



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

**UNIVERSITE SAAD DAHLEB BLIDA 1**

Faculté des Sciences de la Nature et de la Vie

Département de Biologie Physiologie Cellulaire

Laboratoire des Biotechnologies, Environnement et Santé

Projet de fin d'études en vue de l'obtention du diplôme

de Master en science de la nature et de la vie

Filière Sciences Biologiques

**Option** : Génétique et physiologie

**Thème :**

**La position génétique de *Meriones Shawi* dans l'arbre phylogéographique de genre *Meriones*.**

**Réalisé par :** M<sup>lle</sup> AOUABDIA Saida

M<sup>lle</sup> SEMCHA Abir

**Soutenu Le** : 27/09/2020.

**Devant le jury composé de :**

Nom et Prénom	Grade	Lieu	Qualité
M <sup>me</sup> CHELLAL N.	MCB	USDB	*Présidente
M <sup>me</sup> AMOKRANE A.	MAA	USDB	*Examinatrice
M <sup>me</sup> DEROUICHE L.	MCA	ESSAIA	*Promotrice
M <sup>me</sup> EDDAIKRA A.	MCB	USDB	*Co-promotrice

Promotion 2019-2020.

# Remerciements

Tout d'abord, on tient à remercier Dieu, le Généreux, le Tout Puissant qui nous donne la force et le courage, la volonté et les moyens nécessaires pour réaliser ce modeste travail.

Notre première reconnaissance va à notre promotrice M<sup>me</sup> DEROUICHE d'avoir accepté de diriger ce travail, ses apports et ses conseils toujours judicieux et ses encouragements, nous lui disons, Merci.

Nos remerciements aussi à notre Co-promotrice M<sup>me</sup> EDDAIKRA, ainsi que les membres du jury M<sup>me</sup> CHELLAL et M<sup>me</sup> AMOKRANE qui ont accepté d'évaluer et de juger notre travail.

Nos gratitudes vont également pour nos enseignants de master qui nous ont donné l'envie de poursuivre dans cette voie.

Nos remerciements s'adressent aussi aux enseignants et tous les personnels du département de science de la nature et de la vie qui nous ont aidé tout au long de notre cursus universitaire.

Enfin, nos sincères remerciements vont à tous ceux qui ont contribué de près ou de loin à la réalisation de ce mémoire, pour leur soutien et leurs encouragements

Merci à tous.

# *Dédicace*

*Je dédie ce modeste travail à mes parents, mes estime pour eux sont immenses, je vous remercie pour la confiance et l'amour dont vous m'avez toujours entouré, merci de m'avoir aidé à réussir.*

*Que dieu vous préserve. Longue vie heureuse.*

*A tous mes amis.*

*Je vous dédie ce travail et vous souhaite un avenir à la hauteur de vos ambitions.*

*Que notre amitié dure.*

*A Toute ma famille et la famille de mon binôme  
Abir.*

*Tous ceux que j'aime, qui m'aiment et me comblent de conseils.*

*A tous ceux qui, un jour, ont pensé à moi, les plus beaux mots ne sauraient exprimer ma redevance.*

*Saida.*

*Je dédie ce travail*

*A celle que j'ai la plus chère au monde ma mère.*

*Tu m'as donné la vie, la tendresse et le courage  
pour réussir.*

*Une dédicace vraiment spéciale pour une personne  
très chère dans mon cœur :*

*Mon père malgré ton absence mais tu es toujours  
présent dans mon cœur repose en paix  
Incha'allah le paradis sera ton destin.*

*Mes frères Amine et Yahia, ma sœur Imen  
et mon fiancé.*

*Mes cousines Sarah et Amina.*

*Mon binôme Saida et sa famille.*

*Mes amies Achouak, Noussaiba et Abir.*

*Tous les étudiants de l'option génétique promo  
2020 et à tous ceux qui m'aiment*

***Abir.***

# Abréviation

- **ADN** : Acide **Dé**soxyribo**Nuclé**ique.
- **ADNmt** : Acide **Dé**soxyribo**Nuclé**ique **mitochondrial**.
- **ARN** : Acide **Ri**bo**Nuclé**ique.
- **ARNr** : Acides **Ri**bo**Nuclé**iques **ribosomique**.
- **ARNt** : Acides **Ri**bo**Nuclé**iques de **transfert**.
- **ATDB** : Arabidopsis Thaliana **DataBase**
- **Avpr2** : Arginine vaso-**pressin** receptor **2**
- **BEAST** : Bayesian Evolutionary Analysis Sampling Trees
- **BIR** : Break Induced **Re**plication
- **Cytb** : Cytochrome **b**.
- **DAMBE** : Data Analysis in **Molecular Biology** and **Evolution**.
- **DDBJ** : **DNA Data Bank** of **Japan**.
- **D-loop** : Displacement **loop**
- **DnaSP** : **DNA Sequence Polymorphism**.
- **ECDC** : Escherichia coli **Database Collection**.
- **EMBL** : European **Molecular Biology Laboratory**
- **EMBOSS** : European **Molecular Biology Open Software Suite**.
- **GPCR** : **G Protein-Coupled Receptors**
- **Hd** : **Haplotype Diversity**
- **MEGA** : **Molecular Evolutionary Genetics Analysis** using **Maximum Likelihood**.
- **NBRF** : National **Biomedical Research Foundation**.
- **NCBI** : National **Center for Biotechnology Information**.
- **PAML** : **Phylogenetic Analysis by Maximum Likelihood**
- **PAUP** : **Phylogenetic Analysis Using Parsimony** \*and other methods.
- **PCR** : **Polymerase Chain Reaction** / Réaction de polymérisation en chaîne.
- **PDB** : **Protein Data Bank**.
- **PGD** : **Plant Genome Database**.
- **PGD** : **Plant Genome Database**.
- **PHYLIP** : **PHYLogeny Inference Package**.

- **PIR** : Protein Information Resource.
- **PRAP** : Pipeline Resistome Analyse Pan
- **RDP** : Rapid-Onset Dystonia
- **SGD** : Saccharomyces Genome Database.
- **UPGMA** : Unweighted Pair-Group Method using Arithmetic Averages.

## Liste des figures

N° de figure	Titre	Page
01	<i>Meriones shawi</i> (Ahmim, 2019).	02
02	La répartition géographique historique et actuelle de <i>Merione shawi</i> en Algérie (Ahmim, 2019).	05
03	Les trois domaines du vivant (Woese <i>et al.</i> , 1990).	07
04	Schéma illustrant les différentes notions de la représentation d'un arbre phylogénétique (Taguan, 2015)	09
05	Schéma illustrant les notions des différents types de groupes taxonomiques.	10
06	Schéma représentant les différents types de dendrogrammes.	11
07	Les étapes initiales pour récupérer les séquences de la GenBank.	19
08	Les vingt-quatre espèces de genre « Meriones » obtenues dans la GenBank ; ces espèces sont indiquées par des flèches en orange.	20
09	Exemple des séquences nucléotidiques trouvées pour l'espèce <i>Meriones shawi</i> .	20
10	Exemple d'une séquence nucléotidique au format FASTA de « <i>Meriones shawi</i> ».	21
11	Exemple d'un fichier sous format FASTA de « <i>Meriones shawi isis</i> ».	22
12	La page principale de logiciel FaBox.	22
13	La conversion des séquences sous forme Fasta par l'utilisation de Fabox.	23
14	L'obtention des séquences en haplotypes par l'utilisation de Fabox.	23
15	Alignement trimmer (format fasta).	24
16	L'alignement de la séquence la plus courte par « Alignement trimmer ».	24
17	Menu principal de logiciel MEGA version 5.	25
18	Exemple des 3 séquences ajoutées dans le logiciel MEGA.	26
19	L'alignement multiple des séquences du genre « Meriones » en utilisant le programme ClustalW.	26
20	La construction de l'arbre phylogénétique en utilisant la méthode UPGMA.	27
21	Les paramètres utilisés pour la reconstruction phylogénétique.	28

22	Menu principal de logiciel DnaSP 5.10.	29
23	Les paramètres « Haplotype Data File » et « Roehl Data File » ont été utilisés pour avoir les haplotypes, et enregistrer le fichier sous format *.rdf.	29
24	Menu principal de logiciel DAMBE.	31
25	Menu principal de logiciel NETWORK 5.0.	31
26	L'ouverture de fichier sous format rdf produit par logiciel DnaSp.	32
27	L'enregistrement de notre fichier sous format *.out pour la reconstruction de réseau de NETWORK.	32
28	Structure de l'ADN mitochondrial et la position de Cytb ( <b>Rustin et al., 1995</b> ).	35
29	La position de gène Avpr2 sur le chromosome X.	36
30	Les séquences de Cytb ont été convertir sous forme Fasta à partir de bloc note grâce à FaBox.	38
31	Arbre phylogénétique reconstruit par méthodes la UPGMA basées sur 78 séquences de nucléotides de Cytochrome b de genre Meriones en utilisant MEGA5.	40
32	Arbre phylogénétique des séquences Avpr2 du genre Meriones en utilisant la méthode UPGMA par logiciel MEGA5.	42
33	<i>Meriones crassus</i> ( <b>Ahmim, 2019</b> ).	44
34	<i>Meriones libycus</i> ( <b>Ahmim, 2019</b> ).	44
35	Résultat d'alignement des séquences Cytb du genre Meriones en utilisant le logiciel FaBox.	46
36	Les haplotypes des séquences Cytb du genre Meriones en utilisant le logiciel DnaSP.	48
37	Les haplotypes des séquences Avpr2 du genre Meriones en utilisant le logiciel DnaSP.	49
38 (A)	Variations des fréquences GC de Cytb produites à partir de DAMBE.	51
38 (B)	Variations des fréquences d'AT de Cytb produites à partir de DAMBE.	51
39	Variations des fréquences GC (C) et d'AT (D) d'Avpr2 produites à partir de DAMBE.	52
40	Comparaison des fréquences entre Cytb et Avpr2 réalisés par le logiciel DAMBE.	53



<b>41</b>	Réseau haplotypique réalisé par le logiciel NETWORK à partir de 78 séquences de Cytb de genre Meriones.	54
<b>42</b>	La croissance des haplotypes des séquences Avpr2 du genre Meriones en utilisant le logiciel NETWORK.	55

## Liste des tableaux

<b>N° de tableau</b>	<b>Titre</b>	<b>Page</b>
<b>I</b>	Liste des séquences d'ADN trouvées dans la GenBank du genre Meriones.	Annexe
<b>II</b>	Les séquences d'ADN nucléaires et mitochondriales de genre Meriones extraites de la GenBank.	34
<b>III</b>	Comparaison des indices de diversités calculés avec DnaSP sur les échantillons étudiés.	50
<b>IV</b>	Exemple d'une séquence au format GenBank.	Annexe
<b>V</b>	Exemple d'une séquence nucléotidique au format EMBL.	Annexe
<b>VI</b>	Exemple d'une séquence nucléotidique au format DDBJ.	Annexe

# Résumé

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.

Ce mémoire expose l'utilisation des outils de bioinformatique dans les analyses phylogénétiques du genre *Meriones*, tel que la reconstruction de l'arbre phylogénétique, l'identification des haplotypes, et avoir l'évolution des espèces. Nous avons tout d'abord présenté l'espèce *Meriones shawi*, leur morphologie, leur reproduction, leur distribution dans le monde et en Algérie en particulier ; nous avons exposé les méthodes de reconstruction phylogénétique pour conclure les phylogénies de genre *Meriones*.

Dans ce travail, nous avons présenté la méthodologie de l'analyse phylogénétique à l'aide de cinq logiciels : MEGA, FaBox, DnaSP, DAMBE et NETWORK, et nous avons utilisé le programme ClustalW, le programme Bootstrap, la méthode UPGMA et le modèle de Kimura 2-paramètres pour obtenir les arbres phylogénétiques, les haplotypes, les GC skew et les arbres de NETWORK. Ces analyses appliquées sur 78 séquences de Cytochrome b de l'ADN mitochondriale et 07 séquences d'Avpr2 de l'ADN nucléaire de 24 espèces appartiennent au genre *Meriones* prises dans différentes régions géographiques.

**Mot clés :** *Meriones shawi*, bioinformatique, Cytochrome b, Avpr2, phylogénétiques.

# ملخص

إعادة البناء الوراثي هي أداة تستخدم في مجالات متنوعة مثل التصنيف وعلم الأوبئة وعلم الوراثة للسكانيه. إنه إجراء ضروري لتتبع التاريخ التطوري للمنظمات من جميع المستويات التصنيفية.

تتضمن هذه المذكرة عن استخدام مختلف وسائل الاعلام الالي الحيوي دراسة النشوء لجنس في *Meriones*. مثل إعادة بناء شجرة النشوء والتطور، لتحديد الأنماط الفردية، وتطور الأنواع. قدمنا دراسة *Meriones shawi* التحليل النوع وتشكلها وتكاثرها وتوزيعها في العالم والجزائر على وجه الخصوص؛ لقد حددنا طرق إعادة بناء النشوء الأنساب من النوع والتطور لاستنتاج *Meriones*.

في هذه العمل قدمنا منهجية تحليل النشوء والتطور باستخدام خمسة برامج MEGA, FaBox, DnaSP, وبرنامج Bootstrap وطريقة UPGMA وطرز NETWORK وDAMBE واستخدمنا برنامج ClustalW وبرنامج Kimura 2 paramètres لاستنتاج شجرة النشوء والتطور، les GC skew, les haplotypes, و شبكات NETWORK،  
قمنا بمعالجة 78 سلسلة من مورثة Cytb و 07 سلسلة من مورثة Avpr2 لـ 24 نوع من جنس *Meriones*، لمختلف المناطق الجغرافية .

**المصطلحات:** *Meriones shawi*, Cytb, Avpr2, شجرة النشوء. اعلام الالي الحيوي.

# Abstract

The phylogenetic reconstruction is a tool used in the fields as diverse as the taxonomy, epidemiology and genetics of populations. It is a procedure essential to trace the evolutionary history of organizations of all taxonomic levels.

This memory exposes the use of useful of bioinformatic in the analyses phylogenetic of the genus *Meriones*, such as the reconstruction of the phylogenetic tree, to identify haplotypes, and have the evolution of the species. First of all, we have presented the species *Meriones shawi*, their morphology, their reproduction, their distribution in the world and Algeria in particular; we have outlined the methods of phylogenetic reconstruction to infer phylogenies of kind *Meriones*.

In this work, we have presented the methodology of phylogenetic analysis using five software: MEGA, FaBox, DnaSP, DAMBE and NETWORK and we have used the program Clustalw, the Bootstrap program, the UPGMA method and the model of Kimura 2-Parameters to infer the phylogenetic trees, the haplotypes, the GC skew and trees of NETWORK. ; these analysis applied on 78 sequences of Cytochrome b and 07 sequences of Avpr2 of mitochondrial DNA of 24 species belong to the genus *Meriones* taken in different geographical regions .

**Key words:** *Meriones shawi*, bioinformatic, Cytochrome b, Avpr2, phylogenetic.

# Sommaire

<b>INTRODUCTION</b> .....	01.
<b>CHAPITRE I : RAPPELS BIBLIOGRAPHIQUE.</b>	
<b>I-1 : Terminologie et modèle d'étude</b> .....	<b>02.</b>
<i>I.1.1 : Présentation de l'espèce</i> .....	02.
<i>I.1.2 : Classification de l'espèce</i> .....	02.
<i>I.1.3 : Répartition géographique</i> .....	04.
<i>I.1.3.1. Répartition dans le monde</i> .....	04
<i>I.1.3.2. Répartition en Algérie</i> .....	04
<i>I.1.4 : Reproduction de l'espèce</i> .....	05.
<b>I. 2 : Etude phylogénique</b> .....	<b>06.</b>
<i>I.2.1 : Aperçu historique sur la phylogénie</i> .....	06.
<i>I.2.2 : Phylogénie moléculaire</i> .....	07.
<i>I.2.3 : Représentation phylogénique</i> .....	08.
<i>I.2.4 : L'arbre phylogénétique</i> .....	09.
<i>I.2.4.1 : Les groupes taxonomiques</i> .....	10.
<i>I.2.4.2 : Notions de racine et de groupe externe</i> .....	10.
<i>I.2.4.3 : Différentes représentations graphiques pour les arbres</i> .....	11.
<b>I.3 : La phylogénie et l'informatique</b> .....	<b>12.</b>
<i>I.3.1 : Les bases de données</i> .....	12.
<i>I.3.2 : Les banques de données généraliste</i> .....	12.
<i>I.3.2.1: Banque nucléique</i> .....	13.
<i>I.3.2.2 : Banque protéique</i> .....	13.
<i>I.3.3 : Les banque de donnes spécialiste</i> .....	14.
<b>I.4 : les logiciels utilisent dans la phylogénie</b> .....	<b>15.</b>
<i>I.4.1 : Présentations des logiciels d'analyse phylogénétique</i> .....	15.
<i>I.4.1.1 : les méthodes de reconstruction phylogénique</i> .....	15.
<i>I.4.1.1.1: La méthode de parcimonie</i> .....	15.

<i>I.4.1.1.2: La méthode de 'bayesian inference'</i> .....	15.
<i>I.4.1.1.3: la méthode de maximum de vraisemblance</i> .....	16.
<i>I.4.1.2 : les logiciels de nettoyage d'alignement</i> .....	16.
<i>I.4.1.3 : les logiciels de alignement de séquences</i> .....	17.
<i>I.4.1.4 : les logiciels de traitement de séquence</i> .....	17.
<i>I.4.1.5 : les logiciels de visualisation et reconstruction d'arbre phylogénique</i> ...	17.

## **CHAPITRE II : MATERIEL ET METHODES**

<b>II.1. Analyses moléculaires et phylogénétiques</b> .....	<b>18.</b>
<i>II.1.1. Etapes d'Analyses Moléculaires</i> .....	18.
<i>II.1.1.1. Echantillonnage</i> .....	18.
<i>II.1.1.2. Extraction d'ADN</i> .....	18.
<i>II.1.1.3. Amplification de l'ADN par PCR</i> .....	18.
<i>II.1.1.4. La Séquençage</i> .....	19.
<b>II.2. Etude de la GenBank</b> .....	<b>19.</b>
<i>II.2.1. Définition de la GenBank</i> .....	19.
<i>II.2.1.1. Traitement des séquences de la GenBank</i> .....	21.
<i>II.2.1.1.1. FaBox</i> .....	22.
<i>II.2.1.1.2. MEGA</i> .....	25.
<i>II.2.1.1.3. DnaSP</i> .....	28.
<i>II.2.1.1.4. DAMBE</i> .....	30.
<i>II.2.1.1.5. NETWORK</i> .....	31.

## **CHAPITRE III : RESULTATS ET DISCUSSION**

<b>III.1. Résultats de la GenBank</b> .....	<b>33.</b>
<i>III.1.1. Le cytochrome b</i> .....	35.
<i>III.1.2. L'Avpr2</i> .....	35.
<b>III.2. Analyses phylogénétiques</b> .....	<b>36.</b>
<i>III.2.1. Construction des arbres phylogénétiques</i> .....	36.
<i>III.2.1.1. Cytochrome b</i> .....	36.
<i>III.2.1.2. Avpr2</i> .....	41.
<i>III.2.2. Interprétation des relations phylogénétiques entre les différentes espèces de genre Meriones et M. shawi</i> .....	42.

<b>III.3. Détermination des haplotypes.....</b>	<b>45.</b>
<i>III.3.1. Résultats du DnaSP pour le Cytb.....</i>	<i>45.</i>
<i>III.3.2. Résultats du DnaSP pour Avpr2.....</i>	<i>49.</i>
<i>III.3.3. Interprétation des résultats de DnaSP.....</i>	<i>50.</i>
<b>III.4. Analyses statistiques.....</b>	<b>50.</b>
<b>III.5. Résultat de Network .....</b>	<b>53.</b>
<i>III.5.1. Résultat de Cytb.....</i>	<i>53.</i>
<i>III.5.2. Résultat d'Avpr2 .....</i>	<i>55.</i>
<b>CONCLUSION.....</b>	<b>57.</b>
<b>RÉFÉRENCES BIBLIOGRAPHIQUES.....</b>	<b>59.</b>
<b>ANNEXE</b>	



# Introduction

# Introduction

Au cours de ces trente dernières années, la récolte de données en biologie a connu un boom quantitative grâce notamment au développement de nouveaux moyens techniques servant à comprendre l'ADN et d'autres composants d'organismes vivants. Pour analyser ces données, plus nombreuses et plus complexes aussi, les scientifiques se sont tournés vers les nouvelles technologies de l'information. L'immense capacité de stockage et d'analyse des données qu'offre l'informatique leur a permis de gagner une puissance pour leurs recherches et la rencontre entre la biologie et l'informatique, c'est ce qu'on appelle la bioinformatique. L'utilisation du terme bioinformatique est documentée pour la première fois en **1970** dans une publication de **Paulien Hogeweg** et **Ben Hesper** (université d'Utrecht, Pays-Bas).

La bioinformatique fournit des bases de données centrales, accessibles mondialement, qui permettent aux scientifiques de présenter, rechercher et analyser de l'information. Elle propose des logiciels d'analyse de données pour les études de données et les comparaisons et fournit des outils pour la modélisation, la visualisation, l'exploration et l'interprétation des données (**Yvan et al., 2007**).

La notion de phylogénie est indissociable de l'évolution, puisqu'elle part du principe que les organismes possèdent entre eux un lien de parenté. La notion d'évolution, théorisée pour la première fois par **Jean-Baptiste Lamarck** en **1809** sous le nom de transformisme, a mis du temps avant de s'imposer.

Ainsi, il existe inéluctablement une parenté entre les espèces, liées par une histoire évolutive propre à chacune. En **1866**, le biologiste **allemand Ernst Haeckel** invente le terme de « phylogénie » pour désigner cette relation de filiation qui unit les êtres vivants.

Le Mérione de Shaw est l'espèce la plus grande de son genre, elle fréquente des biotopes relativement humides, les Meriones sont des rongeurs de la taille d'un rat qui possède une fourrure épaisse de taille moyenne (**Aulagnier et Thevenot, 1986**).

Le but de ce mémoire repose sur l'utilisation de plusieurs logiciels de bioinformatique afin d'étudier une base des données de biologie moléculaire pour trouver la position génétique de Mérione Shawi dans l'arbre phylogénétique de genre Meriones.



# RAPPELS BIBLIOGRAPHIQUES

## I.1. Terminologie et modèle d'étude :

### I.1.1. Présentation de l'espèce :

*Merione shawi* est l'espèce la plus grande de son genre, c'est une espèce qui fréquente des biotopes relativement humides, les Meriones sont des rongeurs de la taille d'un rat qui possède une fourrure épaisse (Aulagnier et Thevenot, 1986) de taille moyenne, à soles plantaires partiellement velues (nues près du talon), à griffes claires et queue relativement courte, de même couleur que le dos (plus claire sur les côtés), avec un petit pinceau terminal noirâtre. Pelage dorsal doux, fauve à gris piqueté de brun, avec un pelage blanchâtre derrière l'oreille et un pelage gris au-dessus et au-dessous de l'œil. Oreilles à l'extrémité pigmentée. Bulles tympaniques plus longues que le diastème, atteignant le condyle occipital. Mandibule à processus coronoïde court et ramus large ;  $2n = 44$  (Aulanier *et al.*, 2008).



Figure 1 : *Merione shawi* (Ahmim 2019).

### I.1.2. Classification de l'espèce :

Les rongeurs ont une distribution géographique naturelle qui couvre le monde entier (Aulagnier et Thevenot, 1986). Cet ordre est également le plus diversifié, si on doit tenir compte des caractéristiques morphologiques, des aptitudes physiques ainsi que des différents milieux qu'occupent ces animaux (Ouzauit, 2000). La classification des rongeurs repose surtout sur des caractères morphologiques tels que les mensurations corporelles, la dentition, la forme et la structure du crâne et la perforation correspondant au passage de l'artère méningé et autres caractères du pelage (Le Berre, 1990). La fragilité et le traitement fastidieux et délicat du crâne des petits rongeurs conduit souvent à l'utilisation des mesures corporelles comme éléments essentiels de classification (Sicard *et al.*, 1995 ; Bergmans, 1997) ; la position systématique de *Mérione Shawi* est représentée ci-dessous :

<b>Règne</b>	<b>Animalia</b>
<b>Embranchement</b>	Chordata
<b>Sous-embranchement</b>	Vertebrata
<b>Classe</b>	Mammalia
<b>Sous-classe</b>	Theria
<b>Infra-classe</b>	Eutheria
<b>Ordre</b>	Rodentia
<b>Sous/Ordre</b>	Myomorpha
<b>Famille</b>	Muridae
<b>Sous-famille</b>	Gerbillinae
<b>Genre</b>	Meriones
<b>Espèces</b>	<b>Sous espèces</b>
<i>Meriones chengi</i>	
<i>Meriones crassus</i>	
<i>Meriones grandis</i>	
<i>Meriones libycus</i>	<i>Meriones libycus erythrourus</i>
<i>Meriones meridianus</i>	<i>Meriones meridianus penicilliger</i>
<i>Meriones persicus</i>	
<i>Meriones rex</i>	
<b><i>Meriones shawi</i></b>	<b><i>Meriones shawi isis</i></b>
<i>Meriones tamariscinus</i>	
<i>Meriones tristrami</i>	<i>Meriones tristrami blackleri</i>
	<i>Meriones tristrami bodenheimeri</i>
	<i>Meriones tristrami bogdanovi</i>
	<i>Meriones tristrami intraponticus</i>
	<i>Meriones tristrami kilisensis</i>
	<i>Meriones tristrami lycaon</i>
<i>Meriones unguiculatus</i>	
<i>Meriones vinogradovi</i>	
<i>Unclassified Meriones</i>	<i>Meriones sp. Garat An Njila</i>
	<i>Meriones sp. HK-2016</i>
	<i>Meriones sp. n. MAO-2018</i>

La variation de la taille des individus au sein d'une même espèce liée à l'âge et au bien-être de l'animal relance toujours une polémique par rapport à la validité de l'utilisation des paramètres corporels comme base d'identification des rongeurs (Dako *et al.*, 2002). Selon

les mêmes auteurs, au sein des petits rongeurs le problème est encore plus perceptible étant donné que les adultes de certaines espèces possèdent visiblement la même taille. D'où la nécessité d'évaluer alors les mensurations corporelles comme base de classification des petits rongeurs. *Meriones (Pallasiomys) shawi* ou *Meriones shawi*, en français la Mérione de Shaw ou Gerbille de Shaw ; c'est un rongeur du genre *Meriones*.

### I.1.3. Répartition géographique :

#### I.1.3.1. Répartition dans le monde :

*Merione shawi* habite les plaines de l'Afrique du Nord. Il vit dans le sud du Maroc, de l'Algérie, de la Libye et de l'Égypte ; trois sous espèces habitent le Maroc. On constate aussi que ces rongeurs sont incapables d'effectuer des déplacements à longue distance et surtout qu'ils sont totalement incapables de retourner au gîte lorsqu'on les déplace et dans toute son aire de répartition qui s'étend en latitude jusqu'aux abords du Sahara, *M. shawi* est précisément limitée vers le sud par le désert dont elle ne supporte pas les conditions de vie. Elle vit dans les terriers profonds et compliqués où elle se protège de la même façon des pertes d'eau puisque l'humidité relative y est toujours voisine de 100%, et des variations de la température puisqu'elle se trouve à une profondeur où l'amplitude thermique est faible et où la moyenne de température est proche de 20° à 25° (Petter, 1965).

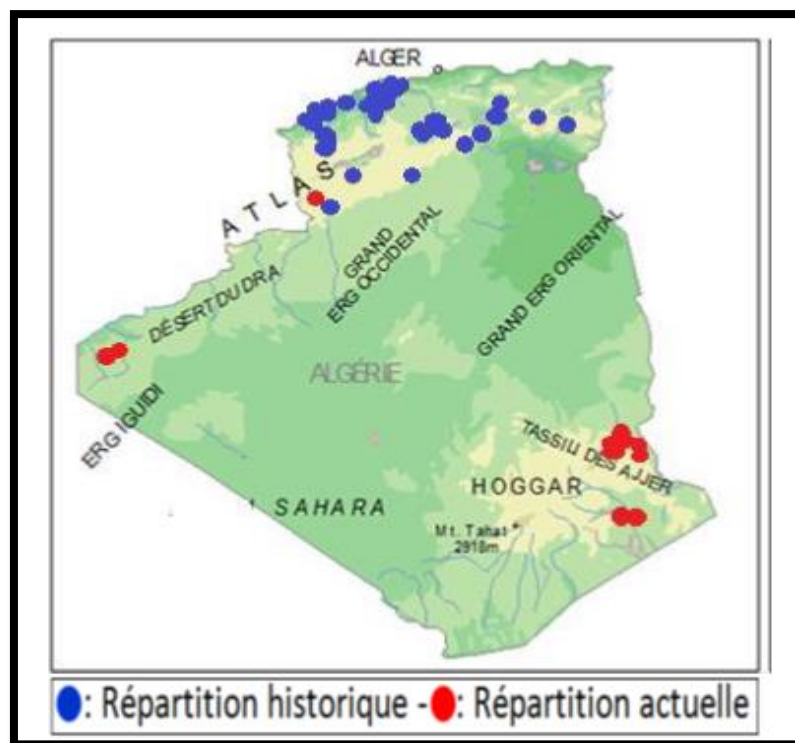
#### I.1.3.2. Répartition en Algérie :

➤ **Répartition historique :**

Habite la zone côtière jusqu'aux hauts plateaux. **Petter (1961)** établit la carte de répartition de cette espèce qui a été signalée en **1840** par **Duvernoy** à Sétif ; en **1841** par **Wagner** à Msila et Mostaganem ; en **1854** par **Gervais** à **Biskra** ; en **1858-1867** par **Loche** à Beni Slimane, Chlef, Sour El-Ghozlane ; en **1882-1885** par **Lataste** à Oued Magra (Msila). Plus tard, elle est signalée en **1913** par **Thomas** à Ain Sefra ; en **1937** par **Laurent** à Aïn Defla, Arrib (Aïn Defla), Beni Rached, Kherba ; en **1947-1948** par **Chaworth et al.** à Guelt Es-stel, Khenchela et Timgad ; en **1961** par **Petter** à 20 km de Mechria ; en **1979** par **Kowalski** à Aïn Berda, Oran, Sidi Ammar (Aïn Temouchent), Tlemcen, Aflou, Batna, Tafraoui et dans des pelotes à Aflou, Frenda, Honaine, Messerghine, Murdjadjo, Sebdou, Timgad.

➤ Répartition et démographie actuelles :

**Adamou et al. (2011)** étudièrent l'amplitude des dégâts provoqués à Tiaret par cette espèce au niveau de la région de Sidi Abdelghani et aussi à Sidi Belabbes. **Chenchouni (2012)** a étudié la faune du lac Ayata du complexe de zones humides d'Oued Righ entre octobre 2009 et juin 2010 et il signala sa présence. **Bachar et Belhamra (2012)** signalèrent la présence de *Meriones shawi* dans quatre stations représentatives des Zibans (Zibans-est : Branis et Sidi-Okba ; Zibans-ouest : Hraichet et Doucen) dans la wilaya de Biskra. **Guerzou et al.** signalèrent la présence de cette espèce dans les pelotes de réjection du grand corbeau en 2012 dans la région de Guelt-es-Stel au Zahrez Gharbi au nord de Djelfa. En 2013, **Adamou-Djerbaoui et al.** étudièrent le régime alimentaire de cette espèce dans la région de Tiaret.



**Figure 2** : La répartition géographique historique et actuelle de *Mérione de shaw* en Algérie (Ahmim, 2019).

#### I.1.4. Reproduction de l'espèce :

L'activité du la Merione de shaw est diurne ou nocturne en période froide et crépusculaire ou nocturne en période chaude. Elle prolifère après une bonne année agricole, lorsque la pluviométrie est importante entre le mois de novembre et le mois d'avril. La période de reproduction commence en fin d'hiver et atteint le maximum au printemps. La durée de gestation est de l'ordre de 20 jours. Le nombre de petits est de 5 à 12 par portée.

L'âge de la reproduction est le suivant : pour le mâle, la maturité sexuelle est atteinte entre 10 et 12 semaines ; pour la femelle, elle se situe entre 10 et 18 semaines. Le poids des animaux est alors de 60 grammes alors que le poids adulte est de 70 à 100 grammes. Les gerbilles se reproduisent durant un an à un an et demi, leur espérance de vie étant de 5 ans. La saison de reproduction se situe durant toute l'année. Les gerbilles sont assez rarement polygames, les couples se forment et durent.

La gestation dure 24 à 26 jours et il y a cinq à six portées par an comprenant chacune de sept à quatorze petits. Le sevrage intervient à 3 semaines, les petits pèsent alors 18 grammes. À la naissance, ils pèsent 3 grammes.

## **I. 2. Etude phylogénique :**

### **I.2.1. Aperçu historique sur la phylogénie :**

Le terme phylogénie (en grec phulon « race, tribu » et genesis « genèse, formation ») inventé par **Ernst Haeckel** en **1866** pour définir l'enchaînement des espèces animales et végétales au cours des temps. Elle consiste à retrouver l'histoire évolutive des organismes reposant sur la connaissance des relations phylogénétiques. La pensée phylogénétique trouve ses racines dans les célèbres arbres évolutifs dressés dans la seconde moitié du **XIX siècle** par **Darwin** et **Haeckel**. Et c'est seulement avec Hennig dans les années 1950 que la pensée « phylogénétique » s'est largement répandue à travers sa mise en pratique dans la systématique éponyme (**Thomas et al., 2016**). Pendant un siècle, la phylogénie est basée sur l'utilisation des caractères morphologique, physiologiques et métaboliques. Le début de la phylogénie moléculaire remonte aux années 60, aux travaux de **Zuckerkandl** et **Pauling (1965)** et **Fitch** et **Margoliash (1967)** basés sur les séquences protéiques du Cytochrome C et de la globine chez les vertébrés. L'objectif de la phylogénie moléculaire est la reconstruction des arbres phylogénétiques par comparaison de l'information génétique présente dans les génomes des êtres vivants. Elle a pour but de classer les espèces selon leurs liens de parenté. La phylogénie



postule que toutes les espèces proviennent d'un ancêtre commun hypothétique (Judd *et al.*, 2002).

Mais la véritable révolution moléculaire en phylogénie commence au début des années 1990 suite à l'invention de la PCR (Polymérase Chain Réaction) et au développement des techniques de séquençage. C'est grâce à l'utilisation de ces séquences qu'il a été possible pour la première fois d'effectuer une véritable classification phylogénétique des microorganismes « les archées », les troisièmes domaines du vivant ont été découvert en se basant sur l'analyse des fragments de l'ARN ribosomique 16S (Figure 3) (Woese *et al.*, 1990).

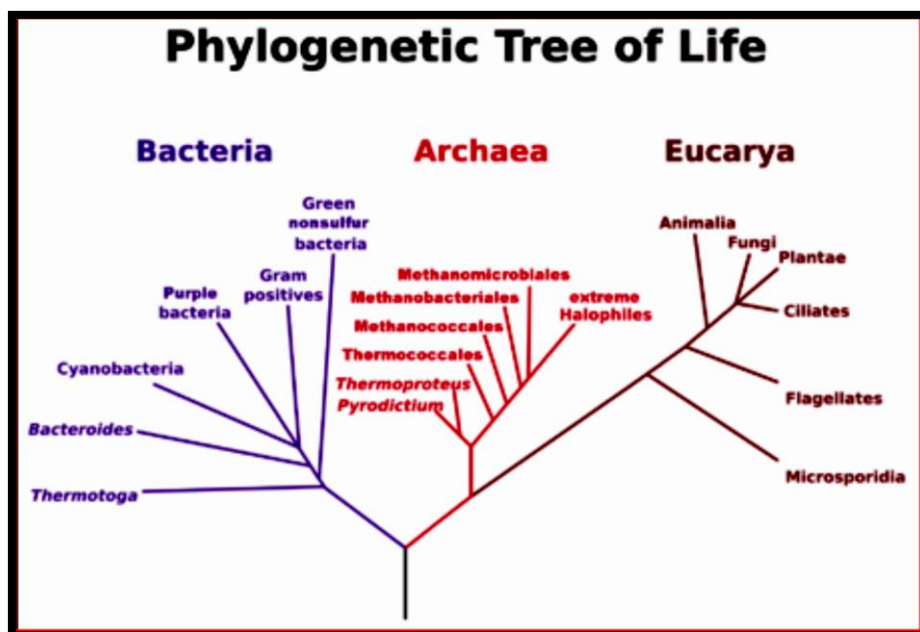


Figure 3 : Les trois domaines du vivant (Woese *et al.*, 1990).

### I.2.2. Phylogénie moléculaire :

La phylogénie moléculaire utilisée pour comparer les séquences des molécules d'ADN ou des protéines des êtres vivants dans le but de déterminer les liens de parenté qui les unissent pour appréhender leur histoire évolutive (phylogénèse). Elle étudie l'histoire évolutive des espèces étudiées à la base d'une portion de leur séquence moléculaire (Diallo, 2009).

La phylogénie moléculaire est basée sur les principes suivants :

- Les mutations se produisent au hasard.
- Les mutations s'accumulent au cours de temps.

- Les mutations se produisant chez un individu et se fixent dans sa descendance.

Il existe aujourd'hui une large gamme de marqueurs moléculaires qui peuvent être utilisés pour accomplir ce type de recherche.

Un marqueur génétique est un caractère mesurable qui peut détecter une variation dans la séquence soit protéique, soit nucléique. Selon le niveau taxonomique étudié, il est essentiel de choisir un type de marqueur qui lui est adapté. Par exemple, si le but est de distinguer des groupes d'individus au niveau de l'espèce, les marqueurs utilisés devront être plus conservés que dans le cas des études de populations (**Beaulieu, 2007**).

Les marqueurs moléculaires concernant la molécule d'ADN elle-même, et, comme tels, sont considérés comme des mesures objectives de la variation (**Vicente et al., 2003**).

✓ **Marqueurs d'ADN nucléaire :**

Un certain nombre de marqueurs sont à présent disponibles pour détecter les polymorphismes d'ADN nucléaire. Dans les études sur la diversité génétique, les marqueurs les plus fréquemment utilisés sont les microsatellites. Les microsatellites sont à présent les marqueurs les plus utilisés dans les études de caractérisation génétique des animaux d'élevage (**Sunnucks, 2001**).

✓ **Marqueurs d'ADN mitochondrial :**

Le génome mitochondrial des animaux est formé par une petite molécule unique d'ADN circulaire dont la taille varie dans une gamme assez étroite chez les vertébrés (16-20kb) (**Avise et al., 1987 ; Boore, 1999**). Sa séquence code pour 37 gènes dont 24 contiennent l'information pour une partie de la machinerie de traduction de la molécule elle-même (ARNt et ARNr), et les 13 autres pour les sous unités de la chaîne de transport des électrons. En plus, il est reconnu une région régulatrice (ou de contrôle) des processus de réplication et de transcription de l'ADNmt d'environ 0.8Kb nommée D-Loop chez les vertébrés (**Avise et al., 1987**). Par ailleurs, il est décrit la présence, dans plusieurs ADNmt, de petits fragments non codants avec des fonctions qui peuvent être aussi régulatrice (**Boore, 1999**).

### **I.2.3. Représentation phylogénique :**

L'élaboration de technologies à haut débit serait inutile si l'on ne disposait pas des capacités d'analyse des données biologiques en croissance exponentielle. Ces données doivent se stocker dans des bases de données électroniques associées à des logiciels spécialement

conçus pour la mise à jour, l'interrogation et l'extraction. Les informations doivent être facilement accessibles et flexibles aux interrogations pour faciliter l'extraction des informations pouvant être analysées pour éclaircir les voies métaboliques et le rôle des protéines et des gènes impliqués (**Rischkowsky *et al.*, 2008**).

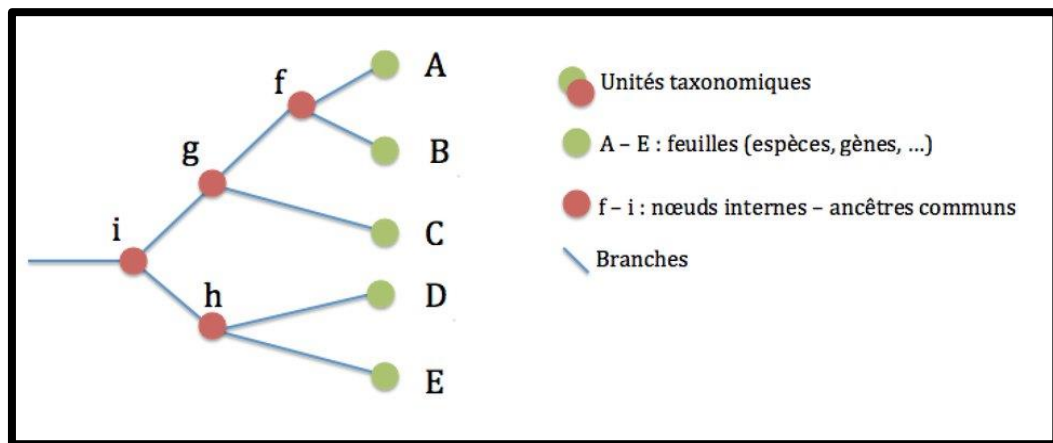
La bioinformatique est fondamentale pour associer les informations provenant de sources différentes et générer une connaissance nouvelle à partir de données existantes. Elle dispose également des potentialités pour simuler la structure, la fonction et la dynamique des systèmes moléculaires et est donc utile pour la formulation des hypothèses et pour la conduite du travail expérimental (**Rischkowsky *et al.*, 2008**).

#### I.2.4. L'arbre phylogénétique :

Un arbre phylogénétique ou dendrogramme est une représentation graphique d'origine biologique, mathématique et informatique d'un groupe de taxon, qui répond à la question : « qui est plus proche parent de qui ? » (**Darlu et Tassy, 1993**) (Figure 4).

Du point de vue mathématique, un arbre est composé de quatre éléments principaux (**Perrière et BrochierArmanet, 2010**) :

- **Les sommets externes degré 1** : Sont appelés **les feuilles** qui représentent les unités taxonomiques opérationnelles (UTOs) comme les organismes ou séquences.
- **Les sommets internes** : Sont appelés **les nœuds** représentant les unités taxonomiques hypothétiques (UTHs).
- **Les arêtes** : Sont appelées **les branches** de l'arbre, elles représentent les lignées évolutives connectant deux UTHs ou une UTH et une UTO.
- **Le sommet degré 2** : Est appelé **racine**, il représente l'ancêtre commun à tous les objets de l'arbre.



**Figure 4** : Schéma illustrant les différentes notions de la représentation d'un arbre phylogénétique (Taguan, 2015).

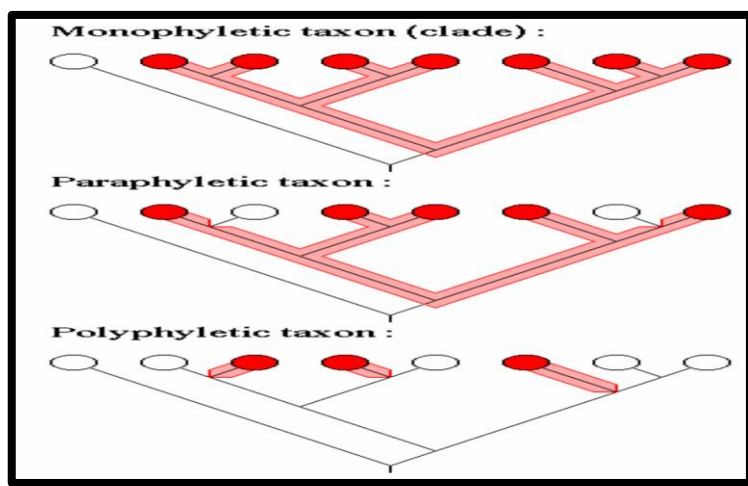
#### 1.2.4.1. Les groupes taxonomiques :

Il existe trois types de groupes taxonomiques. La figure 5 illustre ces types. Ces groupements n'ont pas la même signification biologique (Thomas *et al.*, 2016).

a) **Groupe monophylétique** : Un ensemble de taxon incluant un ancêtre hypothétique et l'intégralité de ses descendants.

b) **Groupe paraphylétique** : Un groupe incluant un ancêtre hypothétique et certains de ses descendants.

c) **Groupe polyphylétique** : Un groupement non apparenté présentant des ressemblances.



**Figure 5** : Schéma illustrant les notions des différents types de groupes taxonomiques.

#### 1.2.4.2. Notions de racine et de groupe externe :

La racine désigne le nœud le plus ancien d'une reconstruction phylogénétique, c'est une représentation de l'ancêtre de l'ensemble de taxons étudiés. C'est également la reconnaissance de l'origine et l'orientation. Une analyse phylogénétique ne peut pas résoudre la position du nœud racine sans l'aide d'un groupe externe qui signifie un groupe de taxons extérieurs.

Toutefois, pour enraciner un arbre, il faut que le groupe externe ne soit pas trop éloigné sur le plan évolutif pour permettre de détecter les caractères homologues. La représentation la plus courante des phylogénies utilise la convention de la trifurcation, trois

branches partent de la base d'un arbre : le premier conduit au groupe externe et les deux autres constituent la première bifurcation au sein du groupe interne (Thomas *et al.*, 2016).

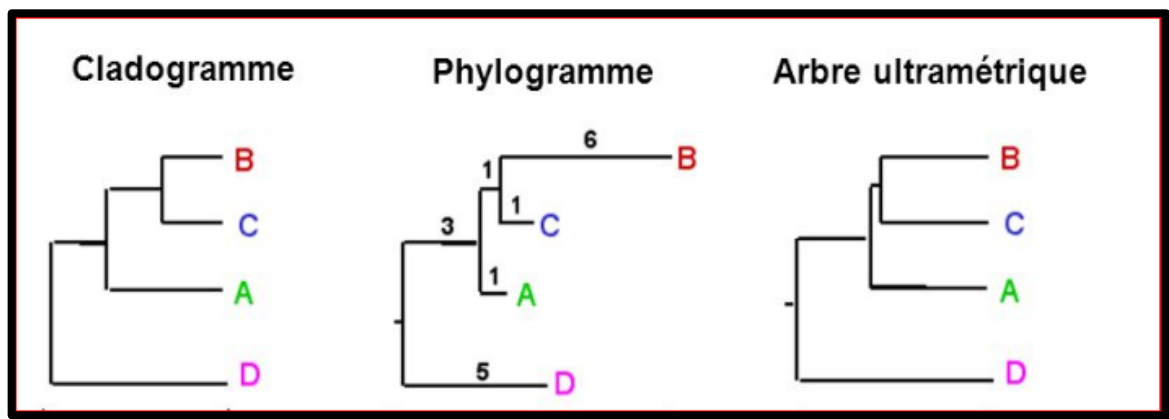
### 1.2.4.3. Différentes représentations graphiques pour les arbres :

Selon la représentation graphique, il est possible de reconnaître le type d'arbre et le type d'information véhiculée (Thomas *et al.*, 2016).

**Le Dendrogramme :** Appliqué à tous les types graphiques arborés (les cladogrammes, les phénogrammes, les phylogrammes ou les arbres ultramétriques tels que les chronogrammes)

(Figur

e 6).



**Figure 6 :** Schéma représentant les différents types de dendrogrammes.

**a) Le cladogramme :** Présentation sous forme de « cladogramme rectangulaire » issu de l'école cladiste (Hennig, 1966). Les taxons terminaux sont équidistants de la racine de l'arbre qui ne représente que les relations de parenté.

**b) Le phylogramme :** Représente sur un même graphe les relations de parenté et le degré de différence entre les nœuds et les taxons terminaux qui ne sont pas tous équidistants de la racine de l'arbre. Les longueurs de branches sont les nombres de changements évolutifs par le temps intervenu entre deux taxons.

**c) L'arbre ultramétrique :** Représente des relations de parenté et des changements évolutifs. Les taxons terminaux sont tous à égale distance de la racine. Les liens entre des nœuds internes n'ont pas nécessairement la même longueur.

**d) Le phénogramme :** Est le produit d'une étude phénétique. Graphiquement, c'est un arbre ultramétrique mais ne représente pas les relations de parenté et regroupe les taxons montrant une plus forte similitude globale (Sneath et Sokal, 1973).

### I.3. La phylogénie et l'informatique :

#### I.3.1. Les bases de données :

Une base de données est un ensemble de données structurées et organisées, permettant le stockage de grandes quantités d'informations afin d'en faciliter leur utilisation.

Il se trouve plusieurs types de banques de données utiles dans le domaine de la génétique entre autres :

- **Ensembl** (European Bioinformatics Institute / Wellcome Trust Sanger Institute)
- **NCBI** (National Center for Biotechnology Information - USA).
- **UCSC** (University of California Santa Cruz).
- **Vista** (University of California).
- **Argo** (BROAD Institute).
- **Mochiview** (University of California Santa Cruz).
- **X: map** (Paterson Institute for Cancer Research).
- **DiProGB** (Leibniz Institute for Age Research).
- **Genatlas** (Université René Descartes - Paris).

**Tagu *et al.* (2010)**, affirment que les banques et les bases de données sont maintenant une source d'information majeure pour la communauté scientifique.

#### I.3.2. Les banques de données généralistes :

Appelées aussi banques primaires : sont les ressources qui collectent, gèrent, archivent et mettent à disposition de la communauté scientifique un ensemble de données primaires, c'est-à-dire obtenues expérimentalement. Elles sont considérées comme banques primaires les

banques généralistes de séquences nucléiques et protéiques bien que la plupart des séquences protéiques ne soient pas obtenues expérimentalement, mais à partir des données de séquences nucléiques, ainsi que les banques qui gèrent les structures tridimensionnelles des protéines.

### I.3.2.1. Banque nucléique :

Selon **Tagu *et al.* (2010)**, il existe trois banques nucléiques internationales :

- ✓ **GenBank** la banque américaine gérée par le National Centre for Biotechnology Information (NCBI).
- ✓ **EMBL** (European Molecular Biology Laboratory databank), le banque européenne maintenue à l'European Bioinformatics Institute (EBI).
- ✓ **DDBJ** (DNA Database of Japan), la banque japonaise.

Ces trois banques gèrent l'ensemble des séquences nucléiques et leurs annotations, elles coopèrent et échangent quotidiennement leurs données afin de garantir une cohérence maximale dans la mise à disposition des séquences de la communauté scientifique, même si chacune de ces banques présente quelques petites spécificités mais la structuration des données y est semblable et leur contenu en séquences nucléiques est strictement identique.

### I.3.2.2. Banque protéique :

D'après **Tagu *et al.* (2010)**, trois banques protéiques aussi coexistent de manière indépendante dont l'objectif est de couverture c'est-à-dire d'exhaustivité et d'annotations :

- ✓ **La banque de données européenne Swiss-Prot** : Qui se caractérise par une excellente qualité d'annotation des données grâce à la contribution d'experts au détriment de l'exhaustivité.
- ✓ **La banque TrEMBL** : Qui contient l'ensemble des séquences protéiques conceptuelles obtenues par traduction automatique des séquences codantes contenues dans EMBL, avec des annotations automatiques non vérifiées, mais avec l'objectif d'obtenir une couverture maximale. De même la banque GenPept correspond à la traduction automatique de l'ensemble des séquences annotées comme codantes (CDS) dans GenBank.

✓ **La banque américaine Protein Information Resource (PIR), à la National Biomedical Research Foundation (NBRF)** : Qui dans les années 1960 fut la première banque de protéines développée. Sa particularité consiste à proposer une classification des séquences protéiques en familles, en fonction de leur degré de similarité, dont l'avantage est de limiter le degré de redondance de la banque d'une part et, d'autre part, de travailler à la standardisation de l'annotation des protéines.

### I.3.3. Les banques de données spécialisées :

Pour pallier ces inconvénients, l'augmentation exponentielle du volume, de la diversité des séquences et la diversité des études, des banques spécialisées ont été développées. Ces développements ont permis l'introduction d'informations spécifiques à chacune permettant ainsi d'avoir des banques adaptées aux besoins des utilisateurs. Elles répondent pour la plupart soit à des besoins ponctuels, soit à des besoins liés à des secteurs d'activité bien précis.

Parmi celles-ci, ont été développées des banques thématiques se consacrant à un domaine bien précis. Ainsi, certaines regroupent des données sur les structures moléculaires tridimensionnelles telles que la PDB (**Berman et al., 2000**). D'autres s'intéressent à la structure en domaine des séquences protéiques comme la banque ProDom (**Servant et al., 2002**). Il y en a également qui centralisent des données sur les signatures, caractéristiques de certaines protéines telle que PROSITE (**Hulo et al., 2006**). D'autres encore traitent des séquences et des structures d'ARN.

- ✓ La RDP est une banque américaine qui contient non seulement les séquences des ARN ribosomiques de la grande et de la petite sous-unité du ribosome, mais aussi des alignements et des arbres phylogénétiques (**Perrière, 2000 ; Guindon, 2003**).
- ✓ ECDC est une banque qui présente une compilation des séquences d'*E. Coli* contenues dans EMBL et dans GenBank.
- ✓ La base SGD « Saccharomyces Genome Database » compile les séquences nucléiques des gènes codant pour les protéines de la levure *Saccharomyces cerevisiae*.
- ✓ PGD « Plant Genome Database » est une base de données qui compile des données sur le génome de plusieurs plantes céréalières. Par ailleurs, ATDB « *Arabidopsis thaliana* DataBase » rassemble toutes les informations sur les données génétiques, le clonage et le séquençage des molécules de la plante crucifère *Arabidopsis thaliana*.



## **I.4. Les logiciels utilisés dans la phylogénie :**

### **I.4.1. Présentation des logiciels d'analyse phylogénétique :**

Il existe une myriade de logiciels, implantés sur différentes machines depuis microordinateur jusqu'aux ordinateurs les plus puissants, écrits dans telle ou telle langue, traitant tel ou tel problème.

Dans le domaine de la reconstruction phylogénétique, il y'a plusieurs programmes traitant les différents aspects de la phylogénie. Pour cela nous allons décrire quelques grandes catégories de logiciels utilisés pour les analyses phylogénétiques selon leurs fonctions.

#### **I.4.1.1. Les méthodes de reconstruction phylogénétique :**

Les méthodes de reconstruction phylogénétique reposent presque toujours sur l'optimisation d'un critère permettant de comparer les phylogénies possibles d'un ensemble de taxons, nous décrivons les principales méthodes de reconstruction phylogénétique, les méthodes de distances, de parcimonie et du maximum de vraisemblance.

##### **I.4.1.1.1. La méthode de parcimonie :**

Les méthodes de maximum de parcimonie ont été initialement développées pour être utilisées sur des données morphologiques (**Hennig, 1966**). La parcimonie, à la différence des méthodes de distances, considère chaque site individuellement pour l'ensemble des séquences. Elle est basée sur la théorie de rasoir d'Occam, qui spécifie que la meilleure hypothèse pour expliquer un processus est celle qui fait appel au plus petit nombre d'évènements (calculer le nombre minimal de substitutions). (**Perrière et Brochier-Armanet, 2010**). Ces parcimonies sont réalisées par des logiciels spécialisés tels que : PHYLIP (Phylogeny Inference Package), PAUP, PAML, EDIBLE, PTP, PHYML, PARDOOT, ALIFRITZ, CRUX, PRAP.

##### **I.4.1.1.2. La méthode 'Bayesian inference' :**

Le principe de cette méthode est relativement simple, une série d'arbre est construite sur le principe que chaque arbre, très voisin du précédent, doit apporter une amélioration à la construction générale. Ainsi, un premier arbre phylogénétique est construit au hasard. Puis un élément de cet arbre est changé, également au hasard. Ce nouvel élément peut être la longueur d'une branche, ou le changement de position d'un taxon puis les arbres successifs sont conservés selon un algorithme particulier (dit de metropolis –Hastings) qui accepte les arbres au *pro-rata* de leur apport : si le nouvel arbre apporte une amélioration, il est toujours accepté ; s'il est moins bon de 10% de l'arbre précédent, il n'est accepté que 90% du temps (Pagel, 2002).

Cette méthode est réalisée par des logiciels spécialisés tels que : PAML, BAMBE, PAL, Vanilla, MrBayes, Mesquite, PHASE, BEAST, MrBayes tree scanners, P4, SIMMAP, IMA2, BALI-Phy.

#### **I.4.1.1.3. La méthode de maximum de vraisemblance :**

Le maximum de vraisemblance a été introduit par Neyman, cependant c'est Felsenstein qui, le premier, a développé un algorithme efficace applicable aux séquences nucléotidiques (Perrière et Brochier-Armanet, 2010). L'analyse du maximum de vraisemblance se compose de trois parties, premièrement un modèle de changement évolutif est spécifié. À partir de ce modèle, différentes hypothèses sur l'histoire de l'évolution sont évaluées en termes de probabilité. Enfin, on choisit l'hypothèse qui présente la probabilité la plus élevée. Le maximum de vraisemblance est la méthode d'estimation la moins affectée par l'erreur d'échantillonnage (Swofford *et al.*, 1996), et donne de meilleurs résultats que les méthodes de distances et de parcimonie, permettant ainsi de comparer facilement les hypothèses. La contrepartie de cela étant bien sûr un temps de calcul bien plus élevé que ces deux méthodes (Strimmer, 1997). Cette méthode est réalisée par des logiciels spécialisés tels que : PHYLIP, DAMBE, EMBOSS, PRAP, PAUP\*, ARB, FastTree, Creux, Vanilla, McRate, DAR.

#### **I.4.1.2. Les logiciels de nettoyage d'alignement :**

C'est l'élimination des régions divergentes et présentant peu de signal phylogénétiques (peu informatives).

- ❖ **Sequencher** : est un logiciel de bioinformatique produit par la société Gene Codes Corporation. Ce logiciel permet de faire l'assemblage de plusieurs séquences d'ADN contiguës relativement courtes afin de créer des séquences plus longues.
- ❖ **BioEdit** : Beaucoup de chercheurs dans le domaine de la biologie moléculaire ont utilisé des modules de BioEdit ; est un éditeur d'alignement de séquences et un programme d'analyse de séquences piloté par la souris, facile à utiliser.

#### **I.4.1.3. Les logiciels d'alignement de séquences :**

L'alignement de deux séquences consiste à mettre en évidence les similitudes et les différences entre les deux séquences. Ces alignements sont réalisés par des logiciels spécialisés tels que : MEGA, MUSCLE, TreeAlign, clustal W, DAMBE, T.coffee, BIRCH, EMBOSS.

#### **I.4.1.4. Les logiciels de traitement de séquences :**

- ❖ Séquence suite manipulation : ensemble de pages html et java script permettant des analyses simples sur les séquences nucléotidiques et protéiques, très bien fait et facile à utiliser (mac, pc, linux).
- ❖ **Fast PCR** : Design des amorces pour PCR ; il faut lire l'aide en ligne pour une prise en main du logiciel ; il permet la prédiction et l'optimisation des amorces.
- ❖ **MB DNA Analysis** : Package permettant de réaliser les cartes de restriction, %GC, recherche d'ORF, dot-plot, alignement multiple, calcul de masse moléculaire, prédiction de structure secondaire.
- ❖ **DRAW32** : Logiciel de dessin de vecteur. Il réalise également des cartes de restriction et propose une électrophorèse virtuelle des fragments obtenus.
- ❖ **Antheprot** : Dédié à l'analyse des séquences protéiques.

#### **I.4.1.5. Les logiciels de visualisation et reconstruction d'arbre phylogénique :**

Actuellement, il existe plusieurs programmes qui permettent la visualisation des arbres, les arbres finalisés ont alors été visualisés et annotés à l'aide des programmes : PhyloDraw,

NJplot, GeneTree, PHYLIP (DRAWTREE et DRAWGRAM), Genedoc, DAMBE, Trecon, TreeView, et Spectrum (**Choi *et al.*, 2000**).

# **MATERIEL ET METHODES**

Cette étude est réalisée au niveau du laboratoire de génétique des populations et biologie de conservation des populations animales de l'Université des Sciences et de la Technologie Houari Boumediene (USTHB) d'Alger.

Dans ce chapitre, nous présenterons quelques fonctions des logiciels utilisés dans notre travail.

### **II.1. Analyses moléculaires et phylogénétiques :**

L'objectif de cette partie repose essentiellement sur l'utilisation de plusieurs logiciels de bioinformatique afin d'étudier une base de données de biologie moléculaire pour trouver l'arbre phylogénétique et la position de l'espèce *Merione shawi* de genre *Meriones*. Pour cela nous décrirons les étapes d'analyse moléculaire et la méthodologie utilisée depuis l'échantillonnage, l'extraction d'ADN, la PCR, le séquençage, jusqu'au traitement des séquences qui se trouvent au niveau de la GenBank en vue de leur utilisation dans l'étude phylogénétique.

#### **II.1.1. Etapes d'analyses moléculaires :**

##### **II.1.1.1. Echantillonnage :**

Les échantillons employés pour extraire l'ADN sont composés de : os, tissu, poils, sang ou matériel fécal.

##### **II.1.1.2. Extraction et quantification de l'ADN :**

L'extraction d'acides nucléiques d'un matériau biologique requiert la lyse cellulaire, l'inactivation des nucléases cellulaires et la séparation de l'acide nucléique souhaité de débris cellulaires. Il existe beaucoup de kits d'extraction et de purification de l'ADN qui sont commercialisés. Cependant il a été montré que la sensibilité de détection par PCR de l'ADN extrait varie selon le kit utilisé (**Yoshikawa *et al.*, 2011**).

##### **II.1.1.3. Amplification de l'ADN par PCR :**

Polymérase Chain Réaction (PCR) (**Kleppe *et al.*, 1971**) est une méthode *in vitro* d'amplification de séquences spécifiques d'ADN en un très grand nombre de copies. Elle permet d'amplifier une séquence spécifique d'ADN dans un mélange réactionnel approprié.

#### II.1.1.4. Séquençage :

Le séquençage de l'ADN est le processus consistant à déterminer l'ordre précis des nucléotides, dans une molécule d'ADN. Il comprend tout procédé ou technologie utilisée pour déterminer l'ordre des quatre bases adénine, guanine, cytosine, thymine le long de la molécule d'ADN. Il est en effet possible de réaliser un diagnostic en séquençant l'ADN. Le travail est réalisé à partir d'ADN purifié. Il existe différentes techniques de séquençage : le séquençage enzymatique, méthode de Sanger, méthode chimique de Maxam et Gilbert, séquençage haut débit.

### II.2. Etude de la GenBank :

#### II.2.1. Définition de la GenBank :

La GenBank est une collection annotée de toutes les séquences d'ADN publiquement disponibles. Cette banque est mise à jour régulièrement grâce à des échanges quotidiens de séquences avec la banque européenne EMBL et la banque japonaise DDBJ.

Toutes les séquences étudiées du genre *Meriones* dans ce travail ont été téléchargées à partir de la GenBank, nous passons par plusieurs étapes distinctes :

1. Construire l'URL suivante : <https://www.ncbi.nlm.nih.gov/genbank/>
2. Choisir recherche par taxonomie.
3. Ecrire le nom de notre genre « *Meriones* » (Figure 7).

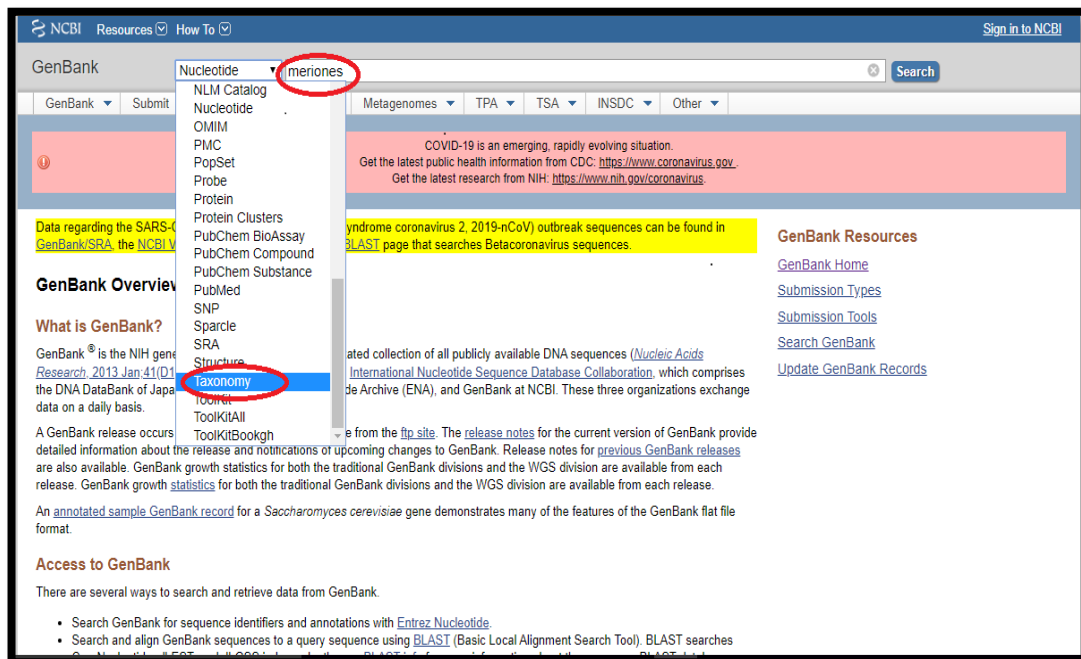


Figure 7 : Les étapes initiales pour récupérer les séquences de la GenBank.

4. En entrant sur Meriones, nous obtenons toutes les espèces qui appartiennent à ce genre (Figure 8).



Figure 8 : Les vingt-quatre espèces de genre « Meriones » obtenues dans la GenBank ; ces espèces sont indiquées par des flèches en orange.

5. Nous avons analysé toutes les espèces une par une pour identifier le nombre des séquences nucléotidiques trouvées dans cette banque (Figure 9).

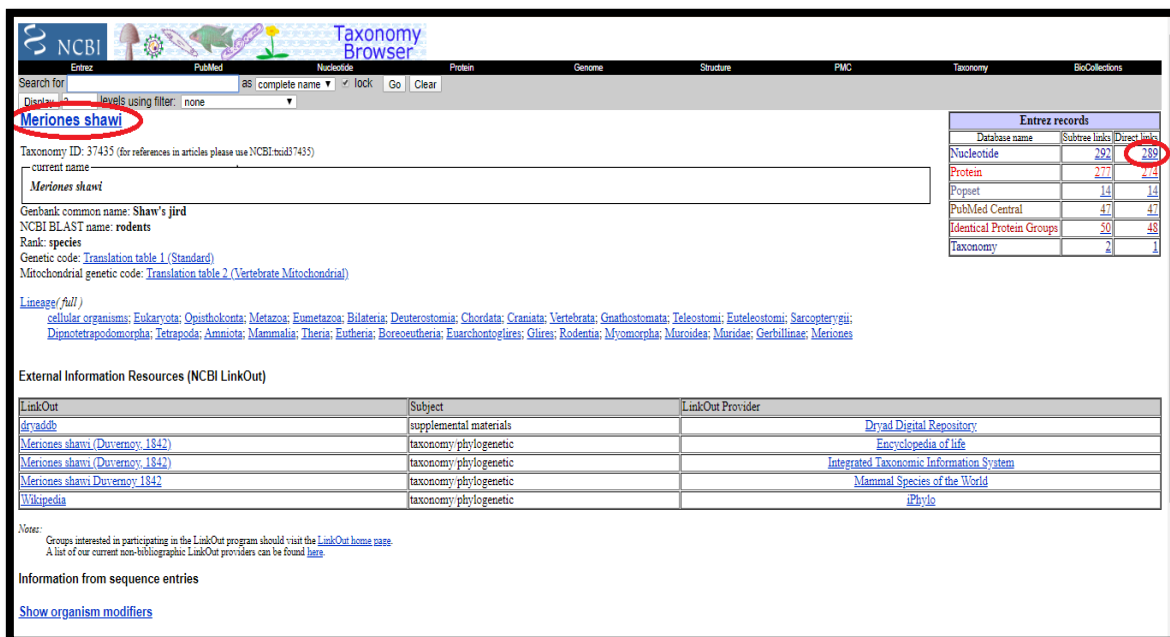


Figure 9 : Exemple des séquences nucléotidiques trouvées pour l'espèce *Meriones Shawi*.



La base de données a été construite pour sélectionner les isolats à étudier et générer les fichiers de sortie correspondant au format informatique « FASTA ». Ces fichiers peuvent, alors être directement utilisés dans nombre de logiciels de traitement et d'analyse de séquences.

### II.2.1.1. 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, Bosque, EMBOSS, phangorn, Bio++, ETE, DendroPy, SeaView, Crux, mais nous avons choisi FaBox, DnaSP5, MEGA5, DAMBE et NETWORK.

Nous avons procédé à la construction des arbres phylogéniques à partir des séquences nucléotidiques de chaque espèce du genre « Meriones » par différentes étapes :

6. Les séquences représentatives de chacune des 24 espèces ont été téléchargées sous format FASTA à partir de la "GenBank" (Figure 10 et 11).

The screenshot shows the NCBI GenBank interface. At the top, there is a search bar with 'Nucleotide' selected. Below the search bar, there is a red banner with COVID-19 information. The main content area displays the FASTA sequence for 'Meriones shawi voucher H-583 cytochrome b (cytb) gene, partial cds; mitochondrial'. The sequence is shown in a monospaced font. To the right of the sequence, there are several interactive elements: 'Change region shown', 'Customize view', 'Analyze this sequence' (with a dropdown menu), 'Run BLAST', 'Pick Primers', 'Highlight Sequence Features', 'Find in this Sequence', 'Related information' (with a dropdown menu), and 'Recent activity' (with a dropdown menu). The 'Recent activity' section shows a list of recent searches, including 'Meriones shawi voucher H-583 cytochrome b (cytb) gene, partial cds; mitochondri' and 'txid37435[Organism: noexp] (289)'. The bottom right corner of the screenshot shows a 'Turn Off' and 'Clear' button.

Figure 10 : Exemple d'une séquence nucléotidique au format FASTA de « Meriones shawi ».

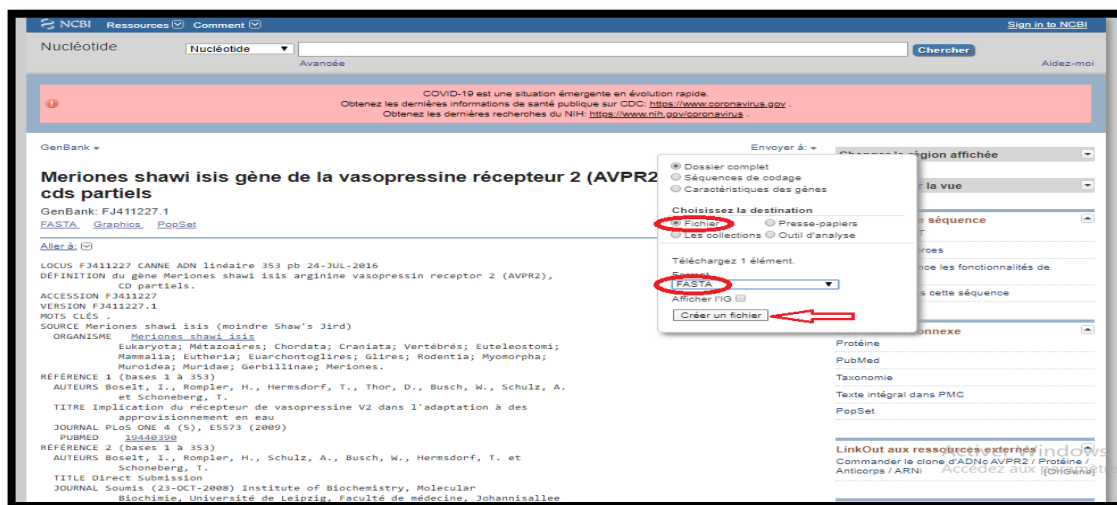


Figure 11 : Exemple d'un fichier sous format FASTA de « Meriones shawi isis ».

### II.2.1.1.1. FaBox :

Le logiciel FaBox par définition est une collection de services de Web simples et intuitifs qui permettent à des biologistes et à des chercheurs médicaux d'effectuer rapidement la tâche typique avec des données d'ordre. Les services le rendent facile à extraire, éditer, et remplacer des en-têtes d'ordre et de joindre ou diviser des ensembles de données basés sur l'information d'en-tête. D'autres services incluent l'effondrement d'un ensemble d'ordres dans des haplotypes et de formatage automatisé des dossiers d'entrée pour un certain nombre de programmes de la génétique de population, tels que l'arlequin, le tcs et les mrbayes (Villesen, 2007) (Figure 12, 13, 14).

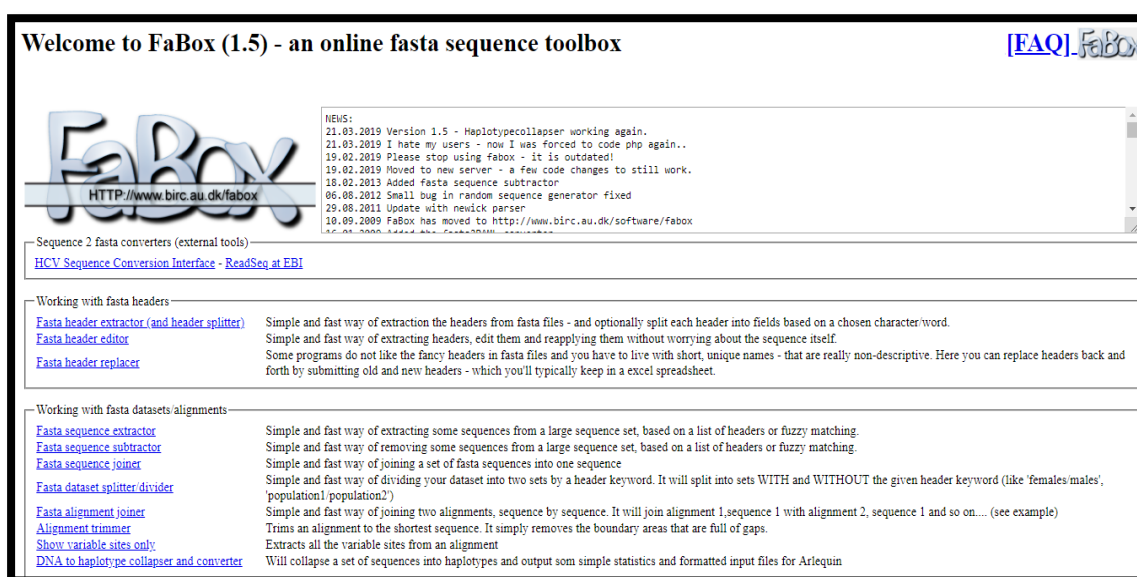


Figure 12 : La page principale de logiciel FaBox.

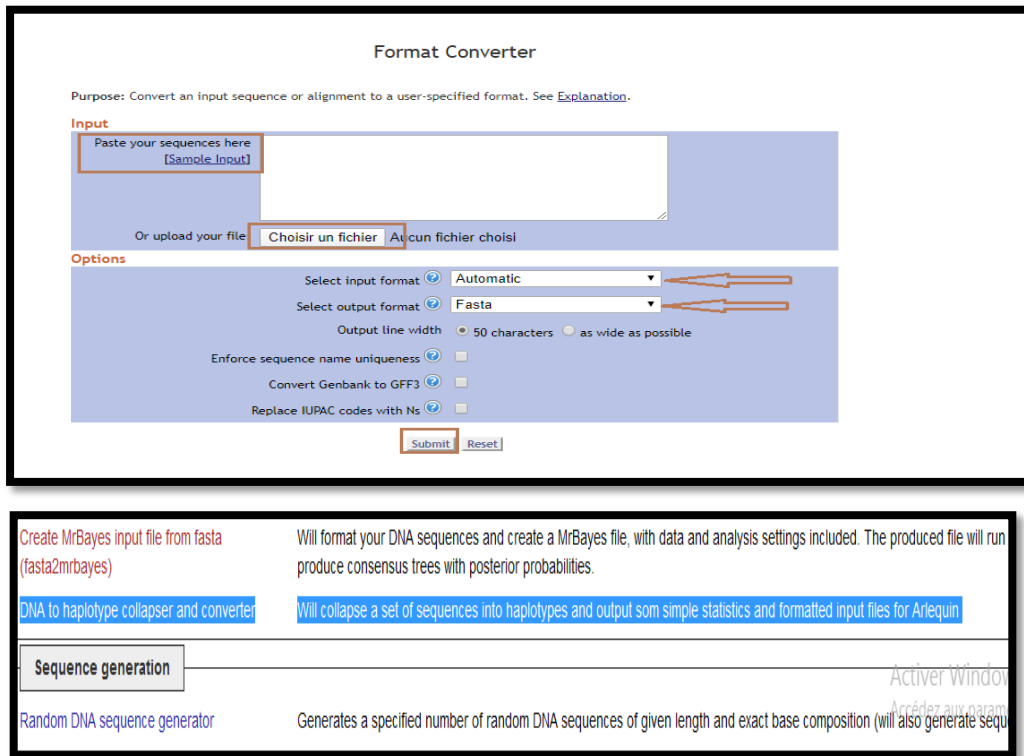


Figure 13 : La conversion des séquences sous forme Fasta par l'utilisation de Fabox.

**Fréquences et rapport des haplotypes**

Numéro d'haplotype	Étiquette	Nombre de séquences	Longueur de l'haplotype (pb)	Séquences appartenant à l'haplotype
1	1	1	1323	meriones_chengi-chine
2	2	1	1323	meriones_crassus-arabie_saoudite
3	3	1	1323	meriones_crassus-Egypte
4	4	1	1323	meriones_crassus-Iran
5	5	1	1323	meriones_grandis-maroc
6	6	1	1323	meriones_ibycus_erythrouus-chine
7	7	1	1323	meriones_persicus-Iran
8	8	1	1323	meriones_persicus-Iran (2)
9	9	6	1323	meriones_persicus-Iran (3) meriones_persicus-Iran (2) meriones_persicus-Iran (3) meriones_persicus-Iran (4) meriones_persicus-Iran (5)
dix	dix	1	1323	meriones_persicus-Iran (6)
11	11	1	1323	meriones_persicus-Iran (7)
12	12	1	1323	meriones_persicus-Iran (8)
13	13	2	1323	meriones_persicus-Iran (9) meriones_persicus-Iran (16)
14	14	1	1323	meriones_persicus-Iran (10)
15	15	3	1323	meriones_persicus-Iran (11) meriones_persicus-Iran (13) meriones_persicus-Iran (15)
16	16	2	1323	meriones_persicus-Iran (12) meriones_persicus-Iran (17)
17	17	1	1323	meriones_persicus-Iran (14)
18	18	2	1323	meriones_persicus-Iran (18) meriones_persicus-Iran (19)
19	19	1	1323	meriones_persicus-Iran (20)
20	20	1	1323	meriones_persicus-Iran (21)
...	...	...	...	...

Figure 14 : L'obtention des séquences en haplotypes par l'utilisation de Fabox.

Alignement cropper va réduire l'alignement (en format Fasta) à la longueur de la séquence la plus courte (suppression les gaps dans le début et la fin de l'alignement). Appuyez sur " show example" pour l'essayer sans soumettre votre propre séquence (Figure 15 et 16).

Working with fasta datasets/alignments

- [Fasta sequence extractor](#) Simple and fast way of extracting some sequences from a large sequence set, based on a list of headers or fuzzy matching.
- [Fasta sequence subtractor](#) Simple and fast way of removing some sequences from a large sequence set, based on a list of headers or fuzzy matching.
- [Fasta sequence joiner](#) Simple and fast way of joining a set of fasta sequences into one sequence
- [Fasta dataset splitter/divider](#) Simple and fast way of dividing your dataset into two sets by a header keyword. It will split into sets WITH and WITHOUT the given header keyword (like 'females/males', 'population1/population2')
- [Fasta alignment joiner](#) Simple and fast way of joining two alignments, sequence by sequence. It will join alignment 1, sequence 1 with alignment 2, sequence 1 and so on.... (see example)
- [Alignment trimmer](#)** Trims an alignment to the shortest sequence. It simply removes the boundary areas that are full of gaps.
- [Show variable sites only](#) Extracts all the variable sites from an alignment
- [DNA to haplotype collapser and converter](#) Will collapse a set of sequences into haplotypes and output some simple statistics and formatted input files for Arlequin

Figure 15 : Alignement trimmer (la forme Fasta).

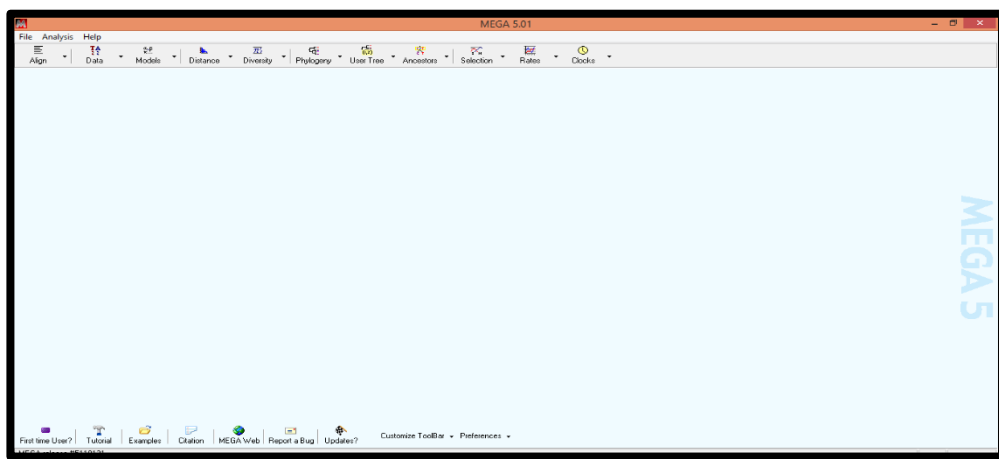
Figure 16 : L'alignement de la séquence la plus courte par « alignement trimmer ».

### II.2.1.1.2. MEGA :

Est un logiciel simple et efficace pour l'extraction des bases de données en ligne ; des séquences pour la construction des alignements et des arbres phylogénétiques, en utilisant des méthodes de la bioinformatique évolutionnaire en biologie fondamentale, la biomédecine, et de l'évolution.

Le plus récent ajout au MEGA5 (Figure 17) est une collection de maximum de vraisemblance (ML) analyse pour inférer des arbres évolutifs, sélection de meilleur ajustement des modèles de substitution (de nucléotides ou d'acides aminés), la déduction et les séquences ancestrales membres (avec des probabilités), et l'estimation des taux d'évolution des sites.

Dans les analyses de simulation informatique des algorithmes d'inférence d'arbres, ML MEGA5 dans la comparaison avec d'autres logiciels en termes d'efficacité de calcul et la précision des estimations d'arbres phylogénétiques, des paramètres de substitution, et le taux de variation entre les sites. Le MEGA interface utilisateur a été améliorée afin d'être axé sur l'activité pour la rendre plus facile pour l'utilisation des scientifiques expérimentés et débutants. Cette version de MEGA est destinée pour la plate-forme Windows, et il a été configuré pour une utilisation efficace sur Mac OS X et Linux (**Tamura, 2011**).



**Figure 17 :** Menu principal de logiciel MEGA version 5.

- Les séquences représentatives de chacune des espèces ont été ajoutées une par une sous format FASTA à partir de "GenBank" (Figure 18).



Figure 18 : Exemple des séquences ajoutées dans le logiciel MEGA.

Après l'intégration de ces séquences dans le programme MEGA5, un alignement multiple a été réalisé en utilisant le programme Clustal W (Figure 19), 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.

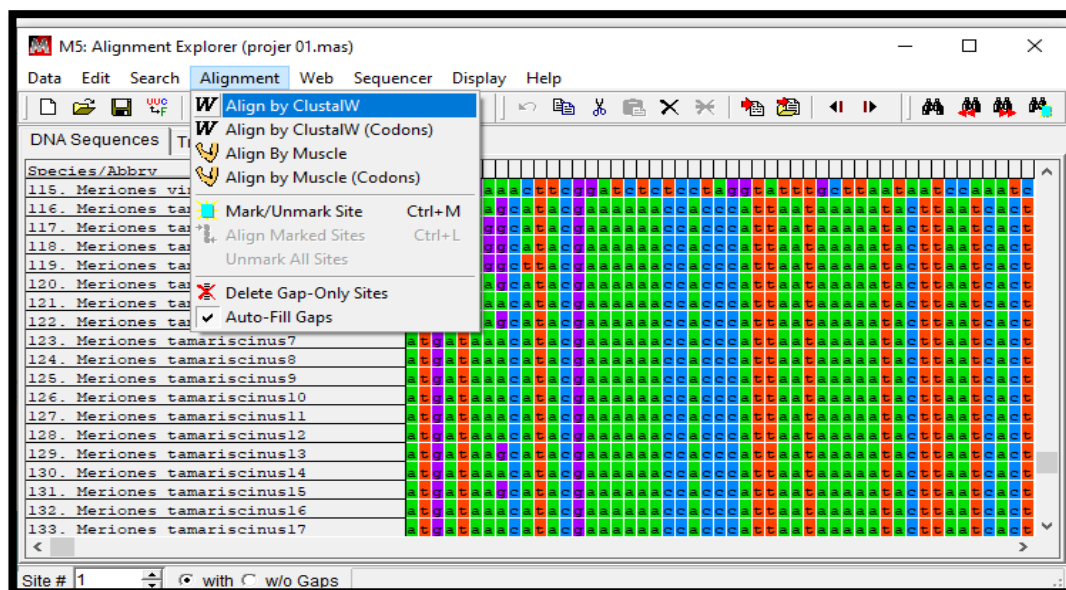
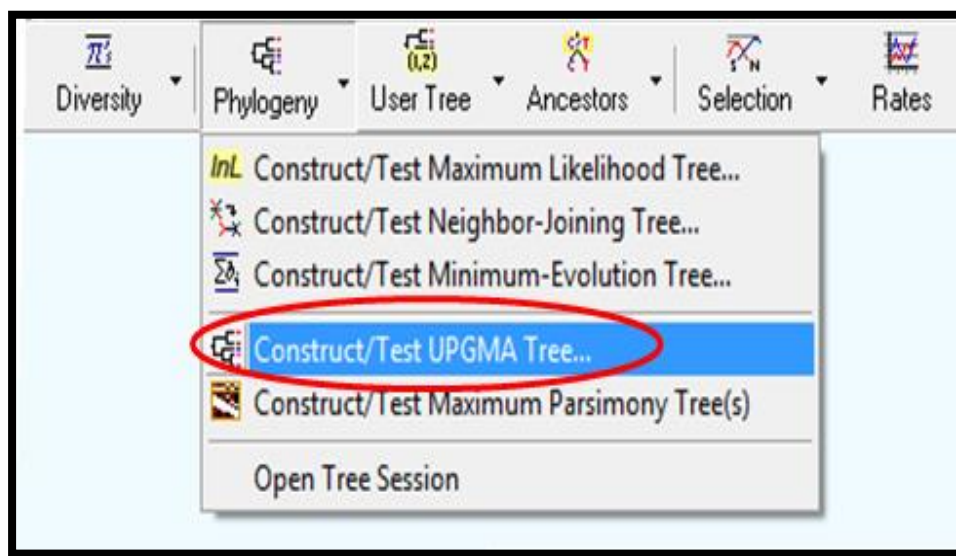


Figure 19 : L'alignement multiple des séquences du genre « Meriones » en utilisant le programme Clustal W.

- Une fois les séquences alignées sont ensuite comparées et servent à la construction d'arbres phylogénétiques ; si les séquences ne sont pas trop divergentes en utilisant l'une des méthodes fondées sur les distances, la méthode UPGMA (Figure 20), UPGMA utilise un algorithme de clusterisation séquentiel dans lequel les relations sont identifiées dans l'ordre de leur similarité et la reconstruction de l'arbre ne se fait pas grâce à cet ordre.



**Figure 20** : La construction de l'arbre phylogénétique en utilisant la méthode UPGMA.

- En utilisant le modèle de « Kimura 2-paramètres », pour appliquer la théorie neutraliste (Kimura) lorsque la plupart des mutations restent neutres, se fixent au hasard (seules les mutations très défavorisant ou létales pour l'individu sont éliminées) et le milieu n'a pas de rôle sélectif.
- Et en utilisant aussi le programme "Bootstrap". Cette méthode, partant du postulat que les caractères évoluent de manière indépendante, est sans doute la plus utilisée en phylogénie. Elle a été inventée par **Bradley Efron** en **1979** et introduite en phylogénie par **Felsenstein** en **1985** dans le package Phylip. Quand on a obtenu un arbre, il faut ensuite évaluer la confiance que l'on peut avoir dans cet arbre. Pour cela nous avons réglée les valeurs de "Bootstrap" pour 1000 (Figure 21) répliqués sont ramenées en pourcentages et ils sont indiqués au niveau des nœuds.

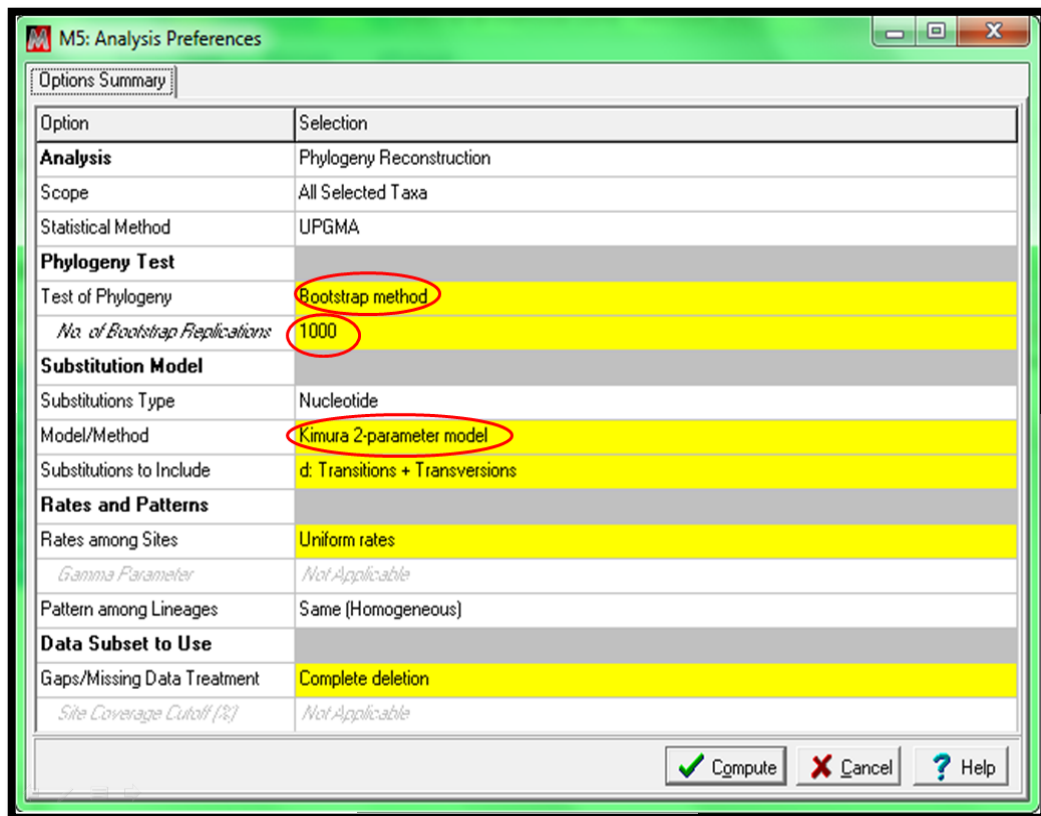


Figure 21 : Les paramètres utilisés pour la reconstruction phylogénétique.

### II.2.1.1.3. DnaSP :

Est un logiciel pour l'analyse du polymorphisme de nucléotide des données alignées d'ordre d'ADN. Il peut estimer plusieurs mesures de variation d'ordre d'ADN dans et entre les populations (dans des sites non codants, ou dans diverses sortes de positions de codons) ainsi que le déséquilibre de liaison, la recombinaison, le flux génique et les paramètres de conversion génique. DnaSP peut également effectuer plusieurs tests de neutralité. En plus il peut estimer les intervalles de confiance de certaines statistiques d'essai par le coalescent. Les résultats des analyses sont présentés sous forme de tableau et de graphique. La version 5 met en œuvre un certain nombre de nouvelles fonctionnalités et des méthodes analytiques permettant de nombreuses analyses du polymorphisme de l'ADN sur de grands volumes de données.

DnaSP (Figure 22) accepte cinq formats de fichier de données d'entrée : FASTA, MEGA, NBRF / PIR, NEXUS et PHYLIP. Les fichiers de données devraient stocker un alignement multiple de séquence d'ADN avec un polymorphisme donné (variation intra-espèces), variation interspécifique des nucléotides (variation entre espèces), ou toute

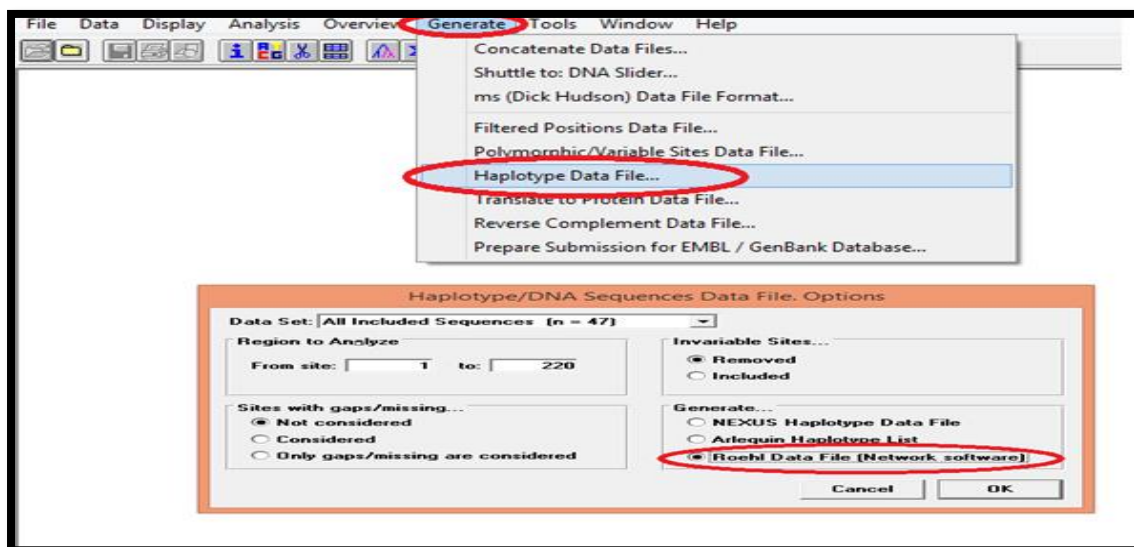


combinaison des deux. Étant donné que tous les formats sont en fichiers ASCII (texte), les fichiers peuvent donc être vu et édité dans n'importe quel éditeur de texte clair. Le logiciel permet l'exportation et la conversion de fichiers de données dans les formats mentionnés ci-dessus, ainsi que dans le format utilisé par Arlequin et NETWORK (**Librado *et al.*, 2009**).



**Figure 22** : Menu principal de logiciel DnaSP 5.10.

Pour avoir les haplotypes de ces séquences, nous avons utilisé les paramètres présentés dans la figure 23 suivantes :



**Figure 23** : Les paramètres « Haplotype Data File » et « Roehl Data File » ont été utilisé pour avoir les Haplotypes, et enregistrer le fichier sous format \*.rdf.

#### II.2.1.1.4. DAMBE :

Nous avons utilisé le logiciel DAMBE pour l'analyse phylogénétique qui peut être classé en six catégories (Figure 24) :

1) Extraction de séquence, l'édition, la manipulation et la conversion entre plus de 20 formats, y compris la séquence standard MEGA, NEXUS, PHYLIP, GenBank, et le nouveau format pour l'interopérabilité, NeXML.

2) Caractérisation et découverte des fonctions telles que la matrice de poids position et Gibbs sampler.

3) Outils d'analyse génomique descriptive avec des versions améliorées de l'indice d'adaptation efficaces, codon nombre de codons, protéines point isoélectrique de profils, de l'ARN et des protéines, la prédiction de la structure secondaire et de calcul de l'énergie, de pliage minimum et génomiques de travers les parcelles avec taille de fenêtre optimisée de la phylogénie moléculaire.

4) Alignement de séquences, y compris la substitution d'essai, la saturation, la base de la distance, la parcimonie maximale et les méthodes du maximum de vraisemblance pour la reconstruction de l'arbre, le test de l'hypothèse de l'horloge moléculaire avec soit une phylogénie ou avec tests de taux relatif, datant de la duplication des gènes et des événements de spéciation, choisir le meilleur ajustement des modèles de substitution, et l'estimation de l'hétérogénéité des taux plus de sites.

5) Basée sur la phylogénie des méthodes comparatives pour les variables continues et discrètes.

6) Y compris les fonctions graphiques d'affichage de la structure secondaire, optimisé en biais, hydrophobicité parcelle, et beaucoup d'autres parcelles de propriétés de l'acide aminé le long d'une séquence de protéine, de l'arborescence et le dessin en faisant glisser des nœuds à l'autre, et la recherche visuelle de l'arbre de parcimonie maximale (Xia *et al.*, 2013).

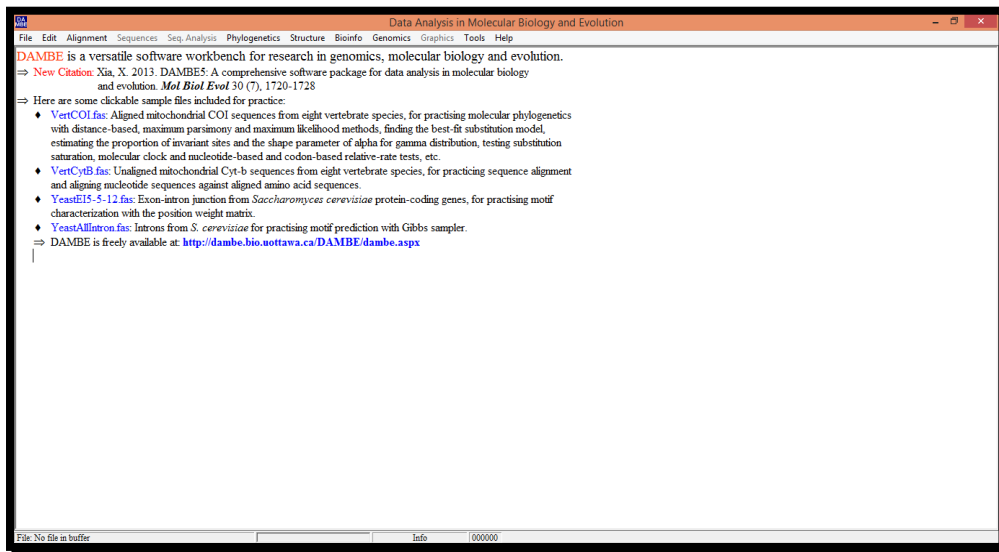


Figure 24 : Menu principal de logiciel DAMBE.

### II.2.1.1.5. NETWORK :

On a terminé notre travail par le logiciel NETWORK (Figure 25), qui est utilisé pour l'alignement de l'ADN est recommandé pour les présentations des séquences d'ADN ou d'acides aminés sous forme de réseau, et éviter les erreurs.

NETWORK est recommandé pour : les présentations graphiques dans des publications de qualité, plus les présentations des diagrammes et secteurs, l'affichage de noms de nœuds plus longs (15 caractères) et de noms de lien (9 caractères), les nœuds proxy de réseau, les mutations par lien et les individus par nœud. Le NETWORK est recommandé pour générer des fichiers rdf (Figure 26, 27) à partir de fichiers de séquençage (Barabasi *et al.*, 2011).

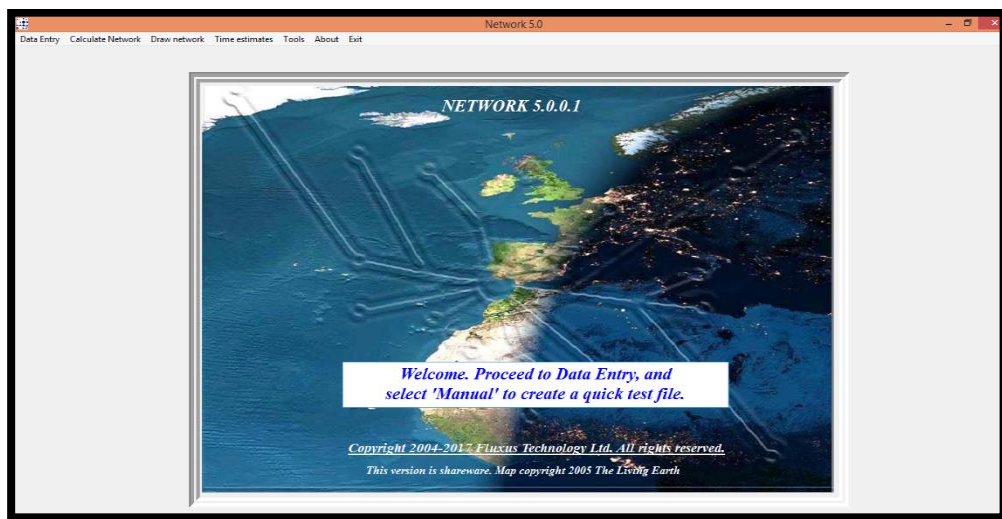


Figure 25 : Menu principal de logiciel NETWORK 5.0.

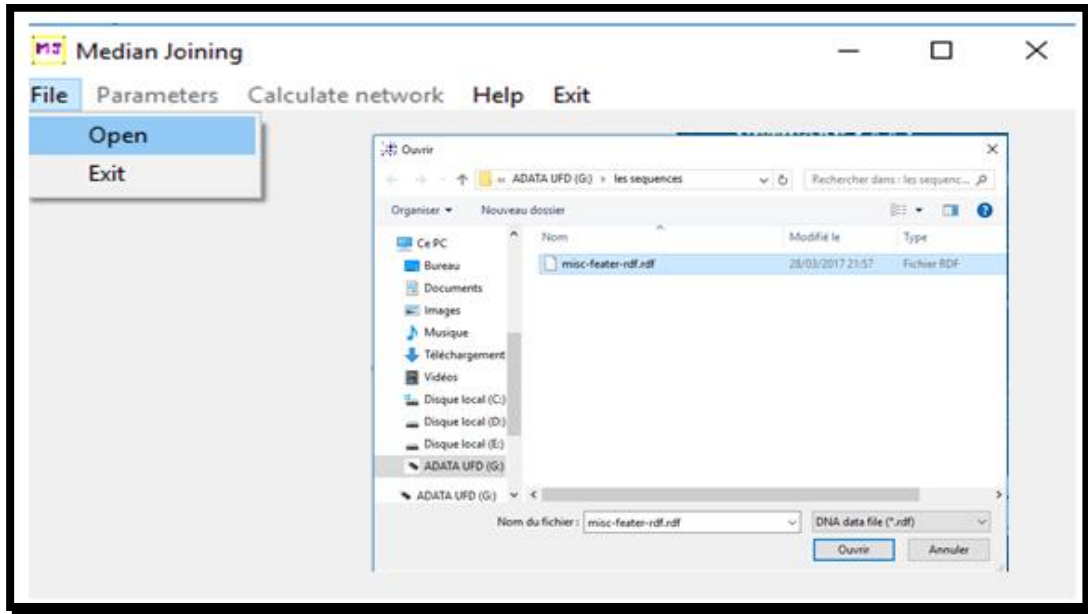


Figure 26 : L'ouverture de fichier sous format rdf produit par logiciel DnaSp.

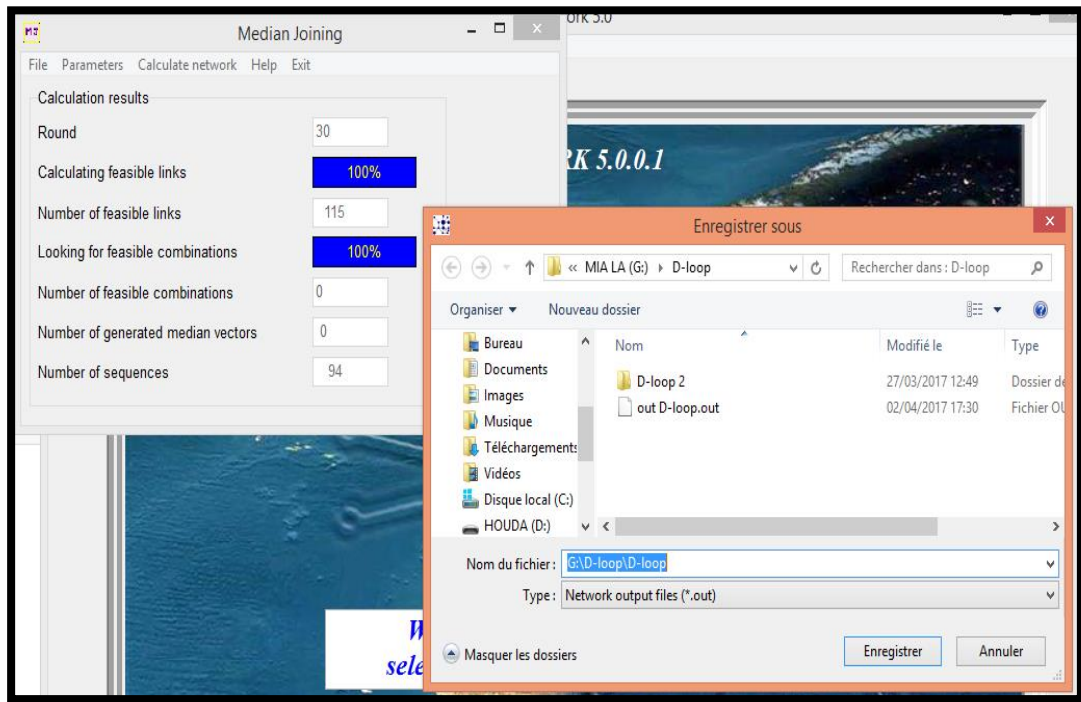


Figure 27 : L'enregistrement de notre fichier sous format \*.out pour la reconstruction de réseau de NETWORK.

# *RESULTATS ET DISCUSSION*

Cette partie est basée sur la comparaison entre les espèces de genre *Meriones* sur le plan moléculaire à partir des données de la GenBank, afin de déterminer la position de l'espèce *Merione shawi* dans l'arbre phylogénétique.

### III.1. Résultats de la GenBank :

24 espèces appartenant au genre *Meriones*, récupérées auprès de la GenBank. Les espèces qui ont été utilisées dans ces analyses phylogénétiques sont : *M. chengi*, *M. crassus*, *M. grandis*, *M. libycus*, *M. libycus erythrorurus*, *M. meridianus*, *M. meridianus pencilliger*, *M. persicus*, *M. Rex*, *M. shawi*, *M. shawi isis*, *M. tamariscinus*, *M. tristrami*, *M. tristrami blackleri*, *M. tristrami bodenheimeri*, *M. tristrami bogdanovi*, *M. tristrami intraponticus*, *M. tristrami kilisisis*, *M. tristrami lycaon*, *M. unguiculatus*, *M. vinogradovi*, *M. sp. Garat An Njila*, *M. sp. HK-2016*, *M. sp. n. MAO-2018*. Localisées dans différentes régions géographiques.

Les séquences de chaque espèce extraite à partir de la GenBank, leur type de gène et leur origine géographique sont mentionnées dans le tableau I en annexe.

Il y a plusieurs séquences identiques à 100%, pour cela nous ne prendrons en considération qu'une seule séquence représentative.

**Tableau II** : Les séquences d'ADN nucléaires et mitochondriales de genre *Meriones* extraites de la GenBank.

Espèce	Nombre total des séquences	Séquences d'ADN nucléaire	Séquences d'ADN mitochondrial	Nombre total de Cytb	Nombre total d'Avpr2	Les séquences non identifiées
<i>M. chengi</i>	3	0	3	1	0	/
<i>M. crassus</i>	8	3	5	3	0	/
<i>M. grandis</i>	2	0	2	1	0	/
<i>M. libycus</i>	156	25	130	105	0	1
<i>M. libycus erythrourus</i>	2	0	2	1	0	/
<i>M. meridianus</i>	675	284	390	167	1	1
<i>M. meridianus pencilliger</i>	1	1	0	0	1	/
<i>M. persicus</i>	80	1	79	78	0	/
<i>M. Rex</i>	4	0	4	2	0	/
<i>M. shawi</i>	289	80	209	205	0	/
<i>M. shawi isis</i>	3	3	0	0	3	/
<i>M. tamariscinus</i>	78	0	77	37	0	1
<i>M. tristrami</i>	22	12	10	9	0	/
<i>M. tristrami blackleri</i>	2	0	2	2	0	
<i>M. tristrami bodenheimeri</i>	2	0	2	2	0	/
<i>M. tristrami bogdanovi</i>	2	0	2	2	0	/
<i>M. tristrami intraponticus</i>	3	0	3	3	0	/
<i>M. tristrami kilisisis</i>	5	0	5	5	0	/
<i>M. tristrami lycaon</i>	8	0	8	8	0	/
<i>M. unguiculatus</i>	109898	726	33	33	4	109139
<i>M. vinogradovi</i>	3	0	3	2	0	/
<i>M. sp. Garat An Njila</i>	2	0	2	1	0	/
<i>M. sp. HK-2016</i>	1	0	1	1	0	/
<i>M. sp. n. MAO-2018</i>	3	0	3	3	0	/
<b>Total</b>	111252	1135	975	671	9	109142

D'après le tableau ci-dessus, nous remarquons qu'il y a :

- + 1135 séquences d'ADN génomique : BRCA1, RAG1, Avpr2...etc.
- + 975 séquences d'ADN mitochondriale majoritairement Cytb, Cox III, D-loop ...etc.

Notre étude est basée sur l'ADN mitochondrial : le Cytochrome b et l'ADN génomique : le gène d'Arginine vaso-pressin receptor 2 (Avpr2), pour établir des liens de parenté entre les espèces du genre *Meriones*.

### III.1.1. Le Cytochrome b :

Le Cytochrome b est un gène présent en nombreuses copies dans la cellule, est considéré comme clonal et rarement ou jamais soumis à la recombinaison. Il est transmis maternellement (hérédité cytoplasmique). Enfin il a un taux de mutation plus élevé que les gènes nucléaires (Figure 28).

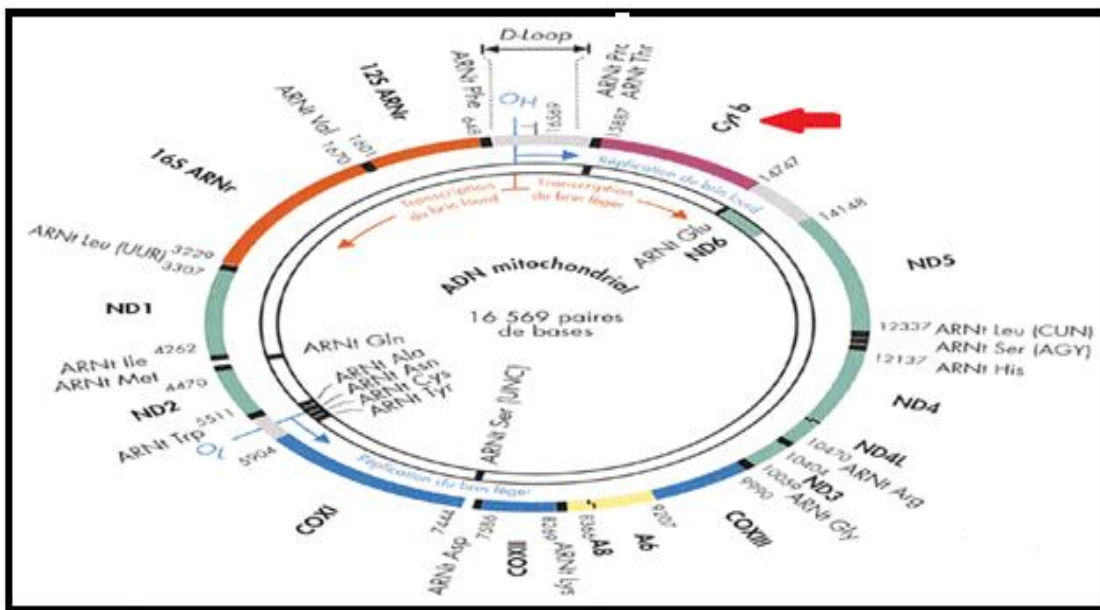


Figure 28 : Structure de l'ADN mitochondrial et la position de Cytb (Rustin *et al.*, 1995).

### III.1.2. L'Avpr2 :

Ce gène code pour le récepteur de la vasopressine de type 2, également connu sous le nom de récepteur V2, qui appartient à la superfamille des récepteurs couplés aux protéines G à sept domaines transmembranaires (GPCR), et se couple à Gs stimulant ainsi l'adénylate cyclase qu'impliqué dans la réabsorption rénale de l'eau. Localisation cytogénétique est Xq28, qui est le bras long (q) du chromosome X en position 28 (Figure 29).



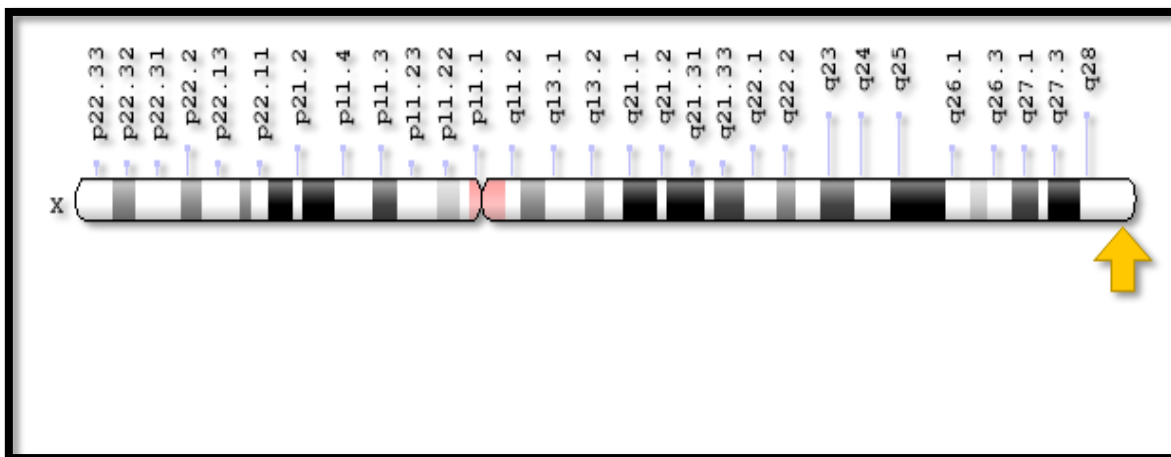


Figure 29 : La position de gène Avpr2 sur chromosome X (Xq28).

### III.2. Analyses phylogénétiques :

#### III.2.1. Construction des arbres phylogénétiques :

La construction de l'arbre phylogénétique a été réalisée à l'aide du programme MEGA5. Premièrement, nous avons choisi les séquences représentatives de Cytb et Avpr2 en se basant sur leurs origines géographiques.

##### III.2.1.1. Cytochrome b :

- *M. Chengi* : Il n'y a qu'une seule séquence de la Chine de taille de 1076 pb.
- *M. crasus* : Nous avons au total 3 séquences dont une séquence est de l'Arabie Saoudite de 1052 pb, une séquence de l'Égypte de taille de 1139 pb, et une séquence de l'Iran de taille de 737 pb.
- *M. grandis* : Une seule séquence du Maroc de taille de 737pb.
- *M. libycus* : Nous avons trouvé 26 séquences, 9 séquences pour la Chine (876pb à 1179pb) nous avons choisi une seule séquence représentative de taille de 1179pb, 2 séquences de Tunisie (1110pb et 1140pb) et de Afghanistan (737pb et 633pb), 9 séquences d'Iran (533pb à 1140pb) nous avons choisi une seule séquence représentative de taille de 1140 pb. Une seule séquence pour : l'Algérie (737pb), Syrie (737pb), Jordanie (737pb), Maroc (737pb).
- *M. libycus erythrourus* : Une seule séquence de la Chine de taille de 1103pb.
- *M. meridianus* : Nous avons trouvé 48 séquences, 9 séquences de la Chine (1099pb à 1140pb), nous avons pris 2 séquences représentatives de taille de 1140pb. 7 séquences de Kazakhstan (674pb à 1140pb) nous avons pris 2 séquences représentatives de taille de 1140pb, 31 séquences de Mongolie (613pb à 1140pb), on a choisi 4 séquences, 2

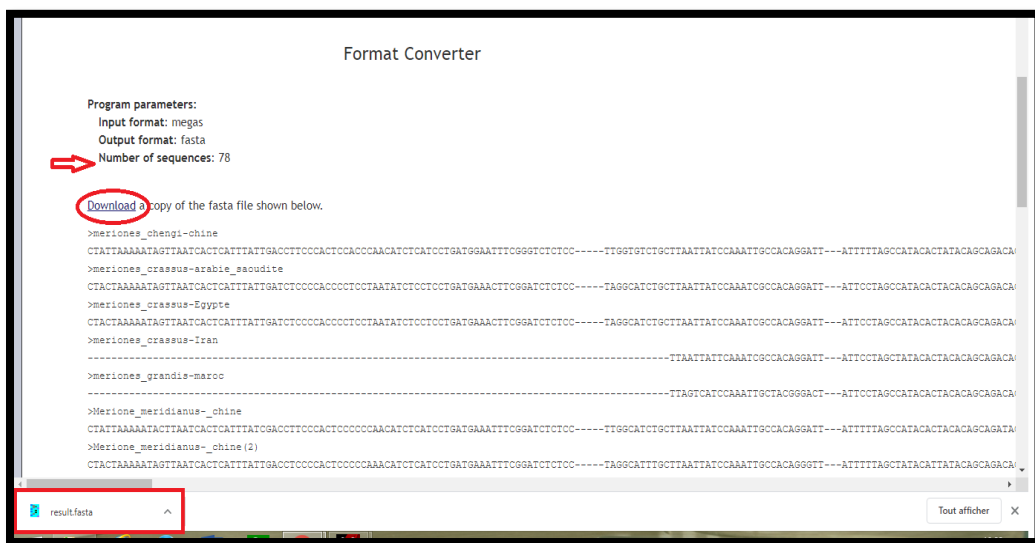
séquences de taille de 1140pb et une séquence de taille de 1128pb et une séquence de taille 1115pb.

- *M. persicus* : Nous avons trouvé 03 séquences au total qui sont d'origine Iranienne de taille de 533pb, 712 pb et 877pb.
- *M. Rex* : Il y'a 2 séquences de l'origine d'Arabie Saoudite de taille de 1140pb et 737pb, nous avons choisi la séquence de taille de 1140pb.
- *M. shawi* : Nous avons trouvé 5 séquences, 3 séquences de Tunisie (1140pb, 1170pb et 737pb), nous avons choisi 2 séquences de taille de 1170pb et 1140pb, une seule séquence du Maroc (1026pb), une seule séquence de taille de (971pb).
- *M. tristrami* : Nous avons trouvé 5 séquences, 4 séquences de l'Iran (878pb ,877pb et 737pb), une seule séquence de Jordanie (1133pb).
- *M. tristrami blackleri* : Il y'a 2 séquences d'origine Turque, l'une de taille de 790 pb et l'autre de 799 pb, nous avons choisi une seule séquence de 799pb.
- *M. tristrami bodenheimeri* : Il y'a une séquence de taille de 789pb.
- *M. tristrami bogdanovi* : On a trouvé une séquence de taille de 789pb.
- *M. tristrami intraponticus* : Il y'a 3 séquences de taille de 795 pb, 810 pb et 795pb, nous avons choisi une seule séquence représentative de taille de 810pb.
- *M. tristrami kilisensis* : 5 séquences de taille de (792pb à 795pb), nous avons choisi la séquence de taille de 792pb.
- *M. tristrami Lycaon* : Au total, il y'a 4 séquences de taille de (857pb à 795pb), on a choisi une séquence de taille de 795pb.
- *M. tamariscinus* : 37 séquences. 33 séquences de taille 1140 pb (on prend une seule séquence représentative), 4 séquences d'origine chinoise (1113 pb, 1104 pb, 1090pb et 1005pb).
- *M. vinogradovi* : On a trouvé 2 séquences d'origine iranienne, une de taille de 878 pb et l'autre de taille de 737 pb.
- *M. sp. Garat An Njila* : Il y'a une seule séquence, d'origine algérienne de taille de 1140pb.
- *M. sp. Hk-2016* : Il y'a une seule séquence d'origine tunisienne de taille de 1140 pb.
- *M. sp. n. MAO-2018* : Il y'a 3 séquences identiques d'origine iranienne de même taille de 533pb. On a présenté ces séquences par une seule séquence.

- *M. unguiculatus* : Au total il y'a 40 séquences de différentes tailles, nous avons choisi (1623. 1745. 1781. 1725. 1675. 2018. 1759. 1196. 2029. 2233. 1820. 2632. 1184. 1069.1844. 2647. 2819. 2626. 1544. 610. 1166. 1116. 1242. 1015. 1088. 1140)pb.

Nous allons étudier ici 78 séquences sélectionnées de Cytb pour faire l'analyse phylogénétique.

Nous avons utilisé le FaBox pour convertir notre fichier de séquences sous le format Fasta afin d'organiser les bases azotiques qui forme une séquence d'une seule ligne (Figure 30).



**Figure 30** : Les séquences de Cytb ont été convertit sous la forme Fasta à partir de bloc note grâce à FaBox.

L'arbre phylogénétique a été construit avec la méthode UPGMA. La distance utilisée est celle du Kimura-2p (= 2 paramètres) qui intègre le fait que les transitions et les transversions ne sont pas équiprobables lors de l'évolution des séquences, notamment pour l'ADN mitochondrial où les transitions peuvent représenter 90% des mutations (**Kimura, 1980**).

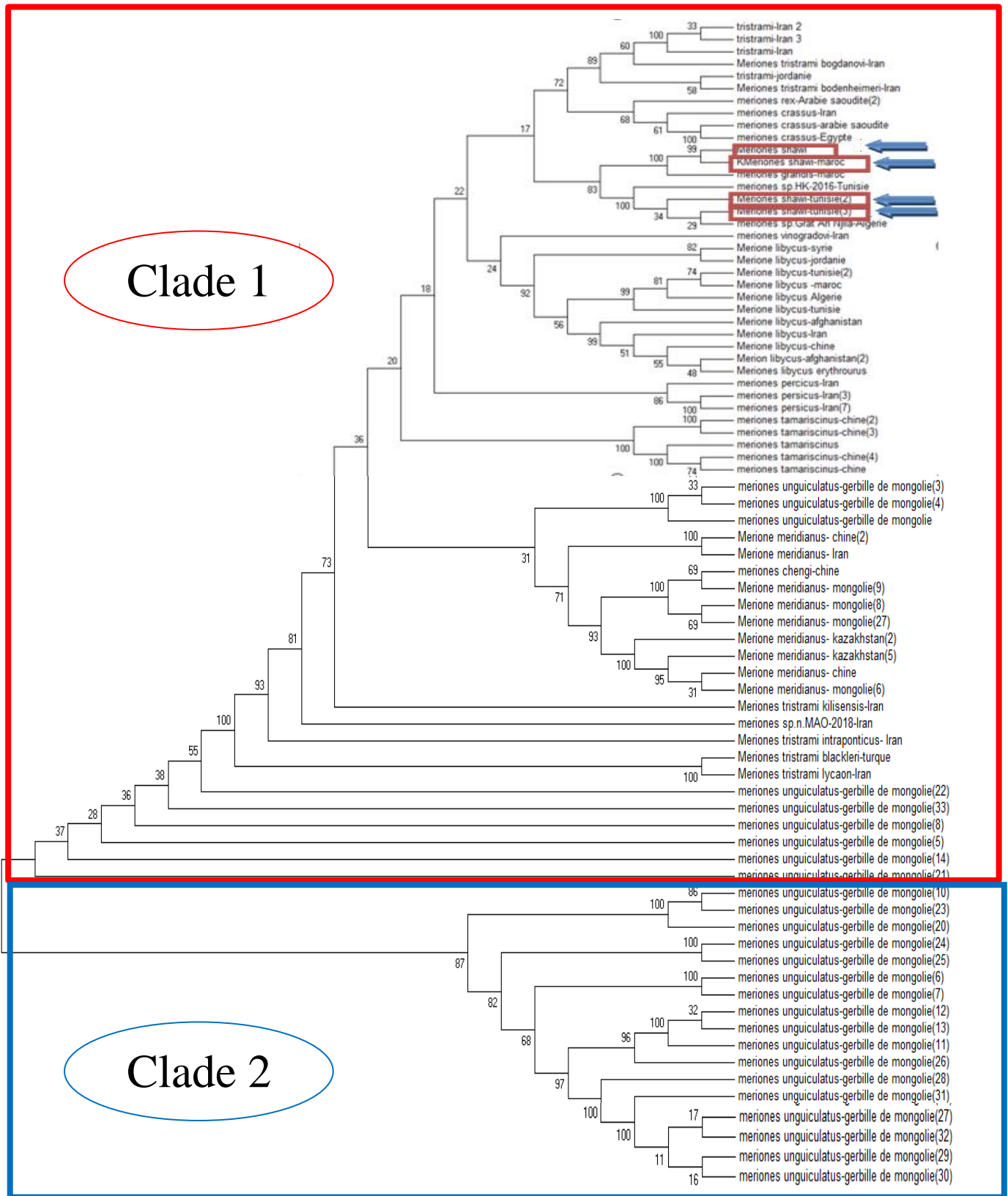
Dans la réalité biologique, les transitions représentent la majorité des substitutions nucléotidiques. Ainsi, lorsque la distance génétique entre deux séquences augmente, le nombre de transitions et de transversions augmente proportionnellement, le nombre de transitions étant toujours supérieur.

Cependant, dans le cas de séquences de plus en plus éloignées, la saturation des substitutions peut être atteinte et les transversions devenir plus nombreuses que les transitions.

Il faut 1000 répliquions de "Bootstrap" pour que cette méthode soit statistiquement valable.

Les relations entre les séquences de Cytb de genre *Meriones* ont été représentées à l'aide d'un arbre réalisé grâce au programme MEGA5.

Le but de visualiser graphiquement est de voir si une proximité génétique existait entre les individus au sein d'un même groupe.



Clade 1

Clade 2

**Figure 31 :** Arbre phylogénétique reconstruit par la méthodes UPGMA basé sur 78 séquences de nucléotides de Cytochrome b de genre *Meriones* en utilisant MEGA5.

Une reconstruction phylogénétique est alors faite à partir de chaque alignement chimérique généré, et la valeur associée à chaque nœud dans l'arbre correspond au nombre de fois sur 1000 répliques où cette même topologie a été trouvée.

D'après l'arbre (Figure 31), les séquences de Cytb se répartissent en 2 clades principaux :

**Clade 1** : Il comprend différentes espèces dont : *M. chengi*, *M. crassus*, *M. grandis*, *M. libycus*, *M. libycus erythrourus*, *M. meridianus*, *M. persicus*, *M. Rex*, *M. shawi*, *M. tamariscinus*, *M. tristrami*, *M. tristrami blackleri*, *M. tristrami bodenheimeri*, *M. tristrami bogdanovi*, *M. tristrami intraponticus*, *M. tristrami kilisesis*, *M. tristrami lycaon*, *M. unguiculatus*, *M. vinogradovi*, *M.sp. Garat An Njila*, *M. sp. HK-2016*, *M. sp. n. MAO-2018*. Ce clade se divise en deux sous clades :

\***Sous clade A** : Il présente : *M. chengi*, *M. crassus*, *M. grandis*, *M. libycus*, *M. libycus erythrourus*, *M. meridianus*, *M. persicus*, *M. Rex*, *M. shawi*, *M. tamariscinus*, *M. tristrami*, *M. tristrami blackleri*, *M. tristrami bodenheimeri*, *M. tristrami bogdanovi*, *M. tristrami intraponticus*, *M. tristrami kilisesis*, *M. tristrami lycaon*, *M. unguiculatus*, *M. vinogradovi*, *M.sp. Garat An Njila*, *M. sp. HK-2016*, *M. sp. n. MAO-2018*.

\***Sous clade B** : Il ne présente que l'espèce *M. unguiculatus*.

**Clade 2** : Il comprend 17 spécimens de *M. unguiculatus* ce clade est divisé en deux sous-clade :

\***Sous clade C** : Il représente 3 spécimens *M. unguiculatus*.

\***Sous clade D** : Il représente 14 spécimens *M. unguiculatus*.

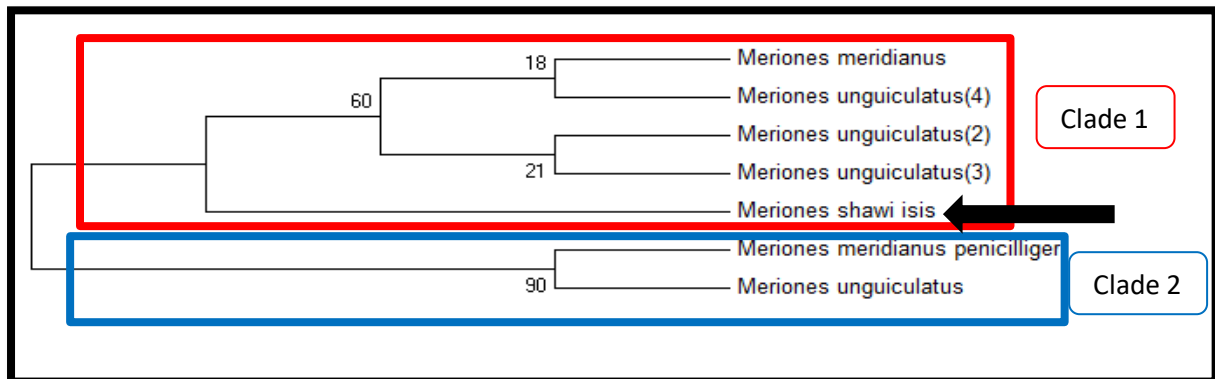
### III.2.1.2. Avpr2 :

- a) *M. meridianus* : Il y a une seule séquence de taille de 1064pb.
- b) *M. meridianus penicilliger* : Nous avons trouvé une seule séquence de Mongolie de taille de 911pb.
- c) *M. shawi isis* : Nous avons trouvé 3 séquences d'origine du Maroc. Une de taille 396pb et 2 de taille de 353pb.

**NB** : On a supprimé les deux séquences de taille de 353pb de *Meriones shawi isis* parce qu'elles sont très courtes par rapport aux autres séquences.

- d) *M. unguiculatus* : On a 4 séquences de différentes tailles (1043pb, 1872pb, 1629pb et 1632pb).

L'arbre phylogénétique de genre *Meriones* a été réalisé à partir des séquences de gène Avpr2 de l'ADN nucléaire (Figure 32).



**Figure 32 :** Arbre phylogénétique des séquences d'Avpr2 du genre *Meriones* en utilisant la méthode UPGMA par logiciel MEGA5.

D'après l'arbre (Figure 32), les séquences d'Avpr2 se répartissent en 2 clades principaux :

**Clade 1 :** Il comprend la majorité des espèces : *Meriones unguiculatus*, *Meriones meridianus*, *Meriones shawi isis*. Ce clade est subdivisé en 2 sous clades :

**\*Sous clade A :** Il comprend 3 spécimens de *M. unguiculatus* et une espèce de *M. meridianus*.

**\*Sous clade B :** Il n'est représenté que par la seule sous-espèce *M. shawi isis*.

**Clade 2 :** Il comprend *M. meridianus pencilliger* et *M. unguiculatus*.

### III.2.2. Interprétation des relations phylogénétiques entre les différentes espèces du genre *Meriones* et *M. shawi* :

D'après les arbres globaux (de Cytb et Avpr2) du genre *Meriones*, nous avons deux clades importants :

(1) **Clade Asiatique-Africain :** *M. chengi*, *M. crassus*, *M. grandis*, *M. libycus*, *M. libycus erythrourus*, *M. meridianus*, *M. persicus*, *M. Rex*, *M. shawi*, *M. shawi isis*, *M. tamariscinus*, *M. tristrami*, *M. tristrami blackleri*, *M. tristrami bodenheimeri*, *M. tristrami bogdanovi*, *M. tristrami intraponticus*, *M. tristrami kilisesis*, *M. tristrami lycaon*, *M. unguiculatus*, *M. vinogradovi*, *M.sp. Garat An Njila*, *M. sp. HK-2016*, *M. sp. n. MAO-2018*.

(2) **Clade Asiatique :** *M. meridianus pencilliger*, *M. unguiculatus*.

Les *Meriones* ont une répartition qui s'étend de l'Afrique du Nord à la Mongolie ; certaines espèces ont une immense aire de distribution (*M. libycus* s'étend du Sahara marocain jusqu'en Afghanistan), tandis que les autres ont une aire plus restreinte. Parmi les rongeurs d'Afrique du Nord, *Meriones shawi* constitue l'une des espèces les plus abondantes de son aire de répartition et elle est présente dans tous les pays du sud du Maroc à l'Égypte (**Aulagnier, 2012**).

Les résultats obtenus de l'arbre phylogénique ont montré que *M. shawi* très proche de *M. grandis* ; d'après **Cabrera (1907)** l'espèce *Meriones grandis* a été décrite comme semblable à *M. shawi*, mais de taille beaucoup plus grande, avec une coloration rouge au niveau du cou et du ventre et une ligne sombre tout au long de la queue. Le crâne de *M. grandis* est long et de largeur faible par rapport à celui de *M. shawi*. *M. grandis* a cependant depuis été mise à plusieurs reprises en synonymie avec *M. shawi*. Une récente étude génétique moléculaire et des populations a décrit deux clades moléculaires proches au Maroc et un en Algérie et ces clades ont été considérés respectivement comme *M. grandis* au Maroc et *M. shawi* en Algérie-Tunisie (**Pavlinov, 2000**). Les deux principaux clades étant sympatriques en la région orientale du Maroc près de la frontière algérienne (**Musser et Carleton, 2000**).

Nous avons trouvé dans l'arbre phylogénique que *Meriones shawi* est proche de *Meriones crassus*, pour vérifier ces résultats et puisque nous n'avons pas trouvé d'études moléculaires nous avons comparé entre ces deux espèces sur le plan morphologique et reproductif.

**Morphologique :** *Meriones crassus* (Figure 33) se caractérise par un pelage doux de teinte fauve clair sur le dos, blanche sur le ventre, la queue de même longueur que le corps, est terminée par un pinceau de poils foncés peu développés. Les soles plantaires sont couvertes de poils blancs, les griffes sont également blanches, les bulles tympaniques sont hypertrophiées c'est par comparaison avec *M. shawi* qui est l'espèce la plus grande du genre, sa queue de même longueur que le corps est terminée par un pinceau réduit, le pelage dorsal varie du fauve au gris piqueté de brun, le pelage ventral est blanc, les ongles sont clairs.





**Figure 33 :** *Meriones crassus* (Ahmim, 2019).

**Reproduction :** La période de reproduction de Merione de shaw commence en fin d'hiver et atteint le maximum au printemps et la durée de gestation est de l'ordre de 20 jours où le nombre de petits est de 5 à 12 par portée par contre la reproduction chez *Meriones crassus* s'observe toute l'année et aussi la gestation dure 21 jours. Les portées comprennent en moyenne 4 jeunes (1 à 5) qui deviennent indépendants vers l'âge d'un mois.

D'après l'arbre phylogénique. Nous avons remarqué que *Meriones shawi* est un peu loin à *Meriones libycus* (Figure 34). Pour vérifier ces résultats nous avons comparé entre ces deux espèces sur le plan morphologique, reproductif et nutritif.

**Morphologique :** Merione de shaw est l'espèce la plus grande du genre, sa queue de même longueur que le corps est terminée par un pinceau réduit, le pelage dorsal varie du fauve au gris piqueté de brun, le pelage ventral est blanc, les ongles sont clairs cependant que *Meriones libycus* se reconnaît par la coloration ocrée et rougeâtre de la base de sa queue terminée par un pinceau de poils noirs. Le pelage dorsal est gris fauve, piqueté de gris, le pelage ventral est blanc. Les ongles sont de couleur grisâtre, parfois noire, les bulles tympaniques sont très développées.



**Figure 34 :** *Meriones libycus* (Ahmim, 2019).

**Nourriture** : Le régime alimentaire de *Merione* de Shaw est très varié ; elle consomme des fruits et légumes au printemps, des céréales en automne et lorsque ces sources de nourriture sont taries en hiver, elle peut consommer des insectes tandis-que *Meriones libycus* se nourrit essentiellement de végétaux : graines, fruits.

**Reproduction** : La reproduction de *Meriones libycus* a lieu d'avril à septembre. La gestation dure 20 à 25 jours on compte 2 à 7 jeunes par portée. Ils restent aveugles 12 à 17 jours. La période de reproduction de *Meriones shawi* commence en fin d'hiver et atteint le maximum au printemps. La durée de gestation est de l'ordre de 20jours. Le nombre de petits est de 5 à 12 par portée.

Ces comparaisons entre ces espèces confirment les résultats de l'arbre phylogénique.

Selon notre résultat ; *Meriones shawi* est proche de *Meriones crassus* et *Meriones libycus* ; *M. shawi* est précisément limitée vers le sud par le désert dont elle ne supporte pas les conditions de vie. Elle y est remplacée par les deux autres espèces, *M. libycus* habitant de préférence les buttes de sable qui sont constituées par le vent au pied de la végétation buissonnante et *M. crassus*, dont les terriers sont au contraire creusés dans des biotopes totalement dépourvus de végétation (Petter, 1961).

### III.3. Détermination des haplotypes :

Nous avons utilisé le DnaSP pour la détermination des haplotypes ; le terme haplotype correspond à une séquence nucléotidique, qui peut être commune à plusieurs individus, mais diffère des autres haplotypes par une ou plusieurs substitutions de nucléotides.

#### III.3.1. Résultats du DnaSP pour le Cytb :

Nous avons utilisé en ligne le logiciel FaBox pour faire l'alignement de la séquence la plus courte. Elle supprime simplement les zones frontalières qui sont pleines de gaps pour le résultat d'alignement dans le logiciel DnaSP (Figure 35).



**Figure 35** : Résultat d'alignement des séquences Cytb du genre *Meriones* en utilisant le logiciel FaBox (Rouge = gaps : La présence des gaps peut être expliquée par 2 possibilités : soit la délétion des bases azotées de séquences qui contiennent les gaps, ou bien l'insertion des bases dans les séquences qui ne contiennent pas de gaps ; Vert = conservés : les zones de similarité entre les séquences).

L'analyse de Cytb de l'ADN mitochondrial révèle une diversité génétique très importante ; les 78 individus étudiés ont permis de caractériser 59 haplotypes différents. 47 haplotypes privés ont été détectés ; un haplotype privé est un haplotype qui ne se retrouve que dans une seule population. Les diversités haplotypique et nucléotidique sont donc très élevées au sein de notre échantillonnage pourtant restreint ( $H_d = 0,9903$ ) (Figure 36).

```

Output of: cropped_alignment.fasta
Haplotype/DNA Sequences Data File

Input Data File: C:\...cropped_alignment.fasta
Number of sequences: 78   Number of sequences used: 78
Selected region: 1-453   Number of sites: 453
Total number of sites (excluding sites with gaps / missing data): 299

Sites with alignment gaps: not considered
Number of variable sites: 299

===== Haplotype Distribution =====
Number of haplotypes, h: 59
Haplotype diversity, Hd: 0.9903

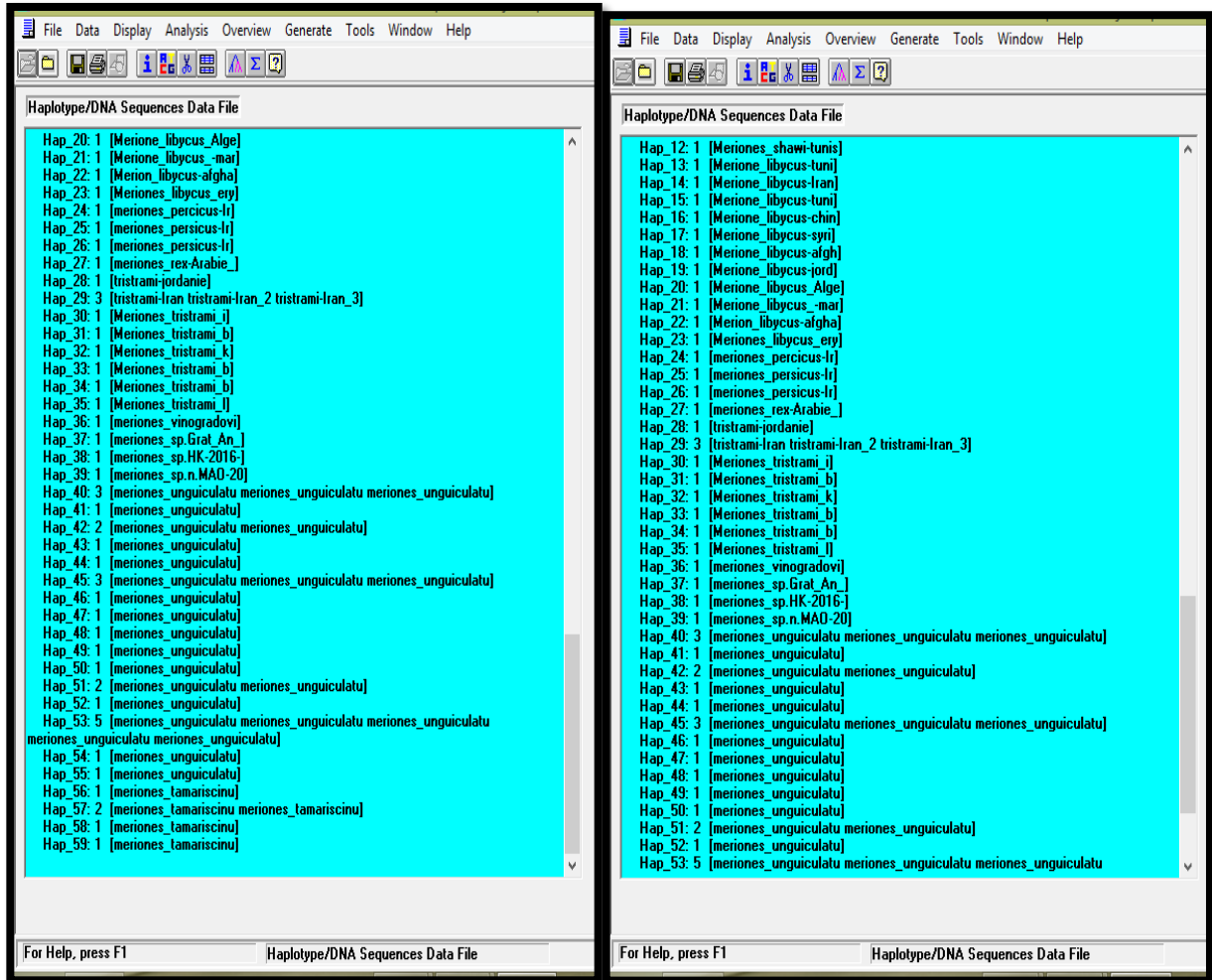
Hap_1: 2 [1 13]
Hap_2: 2 [2-3]
Hap_3: 1 [4]
Hap_4: 1 [5]
Hap_5: 3 [6 10-11]
Hap_6: 1 [7]
Hap_7: 1 [8]
Hap_8: 1 [9]
Hap_9: 2 [12 14]
Hap_10: 2 [15 18]
Hap_11: 1 [16]
Hap_12: 1 [17]
Hap_13: 1 [19]
Hap_14: 1 [20]
Hap_15: 1 [21]
Hap_16: 1 [22]
Hap_17: 1 [23]
Hap_18: 1 [24]
Hap_19: 1 [25]
Hap_20: 1 [26]
Hap_21: 1 [27]
Hap_22: 1 [28]
Hap_23: 1 [29]
Hap_24: 1 [30]
Hap_25: 1 [31]
Hap_26: 1 [32]
Hap_27: 1 [33]
Hap_28: 1 [34]
Hap_29: 3 [35-37]
    
```

```

Haplotype/DNA Sequences Data File

Hap_30: 1 [38]
Hap_31: 1 [39]
Hap_32: 1 [40]
Hap_33: 1 [41]
Hap_34: 1 [42]
Hap_35: 1 [43]
Hap_36: 1 [44]
Hap_37: 1 [45]
Hap_38: 1 [46]
Hap_39: 1 [47]
Hap_40: 3 [48-50]
Hap_41: 1 [51]
Hap_42: 2 [52-53]
Hap_43: 1 [54]
Hap_44: 1 [55]
Hap_45: 3 [56-58]
Hap_46: 1 [59]
Hap_47: 1 [60]
Hap_48: 1 [61]
Hap_49: 1 [62]
Hap_50: 1 [63]
Hap_51: 2 [64-65]
Hap_52: 1 [66]
Hap_53: 5 [67 69-72]
Hap_54: 1 [68]
Hap_55: 1 [73]
Hap_56: 1 [74]
Hap_57: 2 [75-76]
Hap_58: 1 [77]
Hap_59: 1 [78]

Hap_1: 2 [meriones_chengi-chin Merione_meridianus_]
Hap_2: 2 [meriones_crassus-ara meriones_crassus-Egy]
Hap_3: 1 [meriones_crassus-Ira]
Hap_4: 1 [meriones_grandis-mar]
Hap_5: 3 [Merione_meridianus_ Merione_meridianus_ Merione_meridianus_]
Hap_6: 1 [Merione_meridianus_]
Hap_7: 1 [Merione_meridianus_]
Hap_8: 1 [Merione_meridianus_]
Hap_9: 2 [Merione_meridianus_ Merione_meridianus_]
Hap_10: 2 [Meriones_shawi KMeriones_shawi-mar]
Hap_11: 1 [Meriones_shawi-tunis]
    
```



**Figure 36 :** Les haplotypes des séquences Cytb du genre *Meriones* en utilisant le logiciel DnaSP.

○ **47 haplotype privé :**

*M. crassus* Iran, *M. grandis* Maroc, *M. meridianus* Chine, *M. meridianus* Iran, *M. meridianus* Kazakhstan, *M. shawi* Tunisie (02), *M. libycus* Tunisie (02), *M. libycus* Iran, *M. libycus* Chine, *M. libycus* Syrie, *M. libycus* Afghanistan (02), *M. libycus* Jordanie, *M. libycus* Algérie, *M. libycus* Maroc, *M. libycus erythrourus* Chine, *M. persicus* Iran (03), *M. rex* Arabie saoudite, *M. tristrami* Jordanie, *M. tristrami lycaon* Iran, *M. tristrami intraponticus* Iran, *M. tristrami bogdanovi* Iran, *M. tristrami kilisensis* Iran, *M. tristrami blackleri* Turquie, *M. tristrami bodenheimeri* Iran, *M. tristrami vinogradovi* Iran, *M. sp. Garat An Njila* Algérie, *M. sp. Hk-2016* Tunisie, *M. sp. n. MAO-2018* Iran, *M. unguiculatus* Mongolie (11), *M. tamariscinus* Chine (03).

○ **7 haplotype comporte 2 séquences :**

**Hap 1 :** *M. Chengi-chine* et *M. meridianus* Mongolie.

**Hap 2 :** *M. crassus* Arabie saoudite et *M. crassus* Egypte.

**Hap 9:** *M. meridianus* Mongolie (02).

**Hap 10:** *M. shawi* Kazakhstan et *M. shawi* Maroc.

**Hap 42:** *M. unguiculatus* Mongolie (02).

**Hap 51:** *M. unguiculatus* Mongolie (02).

**Hap 57:** *M. tamariscinus* Chine (02).

○ **4 haplotype comporte 3 séquences :**

**Hap 5 :** *M. meridianus* Chine, *M. meridianus* Kazakhstan, *M. meridianus* Mongolie.

**Hap 29 :** *M. tristrami* Iran (03).

**Hap 40 :** *M. unguiculatus* Mongolie (03).

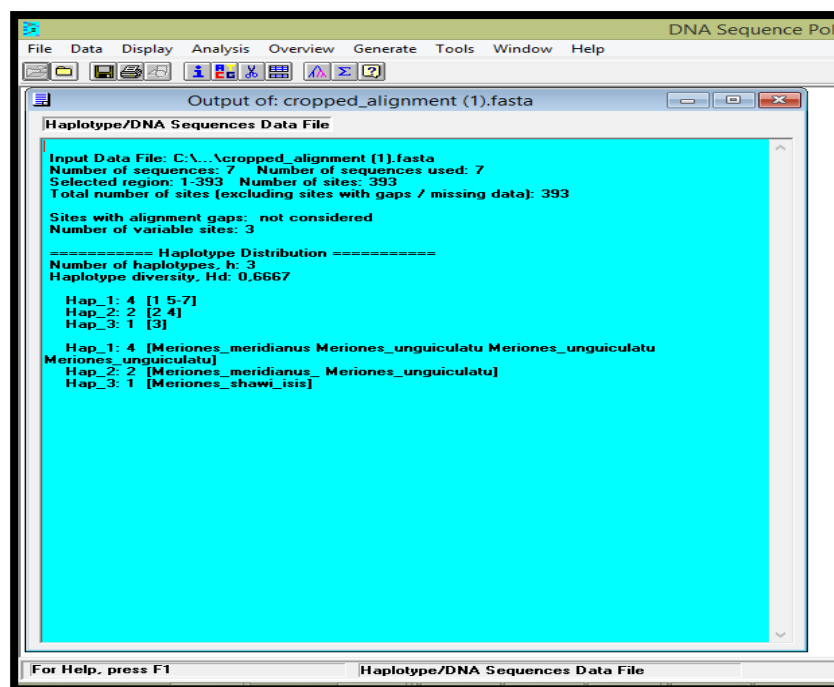
**Hap 45 :** *M. unguiculatus* Mongolie (03).

○ **Un haplotype comporte 5 séquences :**

**Hap 53 :** *M. unguiculatus* Mongolie (05).

### III.3.2. Résultats du DnaSP pour l'Avpr2 :

Sur les 7 séquences d'Avpr2 de genre *Meriones* obtenues, 3 haplotypes ont pu être identifiés, 1 haplotype privé a été détecté (Figure 37).



**Figure 37:** Les haplotypes des séquences d'Avpr2 du genre *Meriones* en utilisant le logiciel DnaSP.

**Hap 1 :** *M. meridianus* et *M. unguiculatus* (03).

**Hap 2 :** *M. meridianus pencilliger* et *M. unguiculatus*.

**Hap 3 :** *M. shawi isis*.

**Tableau III :** Comparaison des indices de diversités calculés avec DnaSP sur les échantillons étudiés.

Population	Clade	Taille de l'échantillon	Nombre d'haplotypes h	Diversité haplotypique Hd	Nombre de sites polymorphes S	Le nombre de site variable
Cytb	1	78	47	0.9903	453	299
	2		12			
Avpr2	1	7	1	0.6667	393	3
	2		2			

### III.3.3. Interprétation des résultats de DnaSP :

Nous avons remarqué que les résultats obtenus par le DnaSP confirment les résultats de MEGA5 ; les relations entre individus, haplotypes ou espèces sont visualisées précédemment via la construction des arbres, par agglomération des plus proches voisins (Saitou *et al.*, 1987). Sachant que plus la diversité haplotypique est élevée au sein de l'échantillonnage, plus il y a de chances d'observer des haplotypes différents lorsque l'on sélectionne 2 individus au hasard (Nei, 1987).

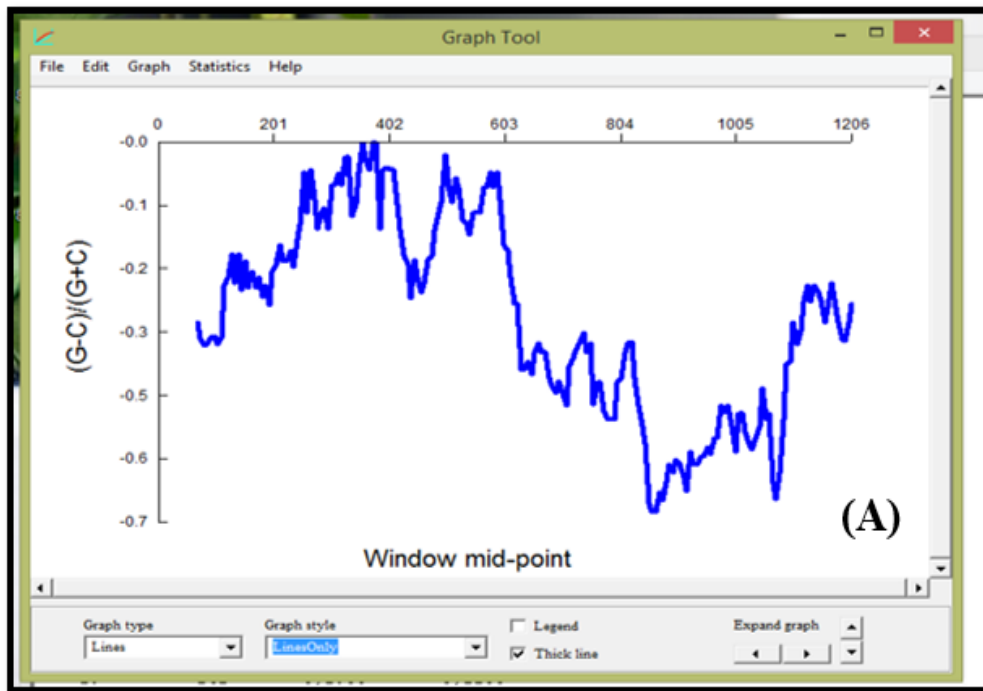
### III.4. Analyses statistiques :

La méthode utilisée par le logiciel DAMBE est basée sur une analyse de la fréquence de GC et AT. Le logiciel GenSkew calcule la normale et le skew (biais) cumulatif de deux nucléotides sélectionnables pour une séquence donnée. Le résultat est montré dans deux graphiques différents (Figures 38, 39). Le degré d'asymétrie compositionnelle, exprimé en termes de biais de GC et d'AT, peut être calculé à l'aide des formules suivantes (Perna *et al.*, 1995) :

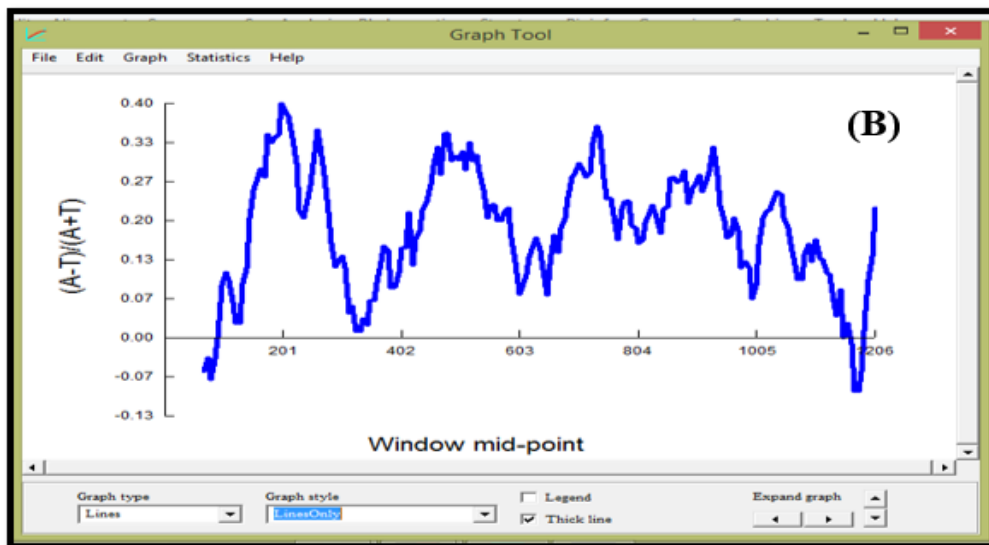
$$\text{Skew} = (\text{nucleotide1} - \text{nucleotide2}) / (\text{nucleotide1} + \text{nucleotide2})$$

- GC skew =  $(G - C) / (G + C)$ .
- AT skew =  $(A - T) / (A + T)$ .

Les méthodes comme GC skew, CGC skew, et Z-curve sont des outils pour mieux étudier le mécanisme de la réplication de l'ADN dans différents organismes.



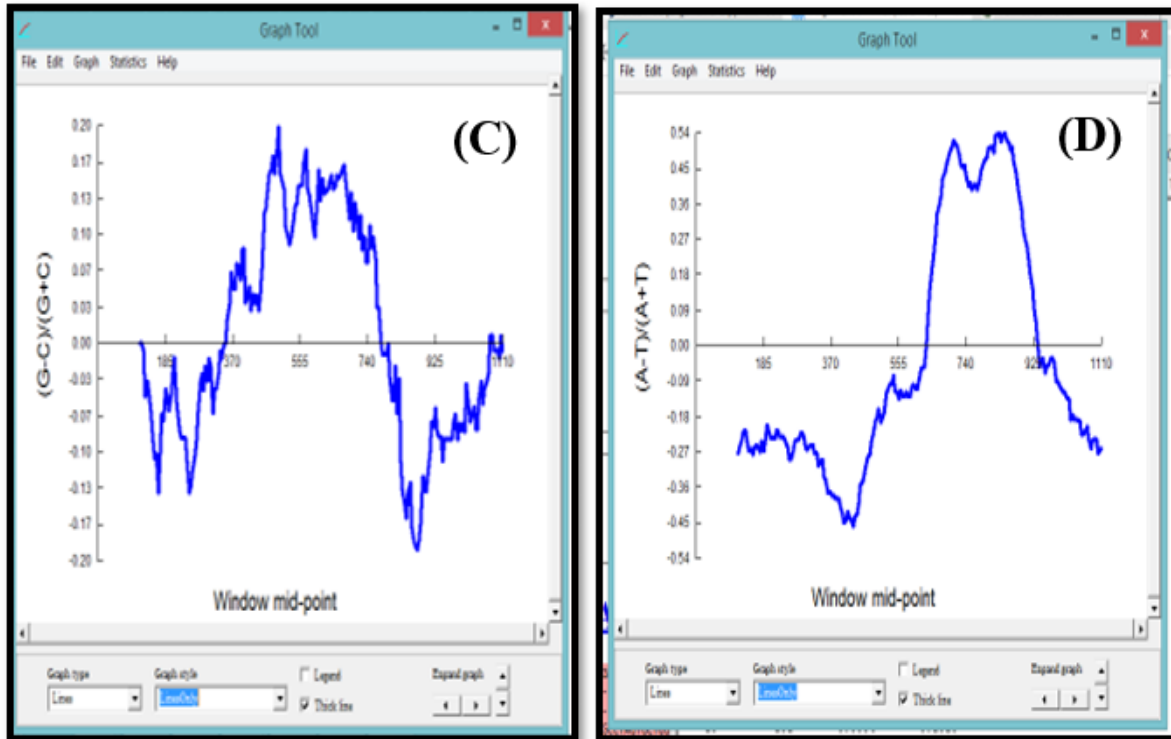
**Figure 38 (A) :** Variations des fréquences GC de Cytb produites à partir de DAMBE.



**Figure 38 (B) :** Variations des fréquences d'AT de Cytb produites à partir de DAMBE.

Nous avons remarqué que le GC skew est négatif (Figure 38 A) cependant l'AT skew est positif (Figure 38 B), sachant que, GC skew positif représente la richesse de G sur C et le GC skew négatif représente la richesse de C sur G (Tillier *et al.*, 2000).



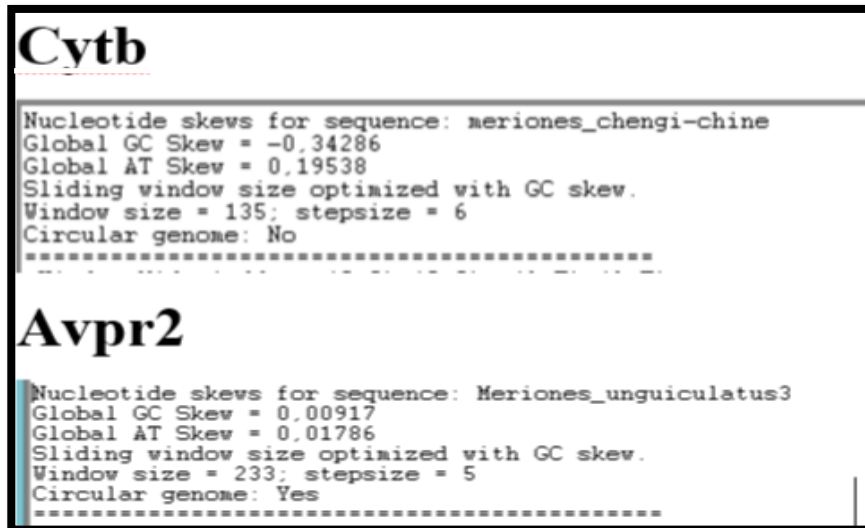


**Figure 39** : Variations des fréquences GC (C) et d'AT (D) d'Avpr2 produites à partir de DAMBE.

Nous avons remarqué que GC Skew est positif (Figure 39 C) et aussi que l'AT Skew est positif (Figure 39 D).

Plus l'ADN est riche en paires G/C plus l'ADN résiste à la dénaturation par l'augmentation de la température. Le taux de G+C d'une molécule d'ADN est la fréquence relative, exprimé généralement en pourcentage. (**Genetik et al., 2006**). Les segments d'ADN composés de plusieurs bases G-C sont plus stables que les séquences composées de plusieurs bases A-T.

A, T, G et C représentent la fréquence d'occurrence de l'équivalent de la base dans une séquence particulière d'une longueur définie (**Lobry et al., 1996**). Nous avons comparé les fréquences de CG et AT des gènes Cytb et Avpr2 (Figure 40).



```
Cytb
-----
Nucleotide skews for sequence: meriones_chengi-chine
Global GC Skew = -0,34286
Global AT Skew = 0,19538
Sliding window size optimized with GC skew.
Window size = 135; stepsize = 6
Circular genome: No
-----

Avpr2
-----
Nucleotide skews for sequence: Meriones_unguiculatus3
Global GC Skew = 0,00917
Global AT Skew = 0,01786
Sliding window size optimized with GC skew.
Window size = 233; stepsize = 5
Circular genome: Yes
-----
```

**Figure 40 :** Comparaison des fréquences entre Cytb et Avpr2 réalisés par le logiciel DAMBE.

Nous avons remarqué qu'il y a une richesse de la cytosine sur la guanine et de l'adénine sur la thymine pour Cytb, il y a une richesse de la guanine sur la cytosine et de l'adénine sur la thymine pour l'Avpr2, donc la fréquence d'Avpr2 est plus élevée par rapport au Cytb (Figure 40), c'est-à-dire le nombre de CG présent dans Avpr2 est plus élevé par rapport au Cytb, parce que cette dernière est moins conservée que l'Avpr2.

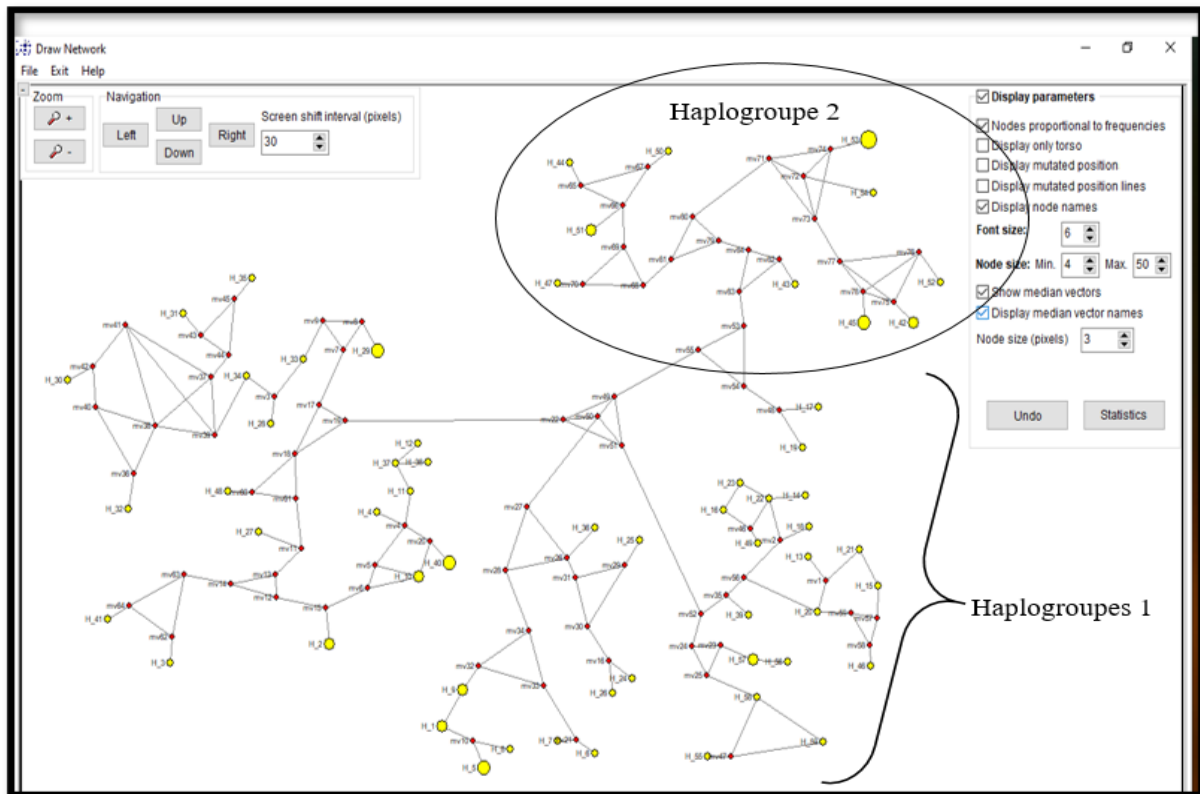
Le gène nucléaire Avpr2 code pour le récepteur-2 de l'arginine vasopressine (V2R), un récepteur couplé à la protéine G. Le gène mitochondrial du Cytb code pour une protéine membranaire intégrale.

En général les taux de mutation de l'ADN mitochondrial sont plus grands que ceux de l'ADN nucléaire (Ballard *et al.*, 2004).

### III.5. Résultat de NETWORK :

#### III.5.1. Résultat de Cytb :

Les haplotypes, utilisable pour des analyses de traçabilité, mais aussi pour étudier l'évolution de l'espèce. Les relations entre les haplotypes peuvent également être appréhendées par la construction de réseaux d'haplotypes. Ces réseaux d'haplotypes sont construits grâce au logiciel NETWORK 5.0 (Figure 41).



**Figure 41** : Réseau haplotypique réalisé par le logiciel NETWORK à partir de 78 séquences de Cytb de genre Meriones.

La construction de ces réseaux est basée sur la théorie de la coalescence (**Kingman, 2000**). Cette théorie consiste en une approche rétrospective qui décrit mathématiquement le processus de fusion binaire de tous les lignages généalogiques d'un échantillon de gènes jusqu'à leur plus proche ancêtre commun.

Dans le réseau d'haplotypes ; chaque cercle correspond à un haplotype ; la taille du cercle est proportionnelle à la fréquence de l'haplotype dans le jeu de données. La longueur des segments entre chaque haplotype est proportionnelle au nombre de mutations qui les sépare.

Sur les 78 séquences de Cytb de genre Meriones obtenues, 59 haplotypes ont pu être identifiés. La structure du réseau apparaît plus complexe et avec de nombreux haplotypes divergents.

- **Haplogroupe 1** : Correspond au groupe Asiatique-Africain : *M. chengi*, *M. crassus*, *M. grandis*, *M. libycus*, *M. libycus erythrourus*, *M. meridianus*, *M. persicus*, *M. Rex*, *M. shawi*, *M. tamariscinus*, *M. tristrami*, *M. tristrami blackleri*, *M. tristrami*

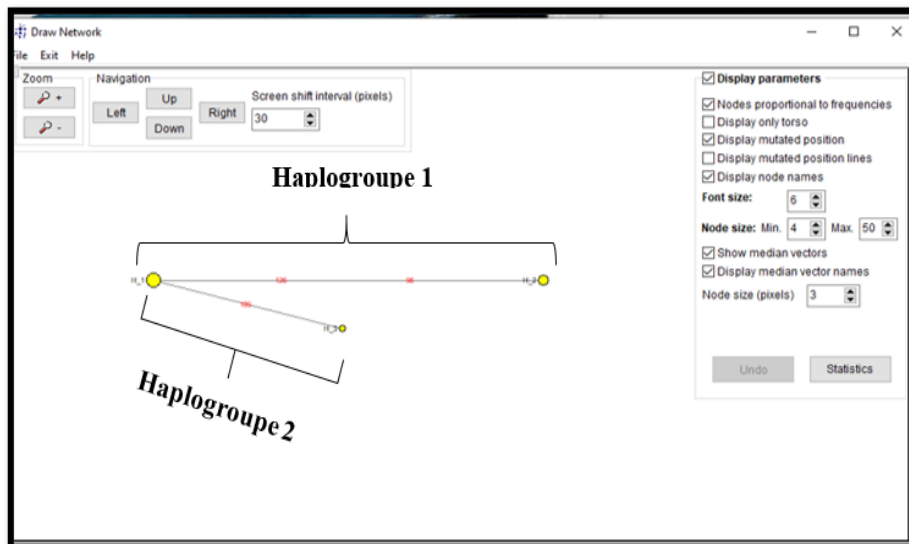
*bodenheimeri*, *M. tristrami bogdanovi*, *M. tristrami intraponticus*, *M. tristrami kilisisis*, *M. tristrami lycaon*, *M. unguiculatus*, *M. vinogradovi*, *M. sp. Garat An Njila*, *M. sp. HK-2016*, *M. sp. n. MAO 2018*. Pour les haplotypes suivants : H1. H2. H3. H4. H5. H6. H7. H8. H9. H10. H11. H12. H13. H14. H15. H16. H17. H18. H19. H20. H21. H22. H23. H24. H25. H26. H37. H28. H29. H30. H31. H32. H33. H34. H35. H36. H37. H38. H39. H40. H41. H46. H48. H49. H55. H56. H57. H58. H59.

- **Haplogroupe 2** : Correspond au groupe Asiatique : *M. unguiculatus* pour les haplotypes suivants : H42. H43. H44. H45. H47. H50. H51. H52. H53. H54.

### III.5.2. Résultat d'Avpr2 :

D'après la structure du réseau (Figure 42), les 07 séquences d'Avpr2 se répartissent en 2 haplogroupes principaux :

- ✓ **Haplogroupe 1** : Correspond au groupe Asiatique-Africain : *M. unguiculatus*, *M. meridianus*, *M. shawi Isis*, pour les haplotypes H2 et H1.
- ✓ **Haplogroupe 2** : Correspond au groupe Asiatique : *M. unguiculatus*, *M. meridianus pencilliger*, pour l'haplotypes H3.



**Figure 42** : La croissance des haplotypes des séquences d'Avpr2 du genre *Meriones* en utilisant le logiciel NETWORK.

Les résultats obtenus des réseaux haplotypiques des deux gènes que ce soit *Cytb* ou *Avpr2* ont montré l'existence de deux haplogroupes major qui représentent le clade d'Asiatique-Africain et le clade d'Asie.

Nous avons observé une distance génétique élevée entre les espèces de genre *Meriones*, cette croissance a lieu sur une période relativement restreinte. D'après les travaux de **Saitou *et al.* (1987)** plus deux séquences homologues sont différentes, plus la distance évolutive qui les sépare est grande, et donc plus les séquences ont divergé il y a longtemps ; à l'inverse, plus deux séquences homologues sont proches, plus la distance évolutive qui les sépare est faible.

A decorative border with a repeating geometric pattern of blue and grey zig-zags surrounds the page.

**CONCLUSION**

# Conclusion

Le genre *Meriones* est un genre de mammifères, des petits rongeurs de la famille des Muridae comportant 15 espèces (*M. chengi*, *M. crassus*, *M. grandis*, *M. libycus*, *M. unguiculatus*, *M. meridianus*, *M. persicus*, *M. Rex*, *M. shawi*, *M. tamariscinus*, *M. tristrami*, *M. vinogradovi*, *M. sp. Garat An Njila*, *M. sp. HK-2016*, *M. sp. n. MAO-2018*), et 8 sous-espèces (*M. libycus erythrourus*, *M. shawi isis*, *M. meridianus pencilliger*, *M. tristrami blackleri*, *M. tristrami bodenheimeri*, *M. tristrami bogdanovi*, *M. tristrami intraponticus*, *M. tristrami kilisesis*, *M. tristrami lycaon* existant en Afrique du Nord et en Asie.

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 espèces et pour comprendre l'évolution adaptative au niveau morphologique et moléculaire.

Les analyses phylogénétiques du genre *Meriones* obtenues dans ce travail ont été élaborées à l'aide de différents logiciels tel que, MEGA5 pour la reconstruction de l'arbre phylogénétique, DnaSP pour la détermination des haplotypes (une séquence nucléotidique, qui peut être commune à plusieurs individus, mais diffère des autres haplotypes par une ou plusieurs substitutions de nucléotides), NETWORK pour avoir l'évolution des espèces, et DAMBE pour faire des calculs statistiques.

Nous avons remarqué que les résultats obtenus par le DnaSP confirment les résultats de MEGA5, les analyses du polymorphisme mitochondrial fournissent des pistes intéressantes sur l'histoire de l'expansion géographique de chaque espèce. Les relations entre individus, haplotypes ou espèces sont visualisées précédemment via la construction des arbres, par agglomération des plus proches voisins.

Nous avons remarqué aussi que les résultats obtenus par le NETWORK confirment les résultats de MEGA5, ils ont montré l'existence de deux haplogroupes pour les gènes Cytb et Avpr2 qui représentent le clade Asiatique-Africain et le clade d'Asie.

Les méthodes d'analyses phylogénétiques, que ce soit en bifurcation, les haplotypes ou en réseau par les logiciels MEGA, DnaSP, NETWORK...., ont montré pour le genre *Meriones* l'existence de deux clades distincts qui sont le clade Asiatique-Africain qui inclut : *M. chengi*, *M. crassus*, *M. grandis*, *M. libycus*, *M. libycus erythrourus*, *M. meridianus*, *M. persicus*, *M. Rex*, *M. shawi*, *M. shawi isis*, *M. tamariscinus*, *M. tristrami*, *M. tristrami blackleri*, *M. tristrami bodenheimeri*, *M. tristrami bogdanovi*, *M. tristrami intraponticus*, *M. tristrami kilisisis*, *M. tristrami lycaon*, *M. vinogradovi*, *M. sp. Garat An Njila*, *M. sp.HK-2016*, *M. sp. n. MAO-2018* et le clade Asiatique qui inclut : *M. unguiculatus*, *M. meridianus pencilliger*.





# *REFERENCES BIBLIOGRAPHIQUES*

## Références bibliographiques

- **Adamou-Djerbaoui M., Denys C., Chaba H., Seid M. M., Djelaila Y. F., Labdelli**
- **Alia Z., Sekour M., Ould El Hadj M. D.** 2012- Importance des rongeurs dans le menu trophique de *Tyto alba* (SCOPOLI, 1759) dans la région de Souf (Algérie). *Revue des BioRessources* Vol 2 N°02.
- **Adamou-Djerbaoui M., Djelaila Y., Labdelli F., Adamou M. S.** 2011- Ecologie et infestation de *Meriones Meriones shawii* DUVERNOY 1842 dans la région de Tiaret. *Revue d'Ecologie et Environnement Univ de Tiaret* N°07.
- **Aulagnier S. L., Granjon G., Amori R., Huttere B., Krystufek N., Yigit G., Mitsain.** 2008- *Meriones shawi*, in: IUCN 2012. IUCN Red List of Threatened Species.
- **Aulagnier S., Thevenot M.** 1986- Catalogue des mammifères sauvages du Maroc. *Travx. Insti. Scie., sér. Zool., Rabat*, 164 p.
- **Aulagnier S., Thevenot M.** 1986- Catalogue des mammifères sauvages du Maroc. Institut Scientifique. Rabat, 116-117-151 p.
- **Avise J. C. S., Arnold J., Ball M. R., Bermingham E., Lamb T., Neigel J. E., Reeb C. A., Saundres N. C.** 1987- "Intraspecific phylogeography: The mitochondrial DNA Bridge Between population genetics and systematics." *Annual review of ecology and systematics*, 489-522 p.
- **Bachar M. F., Belhamra M.** 2012- Contribution à l'étude de la dynamique des populations des rongeurs sauvages dans la zone de Biskra. *Courrier du Savoir – N°13*, 71-81 p.
- **Ballard J. W. O., Whitlock M. C.** 2004- "The incomplete natural history of mitochondria." *Molecular Ecology* 13, 729-744 p.
- **Barabasi A., Gulbahce N., Loscalzo J.** 2011- Network medicine: a network-based approach to human disease. *12(1)*, 56-68 p.
- **Beaulieu M. E.** 2007- Caractérisation moléculaire des champignons ophiostomatoïdes associés à quatre espèces de scolytes de l'écorce colonisant l'épinette blanche au Québec et phylogénie multigénique d'une nouvelle espèce de leptographium. Thèse doctorat, Université Laval, Québec.
- 
-

## Références bibliographiques

- **Bergmans W.** 1997- Les rongeurs du Bénin : Espèces trouvées et espèces attendues pour des raisons biogéographiques (Mammalia, Rodentia). Séminaire national sur les rapports rongeurs/Ophidiens dans les agroécosystèmes du Bénin. 24-28 p.
- **Berman H. M., Westbrook J., Feng Z., Gilliland G., Bhat T. N., Weissig H., Shindyalov I. N., Bourne P. E.** 2000- The Protein Data Bank. *Nucleic Acids Res*, 28(1), 235–42 p.
- **Boore J. L.** 1999- “Animal mitochondrial genome”. *Nucleic acids Res* 27(8), 1767-1780 p.
- **Cabrera-Latorre A.** 1907- Meriones grandis, in Algunos roedores nuevos de Marruecos. *Boletín de la real sociedad española de historia natural*, 175– 177 p.
- **Chenchouni H.** 2012- Diversity Assessment of Vertebrate Fauna in a Wetland of Hot Hyperarid Lands. *Arid Ecosystems*, Vol. 2, N°4, 253–263 p.
- **Choi J. H., Jung H. Y., Kim H. S., Cho H. G.** 2000- A Phylogenetic Tree Drawing System. Department of computer science, Pusan National University, Pusan, Korea. Matériel et méthodes.
- **Dako G. E. A., Codja J. T. C., Bokonon Ganta A. H.** 2002- Evaluation de quelques paramètres corporels pour l’identification des petits rongeurs du Sud Bénin. Acte du séminaire sur la mammalogie et la biodiversité. Abomey-Calavi, Bénin du 30 octobre au 18 novembre, 41-54 p.
- **Darlu P., Tassy P.** 1993- La reconstruction phylogénétique. Concepts et méthodes. Paris, France.
- **Diallo A. B.** 2009- Bioinformatique avancée. Université du Québec à Montréal. Canada.
- **Duvernoy G. I.** 1841- Notes et renseignements sur plusieurs mammifères de l’Algérie. Institut Paris, (9), 400-401 p.
- **Geneti k., Jochen G.** 2006- Auflage, Springer Verlag Berlin Heidelberg.
- **Gervais P.** 1876 - Le genre cténodactyle. *J. Zool. Paris*, (5), 223-232 p.
- **Guerzou A., Boukraa S., Souttou K., Derdoukh W., Guerzou M., Sekour M., Baziz-Neffah F., Doumandji S.** 2012- Diet of Common Raven *Corvus Corax* (Aves, Corvidae) in Algeria. *Advances in Environmental Biology*, 7(7), 1288-1291 p.

## Références bibliographiques

- **Guindon S., Gascuel O.** 2003- A Simple, Fast, and Accurate Algorithm to Estimate Large Phylogenies by Maximum Likelihood (PHYML). *Syst. Biol.* 52(5), 696–704 p.
- **Hennig W.** 1966- Phylogenetic systematic. University of Illinois Press, Urbana.
- **Hulo N., Bairoch A., Bulliard V., Cerutti L., De Castro E., Langendijk-Genevaux P. S., Pagni M., Sigrist C. J.** 2006- The PROSITE database. *Nucleic Acids Res*, 34(Database issue), 227–30 p.
- **Judd W. S., Campbell C. S., Kellogg E. A., Stevens P.** 2002- Botanique systématique : une perspective phylogénétique. De Boeck Supérieur, 466 p.
- **Kimura M.** 1980- A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.*, 16, 111-120 p.
- **Kingman J. F.** 2000- Origins of the coalescent 1974-1982. *Genetics* 156, 1461-1463 p.
- **Kleppe K., Ohtsuka E., Kleppe R., Molineux I., Khorana H. G.** 1971- Studies on polynucleotides. XCVI. Repair replications of short synthetic DNA's as catalyzed by DNA polymerases. *J. Mol. Biol.*, 56, 341-361 p.
- **Kowalski K.** 1979- Terrestrial mammals of Algeria in the collections of the City Museum in Oran. *Afr. Small Mammal Newsl*, (4), 8-10 p.
- **Lataste F.** 1882- Mammifères nouveaux d'Algérie. *Naturaliste*, 4 (1.II.1882), 21-23 p, (15.II.1882), 27 p, (1.III.1882), 36p, (1. V.1882), 69-70 p, (15. V.1882), 77-78 p, (1.VI.1882), 83-85 p, (1.VII.1882), 101-102 p, (15.VIII.1882), 126-127 p.
- **Laurent P.** 1937- Mammifères observés dans la région de Gouraya et du Duperré, *Bull. Soc. Hist. Nat.Afr. N*, 28 (4-5), 226 p.
- **Le Berre M.** 1990- Faune du Sahara – Mammifères. Raymond Chabaud – Lechevalier, T. 2, 360 p.
- **Librado p., Rozas J.** 2009- DnaSP v5: a software for comprehensive analysis of DNA polymorphism data, *Genetics and population analysis*, Vol. 25 N°11, 1451–1452 p.
- **Lobry J. R.** 1996- Asymmetric substitution patterns in the two DNA strands of bacteria. *Molecular biology and evolution* 13, 660-665 p.

## Références bibliographiques

- **Loche V.** 1858- Catalogue des mammifères et des oiseaux observés en Algérie. A. Bertrand, Paris, mammals, 1-32 p.
- **Musser M. D., Carleton.** 2005- Superfamily Muroidea, in: D.E. Wilson, D.M. Reeder (Eds.), Mammal Species of the World, Third Edition, The Johns Hopkins University Press, 894–1531 p.
- **Nei M.** 1987- Molecular Evolutionary Genetics. Columbia University Press, New York.
- **Ouzaouit A.** 2000- La situation des rongeurs au Maroc. Séminaire national sur les surveillances et la lutte contre les rongeurs, Marrakech, 24-30 p.
- **Pagel M.** 2002- Discussion comments on Bayesian phylogenetic inference from animal mitochondrial genome arrangements. journal of the Royal Statistique Society, serie B, vol. 64, Part 4, 681–693p.
- **Pavlinov.** 2000- The contribution to craniometric variation and taxonomy of jirds from the group “shawi–grandis” of the genus Meriones (Gerbilli-dae), Zool. Zh 79, 201–209 p.
- **Perrière G.** 2000- Bases de données et outils d’analyse pour la génomique bactérienne. Mémoire de l’habilitation à diriger des recherches. Univ. Claude Bernard, Lyon 1, 100 p.
- **Perna N. T., Kocher T. D.** 1995- Patterns of nucleotide composition at fourfold degenerate sites of animal mitochondrial genomes. J. Mol. Evol. 41, 353-358 p.
- **Perrière G., Brochier-Armanet C.** 2010- Concepts et méthodes en phylogénie moléculaire. Springer-Verlag France, 250 p.
- **Petter F.** 1961- Répartition géographique et écologique des Rongeurs désertiques de la région paléarctique. Mammalia, 25, Suppl, 222 p.
- **Petter F.** 1961- Répartition géographique et écologique des rongeurs désertiques (du Sahara occidental à l’Iran oriental). Mammalia, Thèse, 222 p.
- **Petter F., Saint Girons M.C.** 1965- Les Rongeurs du Maroc. Travaux de l’institut scientifique Chérifien, série zoologie, V 3, 51 p.
- **Perrière G., Brochier-Armanet C.** 2010- Concepts et méthodes en phylogénie moléculaire. Springer-Verlag France, 250 p.

## Références bibliographiques

- **Rischkowsky B., Pilling D.** 2008- L'Etat des Ressources Zoogénétiques pour l'Alimentation et l'Agriculture dans le Monde, 393-410 p.
- **Rustin P., Rötig A.** 1995- Mécanismes moléculaires des cytopathies mitochondriales. *Med Therap* 1, 435-447 p.
- **Saitou N., Nei M.** 1987- The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol.* 4, 406–425 p.
- **Servant F., Bru C., Carrere S., Courcelle E., Gouzy J., Peyruc D., Kahn D.** 2002- ProDom: automated clustering of homologous domains. *Brief Bioinform*, 3(3), 246–51.
- **Sicard B., Kyelem M., Papillon Y., Diarra W., Keita M.** 1995- Rongeurs nuisibles soudano-sahéliens. Ed. Institut du Sahe CTA-ORSTOM, 54 p.
- **Sneath P. H., Sokal R. R.** 1973- Numerical taxonomy. The principles and practice of numerical classification.
- **Strimmer K. S.** 1997- Maximum likelihood methods in molecular phylogenetics. Herbert Utz Verlag.
- **Sunnucks P.** 2001- Efficient genetic markers for population biology. *Tree*, 15, 199–203 p.
- **Swofford D. L., Olsen G. J., Waddell P. J., Hillis D. M.** 1996- Phylogenetic inference. In D. M Hillis, C. Moritz, and B. Mable (eds.) *Molecular Systematics* (2nd ed.), Sinauer Associates, Sunderland, Massachusetts, 407-514 p.
- **Tagu, D., Risler, J. L.** 2010- Bio-informatique (Principes d'utilisation des outils). Éd. Quae, France, 269 p.
- **Tamura K., Peterson D., Peterson N., Stecher G., Nei M., Kumar S.** 2011- MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Mol. Biol. Evol.* 28, 2731-2739 p.
- **Thomas F., Lefèvre T., Raymond M.** 2016- Biologie évolutive, 2<sup>é</sup>édition. De Boeck supérieur, 959 p.
- **Thomas O., Harter E.** 1913- Expedition to the central western Sahara. Mammals. *Nov. Zoologicae* .Vol XX, 28 p.

## Références bibliographiques

- **Tillier E. R., Collins R. A.** 2000- The contributions of replication orientation, gene direction, and signal sequences to base-composition asymmetries in bacterial genomes. *Journal of molecular evolution* 50, 249-257 p.
- **Vicente M. C., Fulton T.** 2003- Utilisation des marqueurs moléculaires pour l'étude de la diversité génétique végétale. International Plant Genetic Institute for Genomic Diversity Resources Institute.
- **Villesen P.** 2007- FaBox an online for FASTA sequence. *Molecular Ecology Notes*, 7, 965-968 p.
- **Wagner M.** 1841- Reise in der RegenschaftAlgier in den Jahren1836 ,1837 und 1838. L Voss, Leipzeg. Vol. 1.-XXVIII + 482 p.Vol.2.- VI +296 p. Vol 3.-XVI + 296 p.
- **Woes C. R., Kandler O., Wheelis M. L.** 1990- Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proc. Natl. Acad. Sci. U S A*, 87, 4576-4579 p.
- **Xia X.** 2013- DAMBE5: a comprehensive software package for data analysis in molecular biology and evolution. *Mol Biol Evol* . 30(7), 1720-8 p.
- **Yoshikawa H., Dogruman A. F., Dogruman-A. F., Turk S., Kustimur S., Balaban N.** 2011- Evaluation of DNA extraction kits for molecular diagnosis of human Blastocystis subtypes from fecal samples. *Parasitol Res.*; 109:1045-50 p.

# Biblio-web

**NCBI** : <https://www.ncbi.nlm.nih.gov/>

**DAMBE** : <http://dambe.bio.uottawa.ca/DAMBE/dambe.aspx>

**FaBox** : <http://users-birc.au.dk/biopv/php/fabox/>

**MEGA** : <http://mega.software.informer.com/>

**NETWORK** : <http://www.fluxus-engineering.com/NETW5001.exe>



# ANNEXES

# ANNEXES

**Tableau I :** Liste des séquences d'ADN trouvées dans la GenBank du genre *Meriones*.

Espèce	Région	Gène	Taille en pb	Locus
<i>M. chengi.</i>	Chine	Cytb	1076	AB381900
		COX II	684	AB381912
		CO I	645	KF152978
<i>M. Crassus.</i>	Arabie Saoudite	ARNr mitochondrial 12S	951	AJ85125
		Cytb	1052	AJ851267
	Egypte	ACP5	424	KR088990
		Cytb	1139	KR089088
	Israël	IRBP	1272	LN606643
	Turque	IDRADD	501	LN851569
	Iran	COX I	608	MH580694
		Cytb (01)	737	MH580754
<i>M. Grandis.</i>	Maroc	COX I	608	MH580689
		Cytb	737	MH580750
<i>M. libycus erythourus.</i>	Chine	Cytb	1103	AB381902
		COX II	684	AB381913
<i>M. meridianus pencilliger.</i>	Mongolie	Avpr2	911	FJ411224
<i>M. persicus.</i>	Iran	Acp5	404	KR088991
		Cytb	712	KR089030
		Cytb (70)	877	KT949948
		Cytb (07)	533	MH384456
		COX I	608	MH580691
<i>M. Rex.</i>	Arabie Saoudite	ARNr mitochondrial 12S	952	AJ851248
		Cytb	1140	AJ851265
		COX I	608	MH580691
		Cytb	737	MH580751
<i>M. shawi isis.</i>	Maroc	Avpr2	396	Fj411225
		Avpr2 (02)	353	Fj411225
<i>M. Ilbycu.</i>	Chine	COX I	656	KF422715
		COX I	645	KF152970
		COX I (04)	684	KT613846
		ARN ribosomique 16s (03)	545	KT763075
		Cytb (08)	1179	KU156726
		COX I (02)	667	KU182936
		COX I (02)	709	KU182945
		Cytb	876	MT024684
	Tunisie	Cytb (09)	1110	Jq927406

## ANNEXES

		Cytb (09)	1140	KP893540
		IRBP (05)	1262	LN868387
		ARNr mitochondrial 12s	53	KX775971
	Iran	Cytb (04)	1140	KR089029
		Cytb (02)	553	MH384471
		Cytb (32)	731	MH580696
	Algérie	aKr1b7	1316	FN428573
		COX I (07)	608	MH580680
		Cytb (02)	797	MH580736
	Syrie	Cytb	737	MH580741
	Jordanie	Cytb (03)	737	MH580738
	Afghanistan	Cytb (14)	663	MF535299
		Cytb (04)	737	MH580725
	Maroc	Cytb (10)	737	MH580726
		ARNr mitochondrial 12s	951	Aj851250
		Cytb (01)	1140	Aj851266
		Rag2(07)	836	KC480070
		Cytb (08)	941	KC480077
		Gene d'ARN ribosomique M II 16s	571	KF422719
		Génome complet (02).	16341	KR013226
Protéine de liaison aux rétinoïde		958	KT834969	
M. shawi.	Tunisie	BFIBR (20)	721	KM581449
		BFIBR	720	KM581462
		Cytb (25)	1140	MP89952
		Cytb (02)	1170	KP899529
		Microsatellites (07)	59	Kx775964
		IRBP (08)	1262	LN868375
		COX I	608	LN868360
		Cytb	737	MH580751
	Maroc	BFIB (40)	731	MH581467
	Algérie	BFIBR (04)	721	KM581451
		BFIBR (02)	720	KM581467
		Cytb (177)	1062	KM581499
		Cytb	971	KP089031
		COX1	684	U18835
		BRCA1	3092	AF332048
		GHR	898	AF32021
		RAG1	2780	AY294947
	BRCA1	1667	AY29004	

## ANNEXES

		COX1	1294	DQ019083
		GPR33	720	DQ156951
		IRBP	761	KC953400
		Microsatellites (10)	344	KF937056
<i>M. Meridianus.</i>	Chine	ARNr Mitochondrial 12s	951	Aj851252
		Cytb	1140	Aj851268
		Cytb (06)	1107.1121.1120.1113.1125.1124.	AB381894
		COX II	684 (2)	AB381908
		COX II	668	AB381910
		COX II	674	AB381911
		CO I (05)	650.	jQ043496
		CO I (23)	650.657 (10).653 (12)	KF027349
		CO I (06)	645	KF15972
		CO I (02)	697	KT763038
		ARNr 16s (05)	530	KT763073
		Cytb	1099	KU156736
		CO I	647	KU182941
		CO I	700	KU182950
		Cytb (07)	1140	MU204244
	Iran	COX 1	608	MH5880688
		Cytb	731	MH580749
	Kazakhstan	Cytb (09)	1140	MN204228
		Cytb (05)	674.1068.1114.1131.1127	MN204237
		BRCA1 (79)	951	MN204314
		IRBP	954	MN204393
		Cytb	1128	MN204306
	Mongolie	Cytb (30)	1110 (02).1070.1085.1035.1081.1140 (04).1124.1094.1114 (03).1113.1123.1128.1086.1103.1087.1176.1107.1176 (02).613.1118.1112.1079.1125	MN204151
		Cytb (102).	755	Jq065512
		GRB 33	915	AY490641
		GSRT2	503	EV635765
		Avpr2	1064	Fj411223
		TAAR4	900	Fj372495
		D-Loop (138)	756	Jq065374
		IRBP (111)	955	jQ713186
		ZWSS (04)	681.1106.1674.2176	KF027349
		COI (02)	657	Kj466777
Glutamine amino-		2364	Kj160361	

## ANNEXES

		peptidase			
		HIV-A tatinteractive Protéine 2	600	Kj160371	
		ARNm de peptide c	222	Kj160994	
		ARNm1	897	Kj160408	
		Résistance aux inhibiteurs de l'ARNm.	1461	Kj160413.	
		Sphingosine Phosphate lipase (01).	1443	Kj160418	
		ARNm de protéine.	483	Kj160435.	
		MYLK2.	231	Kj721981.	
		VEGF A.	528	Kj789177.	
		Génome complet (02).	16376	KR013227.	
<i>M. Tristrami</i>	Jordanie	Cytb	1193	KU189331	
		ACP5	431	MR088992	
	Iran	Cytb	878	MU561099	
		Cytb (05)	533	MH384466	
		COI	608	MH580695	
		Cytb	737	MH580755	
		Cytb	877	KU189331	
		Gene inconnu	544	GQ506804	
		Gene inconnu	361	GQ506756	
		UCP02	461	GQ506730	
		Protéine 5 associe aux cytosquelette	504	GQ506778	
		Protéine sensible à la puromycine	627	GQ506949	
		Seq sine	504	GQ506694	
		Seq sine	519	GQ506098	
		Gene inconnu	199	GQ506822	
		Proteine5	490	GQ506766	
		Seq sine	425	GQ506668	
		O manosyl N- acétyle glucosaminyl transferase.	261	GQ506703	
	<i>M. tristrami blackleri.</i>		Cytb	799	jQ687393.
			Cytb	790	jQ687394.
<i>M. tamariscinus.</i>	Chine	Cytb	1113	AB381904.	
		Cytb	1104	AB381905.	

## ANNEXES

		Cytb	1090	AB381906.
		Cytb	1005	MT024691.
		COX II	684	AB381916.
		COX I (02)	645	KF152968.
	Mongolie	COX I	661	KX859280
	Chine	COX II	684	AB381916
		Cytb (33)	1140	JN604750
		D-Loop (09)	428	jN604783
		D-Loop (23)	429	jN604792
		Génome complet	16393	KT834971
		Gene de la protéine de liaison aux rétinoïde	942	KT834970
	Génome complet	16389	KX688104	
<i>M. tristrami bodenheimeri.</i>		Cytb (02)	789	jQ 687398
<i>M. tristrami bogdanovi.</i>	Iran	Cytb (02)	789	jQ687399
<i>M. Unguiculatus</i>			1623	XM021635652
			1745.1781.1723.1675 2018.1759.	XM 021636056
			445	XM 021637341
			1196	XM 021639225
			2029	XM 021699660
			2233	XR00476498
			418	XM 021641807
			1820	XM 021643168
			2632	XM0211648188
			1184	XM 021652939
			1069	XM 021659317
			617	XM 021659318
			538	XM 021660025
			672	XM 021660026
			417	XM 021630516
			912	XR002475969
			550	XM021628305
			819	XM 021628776
			737	MH580748
			1844	XM 021635290
			2647	XM 021694324
	2819.2626	XM 021634322		
	1544	XM 021633741		
	445	XM 021632739		
	6866	XM 021631684		
	741	XM 021631234		
	1610	XM 021663265		

## ANNEXES

			1166	XM 021661963
			1116	XM 021661955
			1242	XM 021632724
			1015	KR089033
			1088	AB381901
			1140	AF159405
			1140	AF119264
		Avpr2	1043	FJ411228
			1872	XM021656385
			1632	XM021656387
			1629	XM021656388
<i>M. tristrami kilisensis.</i>		Cytb (03)	792	jQ687380
		Cytb (02)	795	jQ687382
<i>M. tristrami Lycaon.</i>		Cytb (01)	857	jQ687383
		Cytb (03)	792	jQ687389
		Cytb (03)	795	jQ687390
		Cytb (01)	789	jQ687391
<i>M. tristrami intraponticus</i>		Cytb	795	jQ687386
		Cytb	810	jQ687387
		Cytb	792	jQ687388
<i>M. vinogradovi.</i>	Iran	Cytb (01)	878	KU561098
		COX I	608	MH580693
		Cytb (01)	737	MH580753
<i>M. sp. Garat An Njila.</i>	Algérie	ARNr mitochondrial 12S	952	Aj851247
		Cytb (01)	1140	Aj851264
<i>M. sp. Hk-2016 .</i>	Tunisie	Cytb (01)	1140	KP899539
<i>M. sp.n.MAO-2018..</i>	Iran	Cytb (03)	533	MH384463

# ANNEXES

## Exemples des Formats des Banques de données :

Tableau IV : Exemple d'une séquence au format GenBank.

```
GenBank: KF152978.1
Project: GenBank, EBI/ENA

LOCUS KF152978 645 pb ADN ROD linéaire 27-JUL-2015
DEFINITION Meriones chengi voucher JW76 gène de la sous-unité I de la cytochrome oxydase
(COI)
     CDS partiels; mitochondrial.
ADHESION KF152978
VERSION KF152978.1
MOTS CLÉS.
SOURCE mitochondrie Meriones chengi (Cheng's jird)
ORGANISME Meriones chengi
     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
     Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Myomorpha;
     Muridae; Muridae; Gerbillinae; Meriones.
RÉFÉRENCE 1 (bases 1 à 645)
AUTEURS Yin, Y., Jiang, W., Zhang, Z., Li, Y., Twenke, B., Turghan, M., Yang, W.
     et Liu, B.
TITRE DNA Barcode Analysis for Species Identification and Divergence
     Estimates of Glires distribution in Xinjiang, China
JOURNAL Non publié
RÉFÉRENCE 2 (bases 1 à 645)
AUTEURS Yin, Y., Jiang, W., Zhang, Z., Li, Y., Twenke, B., Turghan, M., Yang, W.
     et Liu, B.
TITRE Soumission directe
JOURNAL Soumis (27-MAI-2013) State Key Laboratory of Desert and Oasis
     Ecology, Xinjiang Institute of Ecology and Geography, Chinese
     Academy of Sciences, 818 South Beijing Road, Xinhé District,
     Urumqi, Xinjiang 830001, China
CARACTÉRISTIQUES Localization / Qualificateurs
     source 1..645
         /organism = "Meriones chengi"
         /organelle = "mitochondrie"
         /mol_type = "ADN génomique"
         /specimen_voucher = "JW76"
         /db_xref = "taxon:204208"
         /country = "Chine"
         /lat_lon = "44.21 N,80.41 E"
         /collection_date = "02-juin-2000"
         /collect_by = "Wei Jiang"
         /Identified_by = "Wei Jiang"
     gene
         <1..-> 645
     CDS "COI"
         /gene = "COI"
         /codon_start = 1
         /transl_table = 2
         /product = "sous - unité de cytochrome oxydase I"
         /protein_id = "AB327612.1"
         /translation = "ITGAWAGMAGTALSILRTELGGQNGALLGDDQGTNVIVTAHAIV
MIFTMVMFMMHGGFGNWLVPMLMIGAFDMAMPRMNNMSFWLLPSPFLELLASSMVLVAGA
GTGWTVTYFPPLAGNLAHAGASVDLITFSLHLAGVSSILGAINFITTINMKPPAMTQYQ
TFLFYWSVLITAVLLESLPVLAAGETMLETDRNLNNTTTFDFAGGGDPILYQSLI"
ORIGIN
1  atttcgggg cttagcagg aatagcagga acagccctaa gcacccat tcgaacagaa
61  ttaggacaac cgggcctct actaggagac gatcagattt ataatgcat cgtaacgcc
121  catgctttg ttaataatt cttatagtc atacctataa taatggggg attcggaaac
181  tgaatgccc cactaataa tggcctcc gatatagct tccacgaat aaacaacata
241  agcttgac ttaccccc accattta ctaactag catcctat agtagaagt
301  ggagcagga cagctgaac agtatctt ccttagctg gaaactgac ccagcagga
361  gcatraggg accttaaat ttctcctt caactagtg gatgcac aatccagga
421  ggaataat taacacac tatcattat aaaaacgc cagcataac acaatataa
481  caaccattt ttgatggc ttttaate acagagtc tctatlaet ttaactca
541  gtttagctg caggaatac tatacttc acagaccga acctcaatac aacttttt
601  gaacctgct gaggagggg cctatctc tatcaaac
```



# ANNEXES

Tableau V : Exemple d'une séquence nucléotidique au format EMBL.

```

ID WAC38885; SV 1; linear; genomic DNA; STD; ROD; 1148 BP.
XX
FA KF899533.1
XX
DT 03-09-2015 (Rel. 135, Created)
DT 03-09-2015 (Rel. 135, Last updated, Version 1)
XX
DE Meriones shawi (Shaw's jird) partial cytochrome b
XX
KW .
XX
OS Meriones shawi (Shaw's jird)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;
OC Eutheria; Euarchontoglires; Glires; Rodentia; Myomorpha; Muridae;
OC Gerbillinae; Meriones.
OC Mitochondrion
XX
RN [1]
RP 1-1148
RA Khemiri H., Pimenta J., Acarim A., Chevret P., Nouna S., Lopes A.M.;
RT "Genetic diversity of Tunisian wild Meriones";
RL Unpublished.
XX
RN [2]
RP 1-1148
RA Khemiri H.;
RT ;
RL Submitted (08-MAR-2015) to the DDBJ.
RL Biology, Faculty of Sciences, University of Tunis El Manar, El Manar, Tunis
RL 2092, Tunisia
XX
DR MD5; 159bd874fec2f62391a87853e14a84d5.
XX
FH Key Location/Qualifiers
FH
FT source 1..1148
FT /organism="Meriones shawi"
FT /organeller="mitochondrion"
FT /isolate="Msh-07"
FT /mol_type="genomic DNA"
FT /country="Tunisia:Doubedna"
FT /collected_by="Hanene Khemiri"
FT /db_xref="taxon:37435"
FT CDS KF899533.1:c1...:1148
FT /codon_start=1
FT /transl_table=2
FT /gene="Cytb"
FT /product="cytochrome b"
FT /db_xref="GDN:AM2143177"
FT /db_xref="InterPro:IPR085797"
FT /db_xref="InterPro:IPR085798"
FT /db_xref="InterPro:IPR016174"
FT /db_xref="InterPro:IPR027387"
FT /db_xref="InterPro:IPR038589"
FT /db_xref="InterPro:IPR036158"
FT /db_xref="UniProtKB/TrEMBL:AM2143177"
FT /protein_id="WAC38885.1"
FT /translation="MPLRQNPILLIQANESFDLPTPPNCSQAAWNGSLGGDLVDQI
FT ATGLFLANVTADTTTAFSSVSDICRDVNYGALDRVMWAGAGAFFDCLFDHGRGIVY
FT QSYVITETWADQVLLFAMATAFADVALPAGQSGFAGATVETNLLSAIPYICPTVADW
FT TAGGFSYDQKTLTRTFATHTDLPFDIAVLVWLLFLHETGQANPLGLDQANCKIPNIP
FT YTTADFLQVALLDFFHLLVLTFFDTLGDQDNYTRANPLNTRPHIDQFDVFLFVAIL
FT RSDPNKLGQVLAALVSLDLDFLPLDHTSQGSLDFRPTSQMLVWALDANLLDLTWGG
FT QPVDHFFIITGGAGETVFSDELDFLFDAGIIDDLLKAW"
XX
SQ Sequence 1148 BP; 337 A; 333 C; 147 G; 333 T; 8 other;
atgataaac tacgaataaa caaccacatg taanaatag taaccacac attcattgac 68
ctccaccac caaccacat ctcatcttga tgaacattg ggtctctcct agghatctgc 128
ttagtctctc aactctctac aggatcttc ctggtctac actacacagc ggcacacaca 188
acagcattct cctcagcttc cctatctctg ctgagcgtaa actatggctg acctatctgt 248
tatatacatg caaccggagc ctctatattt ttcatctgac tttctatcca catctggagg 308
ggaatctact aggatctcta tctcttctact gaacatgaa acctggagc aattctttta 368
ttgcgcgtca tagctacagc atttataggc tatgttctac catgaggaca aatctctctc 428
tgagggggca cagctcttcc caacctctc tcaagcaatc catatctggg caaacagta 488
gtagaatgaa tctgaggagg gttctcagta gacaagaca cctaacacag attttttggc 548
ttccattcca tcttaccglt cctatctgca gctctagctc tagttccact ttatctctca 608
catgaaacag gatctaacaa caacttaggc ctgactctca agcagatgaa aattctcattc 668
caaccatctc caacagtcas agacttctta ggcgttggcc tactaatctt ctctctcata 728
cttttagtcc tttctcttcc tgcacacctt gggagccacg caactctacc accagctaac 788
cactaarta ctcaaccca cattaacca gaatgatact tctatctggc atagccatc 848
cttctctcta tcccaacaaa attagggggc gtctctagcc tagtctctac aattctaatc 908
ttactcttc ttctcttaat caacacatc aaacacgaa gctaatctt ccgaccatc 968
tcccaatat tatctaac ccataatcc arctactaa tcttaccata aatttaaat 1028

```

# ANNEXES

Tableau VI : Exemple d'une séquence nucléotidique au format DDBJ

```

LOCUS      KT949992                877 bp    DNA     linear   ROD 02-NOV-2016
DEFINITION Meriones persicus voucher ZMFUM3902 cytochrome b (cytb) gene,
partial cds; mitochondrial.
ACCESSION  KT949992
VERSION    KT949992.1
KEYWORDS   -
SOURCE     mitochondrion Meriones persicus (Persian jird)
  ORGANISM Meriones persicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Myomorpha;
            Muroidea; Muridae; Gerbillinae; Meriones.
REFERENCE  1 (bases 1 to 877)
AUTHORS    Dianat,M., Darvish,J., Cornette,R., Aliabadian,M. and Nicolas
            Colin,V.
TITLE      Evolutionary history of the Persian Jird, Meriones persicus, based
            on genetics, species distribution modelling and morphometric data
JOURNAL     J. Zoolog. Syst. Evol. Res. (2016) In press
REMARK     Publication Status: Available-Online prior to print
REFERENCE  2 (bases 1 to 877)
AUTHORS    Dianat,M., Darvish,J., Aliabadian,M., Nicolas Colin,V. and
            Cornette,R.
TITLE      Direct Submission
JOURNAL     Submitted (21-OCT-2015) Department of Biology, Ferdowsi University
            of Mashhad, Azadi Street, Mashhad, Khorasan Razavi 9177948974,
            Islamic Republic of Iran
COMMENT     ##Assembly-Data-START##
            Assembly Method      :: codoncode aligner v. v4.1.
            Sequencing Technology :: Sanger dideoxy sequencing
            ##Assembly-Data-END##
FEATURES   Location/Qualifiers
     source          1..877
                /organism="Meriones persicus"
                /organelle="mitochondrion"
                /mol_type="genomic DNA"
                /specimen_voucher="ZMFUM3902"
                /db_xref="taxon:1395794"
                /country="Iran"
     gene            <1..>877
                /gene="cytb"
     cds             <1..>877
                /gene="cytb"
                /codon_start=1
                /transl_table=2
                /product="cytochrome b"
                /protein_id="AOZ65632.1"
                /translation="WNFGSLLGVCLIIQIITGLFLAMHYTADTTAFSSVSHICRDV
            NYGMLIRYMHANGASMFICLF IHIGRGIYYGSYIFYETWNIGVILLFTIMATAFMEY
            VLPWQMSFWGATVITNLLSAIPYIGPTIVIEWIWWGGFSVDKATLRRFFAFHFILPFII
            AALVVLVHLLFLHETGSNNPLGLNSNVDKIPFHPYYTMKDFLGVVLLISFFMLMVLFFP
            DLLGDPDNYTPANPLNTPPHIKPEWYFLFAYAILRSIPNKLGGVLLALVLSILILIFLP
            LIHTSKQRSMLFRPIS"
BASE COUNT      251 a                245 c                118 g                263 t
ORIGIN
    1  tgatggaact  tcggatcact  tctaggcgtt  tgcctaatta  tccaaatcat  cacaggatta
    61  ttcttagcta  tacactacac  agcggacaca  acaacagcat  tctcatcagt  ttctcatatc
   121  tgccgagacg  taaactacgg  ctgacttata  cgctatatac  atgcaaacgg  ggccctcaata
   181  ttctttatct  gcctatttat  ccatatcgga  cgaggaaatt  actacggatc  ttacatcttt
   241  tacgaaacat  gaaacattgg  agttattctc  ttattcacta  tcatagccac  tgcattcata
   301  ggatatgttc  taccatgagg  acaaatatcc  ttctgagggg  ccacagttat  caccaacctc
   361  ctctcggcaa  tcccatacat  tgggccaaca  atcgtagaat  gaatttgagg  cgggttttca
   421  gttgataaag  ctacccttac  acgatttttc  gcattccact  ttatcttacc  attcatcacc
   481  gcagccctgg  ttctagtcca  tcttctatcc  ctgcacgaaa  caggatctaa  caaccatta
   541  ggtctcaatt  ctaacgtaga  caaaatcccc  ttccacctcc  actacacaat  aaaagatttc
   601  ctaggagtgc  tcttattaat  ctcttcttc  atacttatag  tcttattctt  tcttgacctc
   661  cttggagatc  cagacaacta  cacaccggcc  aacccaectc  acacccccacc  tcacattaaa
   721  cctgaatgat  acttcttatt  cgcatacgtc  atccttcgtt  ctatcccctaa  taaattagga
   781  ggagtccctag  ccttagtact  ctcaatctta  atcctgattt  tcttaccact  aatccacaga
   841  tccaaacaac  gaagcctaata  attccgacca  atctccc
//

```

# ANNEXES