSEIN Alia** qui a accepté d'être examinatrice de ce mémoire, et de faire partie de ce jury.*

# *Dédicace*

Je dédie ce travail

A toutes les personnes qui, de loin ou de près,  
m'ont aidé à l'élaboration de ce travail

# Sommaire

RESUME

REMERCIEMENTS

DEDICACES

LISTE DES ABREVIATIONS

LISTE DES FIGURES

LISTE DES TABLEAUX

INTRODUCTION.....1

CHAPITRE I : RAPPELS BIBLIOGRAPHIQUE.....2

1.Définition de la phylogénie.....2

1.1 . Les données morphologiques.....2

1.2. La biogéographie.....2

1.3. La paléontologie.....3

1.4. La biologie moléculaire.....3

1.4.1. Les différents types de caractères moléculaires.....3

1.4.1.1. L'hypothèse de l'horloge moléculaire et choix des séquences.....3

a) L'ADN ribosomique .....4

b) L'ADN mitochondrial.....6

c) L'ADN chloroplastique .....6

1.4.1.2. Intérêts des caractères moléculaires.....6

2. La phylogénétiques moléculaires.....7

2.1. Phylogénie et généalogie.....8

2.2. Terminologie phylogénétique.....8

2.2.1. Les arbres phylogénétiques .....	9
a- La méthode des distances génétique .....	11
b- La méthode de maximum de parcimonie.....	11
c- La méthode de maximum de vraisemblance .....	11
3. Le genre de <i>Medicago</i> .....	12
3.1. Generalité .....	12
3.2. <i>Medicago</i> comme plante modèle.....	17
3.3. Intérêt biologique des espèces de <i>Medicago</i> les plus étudié.....	17
3.4. La phylogénie de <i>Medicago</i> .....	18
4. <i>Medicago</i> en Algérie .....	19
<b>CHAPITRE II : MATERIEL ET METHODES .....</b>	<b>21</b>
1. Description de <i>Medicago</i> L. ....	21
2. Étude moléculaire du genre <i>Medicago</i> L. ....	24
2.1. Extraction de l'ADN.....	24
2. 2. Amplification <i>in vitro</i> de l'ADN .....	24
2. 3. Le séquençage.....	25
2.3.1 Principe de séquençage.....	25
2 .3. 2. Le séquençage automatique.....	26
2.3.3 Les marqueurs moléculaires .....	27
a) Marqueurs RFLP.....	27
b) Marqueurs RAPD .....	28
2.3.4. Traitement des caractères de type marqueurs moléculaires.....	28
3. Technique de la bioinformatique .....	29
3.1. base de données .....	29

3.2. Consultation de la GenBank .....	30
2. Traitement des séquences de la GenBank .....	32
3.2. MEGA 5.01 .....	32
3.3. Programme <i>ClustalW</i> .....	33
4. Construction de l'arbre phylogénétique.....	34
a) UPGMA (Unweighted Pair Group Method of Arithmetic Averages).....	34
b) Le Bootstrap .....	35
<b>CHAPITRE III: RESULTATS ET DISCUSSION.....</b>	<b>36</b>
1. Résultats de la GenBank.....	36
1.1. Nombres d'espèces et séquences étudiés dans le genre de <i>Medicago</i> .....	36
a) Intérêt de <i>M. sativa</i> .....	38
b) Intérêt de <i>M.truncatula</i> .....	38
1.2. Le choix du gène utilisé.....	39
2. L'arbre phylogénique obtenu.....	41
3. L'analyse de l'arbre phylogénétique.....	43
4. Comparaison avec d'autres arbres .....	43
4.1. Comparaison avec d'autres arbres de gènes différents.....	43
4.2. Comparaison avec d'autres arbres des gènes chloroplastiques.....	51
Conclusion.....	55

## REFERENCE BIBLIOGRAPHIQUE

## ANNEXE

# **LISTE DES ABREVIATIONS**

**ADN** :Acide Désoxy Ribonucléique.

**DATP** :Désoxyadénosine triphosphate

**DCTP** :Désoxycytidine triphosphate.

**DGTP** :Désoxyguanosine triphosphate.

**DTTP** : Thymidine triphosphate

**EMBL**: European Molecular Biology Laboratory.

**ETS** : External Transcribed Spacer

**FISH** : hybridation fluorescente in situ

**GA3ox1**: gibberellin 3-beta-hydroxylase

**GISH** : Hybridation Génomique In Situ

**ITS** : Internal Transcribed Spacer

**MAPK** : Protéines kinases mitogènes activés.

**MAT** : matière azotée totale

**MEGA**: Molecular Evolutionary Genetics Analysis using Maximum Likelihood.

**NCBI**: National Center for Biotechnology Information.

**NORK**: Nodulation Receptor Kinase

**OTU**: Unité Taxonomique Opérationnelle

**Pb** : Paire de bases.

**PCR**: Polymerase Chain Reaction.

**RFLP**: Restriction fragment length polymorphism.

**RUBISCO** : Ribulose Biphosphate Carboxylase Oxytogenase

**Taq**: Thermus aquaticus.

**UPGMA**: Unweight Pair Group Method with Arithmetic mean.

## Liste des figures

Figure	Titre	Page
1	Organisation de la région ADNr nucléaire. Les flèches indiquent l'orientation et la position approximative des sites d'amorce	4
2	Exemple d'un arbre phylogénétique des êtres vivants	9
3	Organisation d'un arbre phylogénétique	9
4	Illustration des définitions d'espèces soeurs et de classe phylogénétique	10
5	Groupes et associations d'unités taxonomiques dans les arbres	11
6	Carte géographique de la distribution de <i>Medicago</i> dans le monde	13
7	Taxonomie du genre <i>Medicago</i>	16
8	Les différentes parties de <i>MEDICAGO rigidula</i>	21
9	Distrubution de <i>M.ciliaris</i> et <i>M.intertxta</i> en algerie	23
10	Les étapes de l'amplification PCR	25
11	Les étapes initiales pour récupérer les séquences nucléotidiques de la GenBank.	30
12	Quelques espèces du genre « <i>Medicago</i> » obtenues dans la GenBank.	31
13	Exemple des séquences nucléotidiques trouvées pour l'espèce <i>MEDICAGO Arabica</i> .	31
14	Représentation des séquences nucléotidiques des espèces du genre <i>Medicago</i> dans le logiciel <i>MEGA5</i> après l'alignement multiple des séquences	33
15	Les paramètres du logiciel <i>MEGA 5</i> utilisé dans la phylogénie	34
16	Graphe des gènes les plus étudié sur <i>Medicago</i> sur excel 2007	40
17	Arbre UPGMA basée sur des données morphologiques de 43 espèces de genre <i>Medicago</i> .	42
18	Arbre phylogénétique de Steele <i>et al</i> , 2010, basé sur la méthode maximum parcimonie, en utilisant le plastide trnK/matK.	45
19	Arbre phylogénétique de Steele <i>et al</i> , 2010, basé sur la méthode maximum parcimonie, en utilisant le gène GA3ox1	47
20	Arbre phylogénétique de Bena <i>et al</i> , 1997 en utilisant les ITS1 et les ITS2	49
21	Arbre phylogénétique d'Erikson <i>et al</i> , 2012, en utilisant le gène NORK	50
22	Arbre phylogénétique de Valizadeh, 1996	51
23	Arbre phylogénétique de 9 espèces de <i>Medicago</i>	52
24	Arbre phylogénétique de 9 espèces de <i>Medicago</i> de différent pays.	53
25	Arbre phylogénétique de 9 espèces de <i>Medicago</i> d'un seul pays	54

## Liste des tableaux

table	Titres	Page
1	Distribution des populations de <i>M. truncatula</i> collectées dans différents pays et composant les collections mondiales existantes.	14
2	Les espèces spontanées du genre <i>Medicago</i> en Algérie.	22
3	Nombres des séquences nucléotidiques pour chaque espèce étudiée du genre de <i>Medicago</i> .	36
4	Les gènes les plus utilisées dans la GenBank du genre <i>Medicago</i> .	39
5	Liste et les caractéristiques des espèces étudiées de Valizadeh (1996)	51
6	L'origine des espèces étudiées obtenue par la <i>Genbank</i> .	53

La reconstruction phylogénétique est la discipline qui vise à reconstruire l'histoire évolutive des êtres vivants. C'est l'une des disciplines en plein essor des sciences de l'évolution. Elle constitue une branche majeure de la systématique et de la génomique comparative. Assurément, la construction d'arbres généalogiques est aussi vieille que la recherche évolutionniste elle-même. En effet, la reconstruction de l'histoire évolutive des espèces est un défi que les scientifiques de plusieurs disciplines essayent de relever. Cette reconstruction se base sur le fait que certains caractères ont été hérités d'une espèce à l'autre, tandis que d'autres se sont transformés au cours du temps, laissant ainsi des traces de l'évolution. L'évolution des espèces est comprise comme un processus divergent, modélisé le plus souvent sous la forme d'un arbre dont les feuilles représentent les espèces contemporaines et les nœuds, les espèces ancestrales. On parle alors d'arbre d'évolution ou plus communément de phylogénie (**Berry, 1997**). Ainsi, outre l'identification, la description et l'inventaire des êtres vivants, une des principales tâches de la systématique est la classification des espèces à partir d'arbres phylogénétiques. Or, cette étape est capitale car elle seule permet de rendre intelligible l'immense diversité des espèces, étudier l'origine et l'épidémiologie des maladies humaines, connaître l'évolution des traits écologiques et comportementaux à travers le temps, estimer les relations biogéographiques historiques et donner la priorité à la conservation des espèces en danger, et la reconstruction des relations historiques à travers tout le temps (**Hillis, 1996**). A l'heure où la disponibilité des sources de données utilisées pour des analyses phylogénétiques augmente à un rythme exponentiel, les ordinateurs modernes ont stimulé le développement des méthodologies sophistiquées, et rapidement un grand nombre de programmes sont devenus disponibles pour analyser divers types de données (**Nylander, 2001**).

Notre travail est divisé en trois parties, la première parle de la phylogénie, la reconstruction de l'arbre phylogénétique et des généralités sur le genre *Medicago* dont on cherche la relation qui existe entre ces espèces. Suivie de la deuxième partie matériel et méthode ou on a cité les logiciels utilisés dans la reconstruction de notre arbre, puis vient la dernière partie qui est les résultats obtenus des logiciels utilisés.

## **1. Définition de la phylogénie**

Le terme « phylogénie » dont phylum désigne en latin race ou tribu, il fut inventé par Ernst Haeckel en 1866 pour définir l'enchaînement des espèces animales et végétales au cours du temps (**Darlu et Tassy, 2004**). Et il l'a défini comme « l'histoire de développement paléontologique des organismes par analogie avec l'ontogénie ou histoire de développement individuel » (**Gallut, 2001**).

Afin de déterminer les relations entre les espèces, celles-ci sont comparées en fonction de différents caractères. Toute la difficulté de cette comparaison est la distinction entre analogie et homologie. L'homologie est la ressemblance entre des espèces attribuable à un ancêtre commun. Cette notion d'ancêtre commun est importante, car une ressemblance entre deux espèces n'implique pas que ces espèces soient apparentées. En effet, des espèces peuvent acquérir un trait commun par sélection naturelle. Telles espèces sont le fruit d'évolution convergente et les ressemblances, dans ce cas, on parle d'analogies (**Chapus, 2005**).

Plusieurs critères peuvent être utilisés pour la classification des espèces, et ces critères sont présentés comme suit :

### **1.1 . Les données morphologiques**

Elles sont basées sur l'étude et l'observation des caractères physiques de l'espèce. Il peut s'agir de la présence de certains organes, la configuration des ossements, la taille de l'espèce. Il s'agit d'utiliser l'anatomie comparative. Ces les premières données qui ont été utilisées pour la classification et elles existent à nos jours.

### **1.2. La biogéographie**

La distribution géographique des espèces sert à les regrouper c'est-à-dire que les espèces qui vivent dans un même milieu sont plus proches, par contre les espèces, séparées par une grande distance, sont différentes. L'utilisation de la biogéographie est très discutée et généralement utilisée comme complément à d'autres critères.

### **1.3. La paléontologie**

Les fossiles servent à regrouper les espèces dont un ancêtre commun peut être trouvé à travers des fouilles géologiques. Historiquement cette méthode a eu un grand poids dans la taxonomie actuelle.

### **1.4. La biologie moléculaire**

C'est le critère le plus récent et le plus promoteur dans la classification. Il s'agit d'utiliser les informations contenues dans les séquences biologiques (ADN et acides aminés) afin de comparer les espèces.

#### **1.4.1. Les différents types de caractères moléculaires**

##### **1.4.1.1. L'hypothèse de l'horloge moléculaire et choix des séquences**

L'idée de l'existence d'une horloge moléculaire dans l'évolution des séquences biologiques a été émise pour la première fois par **Zuckerlandl et Pauling(1965)**. C'est l'hypothèse selon laquelle les mutations observées au niveau moléculaire (macromolécules d'une même classe fonctionnelle) se fixent à une vitesse constante au cours de l'évolution, indépendamment de l'évolution morphologique (**Robinson, 1997 ; Huelsenbeck, 2000**). L'existence d'une telle horloge peut être examinée en étudiant la variation du nombre de changements moléculaires qui se produisent pendant les périodes évolutives d'égale durée, et en regardant si cette variation est plus grande ou non que celle que peut produire le hasard (**Ayala, 1982**). Pour utiliser cette hypothèse à des fins phylogénétiques, il faut s'adresser à la bonne horloge. On peut utiliser différentes molécules comme on utilise les aiguilles d'une montre pour calibrer l'horloge (**Huelsenbeck, 2000**) :

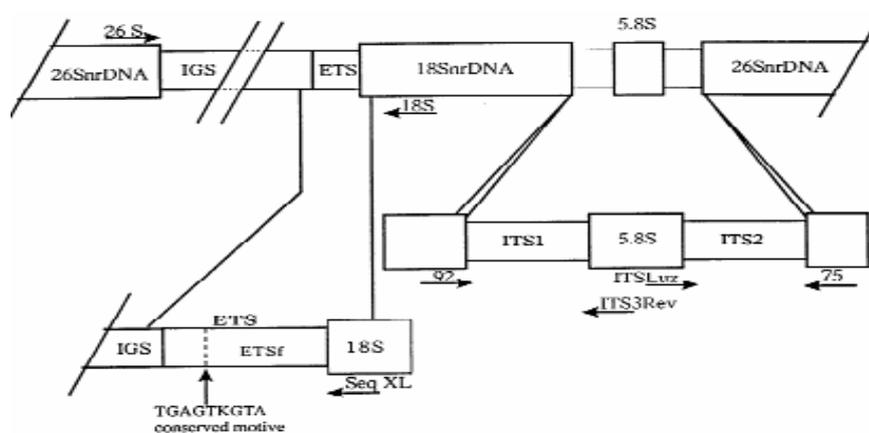
- la trotteuse des secondes (taux de mutation important, par exemple un pseudogène) pour des événements récents (études des populations au sein d'une espèce).
  - l'aiguille des minutes (taux de mutation moyen, par exemple le cytochrome C) pour l'analyse d'un passé proche.
  - l'aiguille des heures (taux de mutations faible : les histones) pour l'étude d'un passé lointain.
- Les conditions minimales que doit remplir une molécule pour pouvoir satisfaire les objectifs de la biologie comparative et la reconstruction phylogénétique sont comme suit (**Mollaret, 2000**) :

- L'ubiquité : la molécule doit être présente chez tous les organismes étudiés.
- L'orthologie : elle doit être qualifiée d'orthologue.
- La conservation : elle doit posséder des régions comparables dont les taux d'évolution ne doivent pas trop différer pour permettre la recherche de sites homologues.

Depuis l'avènement de la PCR (Polymerase Chain Reaction) et le séquençage automatique, les caractères les plus utilisés car plus fiables dans la reconstruction phylogénétique sont les séquences d'ADN notamment : l'ADN ribosomique (ADNr) codant pour les sous-unités des ribosomes, l'ADN mitochondrial et l'ADN chloroplastique. Ces gènes sont ceux qui à ce jour ont fait l'objet de recherche intensive en phylogénie (**Onarici et Sumer, 2002 ; Gagnon, 2004**).

### a) L'ADN ribosomique

L'ADN ribosomique (communément appelé ADNr) codant les ARNs des deux sous-unités des ribosomes et leurs espaces transcrits, ITS (Internal Transcribed Spacer) et ETS (External Transcribed Spacer) utilisés pour un large ensemble d'études phylogénétiques (**Bodo Slotta, 2000 ; Alvarez et Wendel, 2000**). Les premières études ont porté sur les comparaisons des séquences d'ARN qui rentrent dans la constitution des ribosomes, pour ensuite se consacrer à l'ADN codant pour ces ARNs ribosomiques. Les unités de transcription (18S, 5,8S, 26S, ETS, ITS1, ITS2) et les régions non transcrites (NTS) sont présentes dans le génome nucléaire eucaryote en une ou plusieurs centaines de copies (figure 1), elles sont également présentes dans le génome procaryote (**Mollaret, 2000**).



**Figure 1** : Organisation de la région ADNr nucléaire. Les flèches indiquent l'orientation et la position approximative des sites d'amorce (**Bena et al., 1997b**).

Les deux espaceurs transcrits (interne et externe) sont comme leurs noms l'indiquent, des séquences séparant les zones codantes pour les ARNs ribosomiques (5,8S, 18S, et 26S) qui se trouvent transcrits lors de l'expression de ces gènes. Les deux espaceurs ITS1 et ITS2, et l'ADNr 5,8S sont connus sous le nom de la région ITS (**Baldwin *et al.*, 1995 ; Bodo Slotta, 2000**). Quoique l'ITS1 et l'ITS2 appartiennent à l'unité de transcription ribosomale, ces séquences ne sont pas incorporées dans l'ARN ribosomique mur. Cependant, il semble que ces deux régions interviennent dans la maturation des ARNs ribosomiques, la suppression ou la mutation de la région ITS1 peut empêcher la production de la grande et de la petite sous unité ribosomique, la suppression ou la mutation de la région ITS2 peut empêcher la transcription de la grande sous-unité ribosomique (**Bodo Slotta, 2000**). D'après **Baldwin *et al.* (1995); Bodo Slotta (2000) et Alvarez et Wendel (2003)**, plusieurs facteurs font que les deux régions ITS1 et ITS2 de l'ADN ribosomique sont très utilisées dans des analyses phylogénétiques :

- La région ITS et les autres composants de la famille multigène de l'ADN ribosomique sont présents en plusieurs copies dans le génome nucléaire des végétaux, ce qui facilite leur amplification et leur séquençage *in vitro*.
- La famille multigène de l'ADN ribosomique subit une évolution concertée (ordonnée) rapide. Cette propriété de la région ITS est la plus importante du point de vue phylogénétique, car elle favorise la reconstruction précise des rapports phylogénétiques entre les espèces à partir de leurs génomes.
- La région ITS est relativement petite (environ 700 pb) et elle est située entre des régions fortement conservées, l'unité de transcription de l'ARN ribosomique 18s et l'unité de transcription de l'ARN ribosomique 26s. Pour cette raison, des amorces universelles peuvent être utilisées pour amplifier et séquencer la région ITS. Chez la plupart des angiospermes, la région ITS1 est plus grande que la région ITS2 et en conséquence le pourcentage de nucléotides potentiellement informatifs est plus grand dans la région ITS1. La combinaison des deux espaces transcrits (ITS1 et ITS2) dans une seule matrice de caractères donne des arbres phylogénétiques plus robustes et plus fiables (**Bodo Slotta, 2000**).

**b) L'ADN mitochondrial**

L'ADN mitochondrial est présent dans la matrice des mitochondries. C'est une molécule circulaire bicaténaire et sa taille diffère selon les organismes. Tandis que l'ADN mitochondrial animal est constitué d'environ 16 à 18 kilo de paires de bases, l'ADN mitochondriaux des plantes vertes est relativement plus grand et varie entre des limites extrêmes, le plus souvent entre 200 et 2500 kilo de paires de bases (**Oudot-Le seq, 2000**). L'ADN mitochondrial est souvent utilisé dans les études phylogénétiques, tout d'abord parce que le gène est petit, d'isolement facile, à transmission uni-parentale, à taux d'évolution élevé et la relative simplicité des enzymes utilisées comme les endonucléases de restriction. Le fort taux d'évolution serait le résultat d'un appareil de transcription plus laxiste dû à la petite taille du génome, ce taux d'évolution est de 2% par un million d'années. Un constituant de ce génome est le gène de la Cytochrome b oxydase très utilisé dans la systématique moléculaire (**Onarici et Sumer, 2002**).

**c) L'ADN chloroplastique**

D'après Olmstead et al. (1990), parmi les 3 génomes qu'on trouve chez les plantes (le génome nucléaire, le génome mitochondrial et chloroplastique), le génome chloroplastique s'est avéré le plus utile pour les analyses phylogénétiques. En effet, L'ADN chloroplastique est une molécule circulaire relativement petite, sa taille varie peu chez les plantes vertes, elle est de 120 à 217 kilo de paires de base, sa réplication est semi-conservatrice et il est hérité maternellement. Il est présent en nombre élevé de copies (souvent 5000 copies par cellule) le rend relativement facile à extraire. Pour des buts de la systématique, même les extraits de l'ADN total de 2 à 3 grammes de feuilles fraîches du matériel végétal sont riches en ADN chloroplastique (**Onarici et Sumer, 2002**), il possède 40 à 250 gènes très compacts, c'est-à-dire qu'ils comportent très peu de séquences non codantes (**Gagnon, 2004**). Le taux de substitution des nucléotides de l'ADN chloroplastique est relativement faible par rapport à l'ADN nucléaire et mitochondrial animal (**Olmstead et al., 1990**).

**1.4.1.2. Intérêts des caractères moléculaires**

Depuis les années 60, les macromolécules biologiques ont joué un rôle de plus en plus important dans les études évolutives et systématiques. Ces macromolécules biologiques, utilisées comme sources de données moléculaires, sont les protéines et les divers types d'acide désoxyribonucléiques (**Tourasse, 1992; Onarici et Sumer, 2002**). Les protéines

étaient au départ plus faciles à séquencer, mais à présent l'ADN est préféré aux protéines, car il est le véritable support de l'information génétique, celui qui se transmet, permet l'évolution et la fixation des caractères phénotypiques. De plus, les protéines n'expriment pas tout le polymorphisme génétique, elles ne donnent pas accès au polymorphisme des zones non codantes des chromosomes (**Agolin, 2003**). Du point de vue phylogénétique, la définition d'un caractère moléculaire est simple : chaque position sur la molécule, c'est-à-dire chaque résidu (nucléotide ou acide aminé) est un caractère et chaque type de résidu a une position (une des quatre bases azotées ou un des 21 acides aminés) est un état de caractère. L'homologie primaire est établie en postulant que deux protéines ou deux segments d'ADN ayant des séquences, des tailles et souvent des fonctions, mais pas obligatoirement proches sont considérées comme héritées d'un ancêtre commun et ayant divergé pendant l'évolution séparée des deux groupes d'organismes dont ils sont issus. Chaque résidu doit trouver son homologue dans l'autre séquence, ce qui est ajusté lors de la phase d'alignement (**Berry, 1997; Montagnon, 2001**). Par rapport à l'analyse morphologique, l'étude moléculaire permet de travailler sur un nombre de caractères beaucoup plus important, puisque pour un gène de 1500 paires de base, c'est 1500 caractères qui sont analysés (**Pirroux, 2002 ; Schmidt, 2003**). Les autres types de caractères moléculaires sont ceux fournies en utilisant les différents marqueurs moléculaires pour analyser le polymorphisme de l'ADN (RFLP, RAPD...etc.) (**Salemi et Vandamme, 2003**). La RAPD et la RFLP sont les marqueurs les plus utilisés pour comparer des taxa sur la base du polymorphisme des fragments d'ADN.

## 2. La phylogénétiques moléculaires

La phylogénie moléculaire a pour but de reconstruire les relations de parenté entre des séquences de nucléotides ou d'acides aminés (**Lopez et al, 2002**). L'objectif principal de cette étude est de récupérer l'ordre des événements évolutifs et de les représenter dans les arbres comme moyen d'organiser les espèces en fonction de caractères communs, bien avant la proposition de Darwin de l'utiliser comme modèle de l'évolution selon les mécanismes de la sélection naturelle. Ces arbres ont une racine correspondant au concept le plus général dans le cas d'une hiérarchie de concepts, ou à un ancêtre commun à toutes les espèces dans le cas de l'arbre de la vie (**Dowell, 2008 ; Gambette, 2010**).

Le postulat de base de la reconstruction phylogénétique est que la ressemblance est intelligible en termes d'ascendance commune. Le problème général est donc celui de l'inférence sur l'ancêtre et sur ses caractères à partir de l'observation des caractères des taxa

terminaux (**Darlu et Tassy, 1993**). Pour construire une phylogénie, il faut disposer de caractères comparables entre tous les objets (c'est-à-dire gènes ou espèces) que l'on veut analyser. En d'autres termes, les objets analysés doivent être « suffisamment similaires » pour être comparés (**Lopez et al., 2002**); il est impossible de reconstruire la phylogénie des séquences moléculaires, si elles sont toutes semblables ou si elles ne partagent aucun point commun. Il faut qu'elle soit issue d'une même séquence ancestrale (**Ranwez, 2002**).

D'après **Gattolliat (2002)**, les buts d'une reconstruction phylogénétique ne se limitent pas à vouloir reconstruire un arbre aussi proche que possible de l'histoire des taxa, mais également à estimer la manière dont les transformations des caractères se répartissent dans l'arbre.

La construction d'une phylogénie peut répondre à de nombreux besoins, comme par exemple en taxonomie pour tester des classifications existantes et en établir de nouvelles, établir la monophylie d'un groupe et valider des taxa. La systématique phylogénétique refuse de considérer les grades évolutifs comme des taxa classificatoires, sauf s'ils correspondent à des clades (groupes d'espèces appartenant à une même lignée). L'histoire inférée permet aussi d'étudier l'évolution des caractères dans un clade, et enfin, elle donne une base indispensable à la biologie comparative en introduisant une « correction phylogénétique » à l'étude des corrélations entre caractères (**Agolin, 2003**).

### **2.1. Phylogénie et généalogie**

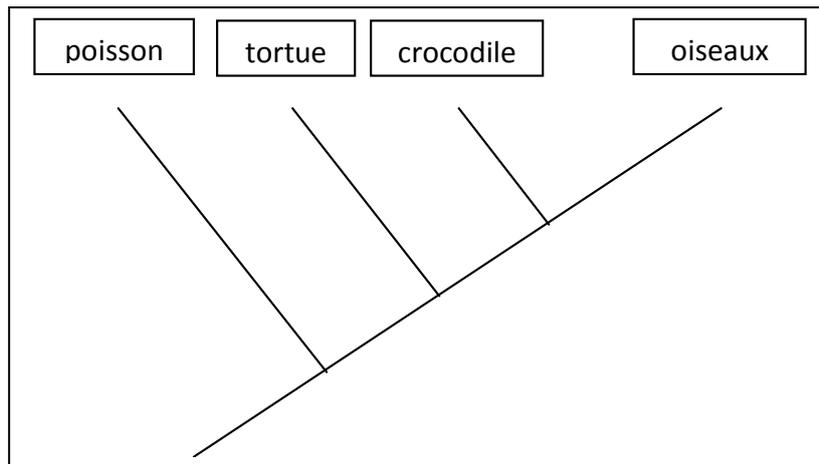
La phylogénie s'apparente à la généalogie et on a d'ailleurs souvent tendance à confondre ces deux termes. Pourtant, la phylogénie est un arbre où l'ascendance est unique, alors que certains arbres généalogiques peuvent présenter une ascendance multiple, c'est le cas d'une généalogie familiale, ou un individu à deux ascendants, à savoir une mère et un père (**Moura, 2002**).

### **2.2. Terminologie phylogénétique**

Nous présentons ci-dessous la définition de quelques termes propres au domaine de la phylogénie que nous emploierons souvent par la suite. Elles concernent les relations phylogénétiques entre les organismes.

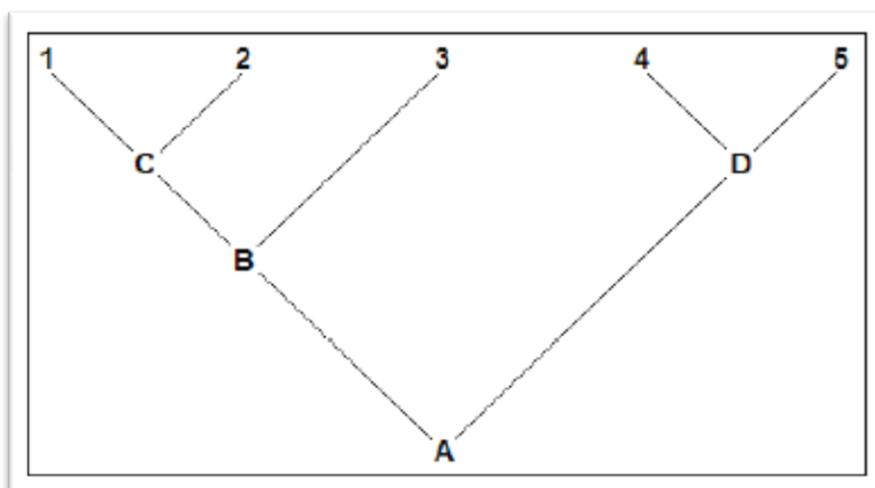
### 2.2.1. Les arbres phylogénétiques

La phylogénie est représentée sous la forme d'un arbre – avec un tronc, des branches, des rameaux – qui se lient de bas en haut. La seule illustration incluse dans l'Origine des espèces est une image de filiation entre espèces hypothétiques situées dans un contexte stratigraphique : le temps géologique se lit de bas en haut (voir la figure 2) (Darlu et Tassy, 2004).



**Figure 2** : exemple d'un arbre phylogénétique des êtres vivants.

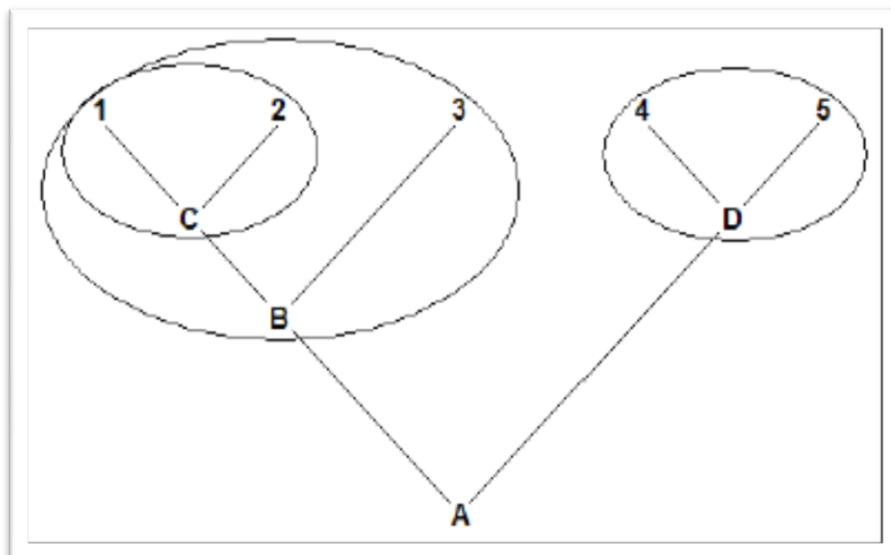
L'arbre présenté en figure 3 est une représentation hypothétique de la phylogénie de 5 phylums. Ces 5 phylums, notés 1 à 5, se situent à l'extrémité des branches dites terminales. Cet arbre est constitué de 3 nœuds, notés B à D, et d'une racine notée A (Vally, 2007).



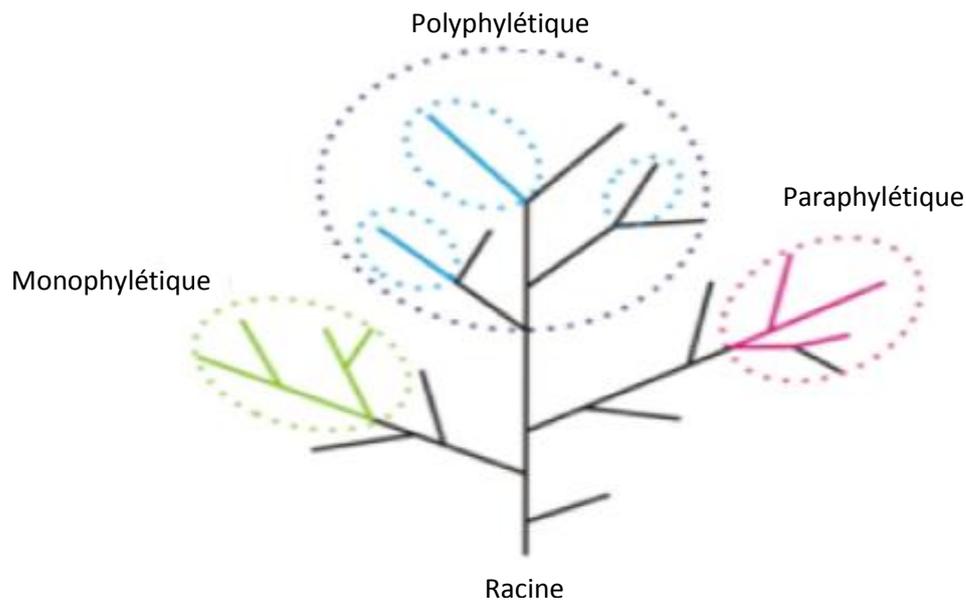
**Figure 3**: organisation d'un arbre phylogénétique (Vally, 2007).

Chaque nœud représente un groupe ou taxon qu'on appelle « événement de spéciation » (Vally, 2007), les nœuds terminaux à l'extrémité des arbres représentent unités taxonomiques opérationnelles (UTO). OTU correspondent aux séquences moléculaires ou taxons (espèces) (Dowell, 2008). L'espèce constitue le taxon de base de la classification. Les nœuds internes représentent les points de divergence des espèces. Ils peuvent aussi correspondre aux derniers ancêtres communs (LCA) (Vally, 2007; Dowell, 2008), L'arrangement des branches est appelé topologie. La longueur des branches est proportionnelle au nombre de changements évolutifs qui se sont produits entre deux nœuds (Vally, 2007).

Les arbres peuvent être constitués d'un seul gène de nombreux taxons (un arbre d'espèces) ou multigénique familles (arbres de gènes), Ces branches peuvent être regroupées de façon différente (Dowell, 2008). Une classe, donc, est composée d'espèces sœurs (ou de phylums frères). On peut ainsi décomposer L'arbre de la figure 4 en 3 classes: les espèces 1 et 2 sont sœurs, l'espèce 3 est quant à elle sœur de la classe formée par les espèces 1 et 2 et leur ancêtre commun C, et la classe formée par les espèces sœurs 4 et 5 et leur ancêtre commun D est sœur de toutes les autres. On dit alors de cette dernière classe qu'elle est la plus basale de l'arbre, les autres étant plus dérivées (Vally, 2007).



**Figure 4 :** Illustration des définitions d'espèces sœurs et de classe phylogénétique (Vally, 2007).



**Figure 5** : Groupes et associations d'unités taxonomiques dans les arbres (Dowell, 2008).

On parle d'un groupe monophylétique celui qui se compose d'un nœud de LCA interne qui en découle (figure 5). Tous les membres au sein de ce groupe sont dérivés d'un même ancêtre et ont hérité d'un ensemble de traits communs uniques. Le groupe paraphylétique est celui qui exclut certains de ses descendants. Et un groupe polyphylétique peut être une collection des OTU lointaines qui sont associés par un caractéristique similaire ou d'un phénotype, mais ne sont pas directement descendus d'un ancêtre commun (Dowell, 2008).

Il y a trois types de méthode de reconstruction phylogénétique qui sont :

- a- La méthode des distances génétique** : calcul de distance génétique (nombre estimé de substitutions nucléotidiques) entre chaque paire d'OTU ; construction progressive de l'arbre phylogénétique à partir de la matrice de distances.
- b- La méthode de maximum de parcimonie** : Recherche de l'arbre phylogénétique dont la topologie demande le nombre le plus faible de changement évolutifs (substitutions nucléotidiques) nécessaire pour expliquer les différences observés entre les OTUs.
- c- La méthode de maximum de vraisemblance** : Recherche de l'arbre phylogénétique dont la topologie est la plus vraisemblable étant donné les séquences nucléotidique des OTUs (les données) et le modèle de l'évolution des séquences nucléotidiques choisi.

La phylogénie est appliquée sur plusieurs espèces animales qu'elle soit ou végétales pour savoir les relations qui se trouvent entre espèces, dans notre projet, on a choisi le genre *Medicago* dont plusieurs chercheurs s'y intéressent grâce à son intérêt sur les plans économiques et génétiques, et contient des espèces modèles.

### **3. Le genre de *Medicago* L.**

#### **3.1. Généralité**

Le nom scientifique du genre *Medicago* n'est pas lié à ses propriétés médicales, mais au fait que, selon Théophraste (372 - 287 av. J.-C.), la luzerne serait originaire de Médie (région allant du Nord-Ouest de l'Iran à l'Azerbaïdjan) (Pierre, 2008). Elle portait alors le nom de *Medicaherb* à « l'herbe de Médie », devenu plus tard le nom de genre : *Medicago* (Messioughi, 2010). Le nom anglais et espagnol alfalfa proviendrait du nom arabe al-fac-façah, qui signifie « le père des aliments » car il servait de fourrage aux chevaux. Quant au nom vernaculaire, il est emprunté à l'occitan luserna, en raison de l'aspect brillant des graines de la plante (Pierre, 2008).

#### **Position systématique d'après Halmi, 2010**

**Règne :** Plantae

**Embranchement :** Spermatophytes

**Sous- embranchement :** Angiospermes

**Classe :** Dicotylédones

**Sous- classe :** Rosidées

**Ordre :** Fabales

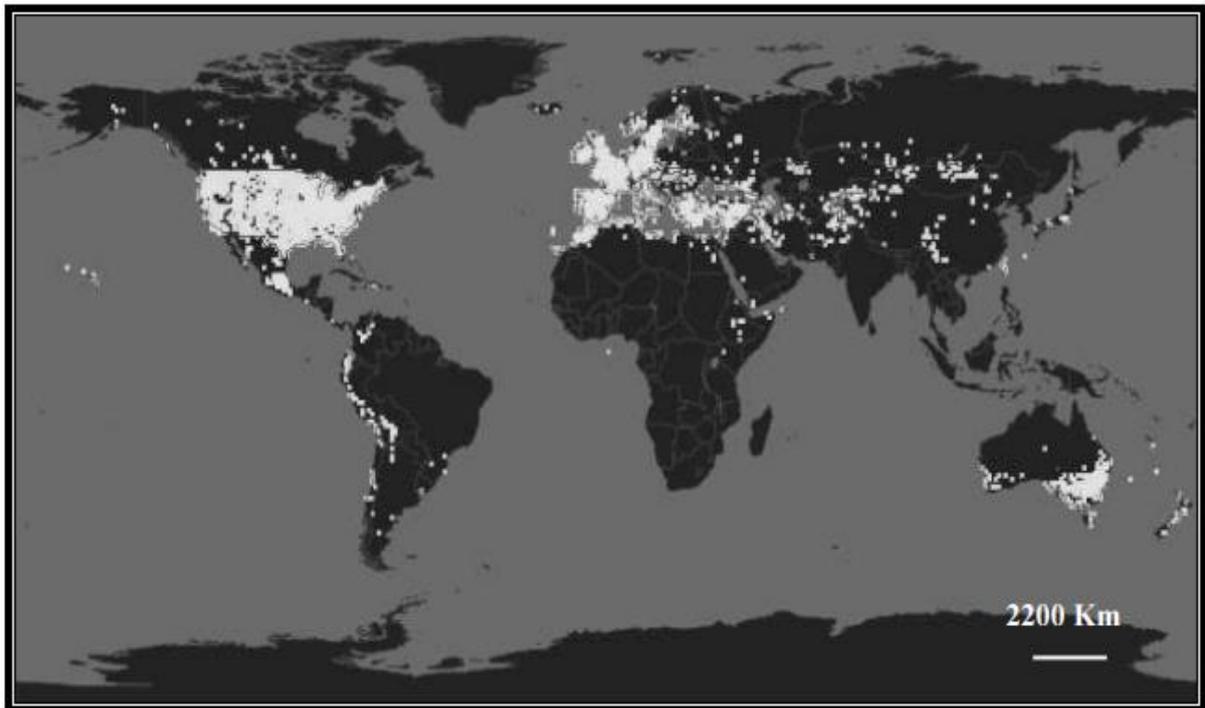
**Famille :** Fabaceae

**Sous-famille :** Faboideae

**Tribus :** Trifolieae

**Genre :** *Medicago* (L.)

Il est admis à la suite de Vavilov que le centre d'origine du genre *Medicago* est le Croissant fertile. La plus vieille référence connue de culture de la luzerne cultivée (*Medicago sativa L.*) date de 1300 ans avant J.C. en Turquie. Mais son extension en Europe n'a débuté réellement qu'avec l'Empire romain, même si les phéniciens l'ont introduit dans le Bassin méditerranéen occidental. Elle se répand ensuite et à la fin du XVIIIe siècle sa zone de culture est mondiale (figure 6) (Prosperiet *al.*, 1993).



**Figure 6:** Carte géographique de la distribution de *Medicago* dans le monde (Halmi, 2010).

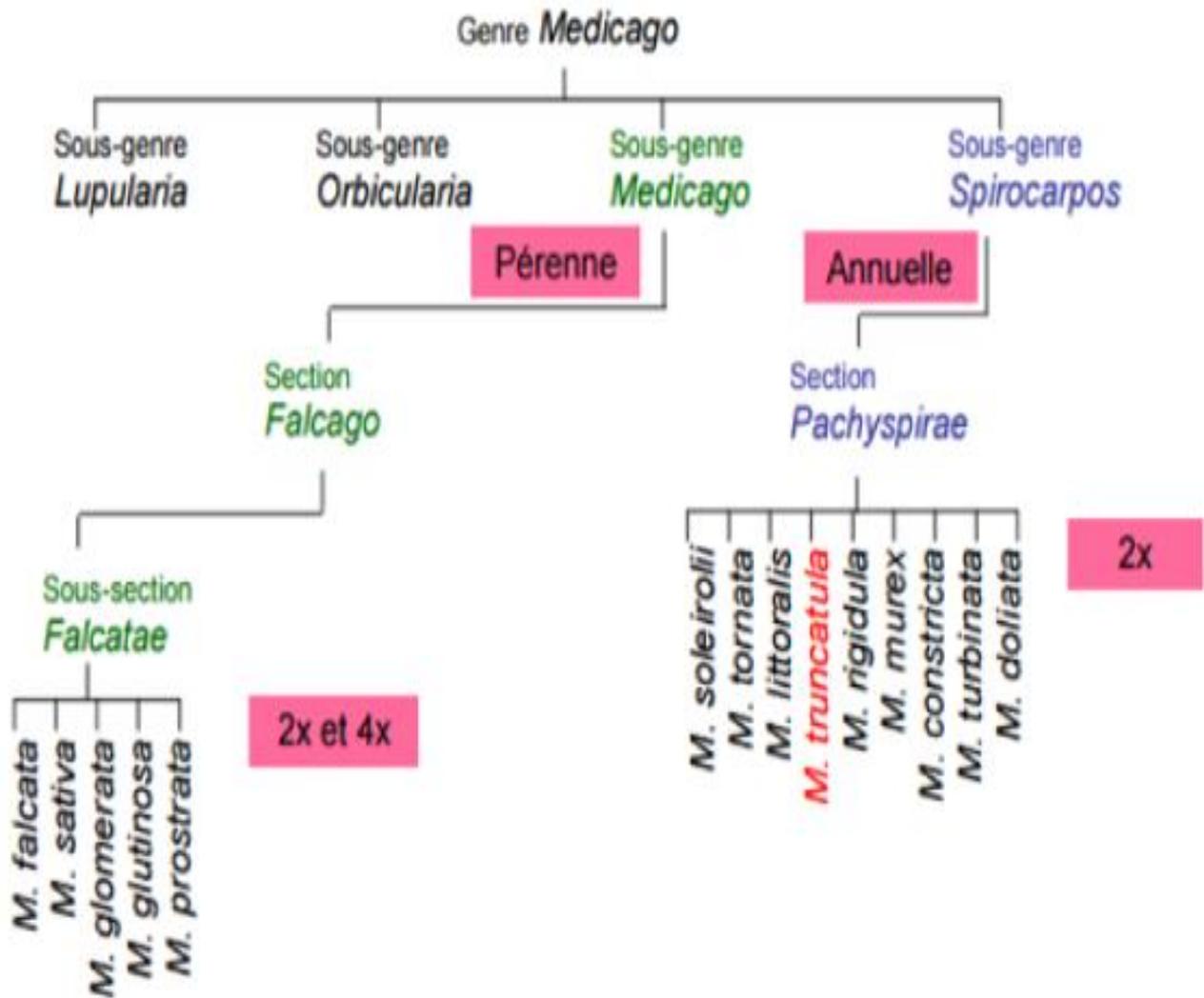
Le tableau suivant montre la distribution de *M. truncatula* qui est une plante modèle dans le monde, elle est fréquente dans la méditerranée et l'australie (figure 7).

**Tableau I:** Distribution des populations de *M. truncatula* collectées dans différents pays et composant les collections mondiales existantes (Source: *Medicago truncatula* handbook, version Novembre 2006).

<b>Pays</b>	<b><i>M.truncatula</i></b>
Algérie	<b>755</b>
Tunisie	<b>951</b>
Maroc	<b>909</b>
Lybie	413
Egypte	2
Jordanie	38
Israël	296
Iran	3
Syrie	3
Türk	50
Cyprès	424
Malta	6
Grèce	355
Italie	292
France	142
Spain	226
Portugal	110
Canarie & Médira	13
Turkménistan et Azerbaïdjan	3
Russie	1
Est de l'Europe	7
Ethiopie	1
Amérique	3
Australie	612
Sud Afrique	17
Total	5632

nombre d'espèces diploïdes, tétraploïdes, exceptionnellement hexaploïdes (*M. arborea*, *M. saxatilis*). Le nombre chromosomique de base est  $x = 8$ , exceptionnellement  $x = 7$  (*M. polymorpha*, *M. praecox*, *M. rigidula* et certaines formes de *M. murex*) (Halmi, S, 2010) ; Le nombre de chromosomes est égal à  $2n=16, 17, 18, 32$  (Heyn, 1963), il appartient à l'ordre des Fabales, superfamille des Légumineuses, famille des *Fabaceae* et sous-famille des *Faboideae*. Le genre *Medicago* est placé dans la tribu des *Trifolieae* qui comprend la majorité des Légumineuses fourragères tempérées comme la luzerne cultivée et le trèfle, et qui est phylogénétiquement proche de la tribu des *Vicieae* incluant les Légumineuses à graines comme le pois (*P. sativum*) et la féverole (*Vicia faba*) (Pierre, 2008).

Le genre *Medicago* .L de la famille des *Fabaceae* (*legumineusae*) est un genre très large et très divers en espèces, environ 87 espèces d'herbes et arbustes répartis de la Méditerranée à l'Asie centrale, comprend le modèle des espèces de légumineuses *M. truncatula*, et nombre d'espèces fourragères importantes telles que *M. sativa* (luzerne), *M. scutellata*, et *M. lupulina* (Steele et al, 2010). Il a été classé en quatre sous-genres sur la base de la morphologie des gousses et des graines : *Lupularia*, *Medicago*, *Orbicularia* et *Spirocarpos* (Figure 8) (Pierre, 2008).



**Figure 7** : Taxonomie du genre *Medicago* (Pierre, 2008).

Selon Pierre, 2008, ces sous-genres comprennent au total 55 espèces dont 21 sont pérennes, 33 sont annuelles et une est de pérennité non définie. Le sous-genre *Spirocarpos* est composé de 29 espèces annuelles (médics) et comporte quatre sections : *Intertextae*, *Leptospirae*, *Pachyspirae* et *Rotatae*. *M. truncatula* appartient à la section *Pachyspirae* aux côtés de huit autres espèces : *M. constricta*, *M. doliata*, *M. littoralis*, *M. murex*, *M. rigidula*, *M. soleirolii*, *M. tornata* et *M. turbinata* (Pierre, 2008).

Les relations au sein du genre ne sont pas encore suffisamment résolues, contribuant à la difficulté à comprendre l'évolution d'un nombre de caractéristiques distinctives telles que l'aneuploïdie et polyploïdie, l'histoire de la vie, de la structure des cotylédons, et le nombre de graines par fruit. Diverses études phylogénétiques du genre *Medicago* ont été menées

(Damerval, 1983; Rose et al, 1988; Brummer et al., 1995; Mariani et al., 1996; Valizadeh et al., 1996), mais seulement quelques espèces annuelles ont été inclus dans chaque étude (Bena et al., 1997).

### 3.2. *Medicago* comme plante modèle

Le genre *Medicago* est très important sur le plan économique particulièrement *M. sativa*, distribue sur tout le bassin méditerranéen. Le genre contient une grande variété de caractères morphologique, plusieurs études phylogénétiques de *Medicago* ont été menées mais seulement quelques espèces annuelles ont été incluses dans chaque étude (Bena et al., 1997). un programme de recherche de l'INRA a permis de définir *Medicago truncatula* comme espèce modèle pour une analyse moléculaire de la résistance aux maladies (Duc et al., 2011).

*Medicago truncatula* présente des caractéristiques permettant et facilitant les analyses génétique et moléculaire (autogamie, faible niveau de ploïdie, génome de petite taille, aptitude à la culture en laboratoire, à l'embryogénèse somatique et à la transgénèse) ainsi qu'une bonne aptitude à entrer en symbiose avec la bactérie *Rhizobium meliloti*, généralement utilisée pour les études portant sur cette spécificité des Légumineuses. Par ailleurs, le temps de génération de graine à graine est relativement court (10 à 12 semaines en conditions contrôlées) et la production importante, ce qui assure une disponibilité rapide du matériel (Brunel, 2008). En raison de ces caractéristiques génétiques, de sa grande biodiversité naturelle, de la facilité de sa transformation et de son interaction avec des microorganismes symbiotiques et pathogènes, *Medicago truncatula* a été choisie comme plante modèle au sein de la famille des Papilionacés (Djebali, 2008).

### 3.3. Intérêt biologique des espèces de *Medicago* les plus étudié

Cultivée pour la production du foin la luzerne est utilisée depuis quelques décennies pour la fabrication d'aliments d'excellente qualité (Fellous, 2011), parce qu'elle est caractérisé par une valeur énergétique moyenne (0.7 UF/Kg de matière sèche), une teneur en protéine élevée et une teneur en matière azotée totale (MAT) importante qui peut varier de 14 à 29% de matière sèche (Messioughi, 2010). Elle est riche en acide amines, vitamines B, C et E, en carotène, provitamine A, favorable à la fertilité, elle s'avère relativement pauvre en sucre et amidon (Kadri et Midoun, 2015).

Les feuilles de la luzerne sont riches en RUBISCO « Ribulose Biphosphate Carboxylase Oxytogenase » qui est une enzyme qui permet la fixation du Co2 et son oxygénation est riche

en acide amines essentiels, et sa composition se rapproche de celles des protéines laitière. (Ben Nja, 2014), elle agirait comme élément protecteur contre le diabète et le cholestérol (Messioughi, 2010), elle contient d'importante quantité de fer, de calcium, de magnésium et de potassium, elle a une action antihémorragique grâce à la vitamine K et anti cholestérol grâce au coumestrol, un œstrogène végétale (Kadri et Midoun, 2015), et elle est utilisé depuis longtemps contre les rhumes, les abcès, les troubles digestifs ainsi elle a des propriétés reminéralisantes, anti infectieuses et desintoxicantes (Ben Nja, 2014).

La luzerne fixe l'azote atmosphérique, mais elle utilise préférentiellement l'azote nitrique du sol, ce qui induit à la diminution de l'azote nitrique du sol, donc elle permet de récupérer et de soustraire au lessivage des surplus de nitrate du sol protégeant ainsi les nappes phréatique, et elle permet ainsi des économies et améliore le bilan environnemental par rapport a d'autre culture fourragère (Messioughi, 2010). Elle joue aussi un rôle important de couverture naturelle du sol pendant plusieurs années, qui le préserve des problèmes d'érosion ou de ruissellement (Kadri et Midoun, 2015).

*M. truncatula* présente plusieurs avantages qui la prédisposent aux études génétiques. C'est une plante annuelle, diploïde ( $2n=16$ ), autogame, et elle possède un génome de faible taille de 5 108 pb, soit environ 4 x celui d'*Arabidopsis*. C'est également une plante apte à la transformation par *AGROBACTERIUM tumefaciens* et à la régénération par embryogenèse somatique (Thomas et al. 1992; Chabaud et al, 1996). Elle forme de petites graines permettant leur culture dans des tubes à essais. *M. truncatula* a permis d'identifier un grand nombre de gènes reliés à la symbiose dont certains ont déjà été isolés, séquencés et leur expression étudiée (Lazrek, 2008).

### 3.4. Phylogénie de *Medicago* :

Vue l'importance du genre *Medicago* plusieurs recherches ont été faite sur la relation familiale de *Medicago*. Afin de mieux comprendre leurs relations taxonomiques. Mais il y a une incertitude dans les arbres phylogéniques en utilisant de différents gènes.

- D'après Valizadeh en 1996, pour savoir la distance entre les espèces, il a utilisé la méthode RFLP (longueur des fragments de restriction) des chloroplastes (Valizadeh, 1996).

- **Marcela Rosato en 2008** à utiliser ADNr afin d'évaluer les niveaux de variation concernant les loci ribosomal sur *M. citrina* et d'évaluer la nature de son polyploïdie. Et pour cela il a utilisé la méthode d'hybridation fluorescente in situ (*FISH*) pour la cartographie physique des séquences 5s et 45s des trois espèces appartenant à la section *Dendrotelis* (*M. arborea*, *M. citrina*, *M. strasseri*) et la méthode *GISH* (Hybridation Génomique In Situ) pour montrer la liaison en elle (**Rosato et al, 2008**).
- Bena et ses collaborateurs ont effectué une étude phylogénique moléculaire basée sur le transcrit ribosomal interne et externe nucléaire (STI et ETS), et ils ont montré que les ETS semblent une source d'information intéressante pour la phylogénie des plantes. Et qui est le même dans 13 espèces étudiées (**Bena et al, 1998b**).
- En **2010, Steele et ses collaborateurs**, ont utilisé les séquences de nucléotide de la région *trnk* plaste/ *mat k* et le gène *GA3ox1* code par le noyau (gibberelline 3- $\beta$ -hydroxylase) en utilisant la parcimonie pour effectuer une relation phylogénétique de 70-73 espèces de *Medicago* (**Steele et al, 2010**).
- **Jonna Eriksson** a travaillé sur le gène nucléaire *NORK*, en examinant s'il y a des anciens événements ou récents qui induirait à une erreur de ressemblance phylogénétique (**Erikson, 2012**).

#### 4. *Medicago* en Algérie

Le genre *Medicago* est représenté par un grand nombre d'espèces en région méditerranéenne. Les espèces pérennes sont *M. falcata* (L.) Lam., *M. lupulina* L., *M. marina* L. et *M. sativa*. Ce dernier taxon, très bien adapté au pâturage, est fréquent en Algérie orientale dans les régions d'altitude (**Abdelguerfi et Abdelguerfi, 2004**).

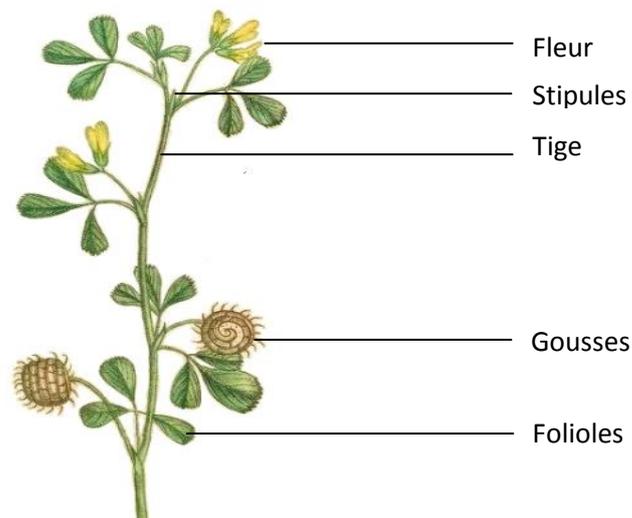
Les espèces du genre *Medicago* possèdent un intérêt agronomique grâce à leur aptitude à la fixation symbiotique de l'azote, permettant une production abondante en protéines végétales, et grâce à leur capacité d'enrichissement des sols en azote organique (la symbiose) ; ils sont souvent utilisés dans les systèmes de rotation avec les céréales. En Algérie, ces plantes assurent l'amélioration de la flore des jachères pâturées et entrent facilement dans la rotation avec les céréales. Par ailleurs, des travaux ont été entrepris pour utiliser des extraits foliaires des luzernes (cas de *M. sativa*) comme aliment destiné à des populations souffrant de carence alimentaire forte, en particulier chez les enfants (**Halimi, 2010**).

Divers travaux sur les genres *Medicago* (**Abdelguerfi et Laouar, 1999**) ont montré une corrélation positive entre les paramètres biogéographiques, particulièrement l'altitude et la pluviosité, et les traits biologiques liés à la croissance et à la reproduction – floraison, inflorescences, gousses et graines. Chez les luzernes, précisément dans le complexe *Medicago ciliaris-intertexta*, cette variabilité soulève aussi des questions d'ordre taxonomique. Les deux espèces, *M. ciliaris*(L.) All. et *M. intertexta* (L.)Mill., ont une répartition à tendance allopathique. La première, dans le Nord-Ouest, montre une tolérance vis-à-vis de la sécheresse; la seconde se rencontre essentiellement dans les régions humides du Nord-Est). Dans les situations de sympatrie, des « populations particulières » se différencient par la forme des épines et la pilosité des gousses. Leur isolement par rapport aux deux autres taxons, *ciliaris et intertexta*, est confirmé par les marqueurs microsatellites et, sans qu'il soit encore possible d'en préciser une nature hybride, leur « originalité taxonomique » est mise en exergue par leur capacité symbiotique spécifique avec *SINORHIZOBIMUM meliloti* (**Amirouche et Misset, 2009**).

## 1. Description de *Medicago* L.

Les plantes du genre *Medicago* sont annuelles ou pérennes, herbacées ou arbustives. Leurs caractères morphologiques sont comme suit :

- Les feuilles sont trifoliées pétiolées et non terminées par une vrille et leurs folioles sont dentées sur la partie distale.
- Deux stipules collées aux pétioles (**Lesins et Lesins, 1979 ; Chakravatry, 1976 et Prosperi et al., 1995**).
- Une corolle papilionacée, caduque, constituée d'un étendard de deux ailes libres et d'une carène formée par deux pétales inférieurs soudés.
- Androcée diadèlphe (9 + 1) : neuf étamines soudées, formant la colonne staminale, plus une dixième étamine libre.
- Le calice est formé de cinq sépales soudés.
- Les inflorescences pédonculées portent jusqu'à 30 fleurs libres.
- La gousse est soit courbée, soit tordue, soit roulée en hélice, épineuse ou non, indéhiscente ou presque (**Quezel et Santa ,1962 ; Prosperi et al., 1995**)
- Le port est généralement dressé mais peut être prostré (collé au sol), les tiges peuvent être poilues.



**Figure 9** : Les différentes parties de *MEDICAGO rigidula*.

( <https://commons.wikimedia.org>)

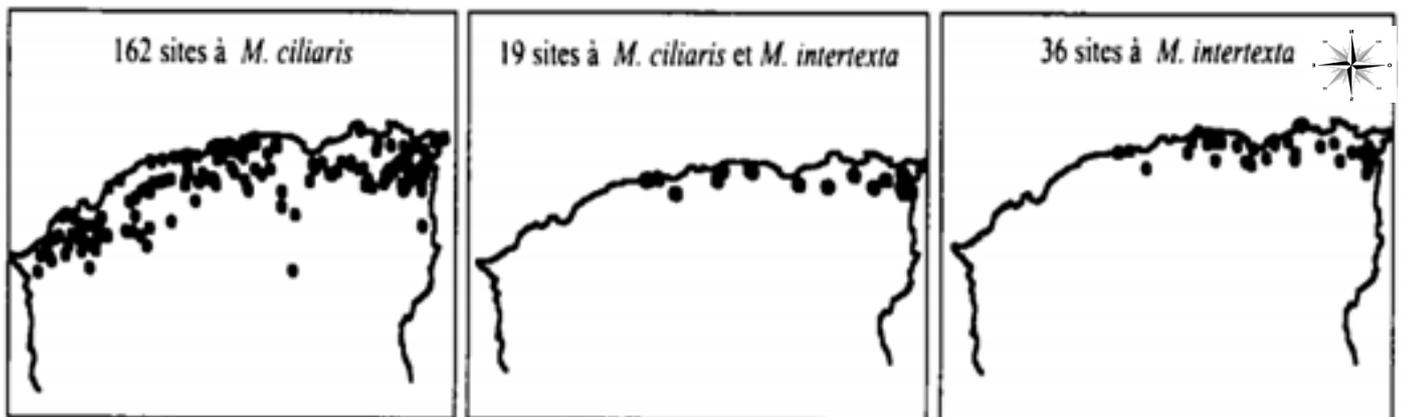
Le genre *Medicago* en Algérie est représenté par de nombreuses espèces spontanées dont les plus importantes sont décrites dans le tableau suivant :

**Tableau II** : les espèces spontanées du genre *Medicago* en Algérie (Halmi, 2010).

Nom d'espèce	type de plante	Gousse	fleur	Habitat
<i>M. intertexta</i>	Annuelle	Glabre de 12 à 15 mm	Jaune	Pâturages, Tell AlgéroConstantinois
<i>M. ciliaris</i>	Annuelle	Velue hirsute, plus ou moins glanduleuses de 10- 15 mm	Jaune	Pâturages, prairies, Tell à sol semi salin
<i>M. falcata</i>	Vivace	Courbée en faucille ou en anneaux, pubescente	Jaune	Pâturages argileux
<i>M. lupulina</i>	Annuelle ou bisannuelle	Réniforme, fortement arquée de 2-3 mm de diamètre	Jaune	Pâturages
<i>M. secundiflora</i>	Annuelle ou bisannuelle	Brun clair de 2-2.5 mm	Jaune	Pâturages
<i>M. marina</i>	Vivace	Contournée en hélice	Jaune	Sables maritimes
<i>M. sativa</i>	Vivace	glabre ou pubescente, à 2-3 tours de spire	Violettes ou jaunes lavées de brun	Cosmopolite
<i>M. scutellata</i>	Annuelle	Hémisphérique à spires emboîtées de 13-15 mm	Jaune	Sols argileux du Tell
<i>M. orbicularis</i>	Annuelle	Ovoïde à 5-6 tours de spires	Jaune	Pâturages du Tell
<i>M. rugosa</i>	Annuelle	Ne dépassant pas de 9 mm	Jaune	Pâturages de Constantine et Alger
<i>M. soleirolii</i>	Annuelle	Glabres	Jaune	Pâturages du Tell Constantinois
<i>M. rigidula</i>	Annuelle	Glanduleuse-fortement pubescente de 10 mm	Jaune	Broussailles, Montagnes
<i>M. tuberculata</i>	Annuelle	Tronquée à sa base, Glabre de	Jaune	Pâturages arides

		10 mm		
<i>M. truncatula</i>	Annuelle	Tronquée à ses deux extrémités De 7-8 mm	Jaune	Tell – Pâturages
<i>M. littoralis</i>	Annuelle	Ovoïde ou globuleuse de 10 mm	Jaune	Pâturages
<i>M. murex</i>	Annuelle	Glabres ou faiblement hispides	Jaune	Tell Algéro- Constantinois
<i>M. minima</i>	Annuelle	Contournée en hélice, pubescente de 3-4 mm	Jaune	Sols pauvres
<i>M. arabica</i>	Annuelle	Glabre de 5-7 mm	Jaune	Tell Algéro- Constantinois
<i>M. laciniata</i>	Annuelle	Ovoïde ou sphéroïdale de 5-6 mm	Jaune	Paturages arides Montagnes
<i>M. hispida</i>	Annuelle	Fortement réticulée- striée de 5-9 mm	Jaune	Tell – Pâturages

Le genre *Medicago* est représenté par un grand nombre d'espèces en région méditerranéenne. Les espèces pérennes sont *M. falcata* (L.) Lam., *M. lupulina* L., *M. marina* L. et *M. sativa* subsp. *Tunetana* Murb. Ce dernier taxon, très bien adapté au pâturage, est fréquent en Algérie orientale dans les régions d'altitude (**Abdelguerfi et Abdelguerfi-Laouar, 2004**) et les espèces les plus connues du genre sont *Medicago ciliaris* et *Medicago intertexta* qui constituent les principales ressources en légumineuses fourragères (**Laouar et Abdelguerfi, 2003**). Dont leur distribution est montrée sur la carte qui suit :



**Figure 10 :** Distribution de *M.ciliaris* et *M.intertexta* en Algérie

L'objectif de cette étude est d'établir les relations phylogénétiques qui peuvent exister entre les espèces du genre *Medicago*, et la construction de l'arbre phylogénique de ce genre qui est réalisé à partir des séquences d'ADN des gènes codant, traité par des logiciels de la bioinformatique.

Cependant ces séquences d'ADN soumises à des différentes techniques de la biologie moléculaires avant de les traiter par la bioinformatique.

## 2. Étude moléculaire du genre *Medicago* L.

### 2.1. Extraction de l'ADN

Pour l'extraction de l'ADN, plusieurs auteurs ont basé sur la méthode de Doyle et Doyle (1987) qui consiste à :

- ❖ La conservation des échantillons dans de l'éthanol et les sécher ;
- ❖ briser les parois cellulaires en broyant l'échantillon en présence de l'azote ;
- ❖ dénaturer les protéines avec de la protéinase K ;
- ❖ Hydrolyser l'ARN avec de l'ARNase ;
- ❖ récupérer le culot de l'ADN avec la centrifugation.

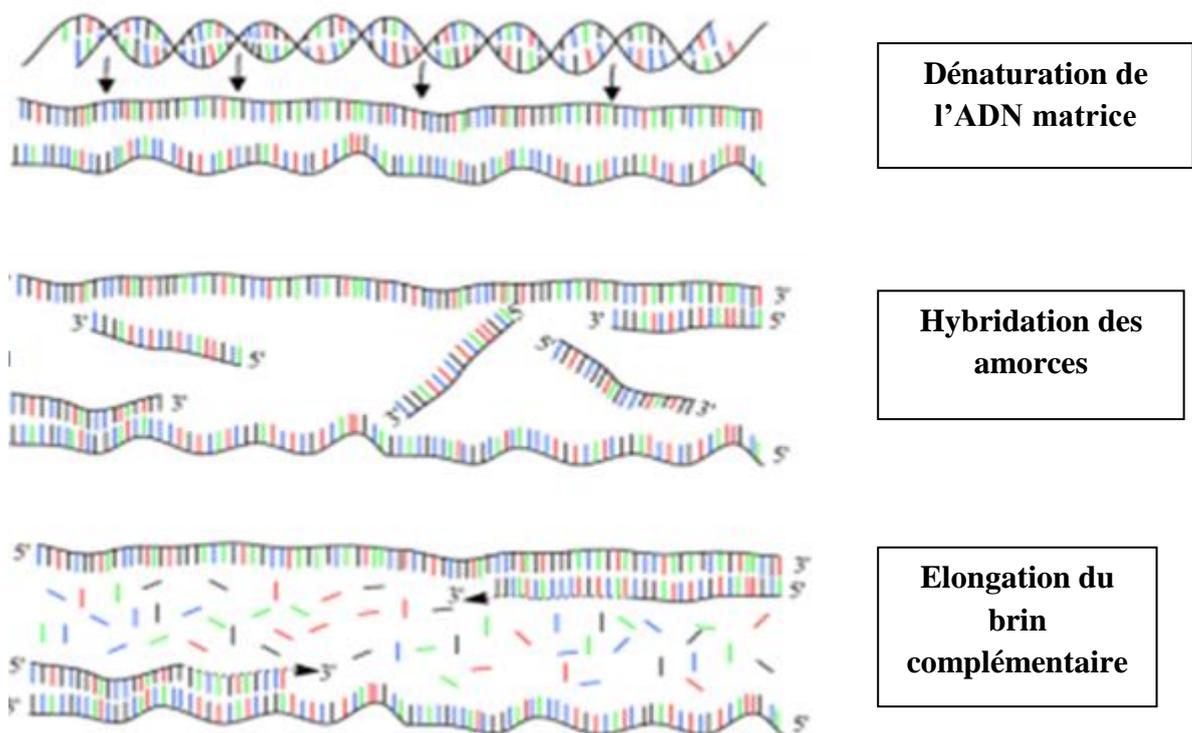
### 2. 2. Amplification *in vitro* de l'ADN

Une fois l'ADN total extrait et ramené à la concentration voulue, on l'amplifie avec la PCR qui est une méthode d'amplification d'un fragment spécifique de la molécule, ce qui permet à la fois de sélectionner la zone du génome à séquencer et de l'obtenir en quantité suffisante pour les étapes ultérieures (Tagu, 1999).

Le cycle d'amplification présente trois étapes (Hochkirch, 2001) :

- **Dénaturation de l'ADN matrice**, en approchant la température de 100°C tout en restant dans le domaine de tolérance de l'enzyme. Pour cela, il faut utiliser une ADN polymérase thermo-résistante, la *Taq* polymérase, issue d'une archaebactérie thermophile (*Thermus aquaticus*), qui supporte la température de 94°C. à l'issue de cette étape, l'ADN est simple brin.

- **Hybridation des amorces** « *annealing* » avec leurs séquences complémentaires sur le brin d'ADN matrice, à la température optimale d'appariement calculée à partir de la composition en bases des amorces (de l'ordre de 50°C).
- **Elongation du brin complémentaire** par la Taq polymérase à partir de l'extrémité 3'OH des amorces, à 72°C, température d'activité optimale de cette enzyme. Au bout d'un cycle, le nombre de copies du fragment est multiplié par deux. Puisque chaque brin nouvellement copié peut servir de matrice pour l'amplification suivante, le nombre de copies du fragment désiré augmente de façon exponentielle (**Figure 11**).



**Figure 11** : Les étapes de l'amplification PCR (M. Somma, M. Querci., 2006).

## 2. 3. Le séquençage

### 2.3.1. Principe de séquençage

Le séquençage de l'ADN est la détermination de la succession des nucléotides le composant, c'est aujourd'hui une technique de routine pour les laboratoires de biologie. Cette technique utilise les connaissances qui ont été acquises depuis une trentaine d'années sur les mécanismes de la réplication de l'ADN. Les techniques de séquençage utilisent des enzymes particulières : les ADN polymérases. Ces enzymes sont capables de synthétiser un brin

complémentaire d'ADN, à partir d'un brin matrice. Ces réactions se font par ajout de désoxyribonucléotides (dNTP : désoxy Nucléotide TriPhosphate) (Tagu, 1999).

Le principe du séquençage selon la méthode de Sanger (1977, *in* Hochkirch, 2001), est une sorte de PCR interrompue aléatoirement. Elle repose sur l'élongation d'un grand nombre de copies d'un brin matrice d'ADN par une polymérase à partir de l'extrémité 3'OH d'une amorce, avec arrêts aléatoires de la copie, provoqués par l'ajout de didésoxy nucléotides (ddNTP). De tels nucléotides n'ont pas de groupement hydroxyle en 3' du désoxyribose, ce qui empêche la poursuite de l'élongation, puisque ce sucre ne peut plus créer de liaison covalente avec le groupement phosphate du nucléotide suivant. La conséquence de ces arrêts aléatoires est la synthèse de brins de toutes les tailles possibles.

Une ADN polymérase synthétise le brin complémentaire de l'ADN à séquencer. Dans le milieu de réaction se trouvent des dNTP en grand nombre, et une faible proportion d'un ddNTP (à adénine, ou guanine, ou thymine, ou cytosine). A un moment totalement aléatoire, un ddNTP sera ajouté à la chaîne en cours de synthèse par l'ADN polymérase. Cette synthèse s'arrêtera donc à cet endroit. Par exemple, si le milieu réactionnel contient une faible proportion de didésoxy ribonucléotide à guanine (ddGTP), on obtiendra à la fin des réactions d'un ensemble de brins d'ADN de tailles variées, selon l'endroit où un ddGTP sera inséré et la réaction sera ainsi stoppée (ce qui correspond, du fait de la complémentarité des bases, à la présence d'une cytosine dans le brin d'ADN séquencé). On répète la même opération avec un milieu contenant du ddATP, un milieu contenant du ddCTP, et un milieu contenant du ddTTP (Hochkirch, 2001).

### **2.3.2. Le séquençage automatique :**

La très grande majorité des séquences réalisées et publiées aujourd'hui sont réalisées sur des séquenceurs automatiques. Ceux-ci sont capables de réaliser les réactions de séquence, puis de les lire : Pour cela, on marque les fragments d'ADN grâce à des marqueurs fluorescents. Une fois la réaction de séquençage terminée, la taille des fragments obtenus est déterminée par une chromatographie. Le séquenceur détecte la fluorescence des fragments sortant (par ordre croissant des tailles) des colonnes de chromatographie, repérant ainsi les fragments d'ADN et leur taille précise. Les systèmes les plus modernes permettent même de lire les quatre nucléotides à partir d'une seule colonne de chromatographie. Le résultat est présenté par la machine sous forme de courbes présentant la fluorescence détectée, et l'interprétation qui en est faite en termes de nucléotides (Tagu, 1999 ; Gagnon, 2004).

### 2.3.3. Les marqueurs moléculaires

Les marqueurs moléculaires correspondent à des différences nucléotidiques existant au niveau de la molécule d'ADN (d'où le terme moléculaire), des techniques de biologie moléculaire permettent de révéler ce polymorphisme de séquences. Ces différences entre allèles peuvent correspondre à des mutations ponctuelles (substitution, insertion, délétion), des réarrangements chromosomiques, ou des mutations silencieuses (sans effet sur l'expression du locus). Les caractères génétiques des marqueurs moléculaires ce sont qu'ils révèlent directement la nature génétique de l'information. Ceci les différencie des caractères morphologiques, physiologiques, et plus généralement de tous les caractères phénotypiques, dus à une addition d'effets génétiques et non génétiques. Cet aspect génétique leur confère un avantage du point de vue de la reconstruction de la phylogénie, par rapport aux caractères classiques utilisés en systématique. Parmi les auteurs qui ont utilisés les marqueurs moléculaires pour construire l'arbre phylogénétique de *Medicago* sont :

- Valizadeh en 1996, qui a utilisé comme méthode pour construire son arbre phylogénétique, la RFLP des chloroplastes (**Valizadeh.M, 1996**).
- Bena et al, 1997 ont utilisé le marqueur moléculaire RADP, pour l'étude de l'évolution du genre *Medicago*.

#### a) Marqueurs RFLP

La technique RFLP repose sur la mise en évidence de la variabilité de la séquence nucléotidique de l'ADN génomique après digestion par des enzymes de restriction. Une multitude de fragments d'ADN de tailles variables est générée par digestion enzymatique, puis séparée sur gel d'agarose et transférée par capillarité sous forme dénaturée sur une membrane de nylon. Cette membrane est mise en contact avec une solution contenant un fragment d'ADN ou sonde qui permet de repérer, par hybridation moléculaire, des fragments d'ADN génomique qui lui sont homologues. La différence entre deux génotypes est révélée par autoradiographie si la sonde est marquée par le phosphore radioactif ou par réaction colorée si elle est associée à un conjugué enzymatique. Le polymorphisme détecté est dû à des mutations au niveau des sites de restriction de l'enzyme (polymorphisme de site de restriction) et/ou à des délétions/insertions d'un fragment d'ADN au voisinage de la zone génomique reconnue par la sonde. C'est le couple enzyme/sonde qui constitue le marqueur.

Bien que cette technique soit co-dominante et permette une analyse génétique complète, elle est lente et laborieuse. Les étapes de transfert et d'hybridation empêchent une automatisation du travail (Najimi *et al.*, 2003).

### **b) Marqueurs RAPD**

Les marqueurs RAPD sont produits par l'amplification des segments anonymes de l'ADN génomique avec de petites séquences d'ADN, environ 10 paires de bases choisies arbitrairement, appelées amorces. Les fragments d'ADN amplifiés sont séparés en fonction de leurs tailles par électrophorèse sur gel d'agarose. Le polymorphisme est détecté comme la présence ou l'absence d'une bande particulière. La méthode est basée sur la probabilité statistique que les sites complémentaires aux nucléotides de l'amorce se reproduisent à plusieurs niveaux dans l'ADN génomique. Il peut y avoir des problèmes avec la répétition des expériences avec une seule amorce, mais ceux-ci peuvent être surmontés en utilisant plusieurs amorces à la fois. Le problème principal est que les marqueurs sont anonymes et on ne peut pas être sûr si les emplacements sont vraiment homologues. Néanmoins, en comparaison avec d'autres méthodes analogues (RFLP), la RAPD est beaucoup plus rapide et plus simple, les marqueurs RAPD sont toujours dominants et ils fournissent plus d'information impliquant le génome entier (Bardakci, 2000, Sepp *et al.*, 2000).

#### **2.3.4. Traitement des caractères de type marqueurs moléculaires**

Dans la reconstruction phylogénétique, le polymorphisme allélique et le polymorphisme de fragments peuvent être traité comme des caractères discrets ou comme des caractères continus. En effet, si l'on ne considère que la présence ou l'absence de telle forme allélique, ou bande dans une espèce, il s'agit d'un caractère discret. En revanche si l'on prend également en considération la fréquence de ces formes alléliques dans les différentes espèces, il s'agit alors d'un caractère continu (Darlu et Tassy, 1993 ; Salemi et Vandamme, 2003). Ce dernier type de caractères ne peut être analysé, dans une perspective phylogénétique, que par des méthodes quantitatives (Onarici et Sumer, 2002).

Pour ce type de caractères moléculaires, il existe actuellement plusieurs logiciels qui permettent d'améliorer la visibilité des bandes, par des traitements numériques des images, en jouant sur la luminosité, le contraste et sur l'intensité relative des couleurs. Certains logiciels, comme Phoretix et TotalLab, permettent la détection automatique de la présence ou d'absence de bandes, le calcul de leurs poids moléculaires par rapport à un standard de poids

moléculaires connus, la génération des matrices des caractères et des matrices des distances, l'inférence des arbres phylogénétiques par différentes méthodes.

Bionumerics /Gel Compar peut être le logiciel le plus puissant dans l'analyse des gels, il est utilisé partout dans le monde : dans les laboratoires de recherches, dans les universités et les hôpitaux. Car, en plus des fonctionnalités citées précédemment, ce logiciel permet de lire les résultats des séquenceurs automatiques, d'explorer les séquences nucléotidiques et protéiques et de les aligner, puis les comparer à une base de données centrale.

### **3. Technique de la bioinformatique**

#### **3.1. Base de données**

Afin de chercher des séquences nucléotidiques en biologie moléculaire, plusieurs bases de données qui comportent toutes les informations génétiques des espèces étudiées ont été créées pour faciliter la tâche des chercheurs et rendre ces séquences accessibles pour les utilisateurs.

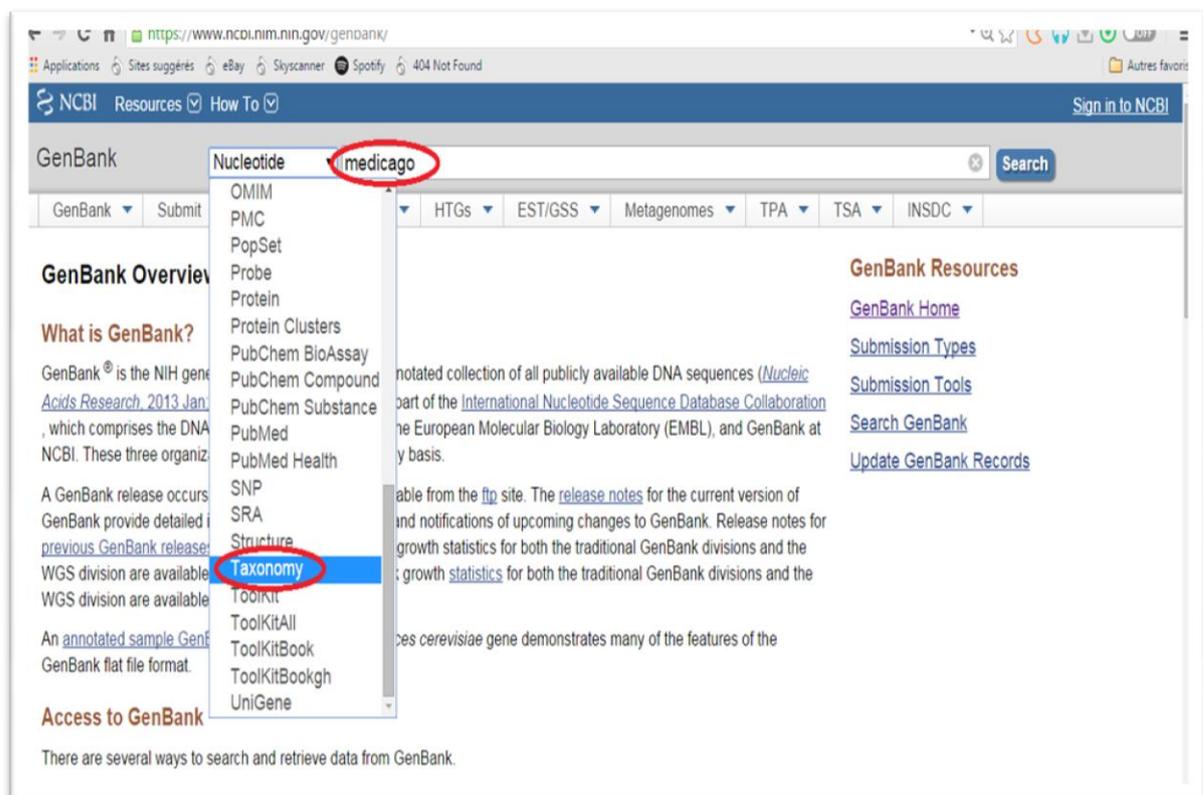
Parmi ces bases de données on trouve :

- **"EMBL" (European Molecular Biology Laboratory)** : Elle a été créée en 1980 et était financée par l'EMBO (European Molecular Biology Organisation) et diffusée par l'EBI (European Bioinformatics Institute, Cambridge, UK).
- **"DDBJ" (DNA Data Bank of Japan)** : Elle a été créée en 1986 et diffusée par le NIG (National Institute of Genetics, Japon).
- **"GenBank"** : qui a été créé en 1982 par la société IG (IntelliGenetics) et est diffusée maintenant par le NCBI (National Center for Biotechnology Information, Los Alamos, USA). (Cheikhrouhou, 2006), elle a été mise en place en 1979 au LANL «Los Alamos National Laboratory», à Los Alamos. Depuis 1992, elle est maintenue au NCBI «National Center For Biotechnology Information», à Bethesda (Perrière, 2000). Le NCBI (National Center for Biotechnology Information) a été créé par le Congrès Américain en 1988 et est chargé spécifiquement de développer des systèmes automatisés pour la biologie moléculaire et la biotechnologie (Pillet, 2000).

### 3.2. Consultation de la GenBank

La GenBank est une collection de toutes les séquences de nucléotides et de protéines disponibles au public dont son URL est le suivant : <http://www.ncbi.nlm.nih.gov/genbank/>, ces enregistrements sont générés à partir des communications directes avec la séquence d'ADN des bases de données des auteurs originaux, qui donnent leurs dossiers pour rendre les données à la disposition de tout le monde ou de le faire dans le cadre du processus de publication (Clark *et al.*, 2015).

Afin de récupérer les séquences nucléotidiques du genre *Medicago* à partir de la GenBank, on procède par la recherche taxonomique de ce genre (figure 12), 85 espèces du genre *Medicago* existent dans cette base de données, après la consultation des séquences nucléotidiques de chaque espèce, on a résumé le tableau de l'annexe.



**Figure 12 :** Les étapes initiales pour récupérer les séquences nucléotidiques de la GenBank.

- En entrant sur *Medicago*, on a obtenus toutes les espèces qui appartiennent à ce genre (Figure 13).

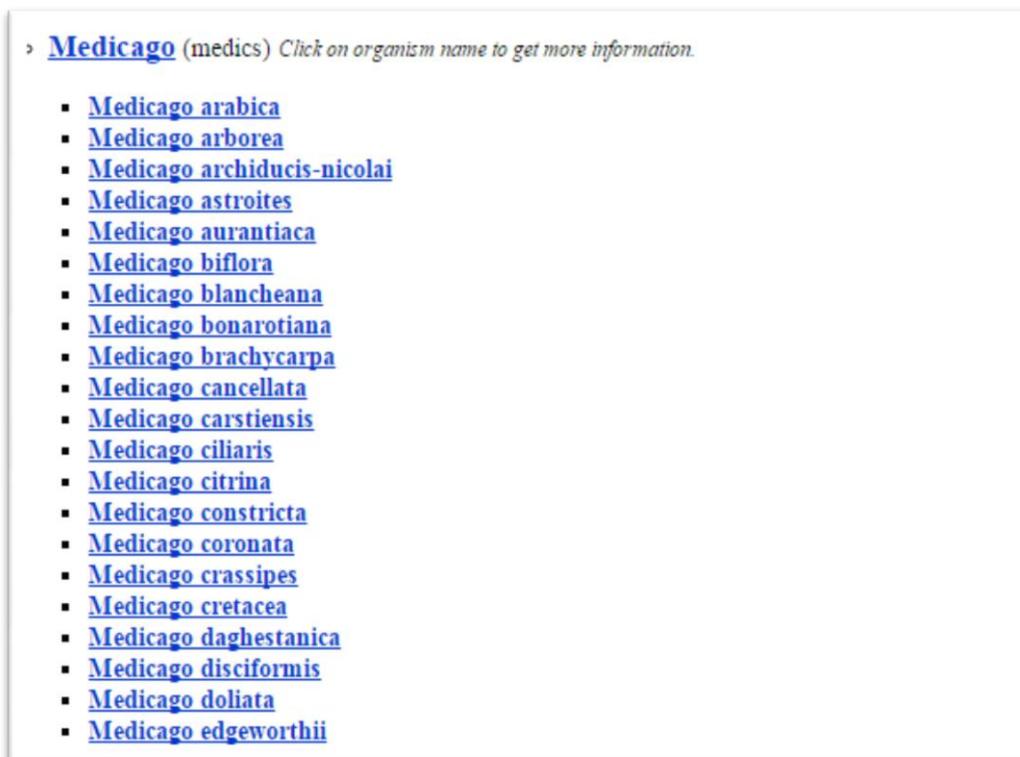


Figure 13: quelques espèces du genre « Medicago » obtenues dans la GenBank.

Nous avons analysé toutes les espèces une par une pour identifier le nombre des séquences nucléotidiques trouvées dans cette banque (Figure 14).



Figure 14: Exemple des séquences nucléotidiques trouvées pour l’espèce *MEDICAGO Arabica*.

Au total, la GenBank contient 333861séquences de 85 espèces du genre «*Medicago* » dont 693 gènes chloroplastiques.

Une fois le gène est choisi, on a pris les séquences nucléotidiques chloroplastique représentatives de chacune des 43 espèces de *Medicago* qu'on a téléchargées sous format FASTA à partir de la "GenBank", et comme *outgroup* on choisi trois espèces du genre *Trigonella*, qui sont *T. glabra*, *T. elliptica* et *T. anguina*, selon la taxonomie classique *Trifolium*, *Trigonella*, et *Melilotus* sont les trois genres les plus proches du genre *Medicago* (Benaet *al.*, 1997).

### 3.3. Traitement des séquences de la GenBank

Pour le traitement des séquences, il existe plusieurs logiciels permettant de générer des arbres phylogéniques environ 367 logiciels de bioinformatique, parmi eux : PHYLIP, PAUP\*, MEGA, Phylo\_win, ARB, DAMBE, PAL, Bionumerics, Mesquite, PaupUp, BIRCH, osque, EMBOSS, phangorn, Bio++, ETE, DendroPy, SeaView, Crux (**Djekoun et Hamidechi; 2010**), mais nous avons choisi MEGA version 5 (MEGA5: **Molecular Evolutionary Genetics Analysis using Maximum Likelihood (Tamura *et al.*, 2011)**).

#### 3.3.1. MEGA 5.01

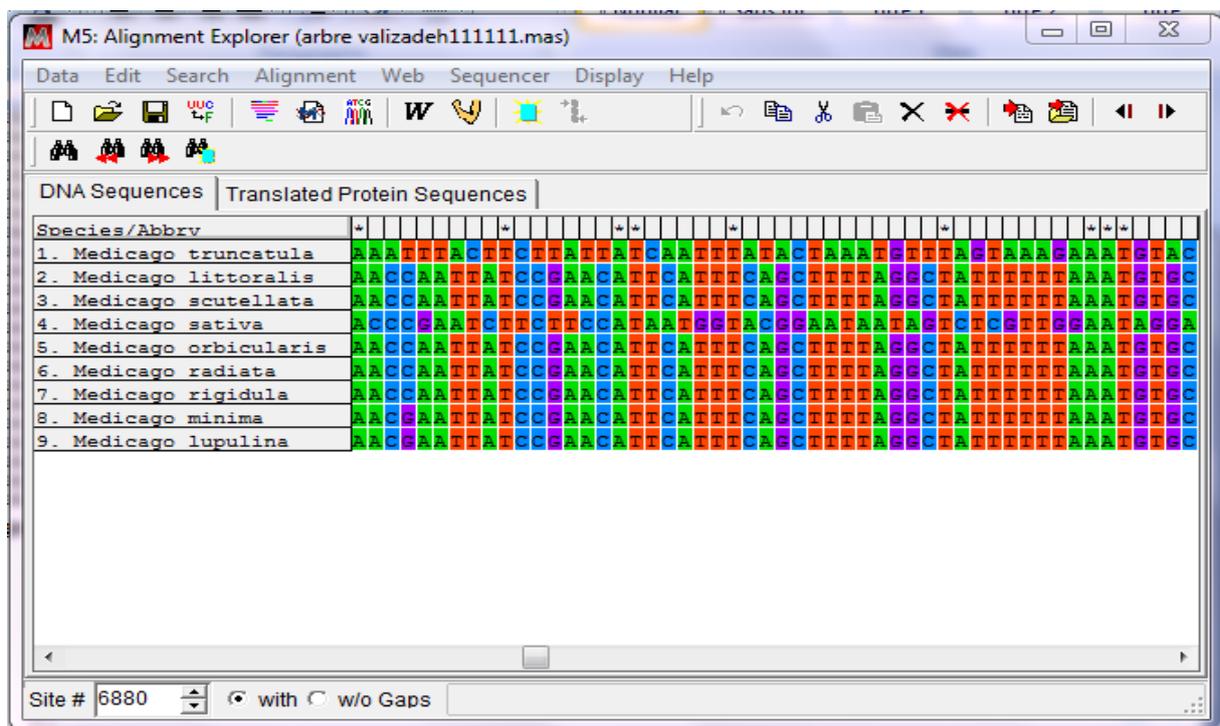
Le Moléculaire Génétique évolutive Analyse (MEGA) est un programme qui permet d'explorer et d'analyser des séquences nucléotidiques et protéiques. En effet, dans les études de l'évolution moléculaire, il est souvent nécessaire de savoir certaines propriétés statistiques de base, telles que les fréquences des différents nucléotides, les fréquences des codons et les rapports de transitions/transversions. Il permet également le calcul des distances évolutives entre les séquences sous différents modèles de substitution (Jukes et Cantor, Kimura à 2 paramètres et Tamura-Nei ) et la reconstruction des arbres phylogénétiques selon trois méthodes, UPGMA, Neighbor-Joining et Parcimonie (**Kumar et al., 2001**).

Avant de passer à l'analyse nucléotidique, un alignement multiple a été réalisé en utilisant le programme ClustalW, nous avons choisi ce type d'alignement par ce qu'il est très utile et permet d'étudier les différences et les similitudes entre les séquences d'un même marqueur moléculaire.

### 3.3.2. Programme *ClustalW*

Clustalw est un programme d'alignements multiples des séquences d'acides nucléiques ou protéiques. Ces séquences sont fournies par l'utilisateur ou proviennent des banques de données. Ce type d'alignement est très utile puisqu'il permet de comparer des séquences provenant de différentes sources et d'étudier les différences et les similitudes entre les séquences de même type mais appartenant à des groupes différents. Ceci permet d'identifier les régions qui sont conservées d'une espèce à une autre (**Thompson et al., 1994**). Les principales étapes d'un alignement multiple avec *CLUSTALW* sont:

- Chaque paire de séquences est alignée séparément selon une méthode approximative globale et rapide (**Wilbur et Lipman, 1983**). Une matrice de distances traduisant la divergence de chaque paire de séquences est ainsi calculée.
- Un dendrogramme est ensuite construit à partir de la matrice des distances en utilisant la méthode de Bootstrap Cet arbre phylogénétique sert de guide qui détermine l'ordre selon lequel l'alignement de l'étape suivante sera réalisé.



**Figure 15:** représentation des séquences nucléotidiques des espèces du genre *Medicago* dans le logiciel *MEGA5* après l'alignement multiple des séquences.

#### 4. Construction de l'arbre phylogénétique

Plusieurs méthodes ont été développées pour construire un arbre phylogénétique à partir d'une matrice de distances (**Golding et Morton, 2003**). Ces méthodes se basent sur la mesure des distances entre les séquences, c'est à dire le nombre de substitutions de nucléotides ou d'acides aminés entre chaque paire de séquences (**Cheikhrouhou, 2006**) parmi ces méthodes la méthode de Neighbor joining (NJ), Maximum Likelihood, Maximum Parsimony et UPGMA, C'est la méthode UPGMA qu'on a utilisé dans le logiciel MEGA 5 (figure 14), afin de construire notre arbre phylogénétique du genre *Medicago*.

##### a) UPGMA (Unweighted Pair Group Method of Arithmetic Averages)

Cette méthode produit des arbres à racines pour lesquelles les longueurs de bord peuvent être considérées comme les temps mesurés par une horloge moléculaire avec un débit constant (**Dowell, 2008**). C'est une méthode très simple et très rapide, développée originellement pour construire des phénogrammes taxonomiques (**Golding et Morton, 2003 ; Salemi et Vendamme, 2003**). Elle se base sur les similarités entre paires de séquences en imposant que les distances soient ultramétriques, donc que les caractères évoluent à une vitesse constante (hypothèse d'horloge moléculaire).

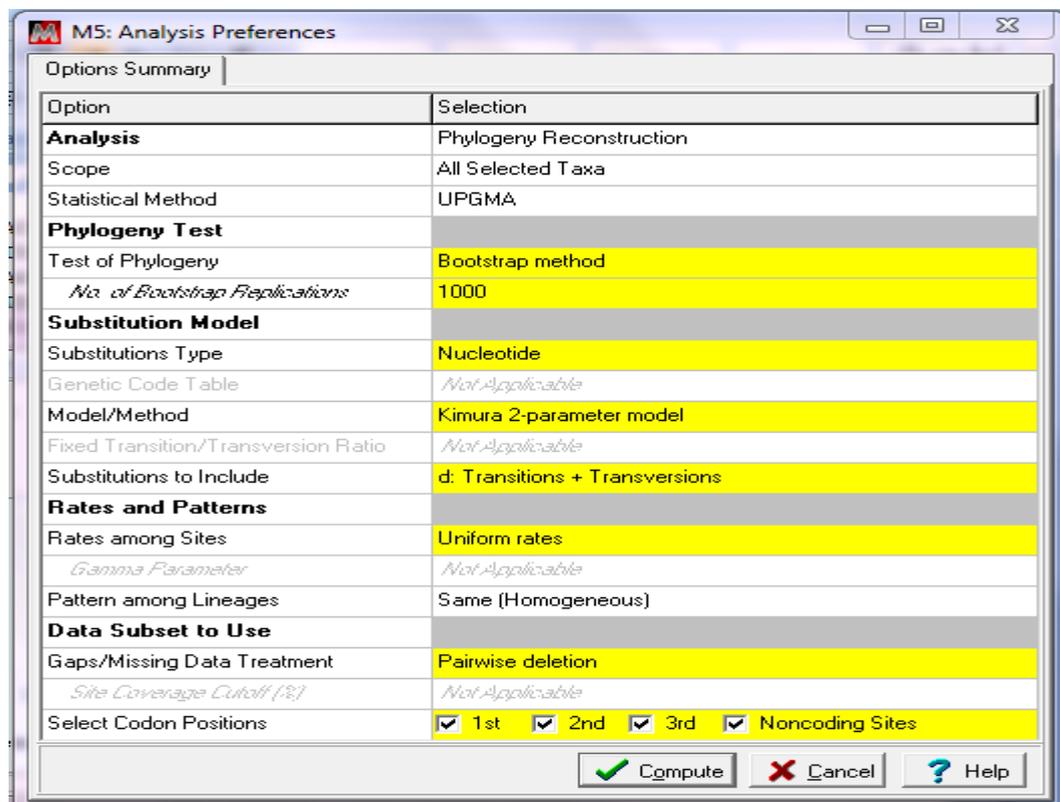


Figure 16 : les paramètres du logiciel MEGA 5 utilisé dans la phylogénie.

La fiabilité des branches des arbres construits a été évaluée à l'aide de la méthode "Bootstrap". Les valeurs de "Bootstrap" pour 1000 répliques sont ramenées en pourcentages et ils sont indiqués au niveau des nœuds.

### **b) Le Bootstrap**

Cette méthode, introduite par Felsenstein (1985), est largement appliquée, avec beaucoup de succès. Elle est utilisée pour tester la fiabilité des branches internes des arbres phylogénétiques. Elle consiste à effectuer un tirage au hasard avec remise des sites. Dans chaque réplique de "Bootstrap", certains sites peuvent être présents plusieurs fois, tandis que d'autres peuvent être absents. Chaque réplique produit un nouvel alignement "artificiel", qui est utilisé pour construire un arbre "artificiel". Pour chaque branche interne on calcule le pourcentage des arbres "artificiels" contenant cette branche. Seules les branches avec un fort % sont considérées comme robustes (**Cheikhrouhou, 2006 ; Riah, 2014**).

## 1. Résultats de la GenBank

### 1.1. Nombres d'espèces et séquences étudiés dans le genre de *Medicago*

Le tableau 3 représente le nombre des séquences nucléotidiques étudiés dans chaque espèce pour le genre *Medicago* prélevé dans la GenBank.

**Tableau III :** Nombres des séquences nucléotidiques pour chaque espèce étudié du genre de *Medicago*.

N°	Espèces	Nombre des séquences nucléotidiques
1	<i>Medicago arabica</i>	32
2	<i>Medicago arborea</i>	15
3	<i>Medicago archiducis-nicolai</i>	3
4	<i>Medicago astroites</i>	4
5	<i>Medicago aurantiaca</i>	4
6	<i>Medicago biflora</i>	12
7	<i>Medicago blancheana</i>	5
8	<i>Medicago bonarotiana</i>	5
9	<i>Medicago brachycarpa</i>	14
10	<i>Medicago cancellata</i>	9
11	<i>Medicago carstiensis</i>	13
12	<i>Medicago ciliaris</i>	16
13	<i>Medicago citrina</i>	7
14	<i>Medicago constricta</i>	12
15	<i>Medicago coronata</i>	6
16	<i>Medicago crassipes</i>	4
17	<i>Medicago cretacea</i>	9
18	<i>Medicago daghestanica</i>	5
19	<i>Medicago disciformis</i>	8
20	<i>Medicago doliata</i>	10
21	<i>Medicago edgeworthii</i>	16
22	<i>Medicago falcata</i>	<b>17643</b>
23	<i>Medicago falcata</i> subsp. <i>falcata</i>	9
24	<i>Medicago fischeriana</i>	9
25	<i>Medicago granadensis</i>	12
26	<i>Medicago heyniana</i>	12
27	<i>Medicago huberi</i>	6
28	<i>Medicago hybrida</i>	9
29	<i>Medicago hypogaea</i>	4
30	<i>Medicago intertexta</i>	10
31	<i>Medicago italica</i>	8
32	<i>Medicago laciniata</i>	16
33	<i>Medicago lanigera</i>	15
34	<i>Medicago laxispira</i>	5
35	<i>Medicago lesinsii</i>	5
36	<i>Medicago littoralis</i>	24
37	<i>Medicago lunata</i>	2

38	<i>Medicago lupulina</i>	86
39	<i>Medicago marina</i>	21
40	<i>Medicago medicaginoides</i>	10
41	<i>Medicago minima</i>	34
42	<i>Medicago monantha</i>	17
43	<i>Medicago monspeliaca</i>	17
44	<i>Medicago murex</i>	25
45	<i>Medicago muricoleptis</i>	8
46	<i>Medicago noeana</i>	17
47	<i>Medicago orbicularis</i>	28
48	<i>Medicago orthoceras</i>	6
49	<i>Medicago pamphylica</i>	2
50	<i>Medicago papillosa</i>	16
51	<i>Medicago persica</i>	1
52	<i>Medicago phrygia</i>	3
53	<i>Medicago pironae</i>	13
54	<i>Medicago platycarpus</i>	14
55	<i>Medicago plicata</i>	5
56	<i>Medicago polyceratia</i>	15
57	<i>Medicago polymorpha</i>	47
58	<i>Medicago popovii</i>	9
59	<i>Medicago praecox</i>	12
60	<i>Medicago prostrata</i>	37
61	<i>Medicago radiata</i>	18
62	<i>Medicago rhodopea</i>	7
63	<i>Medicago rigidula</i>	21
64	<i>Medicago rigiduloides</i>	14
65	<i>Medicago rostrata</i>	2
66	<i>Medicago rotata</i>	10
67	<i>Medicago rugosa</i>	10
68	<i>Medicago ruthenica</i>	104
69	<i>Medicago sativa (alfalfa)</i>	<b>210073</b>
70	<i>Medicago sauvagei</i>	15
71	<i>Medicago saxatilis</i>	9
72	<i>Medicago schischkinii</i>	1
73	<i>Medicago scutellata</i>	15
74	<i>Medicago secundiflora</i>	13
75	<i>Medicago shepardii</i>	12
76	<i>Medicago sinskiae</i>	2
77	<i>Medicago soleirolii</i>	8
78	<i>Medicago sphaerocarpa</i>	3
79	<i>Medicago strasseri</i>	10
80	<i>Medicago suffruticosa</i>	10
81	<i>Medicago syriaca</i>	5
82	<i>Medicago tenoreana</i>	17
83	<i>Medicago tornata</i>	81
84	<i>Medicago truncatula</i>	<b>104852</b>
85	<i>Medicago turbinata</i>	10

Selon le tableau 3 on a constaté que les espèces les plus étudiées sont celles qui ont le plus grands nombres de séquences nucléotidiques, et ces espèces sont *M. sativa* (210073 études) qui est considéré comme plante modèle grâce à son intérêt économique et qui est distribué sur tout le bassin méditerranéen, et *M. truncatula* avec 104852 séquences nucléotidiques qui est une plante modèle aussi pour une analyse moléculaire de la résistance aux maladies (**Jeuffroy et Tivoli, 2011**).

## 1.2. Intérêt des deux espèces

### 1.2.1. Intérêt de *M. sativa*

La luzerne a été qualifiée de « reine des légumineuses » ou même de « reine des cultures fourragères » en raison de ses nombreuses qualités :

- Elle protège les nappes phréatiques, en permettant de récupérer et de soustraire au lessivage les surplus de nitrates présent dans le sol.
- C'est une plante fourragère par excellence, car elle produit des protéines et des carotènes.
- Elle est utile en pharmacie grâce à la Rubisco qui est extraite à partir de ses feuilles et utilisée comme diététique.
- Elle est riche en vitamines du groupe B, C, D, E et A.
- Elle contient des phyto-estrogène qui sont des facteurs de protections contre le cancer, et contre les troubles de la ménopause.
- Elle est utilisée traditionnellement dans le traitement de diabète.
- Elle a des propriétés anti-hémorragiques, anti-fongiques, hémostatiques et régulateur métabolique (**Messioughi, 2010**).

### 1.2.2. Intérêt de *M. truncatula*

- C'est une plante annuelle, qui a un cycle de vie court (cycle de vie de 2 à 3 mois).
- C'est une espèce prolifique, chaque plante produit 500 à 1000 graines.
- Elles ont la capacité à se ressemer naturellement d'une année à l'autre.
- Elle diminue le phénomène d'érosion.
- Elle résiste à la sécheresse, à la salinité et aux maladies.
- Elle possède un génome de faible taille de  $5.10^8$  pb.

- Elle est apte à la transformation pour *AGROBACTERIUM tumefaciens* et à la régénération par embryogenèse somatique.
- Elle forme de petites graines qui permettent leur culture dans des tubes à essais (**Ben friha, 2008**).

### 1.3. Le choix du gène utilisé

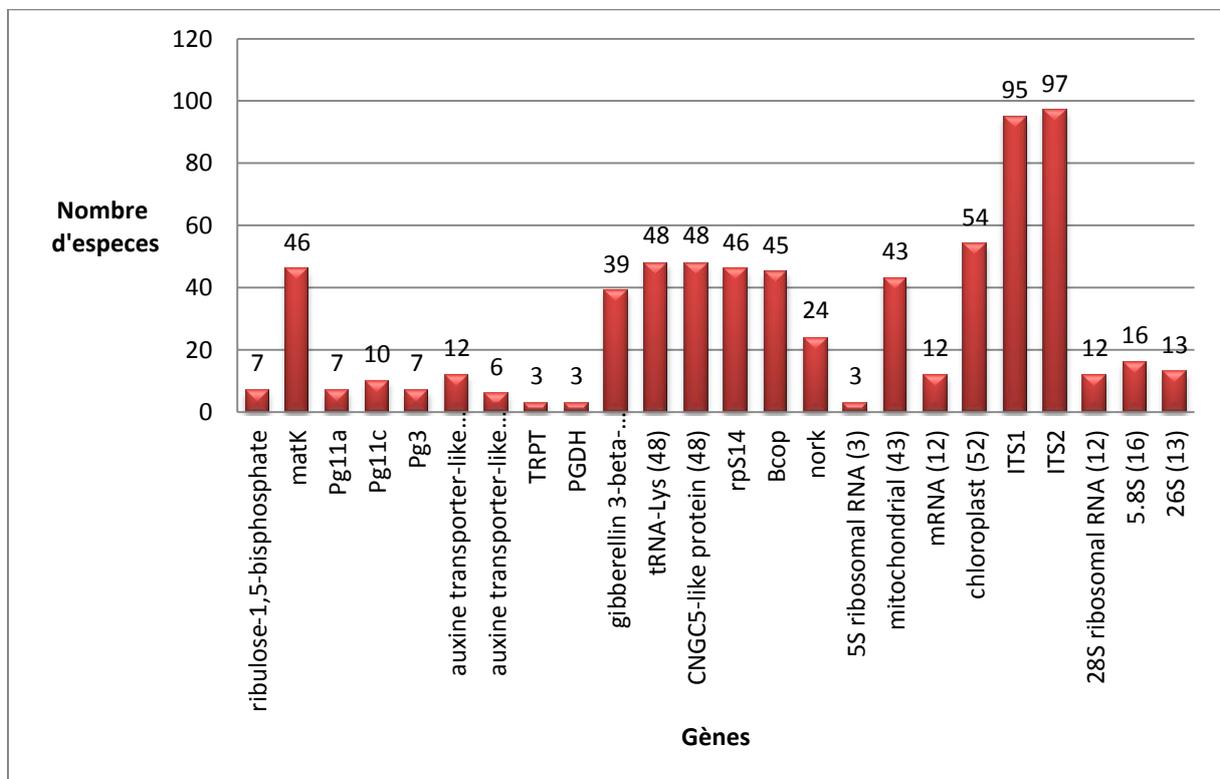
D'après le premier tableau on remarque plusieurs séquences nucléotidiques qui ont été utilisées dans ce genre, on a choisi seulement les séquences des gènes qui se sont répétées à plusieurs reprises, le tableau 4 résume ces 24 gènes et le nombre des espèces dont ils ont travaillé dessus :

**Tableau IV:** Les gènes les plus utilisés dans la GenBank du genre *Medicago*.

N°	Gènes	Nombre d'espèce
1	<b>ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit</b>	7
2	<b>matK</b> (maturase K)	46
3	<b>Pg11a</b> (Polygalacturonase 11a)	7
4	<b>Pg11c</b> (polygalacturonase 11c)	10
5	<b>Pg3</b> (polygalacturonase 3)	7
6	<b>auxine transporter-like protein 2</b>	12
7	<b>auxine transporter-like protein 4</b>	6
8	<b>TRPT</b> (Triose phosphate translocator)	3
9	<b>PGDH</b> (Phosphogluconate dehydrogenase)	3
10	<b>gibberellin 3-beta-hydroxylase</b>	39
11	<b>tRNA-Lys</b>	48
12	<b>CNGC5-like protein</b>	48
13	<b>rpS14</b> (ribosomal protein S14)	46
14	<b>Bcop</b> (Protein complex coatmer beta subunit)	45
15	<b>NORK</b> (nodulation receptor kinase)	24
16	<b>5S ribosomal RNA</b>	3
17	<b>Gene mitochondriale</b>	43
18	<b>mRNA</b>	12
19	<b>chloroplast</b>	43
20	<b>Internal transcribed spacer 1</b>	95
21	<b>Internal transcribed spacer 2</b>	97
22	<b>28S ribosomal RNA</b>	12
23	<b>5.8S</b>	16
24	<b>26S</b>	13

D'après ce tableau, les données moléculaires que nous avons utilisé sont des séquences nucléotidiques du gène de chloroplaste parce qu'il contient des séquences nucléotidiques plus longues que les ITS1 et les ITS2 dont leurs nombres nucléotidiques ne dépassent pas 1000, et il est plus utile pour les analyses phylogénétiques, et ils comportent très peu de séquences non codantes (**Gagnon, 2004**). Ainsi que leur disponibilité dans les banques de données. Toutes les séquences nucléotidiques utilisées dans notre étude ont été récupérées de la banque de données *GenBank*, le mois de juin 2016. Ces séquences sont celles réalisées par plusieurs auteurs : **Neves et Watson, 2004 ; Wang et al., 2008**.

La figure ci-dessous résume le tableau 4, c'est un graphe utilisé avec l'excel 2007.



**Figure 17:** Graphe des gènes les plus étudiés sur *Medicago* sur excel 2007.

Dans ce graphe on remarque 8 gènes qui sont utilisés sur moins de 10 espèces telles que l'ARNr 5S, la Pg3...etc., 5 gènes sont utilisés sur une dizaine d'espèces, 3 gènes sont utilisés pour étudier 3 espèces, 9 gènes qui sont utilisées dans plus de 40 espèces, 95 -97 études sur *Medicago* avec les ITS1 et les ITS2, ces les plus utilisés grâce à leurs petite taille qui est

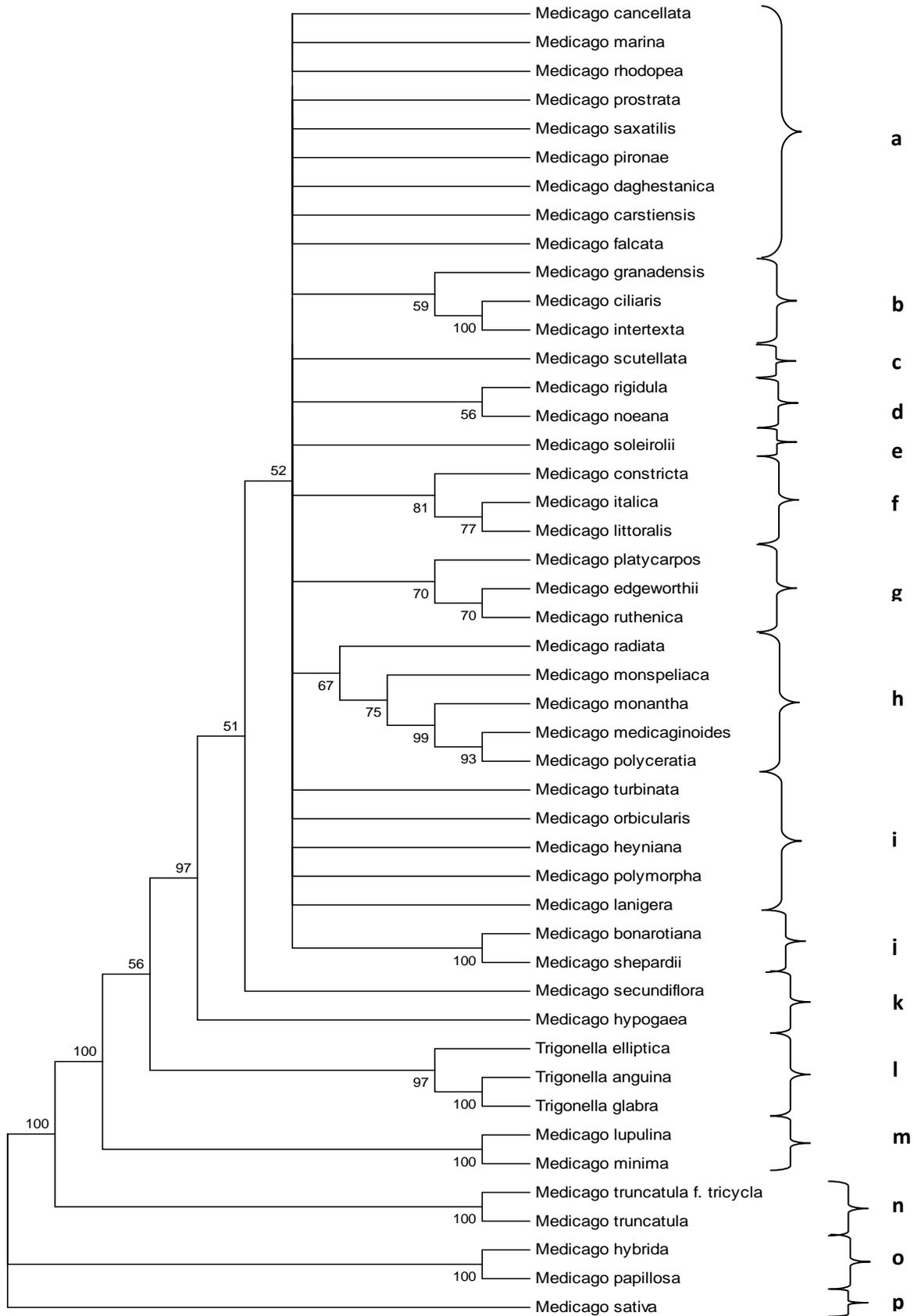
inférieur à 1000 pb, le gène chloroplastique qu'on a choisi est le deuxième après les ITS1 et les ITS2, il est utilisé dans 54 espèces (études).

Les séquences des génomes chloroplastiques et mitochondriaux sont utilisées pour effectuer les analyses phylogénétiques nécessaires aux recherches. Les génomes d'organites, comparativement aux génomes nucléaires, ont l'avantage d'être plus petits ce qui rend leur séquençage plus rapide et moins cher. De plus, ils évoluent plus lentement et sont plus stables. Cela permet de résoudre des phylogénies faisant intervenir des divergences très anciennes (**Gagnon, 2004**). La stabilité selon Palmer est due à la présence des séquences inversées répétées qui stabilise l'ADN chloroplastique et limite l'échange de gènes entre les petites régions simple copie et grande région simple copie (**Palmer, 1983 in Belanger, 2007**).

Dans la plus part des végétaux, l'ADN chloroplastique est transmis maternellement. Chez *Medicago sativa*, la transmission des chloroplastes paternels dépend du génotype du parent femelle qui influence la localisation des chloroplastes soit dans la portion apicale de la cellule œuf, qui forme l'embryon, soit dans la portion basale qui forme le suspenseur. Seuls les chloroplastes présents dans la portion apicale seront transmis (**Gobron, 2012 ; Onarici et Sumer, 2002**).

## 2. L'arbre phylogénique obtenu

L'objectif de cette étape est basé sur le classement des espèces du genre *Medicago* dans un arbre phylogénétique. D'après les résultats qu'on a trouvés dans la GenBank ils sont basés sur le gène chloroplastique, et on a obtenu l'arbre phylogénétique suivant:



**Figure 18 :** Arbre UPGMA basée sur des données morphologiques de 43 espèces de genre *Medicago*.

### 3. L'analyse de l'arbre phylogénétique

Sur cet arbre, en prenant comme référence la sous-section des *Intertextes* groupe le mieux délimité du genre *Medicago* (Bena et al., 1998), nous avons fait ressortir sur cet arbre 29 clusters de niveau de sous-section à section (figure 18). L'arbre est divisé en 3 groupes : le premier contient seulement l'espèce *M. sativa*, le deuxième contient *M. hybrida* et *M. papillosa* et le troisième est divisé en 2 groupes, le premier contient *M. lupulina* et *M. minima* et le deuxième est divisé en deux sous groupes, le premier contient les trois espèces du genre *Trigonella* et le deuxième est divisé en deux sous groupes dont le premier contient une seule espèce qui est *M. hypogaea* et le deuxième contient 22 sous groupes. Tous les groupes ont quelques différences d'agencement des espèces près, se retrouvent dans tous les arbres (figure 19, 20, 21). Comme on peut le constater, il se montre fortement incongruent avec les autres arbres.

### 4. Comparaison avec d'autres arbres

#### 4.1. Comparaison avec d'autres arbres de gènes différents

Si nous comparons la composition en espèces des 29 clusters obtenus dans cette présente analyse avec les arbres de Steele *et al.*, 2010 (figure 19 et 20), basé sur la méthode maximum parcimonie, en utilisant le plastide *trnK/matK* et le gène *GA3ox1*, et l'arbre de Bena et al., 1997 ou ils ont effectué une étude phylogénétique moléculaire basée sur le transcrit interne et externe ribosomal nucléaire (STI et ETS). Trente et un d'espèces annuelle de *Medicago* ont été inclus dans l'étude en utilisant le programme *Clustal W*, ce qui représente plus de la moitié du genre et 85% des plantes annuelles du genre et ils ont rétabli l'arbre de la figure 21. Nous constaterons que, hormis pour la sous-section des *Intertextaes*, tous les autres sous-groupes taxonomiques de *Medicago* se sont montrés polyphylétiques à un moindre degré.

Si on prend l'arbre de Steele et al., 2010, ou ils ont utilisé le gène *GA3ox1* de 70 taxons, on remarque une grande ressemblance :

- *M. cancellata*, *M. saxatilis*, *M. daghestanica* du groupe « a » appartiennent à la section « *Medicago* », selon l'arbre de Steele.
- *M. granadensis*, *M. ciliaris* et *M. intertexta* groupe « b » appartiennent à la sub-section des *intertextaes*.

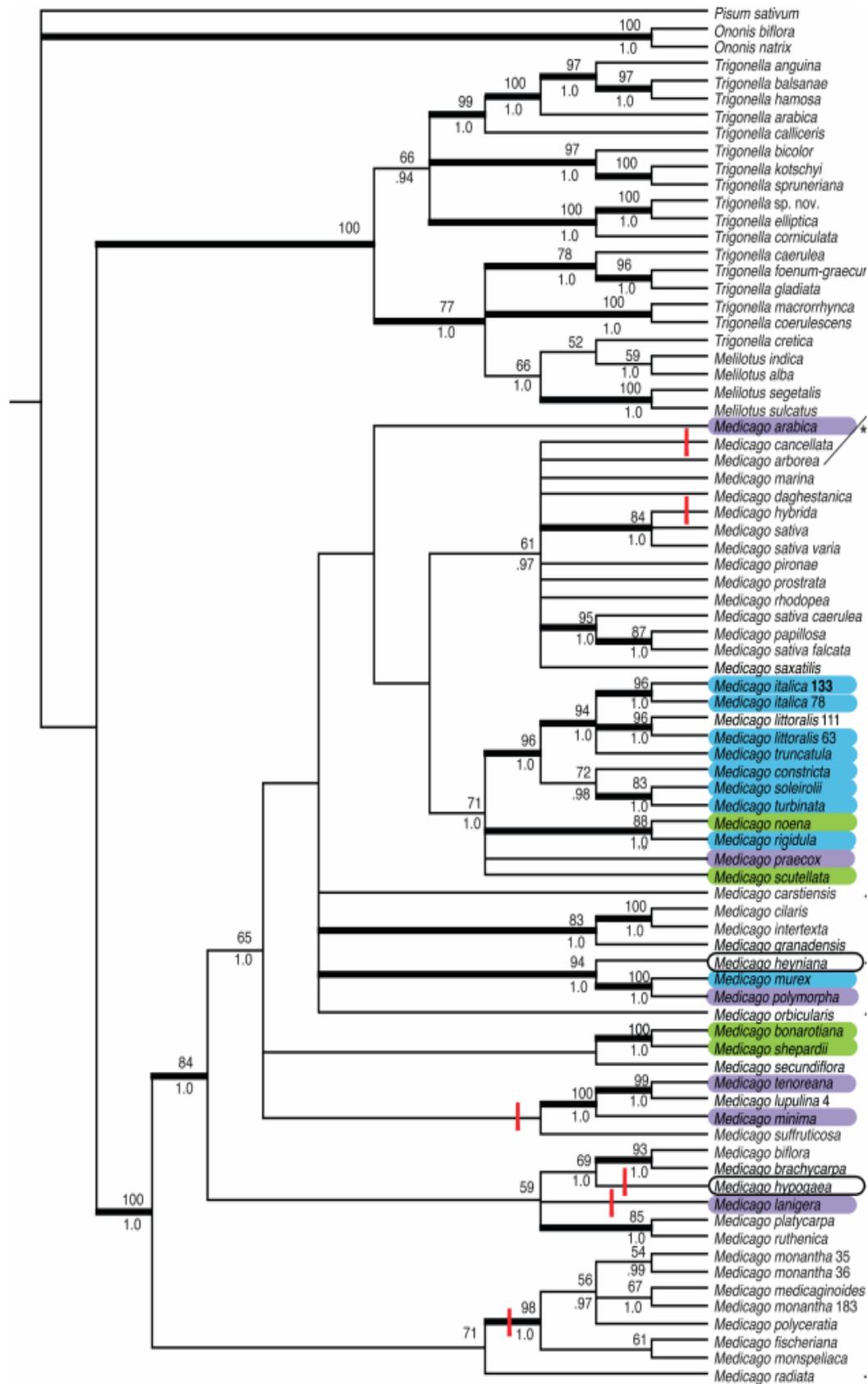
- Les espèces des groupes « d, e et f » dont leurs espèces sont *M. rigidula*, *M. noeana*, *M. soleirollii*, *M. constricta*, *M. italica* et *M. littoralis* appartiennent à la sub-section du clade *Pachyspireae*.
- Trois espèces du groupe « h », qui sont *M. monantha*, *M. medicaginoïdes* et *M. polyceratia* appartiennent dans la section des *Buceras*.
- *M. bonarotiana* et *M. shepardii* du groupe « j », appartiennent au clade *Rotatae*.
- *M. lupulina* et *M. minima* du groupe « m » sont dans le clade des *Leptospireaes*.

Et dans l'arbre du plastide *trnK/matK* de Steele, on remarque une très grande ressemblance par rapport à l'arbre du gène *GA3ox1*, les espèces des groupes « b, d, h et m » dont la valeur du *bootstrap* et comme suit « 59%, 56%, 67% et 100% » sont identiques aux clades de l'arbre de ce gène.

Dans l'arbre de Bena, on remarque une grande différence entre la distribution des espèces, la plus part des groupes de notre arbre sont différents, sauf le groupe « b » de notre arbre qui est identique aux groupes de la section *intertexta*, et les espèces *M. scutellata*, *M. rigidula*, *M. noeana*, *M. soleirollii*, *M. constricta* et *M. littoralis* qui se trouvent dans les groupes c, d, e, f de notre arbre appartiennent à un groupe monophyletique de 81% de *bootstrap*.

Selon la classification de Pierre (2008), les espèces *M. soleirollii*, *M. littoralis*, *M. constricta* et *M. rigidula* appartiennent au sous-genre des *spirocarpos*, section des *pachyspirae*, elles sont annuelles diploïdes, et elles ont des gousses entières à spires serrées. Ces espèces portent les mêmes caractéristiques ce qui explique leurs présences dans un même groupe dans toutes les arbres.

*M. ciliaris*, *M. intertexta*, et *M. granadansis* elles sont annuelles avec des gousses entières et spires serrées elles appartiennent dans une même section des *intertexta*, et elles sont dans un même sub-genre qui contient les 4 sections (*intertexta*, *pachyspirae*, *rotatae*, *leptospirae*).



**Figure 19** : arbre phylogénétique de Steele *et al*, 2010, basé sur la méthode maximum parcimonie, en utilisant le plastide trnK/matK.

Dans la section de *Medicago* on trouve *M. cancellata*, *M. saxatilis*, *M. daghestanica*, *M. falcata* (*sativa*) qui sont des espèces pérennes et elles ont des gousses spiralées.

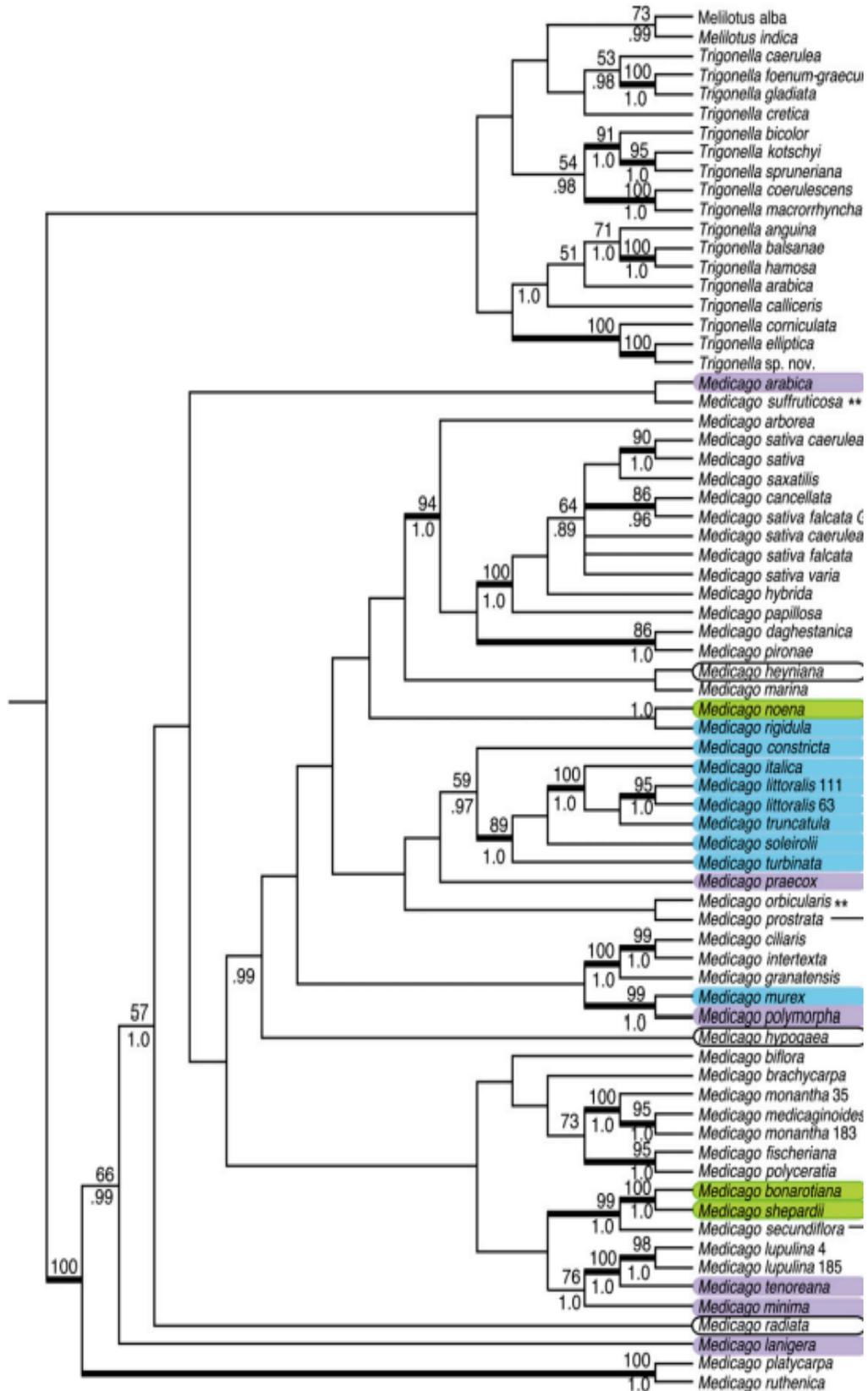
Dans l'arbre de Bena, on remarque une grande différence entre la distribution des espèces, la plus part des groupes de notre arbre sont différents, sauf les deux sections qui comporte que quelques espèces qui sont dans des groupes monophylétiques telle que *M. ciliaris* et *M. intertexta* et *M. daghestanica* de la section *intertexta*, et *M. rigidula*, *M. noeana*, *M. soleirolli* et *M. littoralis* du clade *Pachspireae*.

Cette polyphylie des sous-groupes de *Medicago* serait probablement due au fait que les caractères utilisés pour leurs regroupements, notamment ceux de la gousse, seraient des caractères convergents non hérités d'un ancêtre commun.

*M. granadensis*, *M. intertexta*, *M. ciliaris* et *M. muricolepsis* forment un groupe monophylétique de 99% *Bootstrap*, dans les résultats de Bena, alors que dans notre arbre, *M. intertexta*, *M. ciliaris* et *M. granadensis* appartiennent dans un groupe dont sa valeur *Bootstrap* est de 59%, si on compare ces résultats avec les arbres phylogénétiques de Steele *et al.*, 2010, on remarque que ces trois espèces forment un groupe monophylétique dans les deux marqueurs, dont les valeurs est de 83% dans l'arbre qui est analysé avec le plaste *trnK/matK*, et 100% dans celle qui est analysé avec *GA3ox1*.

Dans ce clade des *intertextas*, il y a une très grande ressemblance entre les deux espèces *M. intertexta* et *M. ciliaris*, leurs valeurs du *bootstrap* est de 99% ou 100% dans toutes les arbres. Elles ont une répartition à tendance allopatrique. La première, dans le Nord-ouest, montre une tolérance vis-à-vis de la sécheresse; la seconde se rencontre essentiellement dans les régions humides du Nord-est (**Amirouche et Misset, 2009**).

Bena *et al.* (1997), en analysant 13 espèces de *Medicago*, n'ont trouvé aucune incongruence entre les analyses séparées ou combinées des séquences nucléotidiques des régions ITS1, ITS2 et EST. Pour expliquer les incongruences des reconstructions basées sur les données moléculaires avec celles basées sur d'autres types de données, Bena *et al.* (1998b) pensent que cela signifie que l'histoire de l'évolution des gènes peut être différente de l'histoire de l'évolution des espèces. Cela s'expliquerait par les différents phénomènes qui surviennent dans les gènes comme par exemple : l'introgession et le transfert horizontal.



**Figure 20:** Arbre phylogénétique de Steele *et al*, 2010, basé sur la méthode maximum parcimonie, en utilisant le gène *GA3ox1*.

Enfin, comme on peut le constater dans tous les arbres construits ici (figure 18,19, 20, 21), ces résultats démontrent la monophylie du genre *Medicago* vis-à-vis de l'extra-groupe *Trigonella*, on remarque que *T. glabra*, *T. elliptica* et *T. anguina* forment un groupe monophylétique dont sa valeur de *bootstrap* est de 97%.

Si on compare nos résultats avec ceux de l'arbre du gène *NORK* d'Erikson , 2012 (figure 22), on remarque une grande différence entre les groupes, sauf *M. ciliaris* et *M. intertexta* qui appartiennent au clade E qui sont dans un groupe de 89%, et *M. sexatilis* et *M. rhodopea* du clade F qui forme un groupe monophylétique de 100%.

Le clade E de l'arbre du gène *NORK*, qui comprend *M. granadensis*, *M. ciliaris* et *M. intertexta* est toujours constant, si on compare ces résultats avec l'arbre des gènes ITS1 et ITS2 de Bena *et al.*, 1997, et les arbres du gène *GA3ox1* et du plastide *trnK/matK* de Steel *et al.*, 2010, on remarque que ces espèces sont toujours dans un même groupe, ce qui confirme les résultats de notre arbre, dont *M. ciliaris* et *M. intertexta* et *M. granadensis* sont dans un groupe de 59%. Ces 3 espèces appartiennent à la section des *Intertextae* du genre *Medicago* (figure 18) avec un niveau de ploïdie  $2n = 2x = 16$  (Halmi, 2010).

Heyn en 1963 a décrit la section des *Intertextaes* comme étant un groupe diploïde et annuel. Les espèces de cette section poussent dans les régions méditerranéennes à étages bioclimatiques allant de l'humide à l'aride supérieur. Ces espèces sont : *M. intertexta*, *M. ciliaris*, *M. muricoleptis*, *M. granadensis*. Les deux dernières espèces ne sont pas signalées en Algérie.

Les principaux caractères des espèces de cette section sont ainsi résumés: Gousses épineuses et enroulées en hélice. Les spires sont orientées, en allant de la base vers l'apex de la gousse, dans le sens des aiguilles d'une montre (Heyn, 1963 ; Lesins et Lesins, 1979). Des nervures radiales anastomosées très saillantes, mais pas de nervures latérales différenciées. Les nervures périphériques rejoignent latéralement la marge de l'épine, et se détachent de la face en formant une aile (Jauzein, 1995). Les graines sont brun-noir à maturité (Small *et al.*, 1989 ; Jauzein, 1995).

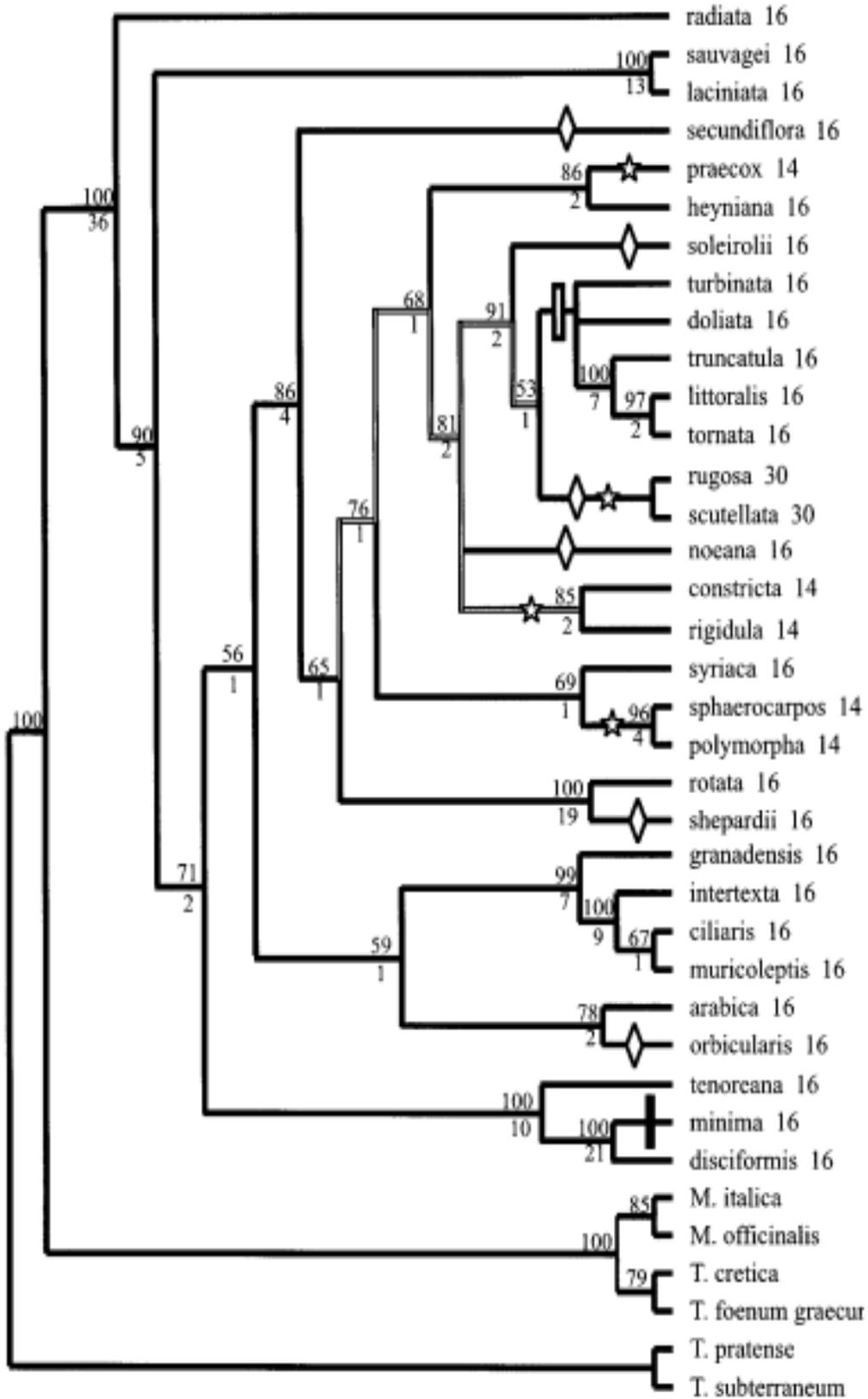


Figure 21 : Arbre phylogénétique de Bena *et al*, 1997 en utilisant les ITS1 et les ITS2.

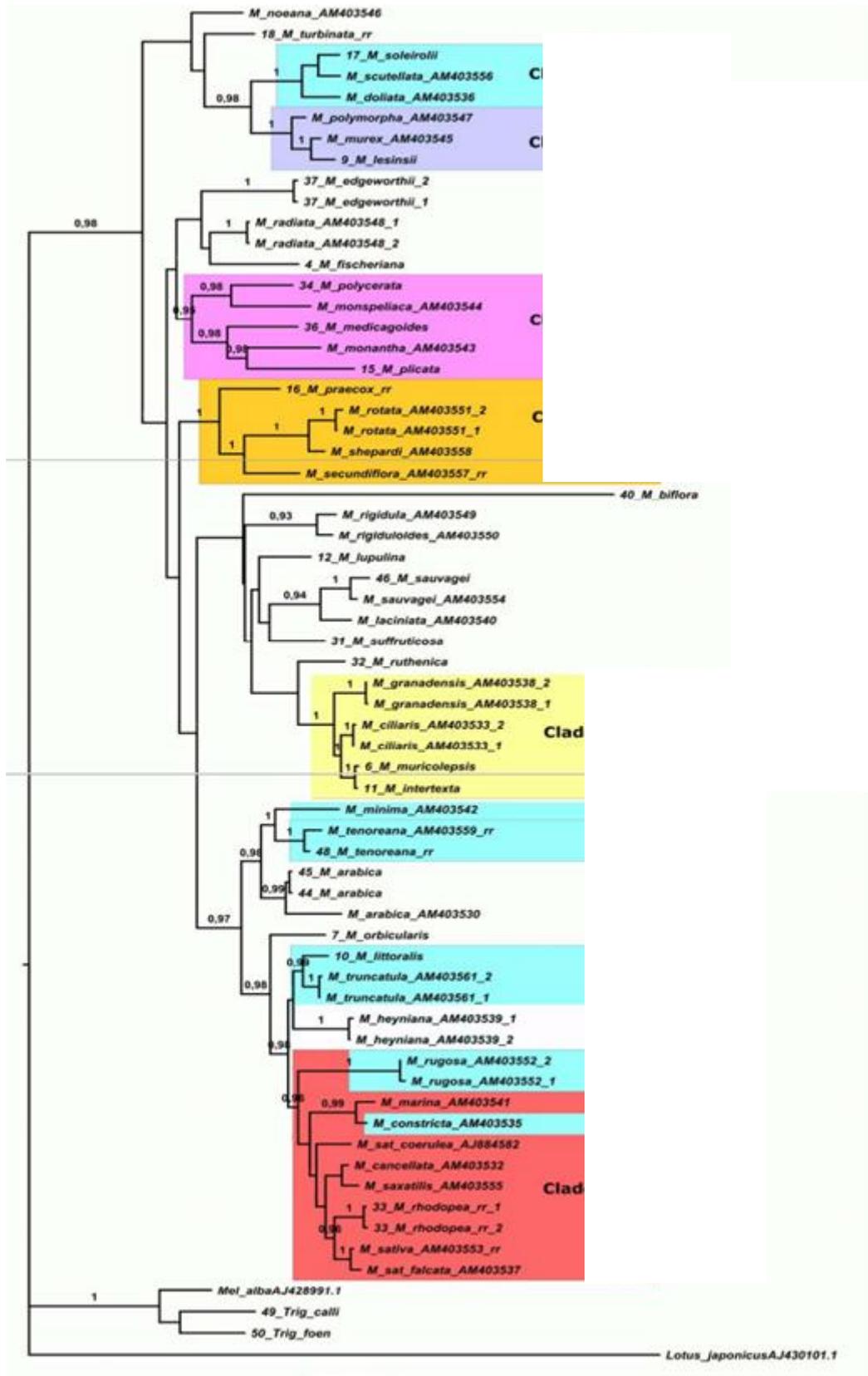


Figure 22: Arbre phylogénétique d'Erikson et al, 2012, en utilisant le gène *NORK*.

#### 4.2. Comparaison avec d'autres arbres des gènes chloroplastiques

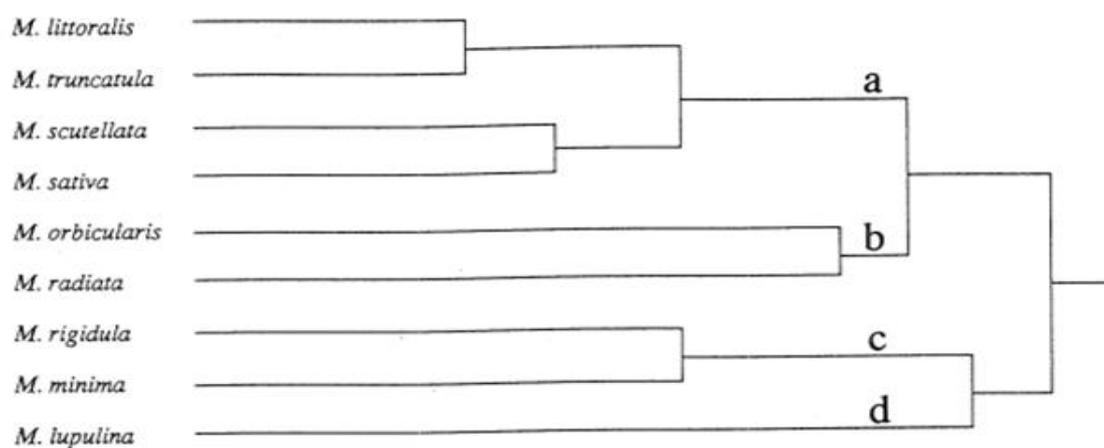
Pour montrer la similarité génétique du genre *Medicago*, Valizadeh (1996) a analysé 9 espèces de *Medicago* qui appartiennent à 4 sous genre de différent continent selon le tableau 5, en utilisant les gènes de chloroplastes par la méthode RFLP et il a obtenu l'arbre 23 qui a 4 clades suivant :

**Le clade a :** *M. littoralis*, *M. truncatula*, *M. scutellata*, *M. sativa*.

**Le clade b :** *M. orbicularis*, *M. radiata*.

**Le clade c :** *M. rigidula*, *M. minima*.

**Le clade d :** *M. lupulina*.



**Figure 23:** Arbre phylogénétique de Valizadeh, 1996.

**Tableau 5 :** Liste et les caractéristiques des espèces étudié de Valizadeh (1996)

Nom	Section	Subsection	Origine
<i>M. orbicularis</i>	<i>Orbicularia</i>	<i>Orbicularis</i>	Azerb. Iran
<i>M. radiata</i>	<i>Orbicularia</i>	<i>Hymenocarpos</i>	Azerb. Iran
<i>M. minima</i>	<i>Spirocarpos</i>	<i>Leptospirae</i>	Azerb. Iran
<i>M. rigidula</i>	<i>Spirocarpos</i>	<i>Pachyspirae</i>	Azerb. Iran
<i>M. littoralis</i>	<i>Spirocarpos</i>	<i>Pachyspirae</i>	Australie
<i>M. truncatula</i>	<i>Spirocarpos</i>	<i>Pachyspirae</i>	Australie
<i>M. scutellata</i>	<i>Spirocarpos</i>	<i>Rotatae</i>	Australie
<i>M. lupulina</i>	<i>Lupularia</i>	–	Azerb. Iran
<i>M. sativa</i>	<i>Medicago</i>	<i>Falcao</i>	Azerb. Iran

En utilisant la méthode RFLP, il a trouvé une grande ressemblance entre *M. littoralis* et *M. truncatula*, et une différence entre *M. lupulina* (Valizadeh, M, 1996).

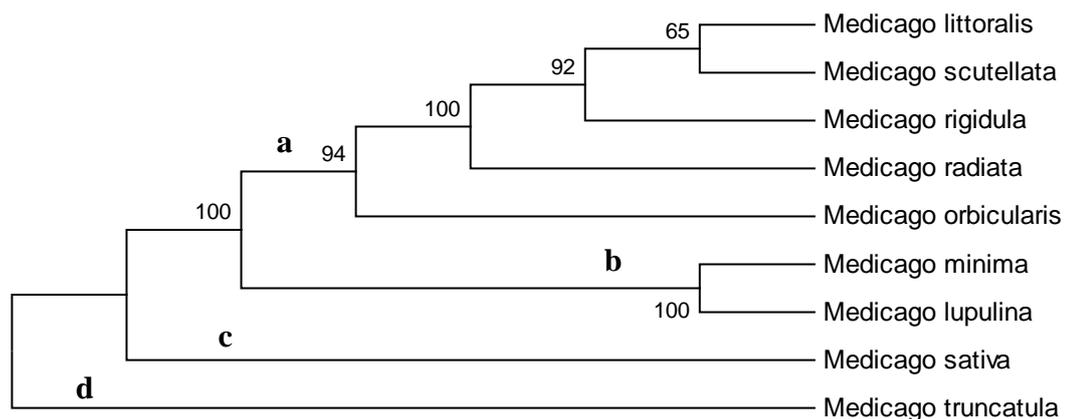
On a choisi les mêmes espèces dont les séquences ont été récupérées par la GenBank, pour construire l'arbre phylogénétique en suivant la même procédure du premier arbre phylogénétique, et on a obtenu les clades suivants (figure 24) :

**Le clade a :** *M. littoralis*, *M. scutellata*, *M. rigidula*, *M. radiata*, *M. orbicularis*.

**Le clade b :** *M. lupulina*, *M. minima*.

**Le clade c :** *M. sativa*.

**Le clade d :** *M. truncatula*.



**Figure 24 :** Arbre phylogénétique de 9 espèces de *Medicago*.

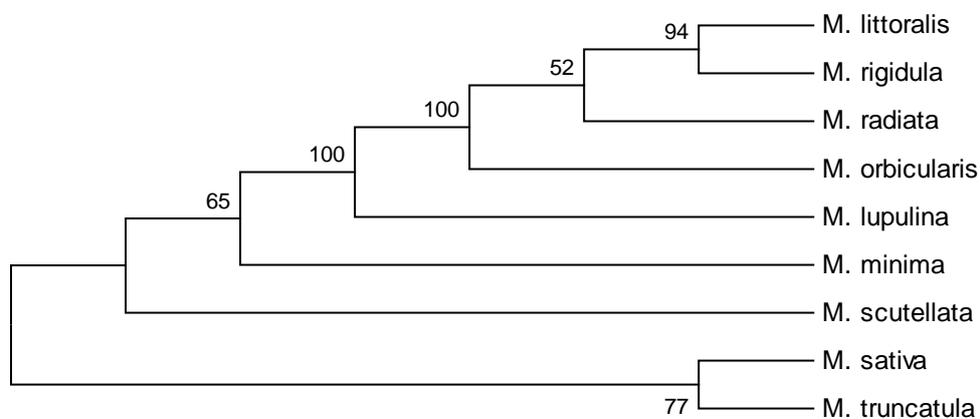
On remarque une ressemblance entre certaines espèces qui sont *M. littoralis*, *M. scutellata*, qui appartiennent dans le clade « a » des deux arbres, et les deux espèces *M. radiata* et *M. orbicularis* qui sont dans le clade « a » de notre deuxième arbre des 9 espèces, sont dans le clade « b » de l'arbre de Valizadeh. Par contre *M. truncatula* est loin de *M. littoralis* ce qui contredit les résultats de Valizadeh qui a trouvé une grande ressemblance entre *M. littoralis* et *M. truncatula*, et une différence entre *M. lupulina*.

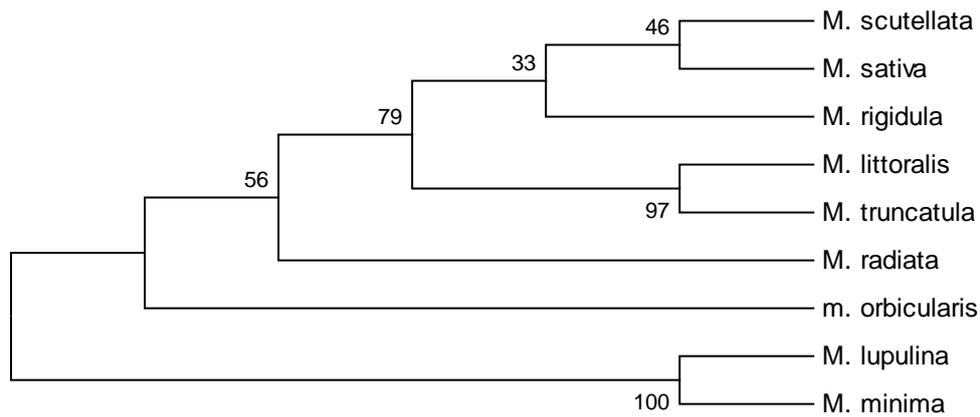
Pour savoir d'où vient cette différence, on a choisi toujours les mêmes espèces de différents continents (tableau 6) pour reconstruire l'arbre de la figure 25.

**Tableau 6 :** L'origine des espèces étudiées obtenue par la *Genbank*

Nom	Pays	Nombre de paire de base	Auteur	Année	Code
<i>M. orbicularis</i>	USA	2354	Steele et al	2016	522101.1
<i>M. radiata</i>	USA	2453	Steele et al	2016	522106.2
<i>M. minima</i>	Pologne	302	Kajtoch et al	2015	746338.1
<i>M. rigidula</i>	USA	2462	Steele et al	2016	159579.1
<i>M. littoralis</i>	USA	2464	Steele et al	2016	159568.1
<i>M. truncatula</i>	Portugal	838	Cunhç et al	2016	858232.1
<i>M. scutellata</i>	Tunisie	533	Zitouna et al	2013	700630.1
<i>M. lupulina</i>	USA	1792	Steele et al	2010	159569.1
<i>M. sativa</i>	Italie	2803	Rosellini	2016	029748.1

D'après les résultats qu'on a obtenue, on remarque que les séquences nucléotidiques des espèces *M. littoralis*, *M. rigidula*, *M. radiata*, *M. orbicularis* et *M. lupulina* des états unis forment un clade monophylétique de 100%, et les espèces *M. sativa* et *M. truncatula* qui sont de l'Italie et le Portugal forment un groupe monophylétique de 77%, *M. littoralis* et *M. truncatula* sont très différent. On a refait un autre arbre (figure 26) avec les mêmes espèces mais leurs séquences nucléotidiques sont présent d'un même pays et on a obtenu l'arbre de la figure 26.

**Figure 25 :** Arbre phylogénétique de 9 espèces de *Medicago* de différents pays.



**Figure 26 :** Arbre phylogénétique de 9 espèces de *Medicago* d'un seul pays.

D'après cette table on remarque une grande ressemblance entre *M. littoralis* et *M. truncatula*, avec un pourcentage élevé de 97%, et *M. lupulina* et *M. minima* avec un pourcentage de 100%, ces résultats confirment les résultats de Valizadeh, cette ressemblance est due à leur présence dans une même région c'est-à-dire la distance génétique entre ces deux groupes.

D'après l'analyse de tous ces arbres, on a déduit que le gène chloroplastique est très utile dans la phylogénie, et qui a donné les mêmes résultats avec quelque gène tel que le gène *trnK/matK* et le gène *GA3ox1*. Et la polyphylie serait probablement due aux caractères convergents non hérités d'un ancêtre commun. Ou à la distance génétique entre ces groupes.

La reconstruction phylogénétique est un outil utilisé dans les domaines aussi divers que la taxonomie, l'épidémiologie et la génétique des populations. C'est une procédure incontournable pour retracer l'histoire évolutive des organismes de tous niveaux taxonomiques. Cette analyse est également d'une grande importance pour clarifier les modèles évolutifs des familles de séquences multigènes et pour comprendre l'évolution adaptative au niveau morphologique et moléculaire.

La connaissance approfondie des données appropriées pour réaliser des analyses phylogénétiques est primordiale. En effet, pour étudier l'évolution et les rapports entre les organismes, différents types de données peuvent être utilisés : données morphologiques, données de polymorphisme de fragments (RAPD, RFLP...etc.) et séquences nucléotidiques ou protéiques. Pour une bonne reconstruction, l'échantillonnage doit être le plus exhaustif possible (recouvrant le maximum de la variabilité des taxa à étudier) et les caractères (morphologiques ou moléculaires) doivent être homologues. Les premières étapes d'une inférence phylogénétique pour tout type de données sont très importantes parce que décisives: dans le cas des données morphologiques, les caractères à utiliser doivent être suffisamment nombreux et bien choisis. Il est ensuite procédé à la polarisation (repérage des apomorphies et des plésiomorphies) des états des de ces caractères selon des critères ontogéniques, paléontologiques ou par comparaison avec un taxon de référence (extra-groupe) dont les degrés d'évolution des états de caractères sont connus. Enfin, la codification de ces états de caractères doit être obligatoirement adaptée au jeu de données de telle sorte qu'elle reflète au mieux leur variabilité et leur degré d'évolution.

Durant ces dernières années, les trois types de séquences nucléotidiques, ADN ribosomique, ADN chloroplastique et ADN mitochondrial sont apparus comme la source d'information de prédilection pour déterminer les liens de parenté entre les taxa au cours de leur évolution, en raison notamment de leurs abondances dans la cellule, de leurs petites tailles et de leurs richesses en gènes avec simple copie (non répétés). Le grand intérêt des systématiciens pour ces séquences, stimulé par les avancées technologiques des techniques d'extraction et de séquençage de l'ADN et le développement des programmes informatiques sophistiqués, s'est traduit par l'apparition d'un grand nombre de banques de données (EMBL, NCBI et DDBJ) destinées au stockage et au traitement des séquences et des phylogénies publiées.

Une fois obtenues, les séquences d'ADN ou de protéines pour les taxa considérés, il est procédé à leur alignement. Cette étape, primordiale, consiste à identifier les positions homologues des résidus (nucléotides ou acides aminés) dans chacune des séquences. L'alignement maximise ainsi la similitude des séquences comparées. Le logiciel le plus utilisé pour ce faire est ClustalW. A l'issue de l'alignement multiple, une matrice de taxa/séquences est obtenue. C'est cette dernière qui servira de base à l'inférence phylogénétique.

Le choix des méthodes de reconstruction d'arbres phylogénétiques repose sur cinq critères : l'efficacité, la puissance, la consistance, la robustesse et la fiabilité.

Pour réaliser la partie pratique de ce travail, nous avons utilisé le logiciel MEGA 5 sur le genre *Medicago*. Les résultats donnés par la méthode UPGMA un arbre phylogénétique qui est soutenu par une forte valeur *bootstrap*. Les groupements taxonomiques générés par les phylogénies obtenues pour 43 espèces sur une combinaison des séquences chloroplastiques sont congruents entre eux mais pas avec ceux d'autres auteurs sur d'autres données et avec la classification traditionnelle excepté pour *Medicago Sect. Intertextae*. En outre, nos résultats démontrent aussi la monophylie du genre *Medicago* vis-à-vis des genres proches *Trigonella*.

- Abdelguerfi et Abdelguerfi-Laouar., (2004).** Les ressources génétiques d'intérêt fourrager et/ou pastoral : Diversité, collecte et valorisation au niveau méditerranéen. Cahiers Options Méditerranéennes; 62 : 29-41.
- Abdelgherfi et Marrakchi., (2000).** Les ressources phylogénétiques fourragères et pastorales : de l'érosion à la conservation. Option méditerranéennes. 45 :15-27.
- Agolin, M., (2003).** Des organismes à l'inférence phylogénétique : Méthodes et techniques en systématique moléculaire. Exemple de la phylogénie des Elapodea (Serpentes). Mémoire de Maîtrise de Biologie des Populations et des Ecosystèmes, Univ. Pierre et Marie Curie, Paris VI. 31p.
- Alvarez, I., Wendel., J.F., (2003).** Ribosomal ITS sequences and plant phylogenetic inference. *Molecular Phylogenetics and Evolution.*, 29 : 417– 434.
- Ayala, F.J., (1982).** *Biologie Moléculaire et Evolutive.* Ed. Masson, Paris., 56-79.
- Baldwin, B.G., Sanderson, M.J., Porter, J.M., Wojciechowski, M.F., Campbell, C.S., Donoghue, M.J., (1995).** The ITS region of nuclear ribosomal DNA: A valuable source of evidence on angiosperm phylogeny. *Ann. Missouri. Bot. Gard.*, 82: 247–277.
- Bardakci, F., (2000).** Random Amplified Polymorphic DNA (RAPD) Markers. *Turk. J. Biol.*, 25: 185-196.
- Bena, G., Lejeune, B., Prosperi, J.M., Olivieri, I., (1998b).** Molecular phylogenetic approach for studying life-history evolution: the ambiguous example of the genus *Medicago* L. *Proc. R. Soc. Lond B. Biology. Sci.*, 265 (1401): 1141-1151.
- Bena, G., Jubier, M.F., Olivieri, I., Lejeune, B., (1998a).** Ribosomal External and Internal Transcribed Spacers: Combined Use in the Phylogenetic Analysis of *Medicago* (Leguminosae). *J. Mol., Evol* 46: 299- 306.
- Bena, G., Lejeune, B., Prosperi, J.M., Olivieri, I., (1998b).** Evolution of Annual Species of the Genus *Medicago*: A Molecular Phylogenetic Approach. *Molecular Phylogenetics and Evolution.*, V.9. (3): 552-559.

- BEN NJA., (2014).** Effet d'un stress salin sur la teneur en polymères pariétaux dans les feuilles de luzerne (*Medicago sativacv* Gabès) et sur la distribution dans les cellules de transfert des fines nervures. These de doctorat en biologie, Université De Carthage., 179p.
- Berry, V., 1997.** Méthodes et algorithmes pour reconstruire les arbres de l'évolution. Thèse de doctorat en Informatique. Dir. Thèse : Gascuel O. Univ. Montpellier II. 1-69. 252p.
- Bodo Slotta, T.A., (2000).** Phylogenetic Analysis of *Iliamna* (Malvaceae) Using the Internal Transcribed Spacer Region. Thesis of Sciences in Biology. Faculty of the Virginia Polytechnic Institute and State University., 1-14. 76p.
- Chapus, C., (2005).** Analyse de données phylogénétiques à l'aide de la signature génomique. Aspects méthodologiques et application aux procaryotes. Thèse de doctorat U.F.R. de Biologie l'université Paris VII., 206p.
- Chakravatry, H.L., (1976).** Plant Wealth of Iraq.Ed.A dictionary of economic plant., 1: 505.
- Cheikhrouhou., (2006).** évaluation des classifications phylogénétiques des bacillaceae basées sur les gènes de l'opéron *rrn* et de gènes de ménage. Université Du Quebec A Montreal., 126p.
- Clark, K., Karsch-Mizrachi, ., Lipman, D.J., Ostell, J and Sayers, E.W., (2015).** GenBank. Nucleic Acids Research, 1 doi: 10.1093/nar/gkv1276. 6p.
- Darlu, P., Tassy, P., (1993).** La reconstruction phylogénétique : concepts et méthodes. Ed. Masson, Paris., 245p.
- Darlu, P., Tassy, P., (2004).** La reconstruction phylogénétique : concepts et méthodes. Ed. Masson, Paris., 258p.
- Dowell., (2008).** Molecular Phylogenetics An introduction to computational methods and tools for analyzing evolutionary relationships., 1-19.
- Erikson, J., (2012).** the phylogeny of *Medicago* (Fabaceae) using the low copy nuclear gene *NORK*. Degree project for bachelor of science in systematic and biodiversity, biology, university of Gothenburg., 1-13.

- Ernest, S., Stephen, R.D., Deborah, S., Erica, J., Heidi, L.Z., (1998).** multiple independent losses of the plastic rpoC1 intron in *Medicago* (Fabaceae) as inferred from phylogenetic analyses of nuclear ribosomal DNA internal transcribed spacer sequences. *Can. J. Bot.*, 76(5): 791-803.
- Gagnon, J., (2004).** Création d'outils pour l'automatisation d'analyses phylogénétiques de génomes d'organites. Mémoire de Maîtrise en Sciences (Bioinformatique). Univ. Laval., 132p.
- Gallut, C., (2001).** Codage et traitement phylogénétique de caractères structuraux de génomes entiers. Thèse Doctorat en Sciences de la Vie. Dir. Thèse : G. Gachelin. Univ. Pierre et Marie Curie, Paris., 6 : 23-32.
- Gattolliat, J. L., (2002).** Etude systématique, cladistique et biogéographique des Baetidae (Ephemeroptera) de Madagascar, Thèse de doctorat en Zoologie et Ecologie Animale. Dir. Thèse : J-M Elouard. Univ. Lausanne., 111-145.
- Gobron., (2012).** Caractérisation génétique et moléculaire d'une incompatibilité nucléocytoplasmique chez *Arabidopsis thaliana*. These de doctorat, univ. Paris sud., 234p.
- Halmi, S., (2010).** Contribution à l'étude cytogénétique de quelques espèces du genre *Medicago* (L). mémoire de magistère en biotechnologies végétales. Université Mentouri Constantine., 97p.
- Heyn, C.C., (1963).** The annual species of *Medicago* Jerusalem: magnes Press, the Hebrew university., vol.XII.145p.
- Huelsenbeck, J.P., Larget, B., Swofford, D., (2000).** A Compound Poisson Process for Relaxing the Molecular Clock. *Genetics.*, 154: 1879–1892.
- Pierre., (2008).** Recherche de d'déterminants génétiques de la date de floraison chez la L'égumineuse modèle, *Medicago truncatula*. Thèse de doctorat. Ecole nationale supérieure agronomique de rennes., 207p.
- Jauzein, P., (1995).** *Flordes champs cultivés* edition INRA, paris., 828p.
- Lapeyronie, A., (1982).** Les productions fourragères méditerranéennes- technique agricole et production méditerranéenne. Maisonneuve et Larose. Paris., 307-315.

- Lazrek., (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. Université de Toulouse III - Paul Sabatier., 254p.
- Lesins et Lesins., (1979).** genus *Medicago* (leguminales) a taxogenetic study. *Plant ecology.*, 50 (2):228-229.
- Hillis, D.M., Moritz, C., Mable, B.K., (1996).** *Molecular Systematics*. 2nd. Ed. Sunderland, Massachusetts: Sinauer.
- Hireche, Y., (2006).** Réponse de la Luzerne (*Medicago sativa* L) au stress hydrique et à la profondeur de semis. Université ai hadj lakhdar. These de magistere . Batna chromosomique chez 35 populations de 17 especes. *Annales de l'institut National Agronomique*. El-Harrach., 12 (1) : 342-354 p.
- Higgins, D.G., Sharp, P.M., (1988).** Clustalw: a package for performing multiple sequence alignments on a microcomputer., *Gene* 73: 237 - 244.
- Mathieu-Demazière., (2014).** Cyclic-AMP signalling in *Sinorhizobium meliloti* and its role in the control of infection during symbiosis with *Medicago*. These de doctorat en biologie, universite de Toulouse 3 Paul Sabatier., 241p.
- Messioughi., (2010).** Analyse des substances actives "les flavonoïdes" et action antibactérienne d'une fabacée à intérêt médicinal "*Medicago sativa*.L." cultivée sur des sols du Nord-Est algérien. Mémoire de magistere en biologie. Université Badji Mokhtar-Annaba., 123p.
- Mollaret. I., (2000).** Phylogénie des Monogènes (Plathelminthes) : apport des caractères spermatologiques et moléculaires. Thèse de doctorat en Sciences Biologiques. Dir. Thèse: J Schrével. Muséum national d'histoire naturelle. Paris., 27-40. 273p.
- Moura. M., (2002).** Méthodes phylogénétiques : de la systématique en biologie et de ses applications en philologie (stemmatologie) et en linguistique historique. Mémoire de licence en Informatique et méthodes Mathématiques. Dir. Mémoire : F. Bavaud. Univ. Lausanne., 8-39. 66p.

- Nylander, J.A.A., (2001).** Taxon Sampling in Phylogenetic Analysis: Problems and Strategies Reviewed. Department of Systematic Zoology Evolutionary Biology Centre. Uppsala University., 28p.
- Olmstead, R.G., Jansen, R.K., Michael, H.J., Downie, S.R., Palmer, J.D., (1990).** Chloroplast DNA and Phylogenetic Studies in the Asteridae. Biological Approaches and Evolutionary Trends in Plants. Academic Press., U.S.A. 120-134.
- Onarici, S.G., Sumer, S., (2002).** Protein and DNA in Systematic Biology., Turk J Biol 27: 47-55.
- Oudot le-secq, M.P., (2000).** Organisation et structure des génomes mitochondriaux des algues brunes *Pylaiella littoralis* et *Laminaria digitata*. Thèse de doctorat en Biologie. Dir. These: S. Loiseaux-de-Göer. Univ. Rennes I., 30-32. 166p.
- Lopez, P., Casane, D., Herve, P., (2002).** Bio-informatique (5) Phylogénie et évolution moléculaires. Collection Médecine/Sciences; 18 : 1146-54.
- Prolea, D., (2002).** Espèces et utilisations, des ressources en protéines à redécouvrir : les plantes fourragères prairiales- la luzerne. Institut du végétal et de l'institut de l'élevage GNIS. Paris., P : 4-7.
- Prosperi, J.M., Guy, P. and Balfourier, F., (1995).** Ressource génétiques des plantes fourragère et à gazon.Ed INRA.Paris., 245p.
- Prosperi, J.M., Oliviers, I., Angevain., Nausat, P., (1993).** Diversité génétique. Conservation et utilisation des ressources génétiques des luzernes méditerranéennes. Le courrier de l'environnement de l'INRA.P., 1-12.
- Prosperi, J.M., Auricht, G.C., Genier, G., Johnson, R.C., (2000).** the genetic Diversity of legumes species in the Mediterranean: Medicago. In the plant genetic resources of legumes in the Mediterranean Maxed N.and Bennett S.Eds Kluwer Academic publishers Netherlands.
- Quezel, P., Sanata., S., (1962).** nouvelle flore de l'Algérie et des régions désertiques méridionales. T. L. Ed. CNRS.Paris., 498-506 p.

- Robinson, M., (1997).** Diversité des Modes d'évolution des Génomes de Rongeurs. Thèse de doctorat en Biologie. Dir. Thèse : M. Dominique. Univ. Claude Bernard - Lyon I., 29-43. 158p.
- Riah, N., (2014).** Thème Diversité et structure génétique des populations de *Rhizobium leguminosarum symbiovar viciae* isolées du pois (*Pisum sativum*) et de la lentille (*Lens culinaris*) cultivés dans deux zones éco-climatiques subhumide et semi-aride de l'est algérien. Thèse de doctorat en Biologie. Univ Constantine., 1. 119p.
- Salemi, M., Vandamme, A.M., (2003).** The phylogenetic Handbook: A Practical Approach to DNA and Protein Phylogeny. Cambridge University Press., 72-133.
- Schmidt, H.A., (2003).** Phylogenetic Trees from Large Datasets Inaugural. Thèse de Doctorat en Mathématique. Dir. Thèse: Von Haeseler A. Univ. Heinrich–Heine–Düsseldorf vorgelegt von Heiko., 123p.
- Schoutteten, F., (2004).** La Luzerne. Fiche technique agro-industrie. Champagen-Ardenne. P1-5.
- Sepp, S., Bobrova, V.K., Troitsky, A.K., Glazunova, K.P., (2000).** Genetic polymorphism detected with RAPD analysis and morphological variability in some microspecies of apomictic *Alchemilla*. *Ann. Bot. Fennici* 37 : 105-123.
- Steele, K., Ickert-Bond, S., Zarre, S., Wojciechowski, F., (2010).** Phylogeny And Character Evolution In *Medicago* (Leguminosae): Evidence From Analyses Of Plastid *Trnk/Matk* And Nuclear *Ga3ox1*. *American Journal of Botany.*, 97(7): 1142–1155.
- Tourasse, N.J., (1992).** Développement d'une distance évolutive entre séquences prenant en compte la variabilité du taux de substitution entre sites et application à la reconstruction de phylogénies moléculaires anciennes. Thèse Doctorat en Génétique et Biologie des Populations. Univ. Claude Bernard – Lyon., 1. 1-61.
- Thompson, J. D., Higgins, D. G. et Gibson, T. J. (1994).** clustalw: improving the sensitivity of progressive multiple alignment through sequence weighting, positionspecific gap penalties and weight matrix choice. *Nucleic Acids Res.*, 22: 4673 - 4680.

**Valizadeh, M., Kwon Kyoo Kang, Akira Kanno., Toshiaki Kameya., (1996).** Analysis of Genetic Distance among Nine Medicago Species by Using DNA Polymorphisms. Breeding science., 46: 7-10.

**Ranwez, V., Gascuel, O., (2002).** TripleML : une amélioration des méthodes de distance pour l'inférence phylogénétique, grâce à une approche locale du maximum de vraisemblance basée sur les triplets. Ed. LIRMM. Montpellier., 37-46.

**Wilbur, W. J., Lipman, D. J., (1983).** Rapid similarity searches of nucleic acid and protein data banks. Proe. Natl. Aead. Sei. USA., 80: 726 - 730.

### **Web bibliographique**

URL:[http://www.noble.org/MedicagoHandbook/pdf/WildAccessions\\_Populations.pdf](http://www.noble.org/MedicagoHandbook/pdf/WildAccessions_Populations.pdf) tableau N1)

Tableau N°1 : les résultats de la GenBank.

Gènes	N°	Espèces	Pb
5.8S rRNA gene (partial), 28S rRNA gene (partial) and internal transcribed spacer 2 (ITS2)	1	<i>Medicago aurantiaca</i>	237 bp
	2	<i>Medicago brachycarpa</i>	236 bp
	3	<i>Medicago huberi</i>	236 bp
	4	<i>Medicago lanigera</i>	236 bp
	5	<i>Medicago medicaginoïdes</i>	236 bp
	6	<i>Medicago monantha</i>	236 bp
	7	<i>Medicago monspeliaca</i>	236 bp
	8	<i>Medicago papillosa</i>	236 bp
	9	<i>Medicago platycarpus</i>	237 bp
	10	<i>Medicago polyceratia</i>	236 bp
	11	<i>Medicago popovii</i>	236 bp
	12	<i>Medicago prostrata</i>	236 bp
	13	<i>Medicago saxatilis</i>	236 bp
ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit	1	<i>Medicago arabica</i>	763 bp
	2	<i>Medicago falcata</i>	658 bp
	3	<i>Medicago lupulina</i>	1406 bp
	4	<i>Medicago minima</i>	658 bp
	5	<i>Medicago murex</i>	533 bp
	6	<i>Medicago polymorpha</i>	1317 bp
	7	<i>Medicago scutellata</i>	533 bp
maturase K (matK)	1	<i>Medicago arabica</i>	916 bp
	2	<i>Medicago archiducis-nicolai</i>	2411 bp
	3	<i>Medicago biflora</i>	2468 bp
	4	<i>Medicago bonarotiana</i>	2221 bp
	5	<i>Medicago brachycarpa</i>	1548 bp
	6	<i>Medicago cancellata</i>	2455 bp
	7	<i>Medicago carstiensis</i>	2209 bp
	8	<i>Medicago ciliaris</i>	2443 bp
	9	<i>Medicago constricta</i>	2477 bp
	10	<i>Medicago cretacea</i>	1521 bp
	11	<i>Medicago daghestanica</i>	2466 bp
	12	<i>Medicago edgeworthii</i>	2359 bp
	13	<i>Medicago falcata</i>	699 bp
	14	<i>Medicago falcata subsp. falcata</i>	723 bp
	15	<i>Medicago fischeriana</i>	1929 bp
	16	<i>Medicago granadensis</i>	2436 bp
	17	<i>Medicago heyniana</i>	2497 bp
	18	<i>Medicago hybrida</i>	2457 bp
	19	<i>Medicago hypogaea</i>	2412 bp
	20	<i>Medicago intertexta</i>	2465 bp

	21	<i>Medicago italica</i>	2487 bp
	22	<i>Medicago laciniata</i>	1521 bp
	23	<i>Medicago lanigera</i>	2456 bp
	24	<i>Medicago littoralis</i>	2464 bp
	25	<i>Medicago lupulina</i>	2386 bp
	26	<i>Medicago marina</i>	2455 bp
	27	<i>Medicago medicaginoïdes</i>	2488 bp
	28	<i>Medicago minima</i>	2436 bp
	29	<i>Medicago monantha</i>	2462 bp
	30	<i>Medicago monspeliaca</i>	2482 bp
	31	<i>Medicago papillosa</i>	2481 bp
	32	<i>Medicago pironae</i>	2420 bp
	33	<i>Medicago platycarpus</i>	2501 bp
	34	<i>Medicago polyceratia</i>	2464 bp
	35	<i>Medicago polymorpha</i>	2473 bp
	36	<i>Medicago prostrata</i>	2449 bp
	37	<i>Medicago radiata</i>	2453 bp
	38	<i>Medicago rhodopea</i>	2435 bp
	39	<i>Medicago rigidula</i>	2462 bp
	40	<i>Medicago ruthenica</i>	2492 bp
	41	<i>Medicago soleirolii</i>	2448 bp
	42	<i>Medicago saxatilis</i>	2444 bp
	43	<i>Medicago scutellata</i>	2466 bp
	44	<i>Medicago secundiflora</i>	2453 bp
	45	<i>Medicago shepardii</i>	2436 bp
	46	<i>Medicago turbinata</i>	2470 bp
polygalacturonase 11a (Pg11a) gene, partial cds	1	<i>Medicago ciliaris</i>	530 bp
	2	<i>Medicago littoralis</i>	824 bp
	3	<i>Medicago littoralis</i>	824 bp
	4	<i>Medicago polymorpha</i>	560 bp
	5	<i>Medicago rigidula</i>	797 bp
	6	<i>Medicago rigiduloides</i>	824 bp
	7	<i>Medicago ruthenica</i>	457 bp
polygalacturonase 11c (Pg11c) (10)	1	<i>Medicago arabica</i>	988 bp
	2	<i>Medicago carstiensis</i>	351 bp
	3	<i>Medicago littoralis</i>	809 bp
	4	<i>Medicago littoralis</i>	813 bp
	5	<i>Medicago marina</i>	644 bp
	6	<i>Medicago polymorpha</i>	957 bp
	7	<i>Medicago rigidula</i>	667 bp
	8	<i>Medicago rigiduloides</i>	984 bp
	9	<i>Medicago ruthenica</i>	823 bp
	10	<i>Medicago sauvagei</i>	824 bp

polygalacturonase 3 (Pg3) gene	1	<i>Medicago laciniata</i>	1173 bp
	2	<i>Medicago marina</i>	1176 bp
	3	<i>Medicago rigidula</i>	796 bp
	4	<i>Medicago rigiduloides</i>	796 bp
	5	<i>Medicago sauvagei</i>	1139 bp
	6	<i>Medicago arabica</i>	946 bp
	7	<i>Medicago carstiensis</i>	790 bp
auxine transporter-like protein 2	1	<i>Medicago arabica</i>	798 bp
	2	<i>Medicago carstiensis</i>	798 bp
	3	<i>Medicago ciliaris</i>	798 bp
	4	<i>Medicago laciniata</i>	785 bp
	5	<i>Medicago littoralis</i>	798 bp
	6	<i>Medicago littoralis</i>	798 bp
	7	<i>Medicago marina</i>	798 bp
	8	<i>Medicago polymorpha</i>	798 bp
	9	<i>Medicago rigidula</i>	798 bp
	10	<i>Medicago rigiduloides</i>	798 bp
	11	<i>Medicago ruthenica</i>	798 bp
	12	<i>Medicago sauvagei</i>	798 bp
auxine transporter-like protein 4	1	<i>Medicago laciniata</i>	1159 bp
	2	<i>Medicago littoralis</i>	1298 bp
	3	<i>Medicago marina</i>	1298 bp
	4	<i>Medicago polymorpha</i>	1298 bp
	5	<i>Medicago rigidula</i>	1298 bp
	6	<i>Medicago rigiduloides</i>	909 bp
triosephosphate translocator (TRPT)	1	<i>Medicago arabica</i>	316 bp
	2	<i>Medicago lanigera</i>	310 bp
	3	<i>Medicago monantha</i>	310 bp
phosphogluconate dehydrogenase (PGDH)	1	<i>Medicago arabica</i>	452 bp
	2	<i>Medicago lanigera</i>	452 bp
	3	<i>Medicago monantha</i>	452 bp
gibberellin 3-beta-hydroxylase	1	<i>Medicago arabica</i>	1403 bp
	2	<i>Medicago arborea</i>	1396 bp
	3	<i>Medicago archiducis-nicolai</i>	1453 bp
	4	<i>Medicago biflora</i>	1458 bp
	5	<i>Medicago bonarotiana</i>	1320 bp
	6	<i>Medicago brachycarpa</i>	1444 bp
	7	<i>Medicago cancellata</i>	1452 bp
	8	<i>Medicago ciliaris</i>	1459 bp
	9	<i>Medicago constricta</i>	1417 bp
	10	<i>Medicago daghestanica</i>	1432 bp
	11	<i>Medicago edgeworthii</i>	1477 bp

	12	<i>Medicago fischeriana</i>	1478 bp
	13	<i>Medicago granadensis</i>	1461 bp
	14	<i>Medicago heyniana</i>	1464 bp
	15	<i>Medicago hybrida</i>	1454 bp
	16	<i>Medicago hypogaea</i>	1465 bp
	17	<i>Medicago intertexta</i>	1450 bp
	18	<i>Medicago italica</i>	1398 bp
	19	<i>Medicago lanigera</i>	1503 bp
	20	<i>Medicago littoralis</i>	1427 bp
	21	<i>Medicago lupulina</i>	1469 bp
	22	<i>Medicago marina</i>	1406 bp
	23	<i>Medicago medicaginoides</i>	1448 bp
	24	<i>Medicago minima</i>	1444 bp
	25	<i>Medicago monantha</i>	1450 bp
	26	<i>Medicago papillosa</i>	1460 bp
	27	<i>Medicago pironae</i>	1462 bp
	28	<i>Medicago platycarpus</i>	1465 bp
	29	<i>Medicago polyceratia</i>	1505 bp
	30	<i>Medicago polymorpha</i>	1480 bp
	31	<i>Medicago prostrata</i>	1481 bp
	32	<i>Medicago radiata</i>	1449 bp
	33	<i>Medicago rigidula</i>	1462 bp
	34	<i>Medicago ruthenica</i>	1479 bp
	35	<i>Medicago soleirolii</i>	1391 bp
	36	<i>Medicago saxatilis</i>	1418 bp
	37	<i>Medicago secundiflora</i>	1462 bp
	38	<i>Medicago shepardii</i>	1457 bp
	39	<i>Medicago turbinata</i>	1417 bp
tRNA-Lys	1	<i>Medicago arabica</i>	2461 bp
	2	<i>Medicago arborea</i>	2453 bp
	3	<i>Medicago archiducis-nicolai</i>	2411 bp
	4	<i>Medicago biflora</i>	2468 bp
	5	<i>Medicago bonarotiana</i>	2221 bp
	6	<i>Medicago cancellata</i>	2455 bp
	7	<i>Medicago carstiensis</i>	2209 bp
	8	<i>Medicago ciliaris</i>	2443 bp
	9	<i>Medicago constricta</i>	2477 bp
	10	<i>Medicago daghestanica</i>	2466 bp
	11	<i>Medicago edgeworthii</i>	2359 bp
	12	<i>Medicago fischeriana</i>	1929 bp
	13	<i>Medicago granadensis</i>	2436 bp
	14	<i>Medicago heyniana</i>	2497 bp

	15	<i>Medicago hybrida</i>	2457 bp
	16	<i>Medicago hypogaea</i>	2412 bp
	17	<i>Medicago intertexta</i>	2465 bp
	18	<i>Medicago italica</i>	2428 bp
	19	<i>Medicago italica</i>	2487 bp
	20	<i>Medicago lanigera</i>	2456 bp
	21	<i>Medicago littoralis</i>	2464 bp
	22	<i>Medicago littoralis</i>	2445 bp
	23	<i>Medicago lupulina</i>	2386 bp
	24	<i>Medicago lupulina</i>	1792 bp
	25	<i>Medicago marina</i>	2455 bp
	26	<i>Medicago medicaginoides</i>	2488 bp
	27	<i>Medicago minima</i>	2436 bp
	28	<i>Medicago monantha</i>	2436 bp
	29	<i>Medicago monantha</i>	2462 bp
	30	<i>Medicago monantha</i>	2322 bp
	31	<i>Medicago monspeliaca</i>	2482 bp
	32	<i>Medicago papillosa</i>	2481 bp
	33	<i>Medicago pironae</i>	2420 bp
	34	<i>Medicago platycarpus</i>	2501 bp
	35	<i>Medicago polyceratia</i>	2464 bp
	36	<i>Medicago polymorpha</i>	2473 bp
	37	<i>Medicago prostrata</i>	2449 bp
	38	<i>Medicago radiata</i>	2453 bp
	39	<i>Medicago rhodopea</i>	2435 bp
	40	<i>Medicago rigidula</i>	2462 bp
	41	<i>Medicago ruthenica</i>	2375 bp
	42	<i>Medicago ruthenica</i>	2492 bp
	43	<i>Medicago soleirolii</i>	2448 bp
	44	<i>Medicago saxatilis</i>	2444 bp
	45	<i>Medicago scutellata</i>	2466 bp
	46	<i>Medicago secundiflora</i>	2453 bp
	47	<i>Medicago shepardii</i>	2436 bp
	48	<i>Medicago turbinata</i>	2470 bp
CNGC5-like protein	1	<i>Medicago arabica</i>	670 bp
	2	<i>Medicago arborea</i>	843 bp
	3	<i>Medicago astroites</i>	794 bp
	4	<i>Medicago biflora</i>	854 bp
	5	<i>Medicago blanchearna</i>	669 bp
	6	<i>Medicago brachycarpa</i>	795 bp
	7	<i>Medicago carstiensis</i>	845 bp
	8	<i>Medicago ciliaris</i>	672 bp
	9	<i>Medicago constricta</i>	845 bp

	10	<i>Medicago coronata</i>	839 bp
	11	<i>Medicago crassipes</i>	801 bp
	12	<i>Medicago cretacea</i>	845 bp
	13	<i>Medicago disciformis</i>	836 bp
	14	<i>Medicago doliata</i>	842 bp
	15	<i>Medicago edgeworthii</i>	795 bp
	16	<i>Medicago falcata</i>	842 bp
	17	<i>Medicago fischeriana</i>	855 bp
	18	<i>Medicago granadensis</i>	671 bp
	19	<i>Medicago heyniana</i>	845 bp
	20	<i>Medicago huberi</i>	795 bp
	21	<i>Medicago intertexta</i>	671 bp
	22	<i>Medicago italica</i>	840 bp
	23	<i>Medicago laciniata</i>	845 bp
	24	<i>Medicago lanigera</i>	851 bp
	25	<i>Medicago laxispira</i>	831 bp
	26	<i>Medicago lesinsii</i>	839 bp
	27	<i>Medicago littoralis</i>	842 bp
	28	<i>Medicago lupulina</i>	842 bp
	29	<i>Medicago marina</i>	822 bp
	30	<i>Medicago minima</i>	841 bp
	31	<i>Medicago monspeliaca</i>	851 bp
	32	<i>Medicago phrygia</i>	852 bp
	33	<i>Medicago pironae</i>	845 bp
	34	<i>Medicago platycarpus</i>	824 bp
	35	<i>Medicago plicata</i>	851 bp
	36	<i>Medicago polymorpha</i>	831 bp
	37	<i>Medicago popovii</i>	841 bp
	38	<i>Medicago prostrata</i>	842 bp
	39	<i>Medicago radiata</i>	800 bp
	40	<i>Medicago rhodopea</i>	841 bp
	41	<i>Medicago rigidula</i>	840 bp
	42	<i>Medicago rotata</i>	670 bp
	43	<i>Medicago ruthenica</i>	841 bp
	44	<i>Medicago soleirolii</i>	787 bp
	45	<i>Medicago sauvagei</i>	845 bp
	46	<i>Medicago secundiflora</i>	842 bp
	47	<i>Medicago shepardii</i>	670 bp
	48	<i>Medicago turbinata</i>	842 bp
ribosomal protein S14 (rpS14)	1	<i>Medicago arabica</i>	963 bp
	2	<i>Medicago arborea</i>	960 bp
	3	<i>Medicago astroites</i>	954 bp
	4	<i>Medicago blanchiana</i>	960 bp

	5	<i>Medicago brachycarpa</i>	960 bp
	6	<i>Medicago carstiensis</i>	955 bp
	7	<i>Medicago ciliaris</i>	959 bp
	8	<i>Medicago constricta</i>	954 bp
	9	<i>Medicago coronata</i>	960 bp
	10	<i>Medicago crassipes</i>	954 bp
	11	<i>Medicago cretacea</i>	961 bp
	12	<i>Medicago disciformis</i>	965 bp
	13	<i>Medicago doliata</i>	964 bp
	14	<i>Medicago edgeworthii</i>	954 bp
	15	<i>Medicago fischeriana</i>	949 bp
	16	<i>Medicago granadensis</i>	960 bp
	17	<i>Medicago heyniana</i>	960 bp
	18	<i>Medicago huberi</i>	960 bp
	19	<i>Medicago intertexta</i>	959 bp
	20	<i>Medicago italica</i>	954 bp
	21	<i>Medicago laciniata</i>	968 bp
	22	<i>Medicago lanigera</i>	648 bp
	23	<i>Medicago laxispira</i>	960 bp
	24	<i>Medicago lesinsii</i>	952 bp
	25	<i>Medicago littoralis</i>	954 bp
	26	<i>Medicago lupulina</i>	961 bp
	27	<i>Medicago marina</i>	959 bp
	28	<i>Medicago minima</i>	958 bp
	29	<i>Medicago monspeliaca</i>	954 bp
	30	<i>Medicago phrygia</i>	954 bp
	31	<i>Medicago pironae</i>	960 bp
	32	<i>Medicago platycarpos</i>	970 bp
	33	<i>Medicago plicata</i>	948 bp
	34	<i>Medicago polymorpha</i>	960 bp
	35	<i>Medicago popovii</i>	957 bp
	36	<i>Medicago prostrata</i>	960 bp
	37	<i>Medicago radiata</i>	960 bp
	38	<i>Medicago rhodopea</i>	960 bp
	39	<i>Medicago rigidula</i>	954 bp
	40	<i>Medicago rotata</i>	960 bp
	41	<i>Medicago ruthenica</i>	953 bp
	42	<i>Medicago soleirolii</i>	954 bp
	43	<i>Medicago sauvagei</i>	968 bp
	44	<i>Medicago secundiflora</i>	643 bp
	45	<i>Medicago shepardii</i>	960 bp
	46	<i>Medicago turbinata</i>	964 bp
protein complex coatmer beta subunit (Bcop)	1	<i>Medicago arabica</i>	536 bp

2	<i>Medicago arborea</i>	542 bp
3	<i>Medicago astroites</i>	541 bp
4	<i>Medicago brachycarpa</i>	544 bp
5	<i>Medicago carstiensis</i>	542 bp
6	<i>Medicago ciliaris</i>	481 bp
7	<i>Medicago constricta</i>	532 bp
8	<i>Medicago coronata</i>	525 bp
9	<i>Medicago crassipes</i>	538 bp
10	<i>Medicago cretacea</i>	540 bp
11	<i>Medicago disciformis</i>	504 bp
12	<i>Medicago doliata</i>	533 bp
13	<i>Medicago edgeworthii</i>	544 bp
14	<i>Medicago fischeriana</i>	538 bp
15	<i>Medicago granadensis</i>	479 bp
16	<i>Medicago heyniana</i>	542 bp
17	<i>Medicago huberi</i>	542 bp
18	<i>Medicago intertexta</i>	481 bp
19	<i>Medicago italica</i>	534 bp
20	<i>Medicago laciniata</i>	560 bp
21	<i>Medicago lanigera</i>	535 bp
22	<i>Medicago laxispira</i>	497 bp
23	<i>Medicago lesinsii</i>	507 bp
24	<i>Medicago littoralis</i>	532 bp
25	<i>Medicago lupulina</i>	548 bp
26	<i>Medicago marina</i>	545 bp
27	<i>Medicago minima</i>	541 bp
28	<i>Medicago monspeliaca</i>	538 bp
29	<i>Medicago phrygia</i>	541 bp
30	<i>Medicago pironae</i>	546 bp
31	<i>Medicago platycarpus</i>	533 bp
32	<i>Medicago plicata</i>	556 bp
33	<i>Medicago polymorpha</i>	497 bp
34	<i>Medicago popovii</i>	535 bp
35	<i>Medicago prostrata</i>	540 bp
36	<i>Medicago radiata</i>	543 bp
37	<i>Medicago rhodopea</i>	543 bp
38	<i>Medicago rigidula</i>	502 bp
39	<i>Medicago rotata</i>	552 bp
40	<i>Medicago ruthenica</i>	535 bp
41	<i>Medicago soleirolii</i>	528 bp
42	<i>Medicago sauvagei</i>	583 bp
43	<i>Medicago secundiflora</i>	532 bp
44	<i>Medicago shepardii</i>	552 bp

	45	<i>Medicago turbinata</i>	533 bp
nork (nodulation receptor kinase)	1	<i>Medicago arabica</i>	1657 bp
	2	<i>Medicago arborea</i>	1191 bp
	3	<i>Medicago cancellata</i>	1753 bp
	4	<i>Medicago ciliaris</i>	1719 bp
	5	<i>Medicago constricta</i>	1782 bp
	6	<i>Medicago doliata</i>	1569 bp
	7	<i>Medicago granadensis</i>	1569 bp
	8	<i>Medicago heyneana</i>	1078 bp
	9	<i>Medicago laciniata</i>	1741 bp
	10	<i>Medicago marina</i>	1819 bp
	11	<i>Medicago minima</i>	1734 bp
	12	<i>Medicago monantha</i>	1538 bp
	13	<i>Medicago monspeliaca</i>	1611 bp
	14	<i>Medicago polymorpha</i>	1667 bp
	15	<i>Medicago radiata</i>	1684 bp
	16	<i>Medicago rigidula</i>	1720 bp
	17	<i>Medicago rigiduloides</i>	1735 bp
	18	<i>Medicago rotata</i>	696 bp
	19	<i>Medicago rugosa</i>	1635 bp
	20	<i>Medicago sauvagei</i>	1736 bp
	21	<i>Medicago saxatilis</i>	1015 bp
	22	<i>Medicago scutellata</i>	1183 bp
	23	<i>Medicago secundiflora</i>	1711 bp
	24	<i>Medicago shepardii</i>	1699 bp
ITS1	1	<i>Medicago arabica</i>	258 bp
	2	<i>Medicago ciliaris</i>	258 bp
	3	<i>Medicago constricta</i>	258 bp
	4	<i>Medicago disciformis</i>	259 bp
	5	<i>Medicago doliata</i>	250 bp
	6	<i>Medicago granadensis</i>	258 bp
	7	<i>Medicago heyneana</i>	258 bp
	8	<i>Medicago intertexta</i>	258 bp
	9	<i>Medicago laciniata</i>	263 bp
	10	<i>Medicago littoralis</i>	258 bp
	11	<i>Medicago minima</i>	259 bp
	12	<i>Medicago polymorpha</i>	262 bp
	13	<i>Medicago radiata</i>	258 bp
	14	<i>Medicago rigidula</i>	258 bp
	15	<i>Medicago rotata</i>	257 bp
	16	<i>Medicago rugosa</i>	258 bp
	17	<i>Medicago soleirolii</i>	252 bp
	18	<i>Medicago sphaerocarpa</i>	260 bp

	19	<i>Medicago syriaca</i>	258 bp
	20	<i>Medicago sauvagei</i>	263 bp
	21	<i>Medicago scutellata</i>	254 bp
	22	<i>Medicago secundiflora</i>	260 bp
	23	<i>Medicago shepardii</i>	258 bp
	24	<i>Medicago turbinata</i>	258 bp
	25	<i>Medicago muricoleptis</i>	258 bp
	26	<i>Medicago noeana</i>	256 bp
	27	<i>Medicago orbicularis</i>	258 bp
ITS2	1	<i>Medicago arabica</i>	236 bp
	2	<i>Medicago ciliaris</i>	236 bp
	3	<i>Medicago constricta</i>	236 bp
	4	<i>Medicago disciformis</i>	238 bp
	5	<i>Medicago doliata</i>	236 bp
	6	<i>Medicago granadensis</i>	236 bp
	7	<i>Medicago heyniana</i>	236 bp
	8	<i>Medicago intertexta</i>	236 bp
	9	<i>Medicago laciniata</i>	236 bp
	10	<i>Medicago littoralis</i>	236 bp
	11	<i>Medicago minima</i>	236 bp
	12	<i>Medicago polymorpha</i>	236 bp
	13	<i>Medicago radiata</i>	236 bp
	14	<i>Medicago rigidula</i>	236 bp
	15	<i>Medicago rotata</i>	236 bp
	16	<i>Medicago rugosa</i>	236 bp
	17	<i>Medicago soleirolii</i>	236 bp
	18	<i>Medicago sphaerocarpa</i>	236 bp
	19	<i>Medicago syriaca</i>	236 bp
	20	<i>Medicago sauvagei</i>	236 bp
	21	<i>Medicago scutellata</i>	239 bp
	22	<i>Medicago secundiflora</i>	236 bp
	23	<i>Medicago shepardii</i>	236 bp
	24	<i>Medicago turbinata</i>	237 bp
	25	<i>Medicago muricoleptis</i>	236 bp
	26	<i>Medicago noeana</i>	236 bp
	27	<i>Medicago orbicularis</i>	236 bp
5S ribosomal RNA	1	<i>Medicago arborea</i>	4777 bp
	2	<i>Medicago citrina</i>	742 bp
	3	<i>Medicago strasseri</i>	1282 bp
mitochondrial	1	<i>Medicago astroites</i>	954 bp
	2	<i>Medicago blanchiana</i>	960 bp
	3	<i>Medicago sativa</i>	2,173 pb
	4	<i>Medicago truncatula</i>	271,618 bp

	5	<i>Medicago brachycarpa</i>	960 bp
	6	<i>Medicago carstiensis</i>	955 bp
	7	<i>Medicago ciliaris</i>	959 bp
	8	<i>Medicago constricta</i>	954 bp
	9	<i>Medicago coronata</i>	960 bp
	10	<i>Medicago crassipes</i>	954 bp
	11	<i>Medicago cretacea</i>	961 bp
	12	<i>Medicago disciformis</i>	965 bp
	13	<i>Medicago doliata</i>	964 bp
	14	<i>Medicago edgeworthii</i>	954 bp
	15	<i>Medicago fischeriana</i>	949 bp
	16	<i>Medicago granadensis</i>	960 bp
	17	<i>Medicago heyniana</i>	960 bp
	18	<i>Medicago huberi</i>	960 bp
	19	<i>Medicago intertexta</i>	959 bp
	20	<i>Medicago italica</i>	954 bp
	21	<i>Medicago laciniata</i>	968 bp
	22	<i>Medicago lanigera</i>	648 bp
	23	<i>Medicago laxispira</i>	960 bp
	24	<i>Medicago lesinsii</i>	952 bp
	25	<i>Medicago lupulina</i>	961 bp
	26	<i>Medicago marina</i>	959 bp
	27	<i>Medicago minima</i>	958 bp
	28	<i>Medicago monspeliaca</i>	954 bp
	29	<i>Medicago phrygia</i>	954 bp
	30	<i>Medicago pironae</i>	960 bp
	31	<i>Medicago platycarpus</i>	970 bp
	32	<i>Medicago plicata</i>	948 bp
	33	<i>Medicago polymorpha</i>	960 bp
	34	<i>Medicago popovii</i>	957 bp
	35	<i>Medicago prostrata</i>	960 bp
	36	<i>Medicago radiata</i>	960 bp
	37	<i>Medicago rhodopea</i>	960 bp
	38	<i>Medicago rigidula</i>	954 bp
	39	<i>Medicago rotata</i>	960 bp
	40	<i>Medicago ruthenica</i>	953 bp
	41	<i>Medicago soleirolii</i>	954 bp
	42	<i>Medicago sauvagei</i>	968 bp
	43	<i>Medicago secundiflora</i>	643 bp
	44	<i>Medicago shepardii</i>	960 bp
	45	<i>Medicago turbinata</i>	964 bp
mRNA	1	<i>Medicago biflora</i>	468 bp
	2	<i>Medicago littoralis</i>	495 bp

	3	<i>Medicago monspeliaca</i>	674 bp
	4	<i>Medicago papillosa</i>	621 bp
	5	<i>Medicago platycarpus</i>	687 bp
	6	<i>Medicago polyceratia</i>	687 bp
	7	<i>Medicago radiata</i>	579 bp
	8	<i>Medicago rugosa</i>	684 bp
	9	<i>Medicago ruthenica</i>	687 bp
	10	<i>Medicago secundiflora</i>	687 bp
	11	<i>Medicago noeana</i>	651
	12	<i>Medicago orbicularis</i>	684 bp
chloroplast	1	<i>Medicago bonarotiana</i>	2221 bp
	2	<i>Medicago brachycarpa</i>	1548 bp
	3	<i>Medicago cancellata</i>	2455 bp
	4	<i>Medicago carstiensis</i>	2209 bp
	5	<i>Medicago ciliaris</i>	2443 bp
	6	<i>Medicago constricta</i>	2477 bp
	7	<i>Medicago cretacea</i>	1521 bp
	8	<i>Medicago daghestanica</i>	2466 bp
	9	<i>Medicago edgeworthii</i>	2359 bp
	10	<i>Medicago falcata</i>	2,471
	11	<i>Medicago fischeriana</i>	1929 bp
	12	<i>Medicago granadensis</i>	2436 bp
	13	<i>Medicago heyniana</i>	2497 bp
	14	<i>Medicago hybrida</i>	125208 bp
	15	<i>Medicago hypogaea</i>	2412 bp
	16	<i>Medicago intertexta</i>	2465 bp
	17	<i>Medicago italica</i>	2487 bp
	18	<i>Medicago laciniata</i>	1521 bp
	19	<i>Medicago lanigera</i>	2456 bp
	20	<i>Medicago littoralis</i>	2464 bp
	21	<i>Medicago lunata</i>	553 bp
	22	<i>Medicago lupulina</i>	2386 bp
	23	<i>Medicago marina</i>	2455 bp
	24	<i>Medicago medicaginoides</i>	2488 bp
	25	<i>Medicago minima</i>	2436 bp
	26	<i>Medicago monantha</i>	2462 bp
	27	<i>Medicago monspeliaca</i>	2482 bp
	28	<i>Medicago murex</i>	533 bp
	29	<i>Medicago papillosa</i>	125203 bp
	30	<i>Medicago pironae</i>	2420 bp
	31	<i>Medicago platycarpus</i>	2501 bp
	32	<i>Medicago plicata</i>	623 bp
	33	<i>Medicago polyceratia</i>	2464 bp

	34	<i>Medicago polymorpha</i>	2473 bp
	35	<i>Medicago prostrata</i>	2449 bp
	36	<i>Medicago radiata</i>	2453 bp
	37	<i>Medicago rhodopea</i>	2435 bp
	38	<i>Medicago rigidula</i>	2462 bp
	39	<i>Medicago rostrata</i>	598 bp
	40	<i>Medicago ruthenica</i>	2492 bp
	41	<i>Medicago soleirolii</i>	2448 bp
	42	<i>Medicago sativa</i> subsp. <i>x</i> <i>tunetana</i>	685 bp
	43	<i>Medicago saxatilis</i>	2444 bp
	44	<i>Medicago scutellata</i>	2466 bp
	45	<i>Medicago secundiflora</i>	2453 bp
	46	<i>Medicago shepardii</i>	2436 bp
	47	<i>Medicago truncatula</i> f. <i>tricycla</i>	123355 bp
	48	<i>Medicago turbinata</i>	2470 bp
	49	<i>Medicago pamphylica</i>	566 bp
	50	<i>Medicago noeana</i>	2,255 bp
	51	<i>Medicago sativa</i>	4,113 bp
	52	<i>Medicago truncatula</i>	123,355 bp
	53	<i>Medicago orbicularis</i>	2,354 bp
internal transcribed spacer 1(68)	1	<i>Medicago arabica</i>	237 bp
	2	<i>Medicago arborea</i>	258 bp
	3	<i>Medicago archiducis-</i> <i>nicolai</i>	594 bp
	4	<i>Medicago astroites</i>	697 bp
	5	<i>Medicago aurantiaca</i>	701 bp
	6	<i>Medicago bonarotiana</i>	258 bp
	7	<i>Medicago brachycarpa</i>	709 bp
	8	<i>Medicago cancellata</i>	258 bp
	9	<i>Medicago carstiensis</i>	258 bp
	10	<i>Medicago ciliaris</i>	237 bp
	11	<i>Medicago constricta</i>	237 bp
	12	<i>Medicago coronata</i>	259 bp
	13	<i>Medicago crassipes</i>	694 bp
	14	<i>Medicago cretacea</i>	257 bp
	15	<i>Medicago daghestanica</i>	616 bp
	16	<i>Medicago disciformis</i>	238 bp
	17	<i>Medicago doliata</i>	230 bp
	18	<i>Medicago edgeworthii</i>	641 bp
	19	<i>Medicago fischeriana</i>	261 bp
	20	<i>Medicago granadensis</i>	237 bp

21	<i>Medicago heyniana</i>	237 bp
22	<i>Medicago huberi</i>	258 bp
23	<i>Medicago hybrida</i>	258 bp
24	<i>Medicago hypogaea</i>	237 bp
25	<i>Medicago intertexta</i>	237 bp
26	<i>Medicago italica</i>	237 bp
27	<i>Medicago laciniata</i>	621 bp
28	<i>Medicago lanigera</i>	641 bp
29	<i>Medicago laxispira</i>	237 bp
30	<i>Medicago lesinsii</i>	237 bp
31	<i>Medicago lunata</i>	709 bp
32	<i>Medicago lupulina</i>	732 bp
33	<i>Medicago marina</i>	616 bp
34	<i>Medicago medicaginoides</i>	642 bp
35	<i>Medicago minima</i>	617 bp
36	<i>Medicago monantha</i>	641 bp
37	<i>Medicago monspeliaca</i>	642 bp
38	<i>Medicago papillosa</i>	258 bp
39	<i>Medicago persica</i>	704 bp
40	<i>Medicago pironae</i>	616 bp
41	<i>Medicago platycarpos</i>	642 bp
42	<i>Medicago plicata</i>	709 bp
43	<i>Medicago polyceratia</i>	643 bp
44	<i>Medicago polymorpha</i>	736 bp
45	<i>Medicago popovii</i>	641 bp
46	<i>Medicago prostrata</i>	258 bp
47	<i>Medicago radiata</i>	641 bp
48	<i>Medicago rhodopea</i>	236 bp
49	<i>Medicago rigidula</i>	237 bp
50	<i>Medicago rigiduloides</i>	258 bp
51	<i>Medicago rostrata</i>	709 bp
52	<i>Medicago rotata</i>	237 bp
53	<i>Medicago rugosa</i>	237 bp
54	<i>Medicago ruthenica</i>	732 bp
55	<i>Medicago sinskiae</i>	237 bp
56	<i>Medicago strasseri</i>	237 bp
57	<i>Medicago suffruticosa</i> <i>subsp. leiocarpa</i>	237 bp
58	<i>Medicago syriaca</i>	237 bp
59	<i>Medicago sauvagei</i>	242 bp
60	<i>Medicago saxatilis</i>	258 bp
61	<i>Medicago scutellata</i>	619 bp
62	<i>Medicago shepardii</i>	237 bp

	63	<i>Medicago turbinata</i>	237 bp
	64	<i>Medicago orthoceras</i>	699 bp
	65	<i>Medicago pamphylica</i>	712 bp
	66	<i>Medicago muricoleptis</i>	237 bp
	67	<i>Medicago noeana</i>	235 bp
	68	<i>Medicago orbicularis</i>	616 bp
internal transcribed spacer	1	<i>Medicago arabica</i>	214 bp
	2	<i>Medicago arborea</i>	236 bp
	3	<i>Medicago archiducis-nicolai</i>	594 bp
	4	<i>Medicago astroites</i>	697 bp
	5	<i>Medicago aurantiaca</i>	701 bp
	6	<i>Medicago bonarotiana</i>	236 bp
	7	<i>Medicago brachycarpa</i>	709 bp
	8	<i>Medicago cancellata</i>	236 bp
	9	<i>Medicago carstiensis</i>	236 bp
	10	<i>Medicago ciliaris</i>	214 bp
	11	<i>Medicago constricta</i>	214 bp
	12	<i>Medicago coronata</i>	236 bp
	13	<i>Medicago crassipes</i>	694 bp
	14	<i>Medicago cretacea</i>	236 bp
	15	<i>Medicago daghestanica</i>	616 bp
	16	<i>Medicago disciformis</i>	216 bp
	17	<i>Medicago doliata</i>	214 bp
	18	<i>Medicago edgeworthii</i>	641 bp
	19	<i>Medicago falcata</i>	445 bp
	20	<i>Medicago fischeriana</i>	214 bp
	21	<i>Medicago granadensis</i>	214 bp
	22	<i>Medicago heyniana</i>	214 bp
	23	<i>Medicago huberi</i>	236 bp
	24	<i>Medicago hybrida</i>	236 bp
	25	<i>Medicago hypogaea</i>	213 bp
	26	<i>Medicago intertexta</i>	214 bp
	27	<i>Medicago italica</i>	214 bp
	28	<i>Medicago laciniata</i>	621 bp
	29	<i>Medicago lanigera</i>	641 bp
	30	<i>Medicago laxispira</i>	214 bp
	31	<i>Medicago lesinsii</i>	214 bp
	32	<i>Medicago lunata</i>	709 bp
	33	<i>Medicago lupulina</i>	732 bp
	34	<i>Medicago marina</i>	616 bp
	35	<i>Medicago medicaginoides</i>	642 bp
	36	<i>Medicago minima</i>	617 bp

	37	<i>Medicago monantha</i>	641 bp
	38	<i>Medicago monspeliaca</i>	642 bp
	39	<i>Medicago papillosa</i>	236 bp
	40	<i>Medicago persica</i>	704 bp
	41	<i>Medicago pironae</i>	236 bp
	42	<i>Medicago platycarpus</i>	642 bp
	43	<i>Medicago plicata</i>	709 bp
	44	<i>Medicago polyceratia</i>	643 bp
	45	<i>Medicago polymorpha</i>	736 bp
	46	<i>Medicago popovii</i>	641 bp
	47	<i>Medicago prostrata</i>	236 bp
	48	<i>Medicago radiata</i>	641 bp
	49	<i>Medicago rhodopea</i>	615 bp
	50	<i>Medicago rigidula</i>	214 bp
	51	<i>Medicago rigiduloides</i>	236 bp
	52	<i>Medicago rostrata</i>	709 bp
	53	<i>Medicago rotata</i>	214 bp
	54	<i>Medicago rugosa</i>	214 bp
	55	<i>Medicago ruthenica</i>	732 bp
	56	<i>Medicago schischkinii</i>	445 bp
	57	<i>Medicago sinskiae</i>	214 bp
	58	<i>Medicago strasseri</i>	214 bp
	59	<i>Medicago suffruticosa</i> <i>subsp. leiocarpa</i>	214 bp
	60	<i>Medicago syriaca</i>	214 bp
	61	<i>Medicago sauvagei</i>	214 bp
	62	<i>Medicago saxatilis</i>	236 bp
	63	<i>Medicago scutellata</i>	619 bp
	64	<i>Medicago shepardii</i>	214 bp
	65	<i>Medicago turbinata</i>	214 bp
	66	<i>Medicago orthoceras</i>	699 bp
	67	<i>Medicago pamphylica</i>	712 bp
	68	<i>Medicago muricoleptis</i>	214 bp
	69	<i>Medicago noeana</i>	214 bp
	70	<i>Medicago orbicularis</i>	616 bp
28S ribosomal RNA (12)	1	<i>Medicago archiducis-nicolai</i>	594 bp
	2	<i>Medicago brachycarpa</i>	709 bp
	3	<i>Medicago edgeworthii</i>	594 bp
	4	<i>Medicago falcata</i>	445 bp
	5	<i>Medicago lunata</i>	709 bp
	6	<i>Medicago lupulina</i>	732 bp
	7	<i>Medicago plicata</i>	709 bp

	8	<i>Medicago polymorpha</i>	736 bp
	9	<i>Medicago rostrata</i>	709 bp
	10	<i>Medicago ruthenica</i>	732 bp
	11	<i>Medicago schischkinii</i>	445 bp
	12	<i>Medicago pamphylica</i>	712 bp
5.8S (16)	1	<i>Medicago arabica</i>	401 bp
	2	<i>Medicago aurantiaca</i>	258 bp
	3	<i>Medicago brachycarpa</i>	258 bp
	4	<i>Medicago fischeriana</i>	261 bp
	5	<i>Medicago huberi</i>	258 bp
	6	<i>Medicago lanigera</i>	258 bp
	7	<i>Medicago medicaginoïdes</i>	259 bp
	8	<i>Medicago monantha</i>	258 bp
	9	<i>Medicago monspeliaca</i>	258 bp
	10	<i>Medicago papillosa</i>	258 bp
	11	<i>Medicago platycarpus</i>	258 bp
	12	<i>Medicago polyceratia</i>	256 bp
	13	<i>Medicago popovii</i>	258 bp
	14	<i>Medicago prostrata</i>	258 bp
	15	<i>Medicago saxatilis</i>	258 bp
	16	<i>Medicago orthoceras</i>	259 bp
26S	1	<i>Medicago arabica</i>	401 bp
	2	<i>Medicago edgeworthii</i>	719 bp
	3	<i>Medicago lanigera</i>	720 bp
	4	<i>Medicago lupulina</i>	3366 bp
	5	<i>Medicago medicaginoïdes</i>	735 bp
	6	<i>Medicago minima</i>	585 bp
	7	<i>Medicago monantha</i>	710 bp
	8	<i>Medicago monspeliaca</i>	709 bp
	9	<i>Medicago platycarpus</i>	748 bp
	10	<i>Medicago polyceratia</i>	715 bp
	11	<i>Medicago popovii</i>	743 bp
	12	<i>Medicago radiata</i>	720 bp
	13	<i>Medicago ruthenica</i>	739 bp

**Tableau 2** : le nombre des genes chloroplastiques et la longueur du gene.

N°	Espece	Nombre de sequences nucleotidiques	Gene de chloroplast	Longueur du gene de chloroplaste (pb)
1	<i>Medicago arabica</i>	32	/	/
2	<i>Medicago arborea</i>	15	/	/
3	<i>Medicago archiducis-nicolai</i>	3	/	/
4	<i>Medicago astroites</i>	4	/	/
5	<i>Medicago aurantiaca</i>	4	/	/
6	<i>Medicago biflora</i>	12	/	/
7	<i>Medicago blanchiana</i>	5	/	/
8	<i>Medicago bonarotiana</i>	5	1	2221
9	<i>Medicago brachycarpa</i>	14	2	1548
10	<i>Medicago cancellata</i>	9	1	2455
11	<i>Medicago carstiensis</i>	13	1	2209
12	<i>Medicago ciliaris</i>	16	1	2443
13	<i>Medicago citrina</i>	7	/	/
14	<i>Medicago constricta</i>	12	1	2477
15	<i>Medicago coronata</i>	6	/	/
16	<i>Medicago crassipes</i>	4	/	/
17	<i>Medicago cretacea</i>	9	3	1521
18	<i>Medicago daghestanica</i>	5	1	2466
19	<i>Medicago disciformis</i>	8	/	/
20	<i>Medicago doliata</i>	10	/	/
21	<i>Medicago edgeworthii</i>	16	1	2359
22	<i>Medicago falcata (sickle medic)</i>	17643	120	
23	<i>Medicago falcata subsp. falcata</i>	9	6	699
24	<i>Medicago fischeriana</i>	9	1	1929
25	<i>Medicago granadensis</i>	12	1	2436
26	<i>Medicago heyniana</i>	12	1	2497
27	<i>Medicago huberi</i>	6	/	/
28	<i>Medicago hybrida</i>	9	3	125208
29	<i>Medicago hypogaea</i>	4	1	2412
30	<i>Medicago intertexta</i>	10	1	2465
31	<i>Medicago italica</i>	8	2	2487
32	<i>Medicago laciniata</i>	16	1	1521
33	<i>Medicago lanigera</i>	15	1	2456
34	<i>Medicago laxispira</i>	5	/	/
35	<i>Medicago lesinsii</i>	5	/	/
36	<i>Medicago littoralis (shore medic)</i>	24	2	2464
37	<i>Medicago lunata</i>	2	1	553
38	<i>Medicago lupulina</i>	86	39	2386
39	<i>Medicago marina</i>	21	1	2455
40	<i>Medicago medicaginoïdes</i>	10	1	2488
41	<i>Medicago minima</i>	34	19	2436 bp

42	<i>Medicago monantha</i>	17	3	2462 bp
43	<i>Medicago monspeliaca</i>	17	1	2482 bp
44	<i>Medicago murex</i>	25	1	533
45	<i>Medicago muricoleptis</i>	8	/	/
46	<i>Medicago noeana</i>	17	1	2,255 bp
47	<i>Medicago orbicularis</i>	28	13	2,354 bp
48	<i>Medicago orthoceras</i>	6	/	/
49	<i>Medicago pamphylica</i>	2	1	566
50	<i>Medicago papillosa</i>	16	5	125203 bp
51	<i>Medicago persica</i>	1	/	/
52	<i>Medicago phrygia</i>	3	/	/
53	<i>Medicago pironae</i>	13	3	2420 bp
54	<i>Medicago platycarpus</i>	14	1	2501 bp
55	<i>Medicago plicata</i>	5	1	623 bp
56	<i>Medicago polyceratia</i>	15	1	2464 bp
57	<i>Medicago polymorpha</i>	47	22	2473 bp
58	<i>Medicago popovii</i>	9	/	/
59	<i>Medicago praecox</i>	12	/	/
60	<i>Medicago prostrata</i>	37	6	2449 bp
61	<i>Medicago radiata</i>	18	1	2453 bp
62	<i>Medicago rhodopea</i>	7	1	2435 bp
63	<i>Medicago rigidula</i>	21	1	2462 bp
64	<i>Medicago rigiduloides</i>	14	/	/
65	<i>Medicago rostrata</i>	2	1	598
66	<i>Medicago rotata</i>	10	/	/
67	<i>Medicago rugosa</i>	10	/	/
68	<i>Medicago ruthenica</i>	104	43	2492 bp
69	<i>Medicago sativa</i> (alfalfa)	210073	/	/
70	<i>Medicago sauvagei</i>	15	/	/
71	<i>Medicago saxatilis</i>	9	3	2444 bp
72	<i>Medicago schischkinii</i>	1	/	/
73	<i>Medicago scutellata</i>	15	7	2466 bp
74	<i>Medicago secundiflora</i>	13	1	2453 bp
75	<i>Medicago shepardii</i>	12	1	2436 bp
76	<i>Medicago sinskiae</i>	2	/	/
77	<i>Medicago soleirolii</i>	8	1	2448 bp
78	<i>Medicago sphaerocarpa</i>	3	/	/
79	<i>Medicago strasseri</i>	10	/	/
80	<i>Medicago suffruticosa</i>	10	/	/
81	<i>Medicago syriaca</i>	5	/	/
82	<i>Medicago tenoreana</i>	17	/	/
83	<i>Medicago tornata</i>	81	/	/
84	<i>Medicago truncatula</i> (barrel medic)	104852	/	/
85	<i>Medicago turbinata</i>	10	1	2470 bp

Figure 1 : Exemple d'une séquence au Format GenBank.

**Medicago sativa trnS-trnG intergenic spacer, partial sequence; chloroplast**

GenBank: HQ199053.1

[FASTA](#) [Graphics](#) [PopSet](#)Go to: 

LOCUS HQ199053 307 bp DNA linear PLN 15-JAN-2013  
DEFINITION Medicago sativa trnS-trnG intergenic spacer, partial sequence;  
chloroplast.  
ACCESSION HQ199053  
VERSION HQ199053.1 GI:343887191  
KEYWORDS .  
SOURCE chloroplast Medicago sativa  
ORGANISM [Medicago sativa](#)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; Gunneridae;  
Pentapetalae; rosids; fabids; Fabales; Fabaceae; Papilionoideae;  
Trifolieae; Medicago.  
REFERENCE 1 (bases 1 to 307)  
AUTHORS Havananda,T., Brummer,E.C. and Doyle,J.J.  
TITLE Complex patterns of autopolyploid evolution in alfalfa and allies  
(Medicago sativa; Leguminosae)  
JOURNAL Am. J. Bot. 98 (10), 1633-1646 (2011)  
PUBMED [21965136](#)  
REFERENCE 2 (bases 1 to 307)  
AUTHORS Havananda,T., Brummer,E.C. and Doyle,J.J.  
TITLE Autopolyploid origins and evolutionary history of the Medicago  
sativa (Fabaceae) species complex inferred from chloroplast DNA  
sequences  
JOURNAL Unpublished  
REFERENCE 3 (bases 1 to 307)  
AUTHORS Havananda,T., Brummer,E.C. and Doyle,J.J.  
TITLE Direct Submission  
JOURNAL Submitted (27-AUG-2010) Plant Biology, Cornell University, 412 Mann  
Library Building, Ithaca, NY 14853, USA  
FEATURES Location/Qualifiers  
source 1..307  
/organism="Medicago sativa"  
/organelle="plastid:chloroplast"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:3879"  
/note="genotype: ABI 408"  
misc\_feature <1..>307  
/note="trnS-trnG intergenic spacer"  
ORIGIN  
1 agggaatccg tcttaattat gcaatagtg catcattta cgataattag agtgttgacg  
61 tgtttcgtaa aggggcccc taatgaaatt tagaaaagg ggcgaatctt attttggat  
121 tttcttttg ctgaaggat aaaaaataaa atcaataatt ctctctttat ttatagataa  
181 taaatcgaaa aaaaaagac atataagata tatcaaatga ctatcaatca atatcaata  
241 agatataaag aaggcttttt tcgagcccta cttttgttc tttttgtta tgcgatgtat  
301 cataagc  
//

**Figure 2: Exemple d'une séquence au Format FASTA.****Medicago minima voucher K P Steele 83 (ASU) tRNA-Lys (trnK) gene, intron; and maturase K (matK) gene, complete cds; chloroplast**

GenBank: HM159571.1

[GenBank Graphics](#)

```
>gi|300390523|gb|HM159571.1| Medicago minima voucher K P Steele 83 (ASU) tRNA-Lys (trnK) gene,
intron; and maturase K (matK) gene, complete cds; chloroplast
```

```
ACTCCATTTTTTACACATTTCTATGAACTAATTGGTTCATCCATACCATCGGTAGGGTTTGTAAAGACCAC
GACTGATCCAGAAAGGAATGAATGGAAAAGCAGCATGTCGTATCAATGGTGAATTTCTAAACATTTTTCT
TGGACCCGGCCAAATTTTCGTTTGAATTGTTGGCTCGGAACAAATGAATTCAGTTGGGTTTAATTAATA
AAAGGATAGAGCTTAGATGCTCCAATTATAGGGAAACAAAAGCAACGAGCTTTCATTTTTTATTTGAAT
GATTACCACATCTAATTTCTACGTTAAAAAAGGATTAGTCTTGTCTGTGGAAAAAGTTTGAATGGATTAAG
GATTTTTGTTATGAGTCCATAATTATTAGCTATTCTCCATTATGTTATGGGGTAGCAATAAATGTGTAGAA
GAAAGAGTATATTGATAAAGATCTTTTTTCCAAAATCAAAGAGCGATTGGTCCAAAAATAAAGGATT
TCTAAGTAGCTTGTGTCTAGAACAAATTAGATGGGAACAGGCGAGCGAGATAGTCCATTAATGGGTTT
TAACCTGTTTTCTAGGTATCTATTCATATTTGTTTTTTTTTCAATTCTAATATATAGAATACCTTGTTT
TGACTGTATCGCACTATGTATCATTGATAATCCAATGAATCTCTGATCCTTTGACCAAATAGAATTTCT
AAAATGAAGAAGAAATATCAAGTATATTTAGAACAAAGCTAGATCTCGCCAAAAGGACTTCTTATACCCAC
TTCTTTTTTCGGGAGTATATTTATGGACTTGCTTATACTCAAATTTGAAGAGATCCATTTTTGTGGAAAA
TATGGGTTCTGACAGTAAATATAGTTTACTAATTGTAAAACGTTAATTACTCGAATGTATCAACAGAAT
CATTTAAGCATTTTGGCTAATGATTGTAAACAAAATCCATTTTGGGGGTATAATGAGAATTTAGATTCTC
AAATAATATCAGAGGGTTTTGCCATCGTCGTGAAATTCATTTTTCTAGAATTAAGCTCTTCCTTAGA
GGAGGCAGAAATCATAAATCTTATAAAAATTTGAGATCAATTCATTCATTTTTCCCTTTTTGGAGGAT
AAATTTACACATTTAAATATGTGTGATACGATATACGAATACCATATCCTATCCATCTGGAATCTTAGTTC
AAATCCTTCGATACTGGGTGAAAGATGCCCTTTTTTTTCAATTTATTACGATTGTTTCTTTATAATTTAG
TAATTTGTAAGTAATCGGAGATTATTTCTTGTTCTCTATAATTTGTATGTATGTGAATATGAATCTATC
TTCCTTTTTCTACGTAATAAATCCTCTCATTACGATTAATACTTTTAGCGTTTTTTTTGAACGAATTT
TTTTTTATGCAAAAAGAGAATCTTTGTAGAAGTTTTTGCTAAGGATTTTTCGTATACTTTAAAATCTT
CAAGGATCCTCTCATTATGTTAGATATCAAGGAAAATGATTCTGGCTTCAAAGAATTCGCCTTTT
TTGATGAATAAATGAAACATTATTTTATTCAATTTATGGCAAGGTTTTTTTTATGTTTGGTCTGAACCAA
GAACTATCAATATAAACGAATTATCCGAACATTCATTTTTCAGCTTTTAGGCTATTTTTTAAATGTGCGGGT
AAATCATTTCAGTGGTACGGAGTCAAATGTTGCAAAATACATTTTAAATCGAAATTTTAGTAAAAAGTT
GATATAATAGTTCCAATTATCCTCTGGTTAGATATTTGGCTAAAGCGAAATTTGTAATGTATTAGGGC
ATCCTATTAGTAAGCCGGTCTGGGCCGATTATCCGATTTTGATATTATTGAACGATTTTTCGAATATG
CAGCAATCTTTGTTATTATTACAATGGATCCTCAAAAAAAAAAAGTTTGTATCAAATAAAATATATACTT
CGGCTTTCTTGATTAATAAATTTGGCTTGTAACACAAAACACTACTGTACGCGCTTTTTTGAACAGATCAG
GTTCAGAAGAATTATTGGAAGGATTTTACAGAGGAAGACCCGATTCTGCCAATAATTGGAATAATGAA
TGCAATTCATTTGATTTTATAGATCCAGATTTCTTCTACTTCGCAAAGGATAAATACAAATCGGGTTTGG
TATTTGGATATCTTTTCAGCAACGATCTGTTCAATGATGAATGATTGGTTATGTTATGATACTTAAATA
TTCTAGATATCGAATTTGATTTTATATATATATAAAAAGATTATGAAATGTTTCATGTAGTTAAGAGTTGA
GTTTCAAGATTCCTCTCTGAAAAGTAACTCAGGTTTAGATGTATACATAGGGAAAGCCGTGTGCAATG
AAAAATGCATGCACGGCTTGGGGAGGATTTTTTATTCTTTTATTTGATTAAGAA
```