

République Algérienne Démocratique et Populaire

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Université de Blida -1-

Faculté de Science de la nature et de la vie

Département de Biologie et Physiologie Cellulaire



Mémoire de fin d'études

En vue l'obtention du Diplôme de MASTER

Option : *bioinformatique*

Thème :

Les méthodes de reconstruction phylogénétique : exploitation des logiciels et application au genre Canis.

Présenté par :

M^{elle} KOBBI Nour El Houda

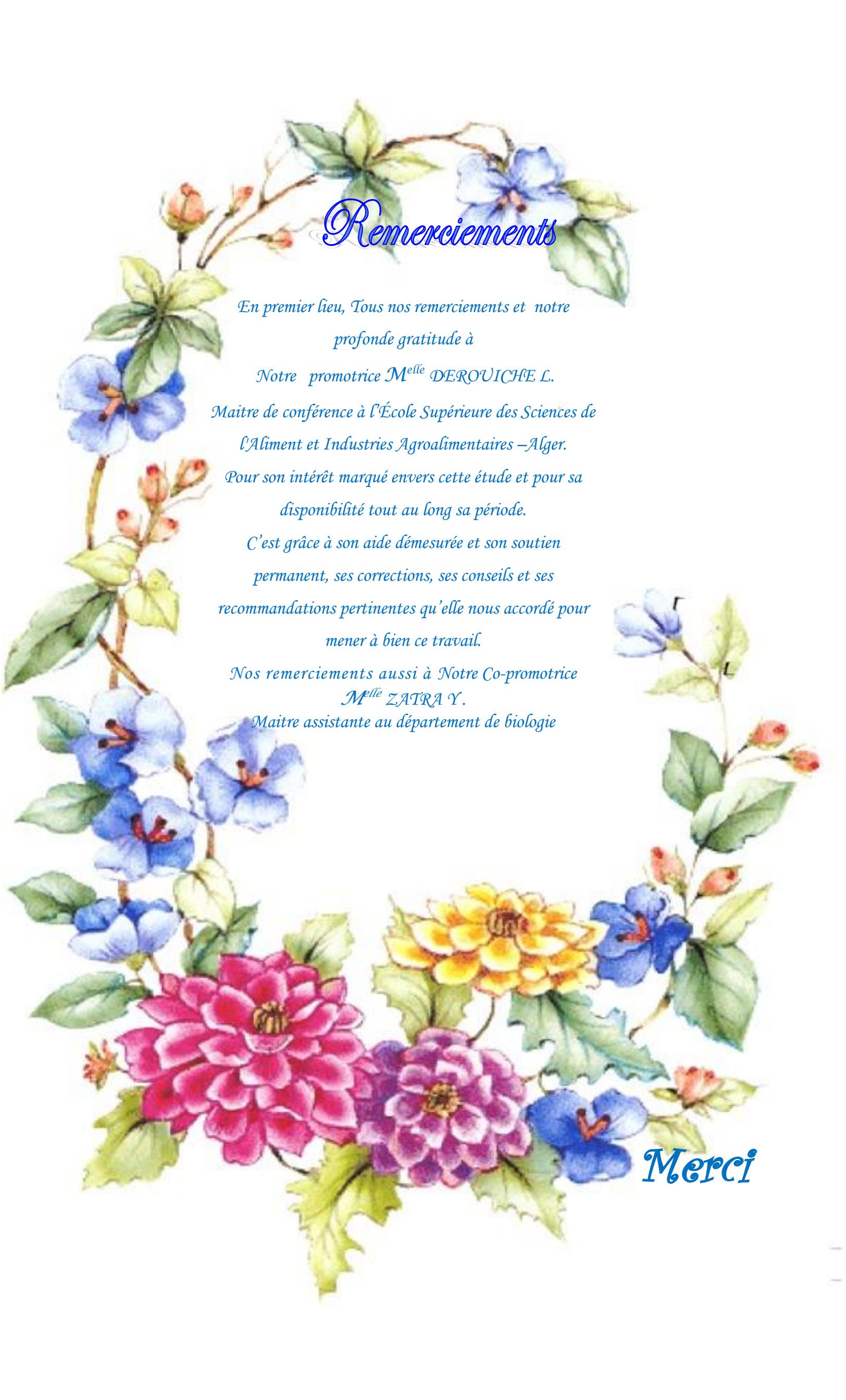
&

M^{elle} SAADI Lamia

Devant le jury composé de :

<i>Présidente</i>	: <i>M^{me} CHAABANE S.</i>	<i>MAA</i>	<i>USDB</i>
<i>Examinatrice</i>	: <i>M^{me} ABDEL HOUSSAIN A.</i>	<i>MCB</i>	<i>USDB</i>
<i>Promotrice</i>	: <i>M^{me} DEROUICHE L.</i>	<i>MCB</i>	<i>ESSAIA</i>
<i>Co-promotrice</i>	: <i>M^{me} ZATRA Y.</i>	<i>MAA</i>	<i>USDB</i>

Année universitaire : 2016/2017



Remerciements

*En premier lieu, Tous nos remerciements et notre
profonde gratitude à*

Notre promotrice M^{elle} DEROUICHE L.

*Maitre de conférence à l'École Supérieure des Sciences de
l'Aliment et Industries Agroalimentaires –Alger.*

*Pour son intérêt marqué envers cette étude et pour sa
disponibilité tout au long sa période.*

*C'est grâce à son aide démesurée et son soutien
permanent, ses corrections, ses conseils et ses
recommandations pertinentes qu'elle nous accordé pour
mener à bien ce travail.*

*Nos remerciements aussi à Notre Co-promotrice
M^{elle} ZATRA Y.*

Maitre assistante au département de biologie

Merci



Remerciements

*Nous exprimons toutes nos reconnaissances aux
membres des jurys :*

*A la présidente M^{me} CHAABANE Maitre assistante
au département de biologie*

*D'avoir bien voulu nous faire l'honneur de présider ce
jury*

*L'examinatrice M^{me} ABDEL HOUSSAIN pour avoir
accepté d'examiner notre travail.*

Pour son encadrement Pour leur aide et l'encouragement

*Nos remerciements s'adressent également à tous
nos enseignants, qui nous ont donné les bases de
la science.*

*Enfin Cette liste de remerciements n'est pas
exhaustive, que tous ceux avec lesquels nous
avons interagi lors de ce travail, sans toutefois
les avoir cités ci-dessus, sachent que leur
contribution à la réalisation de ce mémoire a été
grandement appréciée.*

Merci

Dédicaces



Je dédie ce mémoire à

*Mes très chers parents
qui sont toujours dans mon cœur, qui ont consacré leur vie à
mon éducation et ma réussite, qui m'ont encouragé dans les
moments les plus difficiles de ma vie.
Que Dieu les garde et les protège pour nous.*

*Ma sœur :
RAWANE.*

*Mon frère :
MOHAMED ABD EL WADOUD.*

*Ainsi qu'à mon binôme et ma sœur : LAMIA à qui je
souhaite une vie pleine de bonheur et de succès. Ainsi que sa
famille.*

*Toutes mes amies en particulier : KARIMA, FAIROUZ,
ASSIA, WAFAA.*

*Toutes les étudiantes de l'option bioinformatique promotion
2017.*

Tous ceux qui m'aiment et que j'aime.

NOUREL HOUDA

Liste des abbreviations

A

ADN : Acide Désoxyribonucléique.

ADNmt : ADN mitochondrial.

ATDB : Arabidopsis thaliana DataBase.

C

COI: cytochrome c subunit I.

COII : cytochrome c subunit II.

cyt b : cytochrome b.

C.a.lupaster : Canis aureus lupaster.

D

DDBJ : DNA Data Bank of Japon.

DnaSP : DNA sequence polymorphism.

DAMBE : Data Analysis in Molecular Biology and Evolution.

D-loop : Displacement- Loop.

E

EMBL : European Molecular Biology Laboratory.

EBI : European Bioinformatic Institute.

ECDC : Escherichia coli Database Collection.

H

HMM : Hidden Markov model.

J

JIPID : Japon International Protein Information Database.

L

LANL : Los Alamos National Laboratory.

M

MIPS : Munich Information Center for Protein Sequence.

MCMC : Markov Chain Monte Carlo.

MEGA : Molecular Evolutionary Genetics Analysis using Maximum Likelihood.

mv : vecteurs médians.

N

NCBI : National Center For Biotechnology Information.

NIG : National Institut of Genetics.

NBRF : National Biomedical Research Foundation.

O

OWL : Web Ontology Language.

P

PIR : Protein Information Resource.

PCR: Polymerase Chain Reaction.

PDB : Protein Data Bank.

PGD : Plant Genome Database.

PHYMLIP : Phylogeny Inference Package.

S

SIB : Swiss Institute of Bioinformatics.

SGD : Saccharomyces Genome Database.

U

UPGMA: Unweighted Pair Group Method with Arithmetic Mean.

GLOSSAIRE

- ✚ **ADN mitochondrial** : est présent dans la matrice des mitochondries. C'est une molécule circulaire bicaténaire et sa taille diffère selon les organismes (**Oudot-Le secq, 2000**).
- ✚ **FASTA** : algorithme proposé par Pearson basé sur l'identification rapide de fragments strictement identiques entre la séquence cible et les séquences de la banque (**Comet, 1998**).
- ✚ **Alignement de séquences** : est une des opérations fondamentales dans le processus de la reconstruction phylogénétique, puisque c'est sur elle que se base beaucoup de résultats comme les mesures d'identité et la recherche d'homologies entre séquences (**Comet, 1998**).
- ✚ **UPGMA** : une méthode utilisée pour construire des arbres phylogénétiques si les séquences ne sont pas trop divergentes (**Golding et Morton, 2003**).
- ✚ **Bootstrap** : est une méthode souvent utilisée pour tester la fiabilité des branches internes des arbres phylogénétiques (**Felsenstein, 1985**).

SOMMAIRE

INTRODUCTION.....	1
CHAPITRE I : RAPPELS BIBLIOGRAPHIQUES	
I.1. IDENTIFICATION PHYLOGENETIQUE.....	3
<i>I.1.1. Définition de la phylogénie</i>	<i>3</i>
<i>I.1.2. Phylogénie moléculaire</i>	<i>3</i>
<i>I.1.3. L'arbre phylogénétique</i>	<i>5</i>
I.2. LA PHYLOGENIE ET L'INFORMATIQUE	5
<i>I.2.1. Les bases de données</i>	<i>5</i>
<i>I.2.2. Les banques généralistes.....</i>	<i>6</i>
<i>I.2.2.1. Banques de séquences nucléotidiques</i>	<i>6</i>
<i>I.2.2.2. Banques de séquences protéiques</i>	<i>7</i>
<i>I.2.3. Les banques spécialistes.....</i>	<i>9</i>
I.3. LES LOGICIELS UTILISEES DANS LA PHYLOGENIE	10
<i>I.3.1. Présentation des logiciels d'analyse phylogénétique.....</i>	<i>10</i>
<i>I.3.1.1 Les méthodes de reconstruction phylogénétique et les logiciels utilisés.....</i>	<i>10</i>
<i>I.3.1.1.1. La méthode de parcimonie.....</i>	<i>10</i>
<i>I.3.1.1.2. La méthode de « Bayesian inference ».....</i>	<i>11</i>
<i>I.3.1.1.3. la méthode de Maximum de vraisemblance</i>	<i>12</i>
<i>I.3.1.2. les Logiciels de Nettoyage de l'alignement</i>	<i>12</i>
<i>I.3.1.3. Les logiciels pour l'alignement de séquence.....</i>	<i>13</i>
<i>I.3.1.4. Les logiciels de traitement de séquence.....</i>	<i>13</i>
<i>I.3.1.5. Logiciels de visualisation et reconstruction d'arbres phylogénétiques</i>	<i>13</i>
I.4. REPRESENTATION DES ESPECES DE GENRE CANIS.....	14
<i>I.4.1. Le genre canis, linne 1758: particularités et répartition.....</i>	<i>14</i>
<i>I.4.1.1. Présentation des espèces</i>	<i>14</i>
<i>I.4.1.1.1. Classification.....</i>	<i>14</i>
<i>I.4.1.1.2 Répartition et Morphologie</i>	<i>16</i>

CHAPITRE II : MATERIEL ET METHODES

II.1. ANALYSES MOLECULAIRES ET PHYLOGENETIQUES.....	28
II.1.1. Etapes d'Analyses Moléculaires.....	28
<i>II.1.1.1. Echantillonnage</i>	28
<i>II.1.1.2. Extraction d'ADN</i>	28
<i>II.1.1.3. La PCR.....</i>	28
<i>II.1.1.4. Le Séquençage.....</i>	29
II.1.2. Etude de la GenBank.....	29
<i>II.1.2.1. Définition de la GenBank</i>	29
<i>II.1.2.2. Traitement des séquences de la GenBank</i>	31
<i>II.1.2.2.1. FaBox.....</i>	31
<i>II.1.2.2.2. MEGA.....</i>	33
<i>II.1.2.2.3. DnaSP</i>	37
<i>II.1.2.2.4. DAMBE.....</i>	38
<i>II.1.2.2.5. NETWORK</i>	39

CHAPITRE III : RESULTATS ET DISCUSSION

III.1. Résultats de la GenBank	41
<i>III.1.1. Le génome mitochondrial</i>	41
<i>III.1.1.1. Le cytochrome b.....</i>	42
<i>III.1.1.2. D-loop (Displacement- Loop).....</i>	44
<i>III.1.1.3. Misc-feature.....</i>	46
III.2. Construction des arbres phylogénétiques.....	47
<i>III.2.1. Cytochrome b</i>	47
<i>III.2.2. D-loop.....</i>	50
<i>III.2.3. Misc-feature.....</i>	53
III.3. Résultat des haplotypes	56
<i>III.3.1 Cytochrome b</i>	57
<i>III.3.2 D-Loop</i>	59
<i>III.3.3 Misc-feature.....</i>	60
III.4. Résultats statistiques	62
<i>III.4.1. GC/AT skew.....</i>	62

<i>III.4.1.1. Cytochrome b</i>	62
<i>III.4.1.2. D-Loop</i>	63
<i>III.4.1.3. Misc feature</i>	63
III.5. La construction des réseaux	64
<i>III.5.1. Cytochrome b</i>	65
<i>III.5.2. D-loop</i>	66
<i>III.5.3. Misc-feature</i>	68
CONCLUSION	71
RÉFÉRENCES BIBLIOGRAPHIQUES	72
ANNEXE I	

Au cours de ces trente dernières années, la récolte de données en biologie a connu une **bond** 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. Les travaux de **Charles Darwin**, publié en 1859, vont apporter des éléments solides attestant d'une évolution progressive depuis un hypothétique ancêtre commun universel, à l'origine de toutes les formes de vie actuelles.

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.

Nous avons choisi le genre *Canis* comme un modèle pour une étude phylogénétique à cause de l'apparition d'une nouvelle espèce, le loup d'Afrique (*Lupus lupaster*). Le genre *Canis* est représenté par 34 espèces actuelles de la famille *Canidae*. Ce genre regroupe les loups, les chacals et les coyotes qui sont répartis sur les continents eurasiatique et africain et sur le continent Nord-américain (**Itis, 2010**).

Introduction

Notre objectif de travail est l'application de différents logiciels de bioinformatique pour analyser une base de données a fin de tracer un arbre phylogénétique des espèces de genre *Canis*.

I.1. IDENTIFICATION PHYLOGENETIQUE

I.1.1. Définition de la phylogénie :

Le terme « phylogénie » fut inventé par Ernst Haeckel en 1866 pour définir l'enchaînement des espèces animales et végétales au cours du temps. Jusqu'alors le concept était exprimé par le terme « généalogie ». Ce n'est que dans la dernière édition de l'origine des espèces (1872) que Charles Darwin introduisit le mot phylogeny avec la définition suivante : les lignes généalogiques de tous les êtres organisés. Nous définirons la phylogénie comme «le cours historique de la descendance des êtres organisés».

Haeckel lui-même avait défini la phylogénie comme l'histoire du développement paléontologique des organismes par analogie avec l'ontogénie ou histoire du développement individuel. Les termes « développement » et « évolution » sont tous deux issus de l'embryologie. Pour qualifier les transformations organiques situées dans le temps géologique, le mot «évolution» supplanta progressivement à la fin du XIX^e siècle celui de « développement » Haeckel fut l'un des artisans de ce succès qui se fit au détriment de « transformisme », terme synonyme d' « évolutionnisme » et qui reste le plus souvent associé à l'œuvre de J.-B. Lamarck, quoique ce dernier ne l'utilisât jamais.

Lamarck, en même temps qu'il conçut les bases de la théorie de l'évolution, publia dans sa Philosophie zoologique (1809) un schéma de filiation des animaux. Ce schéma est présenté comme la « distribution générale » des animaux. Ce concept lamarckien exceptionnellement fécond s'oppose explicitement à la classification et se veut une construction qui doit exprimer « l'ordre représentant le plus possible celui même de la nature, c'est-à-dire l'ordre qu'elle a suivi dans la production des animaux et qu'elle a éminemment caractérisé par les rapports qu'elle a mis entre les uns et les autres ». Les « rapports » qu'évoque Lamarck sont les « parentés entre les corps vivants » (Darlu et al., 2008).

I.1.2. Phylogénie moléculaire :

La phylogénie moléculaire consiste à reconstruire les relations de parenté entre les séquences de nucléotides ou d'acides aminés. Elle permet d'étudier l'évolution des séquences et la relation de parenté entre espèces. La nature quantitative des relations,

ainsi établies, a permis le développement de méthodes et de règles rigoureuses pour la construction des arbres phylogénétiques (**Cheikh, 2006**). 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 du temps,
- Les mutations se produisant chez un individu et se fixent dans sa descendance.

Pour reconstituer des phylogénies, on peut utiliser différents gènes, ou les protéines correspondantes, appelés marqueurs phylogénétiques (**Alauzet, 2009**).

Un marqueur génétique est basé sur la variation dans le génome, cette variation est aussi nommée polymorphisme. Selon le niveau taxonomique étudié, il est essentiel de choisir un type de marqueur qui lui est adapté. Par exemple, si le but visé 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 polymorphismes ADN peuvent être détectés dans l'ADN nucléaire et l'ADN d'organites, trouvé dans les mitochondries et les chloroplastes. Les marqueurs moléculaires concernent 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 (**Rischkowsky et al., 2008**).

- **Marqueurs d'ADN mitochondriaux :**

Les polymorphismes d'ADN mitochondrial (ADNmt) ont été largement utilisés lors des analyses de la diversité phylogénétique et génétique. L'ADNmt haploïde transporté par les mitochondries du cytoplasme cellulaire, possède un mode maternel d'hérédité (les animaux héritent l'ADNmt de leurs mères et non de leurs pères) et un taux de mutation élevé ; il ne se recombine pas. Ces caractéristiques consentent aux

biologistes de reconstruire les relations évolutives intra et interraciales par l'évaluation des modèles de mutation de l'ADNmt. Les marqueurs d'ADNmt peuvent également fournir un moyen rapide de détecter l'hybridation entre les espèces et les sous-espèces d'animaux d'élevage (Nijman *et al.*, 2003) ou même sauvage.

I.1.3. L'arbre phylogénétique :

L'arbre phylogénétique est une construction-clé dans l'histoire de la biologie et de la géologie. Son succès opérationnel ne s'est jamais démenti. Imaginons que nous retracions l'histoire de trois espèces choisies arbitrairement (actuelles ou fossiles). En remontant le temps, nous pouvons espérer relier les deux espèces qui dérivent de la même espèce ancestrale; en remontant plus loin encore, nous rencontrons l'espèce ancestrale des trois espèces. Le dessin en deux dimensions de ces relations de parenté est un arbre composé d'une succession de branchements.

Au reste, on n'a jamais trouvé meilleure façon d'illustrer, en deux dimensions, les relations de parenté entre espèces ou groupes d'espèces, en fonction du temps (verticalement) et de la diversité taxinomique (horizontalement). Dans le cas de la diversité, de façon conventionnelle et le plus souvent symbolique, la dimension horizontale permet aussi une représentation du degré de divergence morphologique de deux branches à partir du point de branchement, ce que Darwin nomme la « somme des modifications » (Darlu *et al.*, 2008).

I.2. LA PHYLOGENIE ET L'INFORMATIQUE

I.2.1. Les bases de données :

Depuis que les biologistes travaillent avec des séquences en grande quantité, c'est-à-dire depuis le développement et la généralisation de l'utilisation des méthodes rapides de séquençage, la nécessité d'organiser et d'accéder aisément à ces données s'est fait ressentir. Les premières banques de données qui ont été créées concernent les informations structurales sur les protéines puis, très rapidement après, les séquences protéiques et nucléotidiques (Périère, 2000).

Aujourd'hui on peut distinguer les banques dites généralistes des banques spécialistes (Vergnaud, 2002 ; Pillet, 2000). Les premières ont pour vocation d'être les plus exhaustives possibles, c'est-à-dire rassembler la totalité des séquences ou

informations connues pour tout l'ensemble des espèces avec ou sans expertise particulière. Les secondes se sont constituées autour d'une thématique biologique, afin de réunir les séquences d'une famille protéique pour toutes les espèces. Elles dérivent le plus souvent des banques généralistes et nécessitent l'intervention d'experts qui sont la plus part du temps les auteurs de la banque (**Dorkeld, 1994 ; Blanchet, 1999**).

1.2.2. Les banques généralistes :

Les banques généralistes sur les séquences nucléotidiques ou protéiques à accès libre et rapide, sont devenues des outils indispensables pour les chercheurs biologistes. Celles-ci sont régulièrement mises à jour (**Pillet, 2000**).

Dans cette section, nous présentons les principales collections généralistes de séquences nucléotidiques et protéiques ainsi que les centres de saisie qui leur sont associés.

1.2.2.1. Banques de séquences nucléotidiques :

D'après (**Perrière, 2000 ; Golding et al., 2003**), il existe de par le monde, trois banques généralistes de séquences nucléotidiques accessibles au publique : l'EMBL en Europe, GenBank aux Etats Unis et la DDBJ au Japon.

La base de données EMBL a été créée en 1980 à Hiedelberg (**Perrière, 2000 ; Guindon, 2003**). Elle est gérée et maintenue depuis 1994 par l'Institut Européen de Bioinformatique (EBI) à Cambridge. C'est la première source européenne de données sur les séquences nucléiques (**Pillet, 2000**).

GenBank a été mise en place en 1979 au LANL «Los Alamos National Laboratory», à Los Alamos. Depuis 1992, elle est maintenue au NCBI «National Center For Biotechnology Information», à Bathesda (**Perrière, 2000**). Le NCBI a été créé par le Congrès Américain en 1988 et est chargé spécifiquement de développer des systèmes automatisés pour la biologie moléculaire et la biotechnologie (**Pillet, 2000**).

Enfin, la DDBJ est une banque Japonaise de données créée en 1984 et maintenue au NIG à Mishima (**Perrière, 2000**).

Les données présentes dans ces banques proviennent désormais en quasi-totalité de soumissions directes effectuées par les auteurs, par l'intermédiaire du réseau Internet (**Golding et al., 2003**). Et la plus part des revues de biologie moléculaire n'acceptent de publier des articles se référant à des séquences que si celles-ci sont dotées d'un numéro d'accèsion fourni par les banques. Il convient donc, dès l'obtention d'une nouvelle séquence, de soumettre celle-ci à l'un des trois centres de saisie. Les données restantes sont extraites de la littérature scientifique (essentiellement à partir de documents tels que des livres ou des thèses). Il existe également des procédures de soumissions automatiques pour des séquences provenant des brevets (**Perrière, 2000**).

I.2.2.2. Banques de séquences protéiques :

Tout comme leurs homologues consacrés aux séquences nucléotidiques, les banques dédiées aux protéines ont une organisation centrée autour de types d'informations : les annotations biologiques et biochimiques d'une part (avec les références bibliographiques et les données taxonomiques associées) et les séquences d'une autre part (**Perrière, 2000**).

TrEMBL et GenPept sont des versions protéomiques des deux grandes banques généralistes de séquences nucléotidiques qui sont : EMBL et GenBank. Ainsi, TrEMBL est constituée de séquences protéiques issues de la traduction des séquences codantes de la banque de séquences nucléotidiques de l'EMBL. De son côté, GenPept, distribuée par le NCBI, correspond à la traduction des parties codantes de GenBank (**Blanchet, 1999**).

Ces deux systèmes ne peuvent cependant être considérés comme de véritables banques protéiques, du fait que les annotations qu'ils contiennent sont uniquement celles provenant des séquences nucléotidiques. Les véritables banques de protéines contiennent de nombreuses informations spécifiques (**Perrière, 2000**).

La principale banque de protéines à l'heure actuelle est sans conteste SwissProt, cette banque a été créée en 1986 par Amos Bairoch à Genève, et elle est maintenue et distribuée conjointement par le SIB et l'EBI (**Blanchet, 1999**). Les séquences présentes dans cette banque proviennent de quatre sources : la traduction des gènes annotés dans EMBL, certaines protéines issues d'autres banques protéiques, la

consultation des publications scientifiques et enfin, quelques soumissions directes par les auteurs (**Pillet, 2000**).

Enfin, la banque de séquences protéiques PIR est une des plus anciennes banques de données biologiques puisqu'elle est disponible sur internet depuis 1984. Elle fut mise en place par le NBRF comme succession à l'«Atlas of protein sequence and structure» maintenu depuis 20 ans par Margaret O. DAYHOFF. Depuis 1988, la banque PIR-International fait l'objet d'une collaboration internationale entre le NBRF et le MIPS et le JIPID (**Blanchet, 1999**).

D'autres banques ayant connu un succès bien moindre que SWISSPROT et PIR ont été également développées. On peut ainsi citer OWL, produite à l'université de Leeds en Angleterre (**Pillet, 2000**).

C'est une banque relativement petite en volume qui n'a pas connu de nouvelles versions depuis 1998 et sa maintenance semble désormais abandonnée. Plus récemment, une nouvelle banque a fait son apparition : il s'agit de KIND, une même banque que OWL mais construite à partir de TrEMBL, GenPept, SWISSPROT et PIR. Là encore, la fréquence d'apparition de nouvelles versions semble assez limitée (**Perrière, 2000**).

Les séquences recueillies par les trois centres de saisie : EBI, NCBI et NIG sont échangées quotidiennement, cela de façon à obtenir un ensemble de données aussi complet et cohérent que possible. La conséquence de cela est qu'en pratique, ces trois banques n'en font qu'une car leur contenu est identique à plus de 99,9%. Ceci amène régulièrement les responsables des centres de saisie à se poser la question de l'utilité de la maintenance de trois banques différentes. Il existe ainsi depuis longtemps un projet de fusion d'EMBL et GenBank et de la DDJB en un seul système (**Golding et al., 2003**).

Depuis leur création, la taille de ces banques n'a pas cessé de croître. Si l'on regarde l'allure de cette croissance, on constate qu'elle peut être approximativement assimilée à une exponentielle pour les séquences nucléotidiques. Par contre, le nombre de séquences protéiques croît beaucoup moins vite. D'après (**Perrière, 2000**), l'accélération constatée en 1995 correspond au développement des techniques de PCR, largement utilisée pour l'étude des génomes eucaryotes.

I.2.3. Les banques spécialistes :

Du fait de l'augmentation exponentielle du nombre de séquences nucléotidiques et protéiques publiées, la nécessité de développer des banques spécialisées est rapidement apparue. Ainsi, aujourd'hui, il en existe de nombreuses. Les informations disponibles dans ces banques et leur mise à jour sont très variables (**Blanchet, 1999**).

D'après (**Perrière, 2000**), on distingue deux types de banques spécialisées : les banques thématiques et les banques génomiques.

Les banques thématiques se consacrent à une thématique biologique bien précise (par exemple : banques sur les structures moléculaires des protéines, des complexes protéines-acides nucléiques, de virus et de polysides, banques sur les structures de l'ARN, banques sur des familles de gènes...etc.). Ces banques intègrent donc des séquences et des données qui sont spécifiques à ce domaine. On peut par exemple citer :

La base PDB est la plus connue des banques des données sur les structures moléculaires. Elle rassemble les structures tridimensionnelles des macromolécules obtenues par cristallographie aux rayons x. Elle contient aussi des références bibliographiques et des données sur la structure primaire et secondaire des protéines (**Golding et al., 2003**).

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**).

Les bases de données génomiques sont des banques de données centrées sur les génomes des organismes étudiés en biologie, c'est-à-dire l'ensemble des séquences d'ADN contenues dans les chromosomes d'organismes spécifiques. Ces organismes modèles peuvent être des virus, bactéries, champignons, plantes et animaux. En plus des données de séquences, ces bases rassemblent aussi d'autres types de données (description de phénotypes, mutations, expression de gènes, bibliographies...etc.) (**Pillet, 2000**). Dans cette catégorie on peut citer :

- ✓ 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.3. LES LOGICIELS UTILISEES DANS LA PHYLOGENIE

I.3.1. Présentation des logiciels d'analyse phylogénétique :

Aujourd'hui, il existe une myriade de logiciels, implantés sur différentes machines depuis le microordinateur jusqu'à des ordinateurs les plus puissants, écrits dans tel ou tel langage, traitant tel ou tel problème.

Dans le domaine de la reconstruction phylogénétique, il y a plusieurs programmes traitant les différents aspects de la phylogénie, allant des problèmes de comparaison des séquences, calcul des distances évolutives, reconstruction et visualisation d'arbres phylogénétiques et le Nettoyage de l'alignement.

Pour cela nous allons décrire quelques grandes catégories de logiciels utilisés pour les analyses phylogénétiques selon leurs fonctions.

I.3.1.1 Les méthodes de reconstruction phylogénétique et les logiciels utilisés :

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.3.1.1.1. La méthode de parcimonie :

Les méthodes de parcimonie reposent sur l'idée que l'histoire évolutive la plus plausible est celle qui nécessite un minimum de mutations. Ces méthodes recherchent donc parmi tous les arbres possibles, et toutes les séquences possibles aux nœuds

internes, la combinaison qui minimise le nombre total de mutations requis pour expliquer les données observées.

Les différentes mutations peuvent avoir des coûts différents, et il est possible d'attribuer des poids différents à chacun des sites.

Dans cette approche, les sites peuvent être traités de manière indépendante. Le nombre de mutations correspondant à un jeu de séquences particulier est égal à la somme des mutations requises par chacun des sites. Les coûts attribués aux mutations sont généralement symétriques (**Ranwez et al., 2003**).

Ces parcimonies sont réalisées par des logiciels spécialisés tels que : PHYLIP (Phylogeny Inference Package), PAUP, PAML, EDIBLE, PTP, PHYML, PARDOOT, ALIFRITZ, CRUX, PRAP.

I.3.1.1.2. La méthode de « Bayesian inference » :

La méthode de bayes, plus exactement appelée méthode Bayesian Markov Chain Monte Carlo (MCMC), a récemment apporté un nouveau moyen de construire l'arbre phylogénétique le plus probable. Cette méthode semble donner des résultats prometteurs et son emploi s'est dernièrement généralisé.

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.3.1.1.3. La méthode de Maximum de vraisemblance :

Les méthodes de maximum de vraisemblance utilisent un modèle mathématique du processus d'évolution des séquences pour définir la probabilité qu'une phylogénie puisse produire les séquences observées. Elles cherchent ensuite la phylogénie pour laquelle cette probabilité est maximale. Pour utiliser une méthode de maximum de vraisemblance, il faut donc être capable de choisir un modèle d'évolution, d'estimer ses paramètres, et de calculer la vraisemblance d'un arbre pour ce modèle (**Ranwez, 2013**) .

Cette méthode est réalisée par des logiciels spécialisés tels que : PHYLIP, DAMBE, EMBOSS, PRAP, PAUP*, ARB, FastTree, Creux, Vanilla, McRate, DART.

I.3.1.2. Les logiciels de Nettoyage de l'alignement :

C'est l'élimination des régions divergentes et présentant peu de signal phylogénétiques (peu informatives). Ces régions peuvent ne pas être homologues ou avoir été saturées par des substitutions; pour réaliser cette élimination il faut utilisée Séquencher ou Bio Edit.

Bio Edit est un éditeur d'alignement de séquences et un programme d'analyse de séquences piloté par la souris, facile à utiliser, conçu et écrit par un étudiant diplômé qui sait combien il est frustrant et consommateur de temps de compter sur les processeurs de mots et les programmes de ligne de commande pour la manipulation de séquence. Bio Edit est destiné à fournir un programme unique qui peut gérer les fonctions d'édition et de manipulation de séquences et d'alignement les plus simples que les chercheurs sont susceptibles de faire quotidiennement, ainsi que quelques analyses de séquences de base.

Sequencher est un logiciel de bio-informatique produit par la société Gene Codes Corporation. Ce logiciel permet de faire de l'assemblage de plusieurs séquences d'ADN contigües relativement courtes afin de créer des séquences plus longues (appelées contigs).

Le logiciel accepte les électrophorègrammes des différents contigs issus des séquenceurs de gène et les aligne (appariement), ce qui permet de les comparer les

uns aux autres afin de produire une séquence corrigée portant sur l'ensemble de la portée des contigs mis bout à bout.

I.3.1.3. Les logiciels pour l'alignement de séquence :

L'alignement est une opération qui permet de comparer des séquences de nucléotides pour repérer les éléments correspondants. Cette opération est nécessaire pour toutes les comparaisons de séquences (**Faurnel et al., 1834**).

Ces alignements sont réalisés par des logiciels spécialisés tels que : MEGA, MUSCLE, TreeAlign, clustal W, DAMBE, T.coffee, BIRCH, EMBOSS.

I.3.1.4. Les logiciels de traitement de séquence :

Quelques logiciels généralistes gratuits et téléchargeables pour le traitement des séquences biologique.

- 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).
- FastPCR : 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, calcule 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.3.1.5. Logiciels de visualisation et reconstruction d'arbres phylogénétiques :

Historiquement, les premières personnes ayant démocratisé la visualisation sous forme d'arbre sont les généalogistes et les naturalistes. De plus, **Charles Darwin** les réutilisa allègrement en 1859 dans son ouvrage devenu référence De l'origine des

espèces. Mais si l'on cherche le précurseur historique, c'est en réalité **Willi Hennig**, père de la cladistique, qui a réussi à démocratiser les arbres phylogéniques.

Comme en généalogie où les arbres sont utilisés pour visualiser les relations de parenté, les arbres sont un très bon moyen de faire ressortir l'évolution d'une ou plusieurs espèces.

Actuellement, Il existe plusieurs programmes qui permettent la visualisation des arbres phylogénétiques, par exemple, les programmes : PhyloDraw, NJplot, GeneTree, Phylip (DRAWTREE et DRAWGRAM), Genedoc, Dambe, Treecon, TreeView, et Spectrum (**Choi et al., 2000**).

I.4. REPRESENTATION DES ESPECES DE GENRE CANIS

I.4.1. Le genre *Canis*, Linne 1758: particularités et répartition

Canis est un genre de mammifères carnivores de la famille des canidés qui regroupe les chiens, les dingos, les loups, les coyotes et les chacals. Il est, comme tous les canidés, pourvus d'un excellent odorat (**Linnæus, 1758**). Ce genre est réparti sur les continents eurasiatique et africain et sur le continent Nord-américain. Rappelons que les chiens dont la répartition n'est pas liée à un biotope mais à la présence humaine sont, au vue des analyses génétiques, une sous-espèce de loup dont le nom scientifique valide est *Canis lupus familiaris* (**Tedford, 2009**).

I.4.1.1. Présentation des espèces :

I.4.1.1.1. Classification :

La classification systématique de ce mammifère est comme suit :

Tableau I : La classification de toutes les espèces de genre *Canis*.

Règne	Animales
Embranchement	Chordates
Sous Embranchement	Vertébrés
Classe	Mammifères
Ordre	Carnivores
Famille	Canidae
Genre	<i>Canis</i>

Espèce	<i>Canis adustus</i>
	<i>Canis aureus</i>
	<i>Canis latrans</i>
	<i>Canis mesomelas</i>
	<i>Canis simensis</i>
	<i>Canis anthus</i>
	<i>Canis himalayensis</i>
	<i>Canis indica</i>
	<i>Canis rufus</i>
	<i>Canis lycaon</i>
	<i>Canis environmental sample</i>
	<i>Canis lupus</i>
Sous-espèce de <i>Canis lupus</i>	<i>Canis lupus campestris</i>
	<i>Canis lupus chanco</i>
	<i>Canis lupus desertorum</i>
	<i>Canis lupus dingo</i>
	<i>Canis lupus hattai</i>
	<i>Canis lupus hodophilax</i>
	<i>Canis lupus labradorius</i>
	<i>Canis lupus laniger</i>
	<i>Canis lupus lupus</i>
	<i>Canis lupus mogollonensis</i>
	<i>Canis lupus pallipes</i>
	<i>Canis lupus signatus</i>
	<i>Canis lupus variabilis</i>
	<i>Canis lupus arctos</i>
	<i>Canis lupus baileyi</i>
	<i>Canis lupus lupaster</i>

I.4.1.1.2 Répartition et Morphologie :

La répartition et les caractères morphologiques des espèces de genre *Canis* ont été largement étudiés par des nombreux auteurs (**Philip G., 2012 ; Sillero-Zubiri C, Gottelli D., 1994**).

Parmi les 33 espèces de genre *Canis*, les espèces *Canis aureus* et *Canis lupus* qui est représenté par le sous espèce *Canis lupus lupaster* seraient les seules espèces de genre *Canis* qui existent en Algérie.

Tableau II : La répartition et la morphologie des deux espèces en l'Algérie et dans le monde.

Espèce	Répartition dans le monde	Répartition en l'Algérie	Morphologie
<i>Canis aureus</i> (le chacal doré), (Linnaeus, 1758)	Il se produit dans le nord d'Afrique (Egypte, Libye, Tunisie, Algérie, Maroc, Sénégal, Sahara occidental), l'est Africain (Kenya, Ethiopie) et le sud-est de l'Europe au Moyen-Orient (Iran, Turquie, Jordan, Israël, Grèce) et jusque au sud de l'Asie (la Birmanie), (Dorst et al., 1976) .	-Il serait la seule espèce de chacal existant en Algérie, elle est représentée par l'unique sous espèce <i>canis aureus algirensis</i> , (Wagner 1841) . -Le chacal doré est très répandu dans tout le pays, depuis le littoral jusqu'aux limites méridionales. -On le rencontre dans les montagnes du centre Sahara (Hoggar, Tassili N'Ajjer) et même à une centaine de kilomètres des points d'eau dans les zones désertiques (Régnier, 1960 ; Kowalski et al., 1991) .	-Longueur du corps est de 70 à 85 cm. -Longueur de queue environ de 25 cm. -Taille debout est approximativement 40 cm. -Fourrure est généralement brute et pas très longtemps. -Manteau est habituellement jaune pour pâlir l'or et brun-incliné, mais la couleur peut varier avec la saison et la région.

<i>Canis lupus lupaster</i> (le loup Afriacain), (Hemprich et al., 1833).	Afrique du nord-est (Sénégal, Algérie) ; Egypt, sud Sinäi, nord de ville Sharm el Shikh.	-Il a déjà été observée dans le nord-est de l'Algérie, il s'étendant des régions méditerranéennes, côtières et accidentées (terres cultivables protégées y compris, maquis, bois du pin et forêts de chêne), région côtière entre Skikda et El-Kala.	-Un grand corps mince. -Une tête plus large. -Un cou large et puissant. -Il porte un manteau sombre. -La queue est courte. (Philippe Gaubert et al., Août 2012).
--	--	--	--

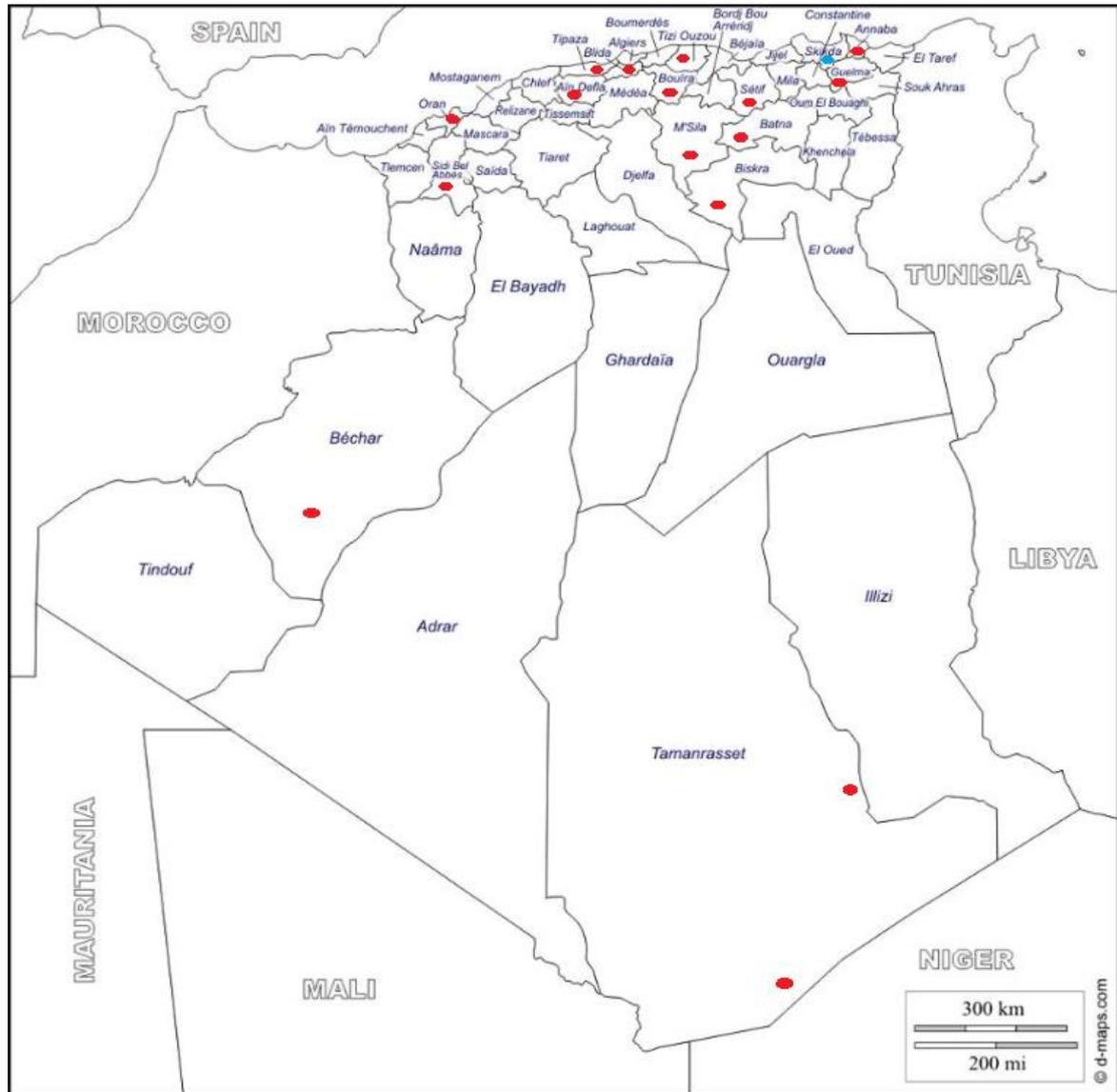


Figure 1 : Carte de distribution du chacal doré (*Canis aureus*) (Kowalski et al., 1991) et du loup africain (*Canis lupus lupaster*) (Philippe Gaubert et al., 2012) en Algérie.

● *Canis lupus lupaster*

● *Canis aureus*

Tableau III : Répartitions géographiques et description morphologique des 31 espèces existant dans le monde de genre *Canis*.

Espèce	Répartition	Morphologie
<i>Canis anthus</i> (loup Africain doré) (Cuvier, 1820).	Le nord-ouest et le nord-est d'Afrique, du Sénégal à l'Égypte à l'est, dans une région incluant le Maroc, et la Libye dans le nord au Nigéria, au Tchad et en Tanzanie au sud et dans les montagnes de l'Atlas.	<ul style="list-style-type: none"> -Les deux sexes pesant 7-15 kg et 40 cm de hauteur. -Il a un museau relativement long. -La queue est relativement courte, mesurant 20 cm de longueur. -La couleur de la fourrure varie individuellement, saisonnièrement et géographiquement. -La coloration typique soit de couleur jaunâtre à gris argenté, avec des membres légèrement rougeâtres et des taches noires sur la queue et les épaules. -La gorge, l'abdomen et les marques faciales sont habituellement blanches. -Les yeux sont de couleur ambrée.
<i>Canis himalayensis</i> (Aggarwal et al., 2007).	Jammu et Cachemire (Inde).	<ul style="list-style-type: none"> -Il mesure environ 70 cm au garrot à l'âge adulte pour un poids allant de 48 à 55 kg, -Il a un couleur fauve, on le trouve également portant un pelage noir (R. K. Aggarwal et al., 2007).
<i>Canis indica</i> (Aggarwal et al., 2007).	Inde.	-Il est relativement petit de taille par rapport <i>Canis lupus pallipes</i> (R. K. Aggarwal et al., 2007).
<i>Canis lupus</i> <i>Canis lupus arctos</i> (loup arctique) (Pocock, 1935)	IceLand, l'Amérique arctique.	<ul style="list-style-type: none"> -Le palais et le museau sont une bagatelle plus étroite, les largeurs à travers le zygomata et les processus post orbital sont relativement grand. -Le front est plus élevé (R. I. Pocock et al., 1935).
<i>Canis lupus baileyi</i> (loup gris Mexicain) (Nelson et al., 1929).	Nord du Mexique, Sud-ouest USA.	-Ils sont l'un des plus petites sous-espèces de loup gris en Amérique du Nord, mais il y a de grandes portées en taille chez les loups mexicains et autres sous-espèce (Philip et al., 2000).

<i>Canis lupus campestris</i> (loup de steppe) (Dwigubski, 1804).	Ukraine, dans l'Asie centrale, sud-est du Sibérie, la Chine.	-Il est plus petit que le loup gris commun (<i>canis lupus</i>). -Il a un poil plus dur, plus court et plus clair. -Son pelage est gris avec une teinte ocre.
<i>Canis lupus chanco</i> (loup Mongol) (Hodgson, 1847).	Centre d'Asie (la Chine, Mongolie, Tibet, Cachemire, Népal).	-La couleur générale est plus riche en ocre que dans le type de <i>chanco</i> , est un corps plus petit et des jambes plus courtes. -Le manteau complet d'hiver est plus court qu'en long. -Les poils de contour du dos étant d'environ 45 mm. -Les oreilles et la couronne sont brun foncé. -Le front et le museau sont bruns. -Pas de patch pâle sur les yeux. -Les joues au coin de la bouche gris. -Les cheveux étant à tête noire (R.I.Pocock et al., 1935).
<i>Canis lupus desertorum</i> , (Bogdanov, 1882).	Altai dans Shinjang, Chine.	-Il a un poil court et dur. -Le pelage plus clair que celui du loup des forêts d'Europe et d'Amérique du nord. -Son pelage est de couleur gris sombre avec une teinte ocrée (Proche de <i>Canis lupus campestris</i>) (Honghai Zhang et al., 2013).
<i>Canis lupus dingo</i> (Meyer, 1793).	l'Australie et Sud-Est d'Asie (sud-est de la Chine, Laos, Malaisie, Thaïlande, Indonésie, en Bornéo.	-Les mâles adultes sont généralement plus grands que les femelles pèsent entre 11,8 et 19,4 kg, la longueur du corps moyenne de 920 mm. -Les femelles pèsent entre 9,6 et 16,0 kg et une moyenne 885 mm de longueur du corps. -Les dingos sont en général de couleur gingembre avec les points blancs (Oakman, 2004; Corbett, 1995).
<i>Canis lupus hattai</i> (Ezo loup) (Kishida, 1931).	Le Hokkaido et le Sakhaline, les îles de Kourile (Japon).	-De 70 à 80 cm à la garenne. -Il a une énorme tête et une bouche armée d'énormes crocs et dents. -Il est généralement très maigre mais extrêmement musclé.

		<p>-La couleur grise en été et blanc grisâtre en hiver.</p> <p>-Sa fourrure est épaisse et longue.</p> <p>-Ses pieds sont remarquables pour leur taille, trois ou quatre fois plus grandes que les pieds du plus grand chien qu'ils ressemblent à la forme, seules les griffes sont beaucoup plus longues. Leurs gros pieds leur permettent de voyager rapidement sur une neige profonde.</p>
<i>Canis lupus hodophilax</i> (loup Japonais) (Temminck, 1839).	Japon (Honshū, Kyushu et Shikoku).	<p>-Un corps et des jambes plus petites.</p> <p>-Le manteau complet d'hiver est plus court qu'en long.</p> <p>-Les poils de contour du dos étant d'environ 45 mm.</p> <p>-Les oreilles et la couronne brun foncé.</p> <p>-Le front et le museau brun, pas de patch pâle sur les yeux.</p> <p>-Les joues au coin de la bouche gris.</p> <p>-Les cheveux étant à tête noire (R. I. Pocock et al., 1935).</p>
<i>Canis lupus labradorius</i> (loup de labrador) (Goldman, 1937).	Nord-Américain, et Nord Canada.	-Il a une couleur noir, gris grizzly au presque blanc.
<i>Canis lupus laniger</i> (loup tibétain) (Hodgson, 1847).	Sud-est d'Asie (Chine "Sunbio, Beijing", Mongolie, Tibet).	<p>-Son manteau d'hiver est plus épais et plus laineux.</p> <p>-La laine étant en règle générale moins cachée par les poils de contour, et aussi par la longueur plus courte du crâne, bien que cette différence, bien marquée par rapport au loup typique, à la couleur générale extrêmement variable saisonnièrement (R. I. Pocock et al., 1935).</p>
<i>Canis lupus lupus</i> (loup eurasiens) (Linnaeus, 1758).	Eurasie, nord et centre d'Europe (Suède, Italie, France).	<p>-Le manteau, bien que bien épaissi avec le sous-bois n'est pas très long.</p> <p>-La laine est d'environ 45 ou 50 mm.</p> <p>-La couleur générale n'est pas fauve, les côtés du corps étant clairsemés, ou peut-être obscurément gris vif, saupoudrés de pointes de cheveux noirs.</p>

		<p>-Le dos, le côté supérieur du cou et de la queue fortement marbré ou strié.</p> <p>-Les bandes noir et blanc des poils de contour, les oreilles sont brun foncé.</p> <p>-La couronne et la moitié supérieure du museau sont gris foncé (R. I. Pocock et al., 1935).</p>
<p><i>Canis lupus mogollonensis</i> (loup de montagne Mogollon) (Goldman, 1937).</p>	<p>Les états Américains de l'Arizona et du Nouveau-Mexique.</p>	<p>-Il mesurait de 1,35 à 1,50 m de long pour un poids allant de 27 à 36 kg.</p> <p>-La couleur de son pelage était un mélange de noir et fauve, avec une bande de couleur cannelle le long du dos.</p>
<p><i>Canis lupus pallipes</i> (loup indien) (Sykes, 1831).</p>	<p>Moyen-oriental (Iran, Palestine) qui s'étend du sous-continent indien vers la péninsule Arabe, l'Inde ouest et le Pakistan</p>	<p>-Le manteau complet d'hiver est relativement court et mince.</p> <p>-La queue n'est pas pleine.</p> <p>-Les cheveux de contour mesurant 67 mm au dos.</p> <p>-La couleur générale de la laine est Buffy sur les flancs, la nuque Buffy.</p> <p>-Le front est noir, avec de petites taches pâles au-dessus des yeux (R. I. Pocock et al., 1935).</p>
<p><i>Canis lupus signatus</i> (loup ibérique) (Cabrera, 1907).</p>	<p>Afghanistan, Arabie Saoudite, Nord du Portugal, ouest-nord d'Espagne.</p>	<p>-C'est un loup plutôt petit.</p> <p>-Son manteau d'hiver complet est plus court.</p> <p>-La couleur général est un brun grisâtre noir, à peine soulagé par les zones pâles des cheveux de contour.</p> <p>-L'oreille est rouillée, Il a un grand morceau blanc de chaque côté du museau à l'avant, mais la partie de la lèvre et la totalité de la joue jusqu'à l'œil sont gris, mouchetées.</p> <p>-L'ensemble de la gorge, à l'exception de sa partie antérieure blanchâtre, est gris sombre.</p> <p>-La queue est comme le dos, avec une pointe noire (R. I. Pocock et al., 1935).</p>

<p><i>Canis lupus variabilis</i> (Pei, 1934).</p>	<p>Il provient de la rive droite du bas Rivière d'Adycha (bassin de Yana River) à Ulakhan-Sullar, (est/Asie du sud, le Proche-Orient, la Sibérie).</p>	<p>-Il a une taille de corps relativement petite. -Un museau allongé et mince. -Un métastyle relativement long (Esther et al., 2015).</p>
<p><i>Canis rufus</i> (loup rouge) (Audubon et al., 1851).</p>	<p>Sud-Est des Etats-Unis, des secteurs de leur gamme historique-rivière d'alligator en nord du Caroline.</p>	<p>-Il a une petite taille. -Longues jambes et oreilles. -Sa fourrure plus courte. -Ils ont une longueur totale entre 1000 et 1300 mm. -La longueur de queue de 300 à 420 mm. -La taille d'épau de 660 à 790 mm. -La couleur fauve, gris ou noir. -Le dos est noirâtre. -Le museau et les membres sont fauves. -La queue est inclinée (Nowak et al., 1991; Wayne, 1995).</p>
<p><i>Canis lycaon</i> (eastern Canadian wolf) (Schreber, 1775).</p>	<p>Nord-est d'Amérique, Eurasie, l'est du Canada.</p>	<p>-Il est plus petit que le loup européen. -Les oreilles plus larges et largement séparées. -La queue n'est pas si grosse que chez les loups ordinaires. -La peau est noire. -Elle à des lèvres grises et un patch blanc sur la poitrine devant les pattes avant (R. I. Pocock et al., 1935).</p>
<p><i>canis adustus</i> (Sundevall, 1847).</p>	<p>Ils sont trouvés en Afrique tropical s'étendant de 15 degrés de nord à 23 degrés de l'attitude de sud.</p>	<p>-Des jambes et des oreilles plus courtes. -Ils tendent à être gris-clair pour se bronzer et sont perceptibles par une astuce blanche sur leurs queues relativement foncées. -Ils tendent à avoir une rayure blanche de coude aux rayures latérales de hanche et de noir qui ne sont pas</p>

	Et en Afrique centrale, et ont été connus pour habiter des secteurs aussi hauts que 2.700 mètres.	toujours remarquables. -Les mâles sont légèrement plus grands que les femelles ; les mâles s'étendent de 7,3 à 12 kg, tandis que les femelles sont rarement connues pour peser plus de 10 kg (Kingdon, 1977).
<i>Canis latrans</i> (les coyotes) (Say, 1823).	Ils sont indigènes à la région de Néarctique. Ils sont trouvés dans l'ensemble du nord de l'Amérique centrale. Ils s'étendent du Panama, Mexique, Etats-Unis, Canada.	-La couleur varie du brun grisâtre à un gris jaunâtre sur les parties supérieures. -La gorge et le ventre sont blanchâtres. -Les pattes antérieures, les côtés de la tête, le museau et les pieds sont brun-rougeâtre. -Le dos a sous-pelage coloré fauve et les longs, noir-inclinés poils de garde qui produisent une rayure dorsale noire et une croix foncée sur le secteur d'épaule. -La queue, qui est moitié de la longueur du corps, noire. -Plus petits que les loups gris et beaucoup plus grands que les renards. -Le nez est noir et habituellement moins d'un pouce de diamètre. -Les oreilles sont grandes par rapport à la tête. -Le museau est long et mince. -Les pieds sont relativement petits pour la taille du corps (Banfield, 1974).
<i>Canis mesomelas</i> (le chacal à dos noir), (Schreber, 1775).	Il se trouve seulement en Afrique. Les espèces vit dans deux secteurs discrets séparés approximativement	-La caractéristique principale du chacal à dos noir, qui lui donne son nom, est : -Les cheveux noirs de derrière du cou à la queue (Van Valkenburgh, 1994). -Les adultes tiennent environ 38 cm à l'épaule et sont presque un mètre long dans la longueur. -La tête est comme un chien.

	de 900km. Afrique du sud, la Namibie, le Botswana, et le Zimbabwe ; Kenya, la Somalie, et l’Ethiopie (Smithers, 1983).	-Le museau aigu et de hautes oreilles aiguës. -Le manteau d’hiver des adultes masculins développe un rougeâtre à une couleur rousse presque profonde. (Ginsberg, 1990).
<i>Canis simensis</i> (loup éthiopien) (Ruppell, 1840).	On le trouve seulement dans six ou sept chaînes de montagne de l’Ethiopie. Ceci inclut les montagnes d’Arssi et de balle de l’Ethiopie du sud-est, les montagnes de Simien, du Shoa du nord-est, du Gojjam, et du Mt. Guna (Ginsberg et al., 1990; Sillero-Zubiri et al., 1995).	-Ils ont un manteau rougeâtre avec tache blanche sur les jambes, le bas ventre, la queue, le visage, et le menton. -La frontière entre la fourrure rouge et blanche est tout à fait distincte. -Les taches blanches sur le visage incluent un croissant blanc caractéristique au-dessous des yeux et une tache blanche sur les joues. -Le menton et la gorge sont également blancs. -La queue est identifiée par une rayure noire indistincte en bas de sa longueur et une brosse des poils noirs à l’astuce. -Les oreilles sont larges et aigu et le nez, les gorges, et le palais sont noirs. -Les mâles mesurent de 928 à 1012 mm et les femelles de 841 à 960 mm. -La queue est de 270 à 396 mm de longueur.
<i>Canis sp</i>	Japon, Allemagne, Canada, USA, Angleterre, Palestine, Chine,	
<i>Canis sp. Belgium/36,000</i>	Belgique	
<i>Canis sp. CANInt1</i>	Canada	

<i>Canis sp. Russia/33,500</i>	Russie	
<i>Canis environmental sample</i>	USA	
<i>Canis lupus x Canis lupus familiaris</i>	Estonie	
<i>Canis sp. KEB-2016</i>	Nord-Américain	

Remarque : Pour les 7 dernières espèces il n'y a pas de données sur leurs morphologique.



Figure 2 : Le chacal doré *Canis aureus* (le parc national de Ben Aknoun, 2017).



Figure 3 : Le coyote *Canis latrans* (le parc national de Ben Aknoun, 2017).



Figure 4 : Le loup gris *Canis lupus* (le parc national de Ben Aknoun, 2017).

II.1. ANALYSES MOLECULAIRES ET PHYLOGENETIQUES

L'objectif de cette partie repose sur le positionnement de différentes espèces de genre *Canis* dans un arbre phylogénétique, réalisé à partir des séquences d'ADN des gènes traités par des logiciels de la bio-informatique. Pour cela nous décrirons la méthodologie utilisée depuis l'échantillonnage, l'extraction, la PCR, le séquençage et l'analyse des séquences d'ADN.

Notre étude sera basée sur le traitement des séquences d'ADN qui se trouvent au niveau de la GenBank et leur utilisation dans une étude phylogénique.

II.1.1. Etapes d'Analyses Moléculaires :

II.1.1.1. Echantillonnage :

Consiste essentiellement à tirer des informations d'une fraction d'un grand groupe ou d'une population, de façon à en tirer des conclusions au sujet de l'ensemble de la population. Son objet est donc de fournir un échantillon qui représentera la population et reproduira aussi fidèlement que possible les principales caractéristiques de la population étudiée. Afin d'obtenir une représentation correcte de l'histoire à inférer (**Rannala et al., 1998 ; Agolin, 2003**).

II.1.1.2. Extraction d'ADN :

L'extraction permet d'obtenir un ADN pur et intacte ; l'extraction d'ADN à partir d'un tissu est réalisée grâce à des kits qui facilitent le travail.

II.1.1.3. La PCR :

La PCR (Polymérase Chain Reaction) est une technique mise au point en 1985 par **Kary Mullis** et ses collaborateurs. Elle permet d'amplifier *in vitro* une séquence spécifique d'un acide nucléique et d'obtenir une quantité importante de cette séquence, afin de la détecter et de la manipuler. Elle permet donc d'obtenir, à partir d'un échantillon complexe et peu abondant, d'importantes quantités d'un fragment d'ADN spécifique et de longueur définie. Dans un laps de temps court (deux heures pour une séquence de 1000 pb), environ un million de copies de la séquence souhaitée peuvent être obtenues. La quantité d'amplifiât obtenue est généralement suffisante pour une utilisation ultérieure.

Le principe et les conditions expérimentales de la PCR sont simples. Il s'agit de réaliser une succession de réactions de répllication d'une matrice double brin d'ADN. Chaque

réaction met en œuvre deux amorces oligonucléotidiques orientées en sens inverse. Les amorces définissent alors, en la bornant, la séquence à amplifier.

II.1.1.4. Le Séquençage :

Le séquençage consiste à retrouver la suite de nucléotides correspondant au brin d'ADN étudié. La méthode utilisée aujourd'hui est celle proposée par Frederick Sanger (**Sanger et al., 1977**). Depuis 1977, la méthode a considérablement évolué grâce à la mise au point de séquenceurs automatiques et du marquage des nucléotides à l'aide de fluorochromes. Elle est devenue aujourd'hui une technique rapide et fiable utilisée fréquemment dans le diagnostic des maladies héréditaires.

II.1.2. Etapes d'Analyses phylogénétique :

II.1.2.1. Etude 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.

Afin d'une reconstruction phylogénétique du genre *Canis*, nous récupérons les séquences qui se trouvent au niveau de la GenBank, nous passons par plusieurs étapes distinctes :

1. Nous construisons l'URL suivante : <http://www.ncbi.nlm.nih.gov/Genbank/>
2. Nous avons choisi recherche par taxonomie.
3. Nous avons écrit le nom de notre genre « *Canis* », (**Figure 5**).

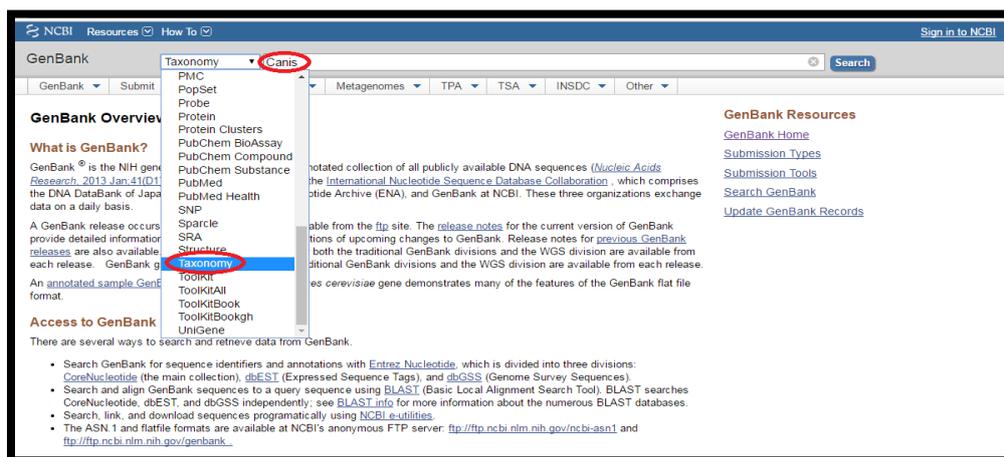


Figure 5 : Les étapes initiales pour récupérer les séquences de la GenBank.

4. En entrant sur *Canis*, nous obtenons toutes les espèces qui appartiennent à ce genre (**Figure 6**).

▪ <i>Canis himalayensis</i> (Himalayan wolf)
▪ <i>Canis indica</i>
▪ <i>Canis latrans</i> (coyote)
○ <i>Canis lupus</i> (gray wolf)
▪ <i>Canis lupus arctos</i> (Arctic wolf)
▪ <i>Canis lupus baileyi</i> (Mexican gray wolf)
▪ <i>Canis lupus campestris</i> (steppe wolf)
▪ <i>Canis lupus chanco</i> (Mongolian wolf)
▪ <i>Canis lupus desertorum</i>
▪ <i>Canis lupus dingo</i> (dingo)
▪ <i>Canis lupus familiaris</i> (dog)
▪ <i>Canis lupus hattai</i> (Ezo wolf)
▪ <i>Canis lupus hodophilax</i> (Japanese wolf)
▪ <i>Canis lupus labradorius</i> (Labrador wolf)
▪ <i>Canis lupus laniger</i> (Tibetan wolf)
▪ <i>Canis lupus lupaster</i> (African wolf)
▪ <i>Canis lupus lupus</i> (Eurasian wolf)
▪ <i>Canis lupus megalonyx</i>
▪ <i>Canis lupus pallipes</i> (Indian wolf)
▪ <i>Canis lupus signatus</i> (Iberian wolf)
▪ <i>Canis lupus varabilis</i>
▪ <i>Canis lupus × Canis lupus familiaris</i> (wolf/dog)
▪ <i>Canis lycaon</i> (eastern Canadian wolf)
○ <i>Canis mesomelas</i> (black-backed jackal)
▪ <i>Canis mesomelas elongae</i> (eastern African black-backed jackal)
▪ <i>Canis rufus</i> (red wolf)
▪ <i>Canis simensis</i> (Ethiopian wolf)
▪ <i>Canis sp.</i>
▪ <i>Canis sp.</i> Belgium/36,000
▪ <i>Canis sp.</i> CAN101
▪ <i>Canis sp.</i> KEB-2016
▪ <i>Canis sp.</i> Russia/33,500
○ environmental samples
▪ <i>Canis environmental sample</i>

Figure 6 : Les 34 espèces de genre « *Canis* » obtenues dans la GenBank.

5. Nous avons analysé toutes les espèces une par une pour identifier le nombre des séquences nucléotidiques trouvées dans cette banque (**Figure 7**).

Search for: **Canis adustus**

Display: 13 levels using filter: none

Taxonomy ID: 68723
Genbank common name: side-striped jackal
Inherited blast name: carnivores
Rank: species
Genetic code: Translation table 1 (Standard)
Mitochondrial genetic code: Translation table 2 (Vertebrate Mitochondrial)

Entrez records

Database name	Direct links
Nucleotide	80
Protein	63
Popset	30
PubMed Central	19
Taxonomy	1

External Information Resources (NCBI LinkOut)

LinkOut	Subject	LinkOut Provider
Canis adustus	taxonomy/phylogenetic	Animal Diversity Web
Canis adustus Sundevall, 1846	taxonomy/phylogenetic	AnimalBase
Canis adustus Sundevall, 1847	taxonomy/phylogenetic	Arctos Specimen Database
Canis adustus taxonomy	taxonomy/phylogenetic	Dryad Digital Repository
dryaddb	supplemental materials	Encyclopedia of life
dryaddb	supplemental materials	Global Biotic Interactions
Canis adustus Sundevall, 1847	taxonomy/phylogenetic	Integrated Taxonomic Information System
Show Biotic Interactions	taxonomy/phylogenetic	
Canis adustus Sundevall, 1847	taxonomy/phylogenetic	

Figure 7 : Exemple des séquences nucléotidiques trouvées pour l'espèce « *Canis adustus* ».

Au total, la GenBank contient 12 179 séquences de 33 espèces du genre « *Canis* ».

- 2423 Séquences d'ADN nucléaire.
- 9756 Séquences d'ADN mitochondrial.

II.1.2.2. 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 (Djekoun et Hamidechi ; 2010), mais nous avons choisi FaBox, DnaSP5, MEGA5, DAMBE, Network.

Nous pouvons procéder à la construction des arbres phylogénétiques à partir des séquences nucléotidiques de chaque espèce du genre « *Canis* » par différents étapes :

Les séquences représentatives de chacune des 33 espèces ont été téléchargées sous format FASTA à partir de "GenBank" (Figure 8).

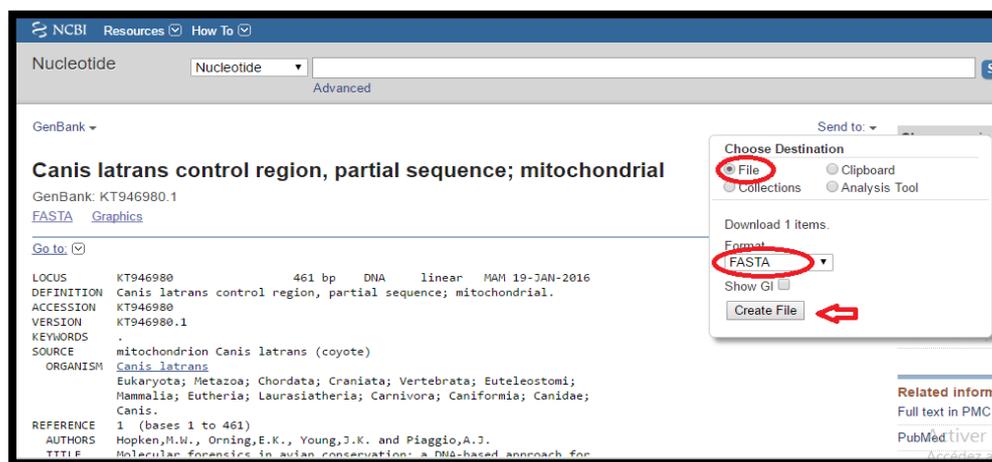


Figure 8 : Exemple d'un fichier sous format Fasta de « *Canis latrans* ».

II.1.2.2.1. 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 rend facile d'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 s'effondrer 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 ; 2010).

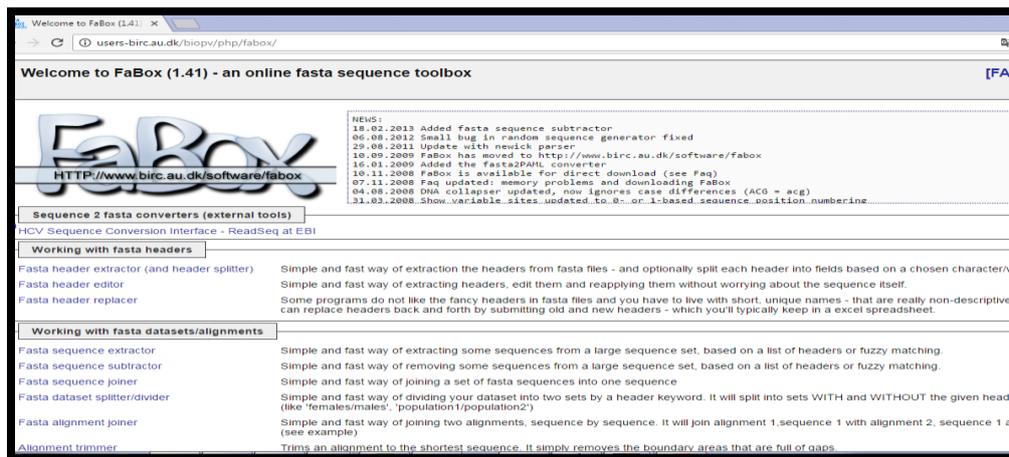


Figure 9 : Menu principal de logiciel FaBox.

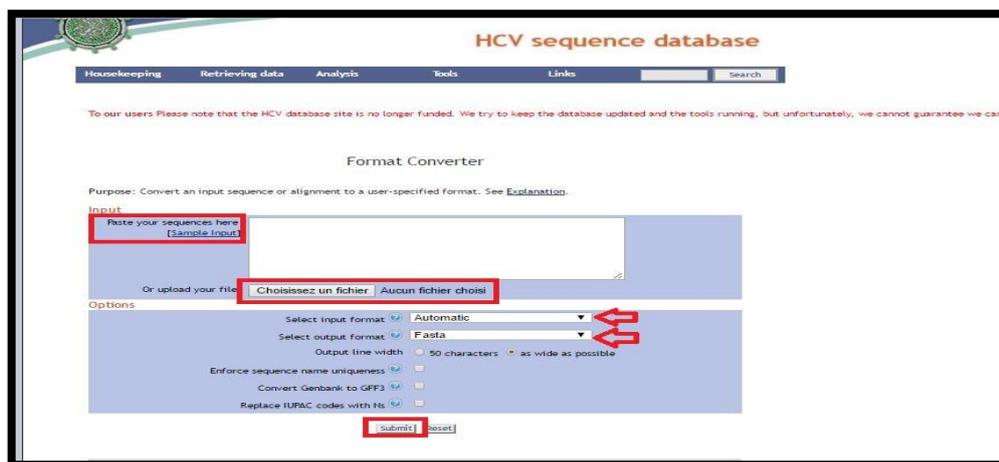
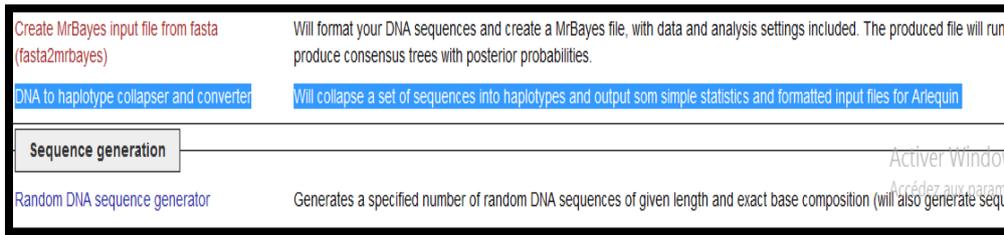


Figure 10 : L'outil « format convertir » de FaBox été utilisé pour convertir les séquences Fasta à partir d'un bloc note.



Frequencies and report of haplotypes

Haplotype number	Label	Number of sequences	Length of haplotype (bp)	Sequences belonging to haplotype
1	1	1303	1303	canis himalayensis- Inde
2	2	1	1310	Canis indica- Inde
3	3	1	460	canis aureus- Lituanie
4	4	1	406	canis aureus- Grèce
5	5	1	406	canis aureus- Ukraine
6	6	1	450	canis aureus- Italie
7	7	1	1320	canis aureus- inde
8	8	1	289	canis aureus- Sénégal
9	9	1	664	canis aureus- Croatie
10	10	1	325	canis rufus- Caroline du Nord (USA)
11	11	1	420	canis rufus- USA
12	12	1	233	canis simensis- USA
13	13	1	306	canis sp-KEB-2016- USA
14	14	1	218	canis environmental sample- USA
15	15	1	1268	canis latrans- USA
16	16	1	521	canis latrans- Canada
17	17	1	1489	canis latrans- Chine
18	18	1	444	canis latrans- Caroline
19	19	1	329	canis latrans- Caroline du Nord
20	20	1	423	canis latrans- Sud-est des États-Unis
21	21	1	291	canis sp- Allemagne
22	22	1	591	canis sp- USA
23	23	1	610	canis sp- Roumanie
24	24	1	583	Canis lupus chanco- China
25	25	1	425	Canis lupus dingo- Taiwan

Figure 11 : L’outil « DNA to haplotype collapser and converter » pour l’obtention des séquences en haplotypes.

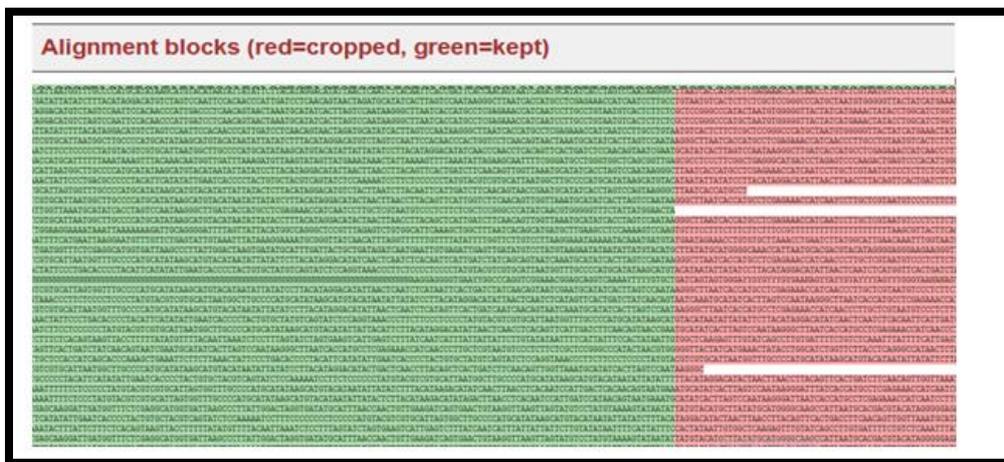


Figure 12 : L’outil « alignement trimmer » utilisé pour l’alignement de la séquence la plus courte ; elle supprime les zones qui sont plein de gaps.

II.1.2.2.2. MEGA :

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

plus récent ajout au MEGA5 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, ML5 MEGA 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).

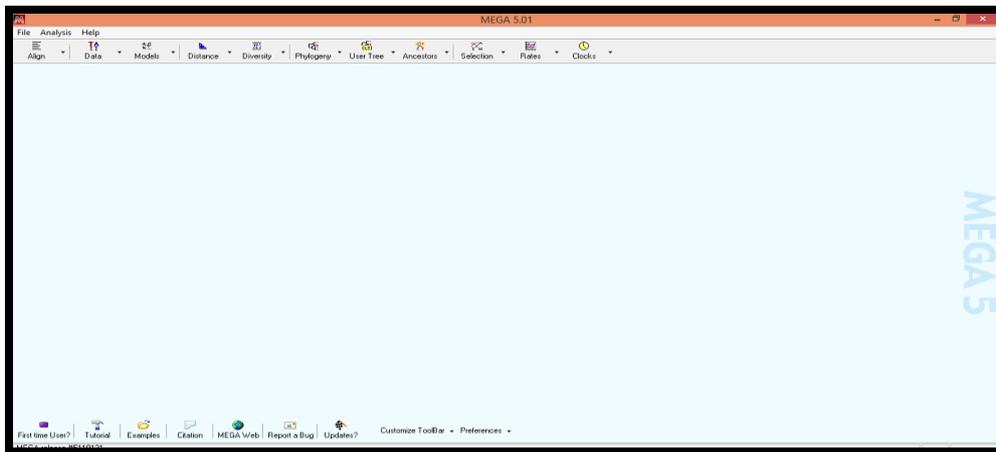


Figure 13 : Menu principal de logiciel MEGA version 5.

- Les séquences représentatives de chacune des 33 espèces ont été ajoutées une par une sous format FASTA à partir de "GenBank" (Figure 14).



Figure 14 : Exemple des 3 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 ClustalW (**Figure 15**), 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.

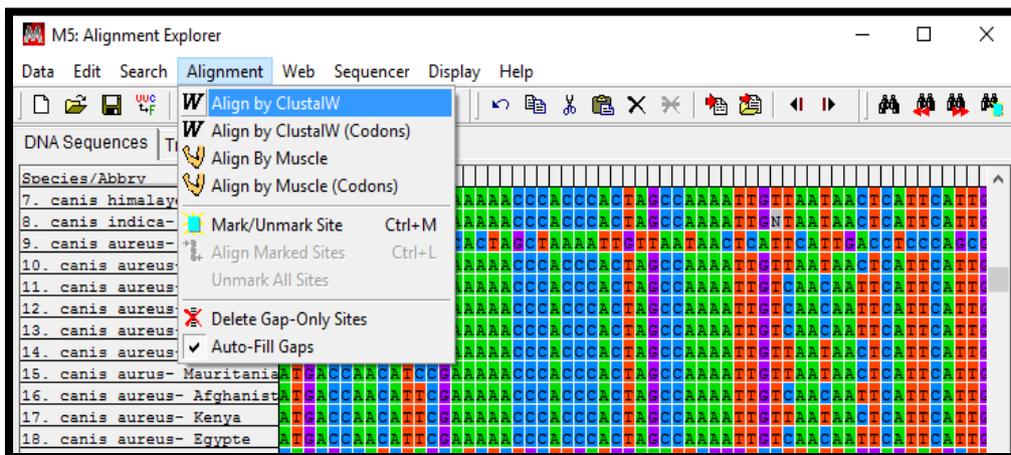


Figure 15 : L'alignement multiple des séquences du genre «*Canis* » en utilisant le programme ClustalW.

- 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 16**), 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 se fait pas à pas grâce à cet ordre.

- En utilisant le modèle de « Kimura 2-paramètres » (**Figure 17**) 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 17**) répliqués sont ramenés en pourcentages et ils sont indiqués au niveau des nœuds.

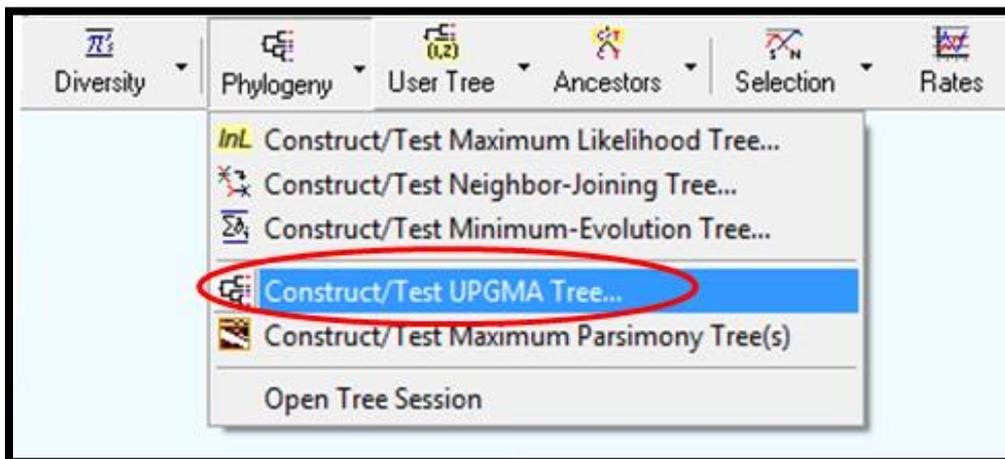


Figure 16 : La construction de l'arbre phylogénétique en utilisant la méthode UPGMA.

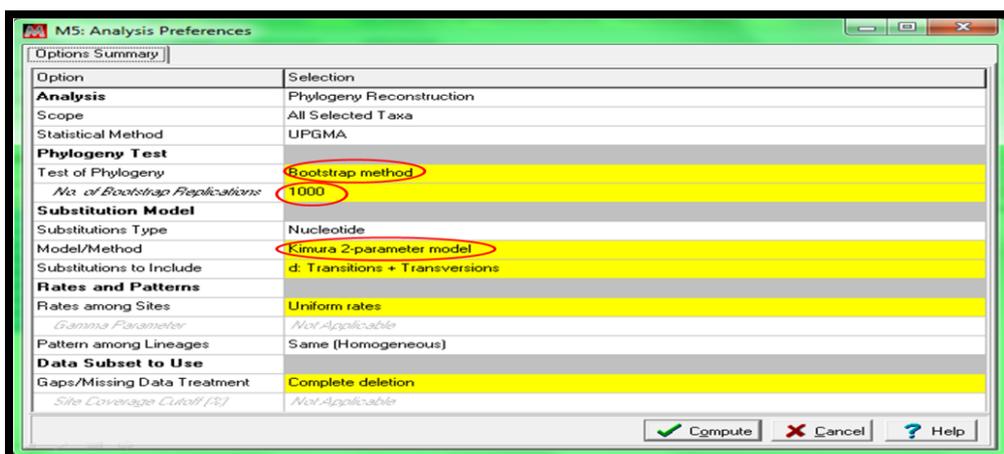


Figure 17 : Les paramètres utilisés pour la reconstruction phylogénétique.

II.1.2.2.3. 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 divers 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 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 18 : Menu principal de logiciel DnaSP 5.10.

Pour avoir les haplotypes de ces séquences, nous avons utilisées les paramètres suivants :

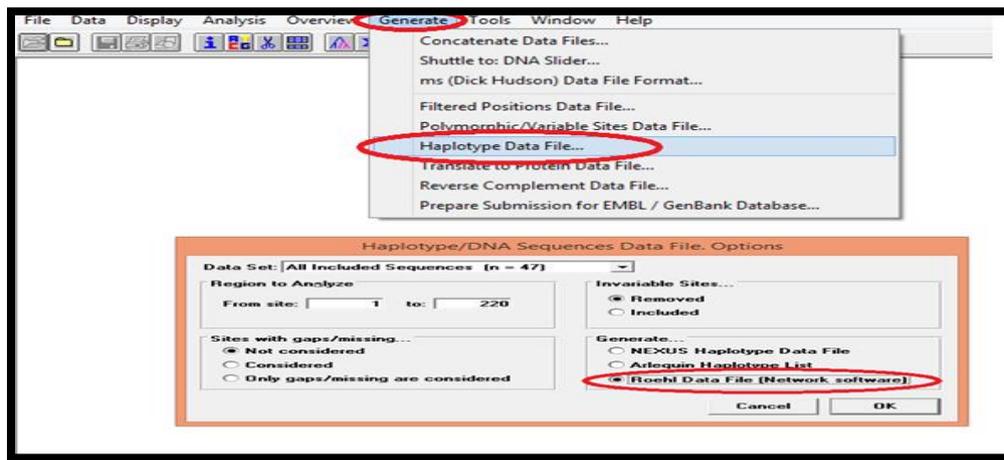


Figure 19 : Les paramètres « Haplotype Data File » et « Roehl Data File » ont été utilisés pour avoir les Haplotypes, et enregistrer le fichier sous format *.rdf.

II.1.2.2.4. DAMBE :

Nous avons utilisé le logiciel DAMBE pour l'analyse phylogénétique qui peut être classés en six catégories :

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).

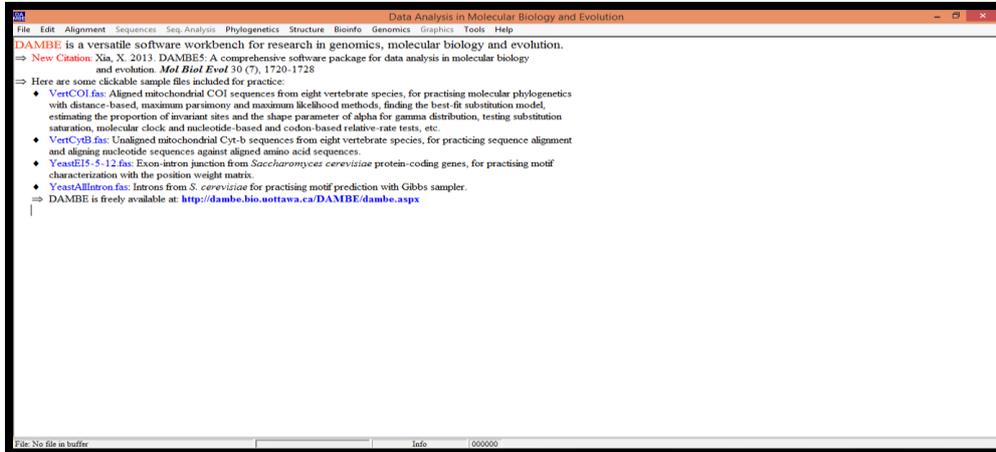


Figure 20 : Menu principal de logiciel DAMBE.

II.1.2.2.5. NETWORK :

On a terminé notre travaille par le logiciel **NETWORK**, 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 22, 23**) à partir de fichiers de séquençage (Barabasi *et al.*, 2011).



Figure 21 : Menu principal de logiciel NETWORK 5.0.

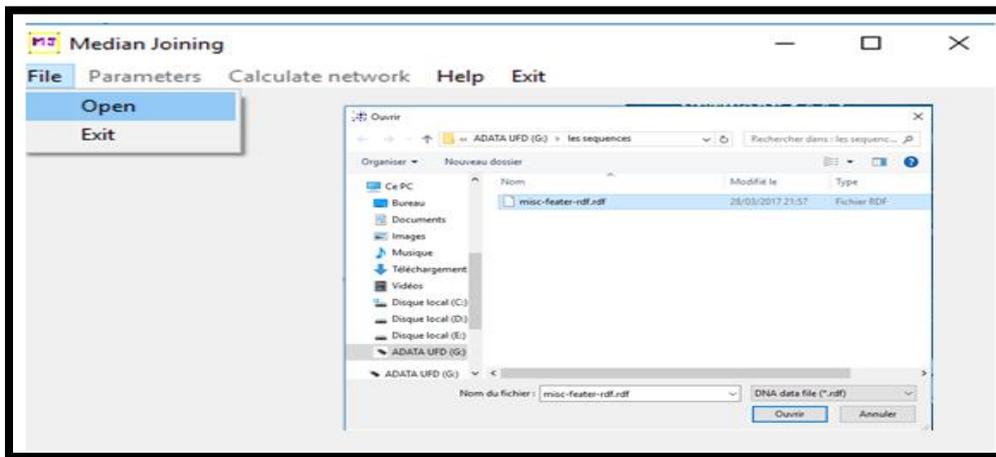


Figure 22 : L'ouverture de fichier sous format rdf produit par logiciel DnaSp.

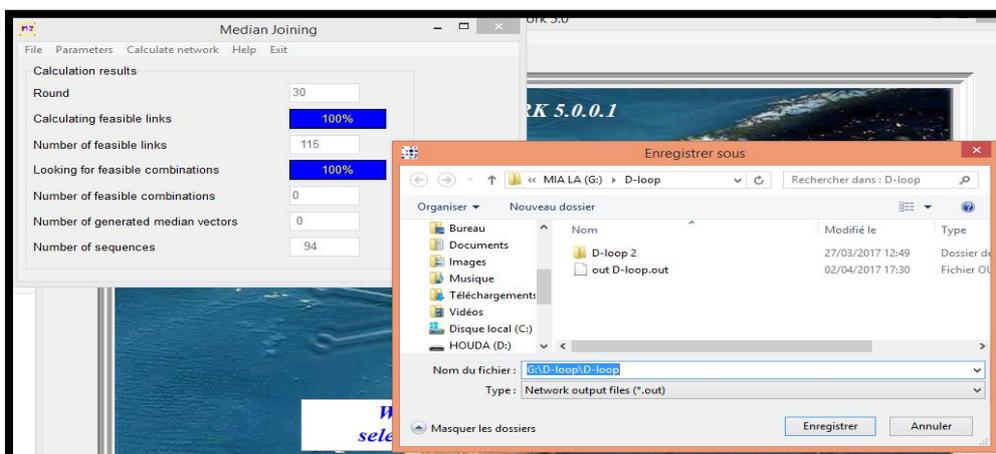


Figure 23 : L'enregistrement de notre fichier sous format *.out pour la reconstruction de réseau de NETWORK.

Les résultats rapportés dans ce travail, concernant la position de l'espèce du l'Algérie dans l'arbre phylogénétique par rapport aux autres espèces qui sont localisés dans le monde et qui appartient au même genre « *Canis* ».

III.1. Résultats de la GenBank :

Notre étude est basée sur la reconstruction phylogénétique à partir de 388 séquences (gène : *cytb*, *IRBP*,.....ect) de 33 espèces appartenant au genre *Canis*, récupérées auprès de la GenBank. Les espèces qui ont été utilisé dans ces analyses phylogénétiques sont : *Canis adustus*; *Canis anthus*; *Canis aureus*; *Canis himalayensis*; *Canis indica*; *Canis latrans*; *Canis lupus arctos*; *Canis lupus baileyi*; *Canis lupus campestris*; *Canis lupus chanco*; *Canis lupus desertorum*; *Canis lupus dingo*; *Canis lupus hattai*; *Canis lupus hodophilax*; *Canis lupus labradorius*; *Canis lupus laniger*; *Canis lupus lupus*; *Canis lupus mogollonensis*; *Canis lupus pallipes*; *Canis lupus signatus*; *Canis lupus variabilis*; *Canis lupus x Canis lupus familiaris*; *Canis lycaon*; *Canis mesomelas*; *Canis rufus*; *Canis simensis*; *Canis sp.*; *Canis sp. Belgium/36,000*; *Canis sp. CANInt1*; *Canis sp.KEB-2016*; *Canis sp. Russia/33,500*; *Canis environmental sample*.

Notre travail est basé sur les gènes mitochondriales extraites à partir de la GenBank. Annexe I.

III.1.1. Le génome mitochondrial :

La mitochondrie possède son propre génome sous forme d'un ADN bicaténaire et circulaire. L'ADN mitochondrial représente environ 1% de l'ADN cellulaire total. Chez l'homme, cet ADN est composé de 16569 paires de bases (pb) et ne comporte pas d'introns (**Anderson et al., 1981**). L'ADN mitochondrial contient 37 gènes qui codent pour 13 protéines, 22 ARN de transfert et les ARN ribosomiaux 12S et 16S (**figure 24**). Toutes les protéines codées sont des sous-unités de complexes enzymatiques du système de la phosphorylation oxydative. Les 2 ARN ribosomiaux, les 12 protéines et les 14 ARN de transfert (ARNt) sont codés par le brin lourd. Le brin léger code quant à lui 8 ARN de transfert et une protéine.

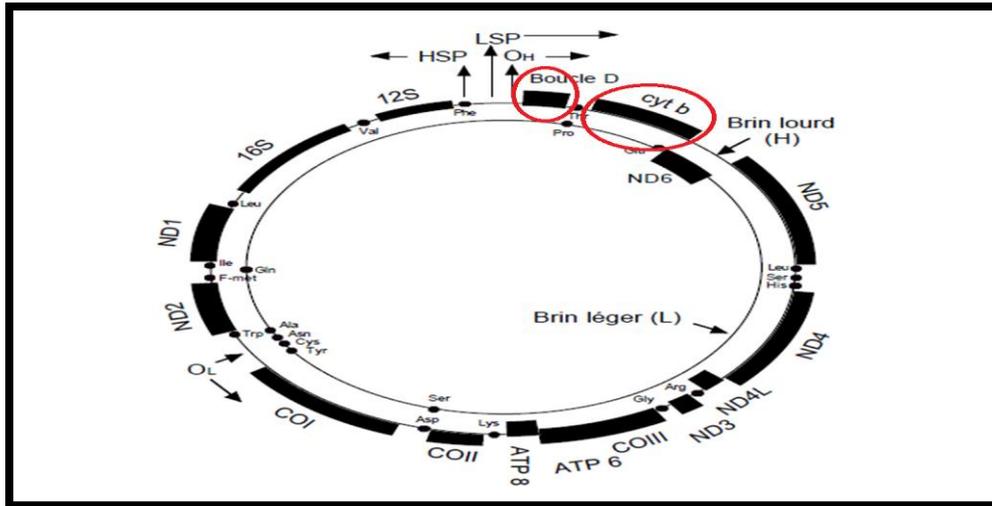


Figure 24 : Organisation de l'ADN mitochondrial humain (Anderson et al, 1981).

HSP : promoteur du brin lourd, LSP : promoteur du brin léger. Boucle D : boucle de déplacement ou D-Loop. Cyt b : gène du cytochrome b (complexe III). CO : gènes des sous-unités du complexe IV. ATP : gènes de sous-unités de l'ATPase. 12S et 16S : gènes des ARN ribosomiaux.

Parmi les gènes mitochondriaux notre choix s'est porté sur le cytochrome b (cytb), D-loop, (figure 1), et misc-feature pour établir des liens de parenté entre les espèces du genre *Canis*, parce que la plupart des études moléculaires réalisées à ce jour sur les *Canidae* ont également utilisé le cytb et la D-loop, (R. K. Aggarwal et al., 2007 ; Philippe Gaubert et al., 2012 ; Klaus-Peter Koepfli et al., 2015).

Nous avons 829 séquences d'ADN mitochondriale :

- ✓ 222 séquences de cytb.
- ✓ 299 séquences de D-loop.
- ✓ 308 séquences de misc-feature.

III.1.1.1. Le cytochrome b :

Le cytochrome b est un gène présent en de nombreuses copies dans la cellule, est considéré comme clonal et rarement ou jamais soumis à la recombinaison (Le B. M., 1990 et al., 2000). Il est transmis maternellement (hérédité cytoplasmique) (Gyllensten et al., 1991). Enfin il a un taux de mutation plus élevé que les gènes nucléaires.

Le tableau IV présente les séquences de genre *Canis* qui sont téléchargés de la GenBank, ces séquences sont classées selon leur répartition géographique et leur taille.

Tableau IV: Les séquences d'ADN mitochondrial (cyt b) extraites de la GenBank pour le genre *Canis*.

Espèce	Répartition	Taille	Locus
<i>Canis anthus</i>	Maroc	1139	KT378607
<i>Canis himalayensis</i>	Inde (19)	1140	AY291431
<i>Canis indica</i>	Inde (4)	1140	AY291432
<i>Canis aureus</i>	USA (2)	385	AF028138
	Sénégal (2)	402	JQ088656
	Palestine (22)	1140	KT447731
	Inde (9)	1140	AY291433
	Serbie (10)	1140	KT447758
	Maroc (2)	1140	KT447759
	Mauritanie (2)	1140	KT447761
	Afghanistan	1140	KT447748
	Kenya (19)	1140	KT447733
	Égypte	1140	KT447732
	Bulgarie	332	KT343786
<i>Canis mesomelas</i>	USA (4)	396	AF028143
<i>Canis simensis</i>	Éthiopie (10)	1140	KT447693
<i>Canis sp Belgium</i>	Belgique	1102	KF661079
<i>CanisSp Russia</i>	Russie	1139	KF661092
<i>Canis latrans</i>	Chine	1139	EU789789
	USA (59)	1140	KT447695
	USA (8)	16732	DQ480509
<i>Canis sp.</i>	Canada (3)	421	KT946978
<i>Canis lupus baileyi</i>	USA	1140	HM222711
<i>Canis lupus campestris</i>	Mongolie	1139	KC896375
<i>Canis lupus chanco</i>	Chine	1139	EU442884
	Chine	16709	GQ374438
	Himalaya	385	AY333748
<i>Canis lupus desertorum</i>	Chine	1139	KC461238
<i>Canis lupus dingo</i>	Australie (7)	940	JX094432
<i>Canis lupus laniger</i>	Chine	1139	KF573616
	Chine (2)	16723	NC_011218
	Éthiopie	1140	HQ845258
	Maroc	474	KM670014

<i>Canis lupus lupaster</i>	Eurasie	189	KT447711
	Mali (7)	402	JQ088665
<i>Canis lupus lupus</i>	Suède (2)	1139	NC_009686
<i>Canis lupus pallipes</i>	Himalaya	385	AY333749
<i>Canis lycaon</i>	Nord-Américain	593	JF342907
<i>Canis adustus</i>	Sud-Américain (2)	385	AF028136
	Guinée(6)	402	JQ088650
	Éthiopie	1121	KP874955
	Eurasie	15263	KT448271

III.1.1.2. D-loop (Displacement-loop) :

C'est une région contrôle de 1121 pb (position nucléotidique : 16024-576) non codante appelée D-loop, qui contient l'origine de réplication du brin lourd et les promoteurs des brins lourds et légers, ainsi qu'une séquence de 30 pb contenant l'origine de réplication du brin léger (**Anderson et al., 1981**). Cette dernière est un site essentiel pour la réplication et la transcription du génome mitochondrial. Elle contient, en effet, le point d'origine de la synthèse du brin lourd par où commence la réplication de l'ADNmt, ainsi que la majorité des promoteurs de la transcription de ce génome et plusieurs sites de fixation à certains facteurs de transcription.

Tableau V : Les séquences d'ADN mitochondrial (D-loop) extraites de la GenBank pour le genre *Canis*.

Espèce	Répartition	Taille	Locus
<i>Canis anthus</i>	Maroc (3)	1336	KT378607
<i>Canis himalayensis</i>	Inde	664	AY289977
<i>Canis indica</i>	Inde	666	AY289973
<i>Canis aureus</i>	Turquie (4)	439	KT988009
	Bulgarie	442	KT343803
	Palestine 1	1267	KT448274
	Palestine 2	438	KT343802
	Inde (16)	671	AY289996
	Serbie	305	HQ845260
	Ukraine	640	KT268319
	Pologne	640	KT268318
Espagne	659	KJ490946	

	Italie	595	AF184048
	Kenya	1269	KT448273
<i>Canis rufus</i>	USA	341	HM755703
<i>Canis simensis</i>	Éthiopie	310	HQ845261
	Éthiopie	234	AY551930
<i>Canis sp Belgium</i>	Belgique	633	KF661079
<i>Canis sP Russia</i>	Russie	1266	KF661092
<i>Canis latrans</i>	Chine	735	EU789789
	USA (8)	1264	DQ480509
	USA (64)	538	EU400578
	Canada (22)	305	GQ863734
	Canada (6)	702	AF098154
	Canada	398	AF020700
	Nord-Américain (61)	421	FM209368
	Nord-Américain (2)	708	AF008158
	Centre-Nord de Texas (24)	229	GQ981348
	Corée (2)	1266	AY172674
<i>Canis lupus chanco</i>	Népal	220	KT321360
	Chine (13)	1261	EU442884
	Russie (2)	598	AB480743
	Himalaya (4)	335	AY333741
	Mongolie (5)	673	AB007378
<i>Canis lupus hattai</i>	Japon	599	AB500701
<i>Canis lupus lupaster</i>	Maroc (2)	306	KM670012
<i>Canis lupus lupus</i>	Mongolie	673	AB007372
<i>Canis lupus pallipes</i>	Himalaya (4)	336	AY333746
	Mongolie	672	AB007374
<i>Canis signatus</i>	Portugal (5)	326	JX845625
	Espagne (4)	647	EF380229
<i>Canis lycaon</i>	Nord-Américain (6)	1286	JF342907
<i>Canis adustus</i>	Éthiopie	303	KP874956
<i>Canis lupus campestris</i>	Mongolie	1258	KC896375
<i>Canis lupus desertorum</i>	Chine	1309	KC461238
<i>Canis lupus hodophilax</i>	Japon (14)	598	AB480736
<i>Canis lupus laniger</i>	Chine (3)	1270	KF573616

III.1.1.3. Misc-feature :

Région d'intérêt biologique qui ne peut être décrite par toute autre clé de fonction; une caractéristique nouvelle ou rare.

Tableau VI : Les séquences d'ADN mitochondrial (misc-feature) extraites de la GenBank pour le genre *Canis*.

Espèce	Région	Taille (pb)	Locus
<i>Canis himalayensis</i>	Inde (18)	1302	AY289977
<i>Canis indica</i>	Inde (9)	1309	AY289973
<i>Canis aureus</i>	Lituanie	460	KT123040
	Grèce	406	KT362174
	Ukraine	406	KT362176
	Italie	450	KF588364
	Inde	1319	AY289996
	Sénégal	289	JQ088675
	Croatie (3)	664	JX220168
<i>Canis rufus</i>	Nord caroline (USA)	325	AY280913
	USA	420	GU903017
<i>Canis simensis</i>	USA	233	AY562100
<i>Canis sp.KEB-2016</i>	USA	306	KU556391
<i>Canis environmental sample</i>	USA (60)	218	JF298093
<i>Canis latrans</i>	USA	1268	AY240094
	USA	461	KT946980
	USA	402	GQ849379
	USA	223	GU062764
	Canada	225	FJ889993
	Canada	520	JN036836
	Chine (12)	1489	HQ389389
	Caroline (44)	444	EF508155
	Caroline du Nord (27)	329	AY280925
	Sud-est des États-Unis	423	AF541876
<i>Canis sp</i>	Allemagne	291	AH001620
<i>Canis lupus chanco</i>	Chine (16)	591	KC414568
<i>Canis lupus dingo</i>	Taiwan (8)	610	KC106733

	Polynésie (20)	583	AY660656
<i>Canis lupus labradorius</i>	Nord-Américain	425	GQ849370
<i>Canis lupus lupaster</i>	Algérie (7)	290	JQ088679
<i>Canis lupus signatus</i>	Portugal (11)	423	KT967960
	Portugal (26)	399	GQ366772
<i>Canis lupus variabilis</i>	Russie	277	KJ909853
<i>Canis lupus x Canis lupus familiaris</i>	Estonie (17)	1352	JN182137
<i>Canis lycaon</i>	Canada (3)	224	AY267720
	Nord-américain	407	GQ849369
<i>Canis adustus</i>	Bénin	291	JQ088674
	Guinée (5)	287	JQ088669

III.2. Construction des arbres phylogénétiques :

La construction des arbres phylogénétiques ont été réalisées à l'aide du programme MEGA 5. Tout d'abord, nous avons choisi les séquences représentatives de cytb, D-loop et misc-feature en se basant sur leurs origines géographiques.

III.2.1. Cytochrome b :

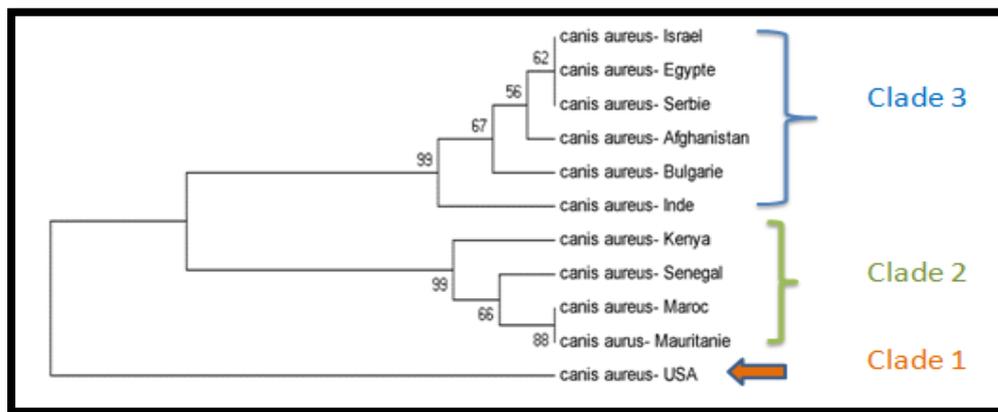


Figure 25 : L'arbre de l'espèce *Canis aureus* réalisé à l'aide du programme MEGA 5.

Canis aureus : cet arbre est divisé en trois clades, le premier clade contient que la séquence d'USA, le clade 2 inclus les séquences des pays Africain : Kenya, Sénégal, Maroc, Mauritanie, et le dernier clade contient les séquences de : Palestine, Egypte, Serbie, Afghanistan, Bulgarie, et l'Inde.

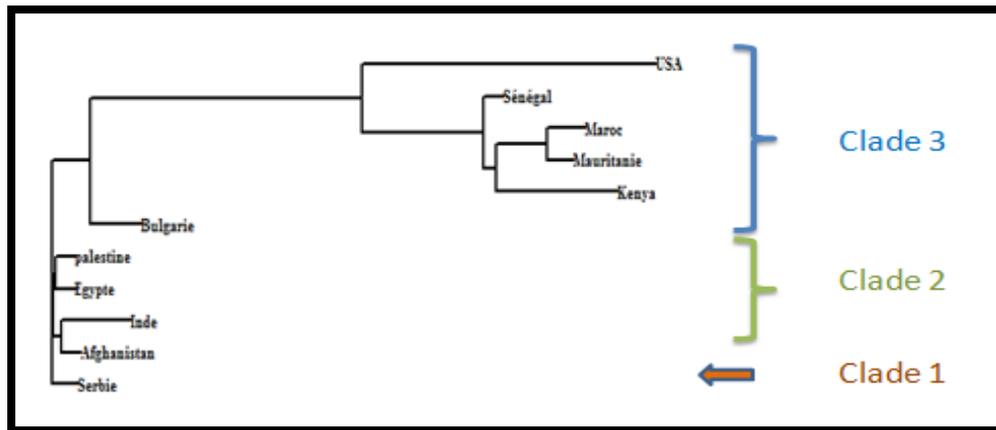


Figure 26 : L'arbre de l'espèce *Canis aureus* réalisé à l'aide du programme DAMBE.

Cet arbre est divisé en trois clades, le premier clade inclus la séquence de Serbie, le deuxième est divisé en deux sous-clades l'un comporte les séquences de l'Inde et Afghanistan et l'autre comporte les séquences de l'Egypte et Palestine, et le troisième clade est subdivisé en quatre sous-clades contient les séquences de l'Amérique, Sénégal, Maroc, Mauritanie et Kenya.

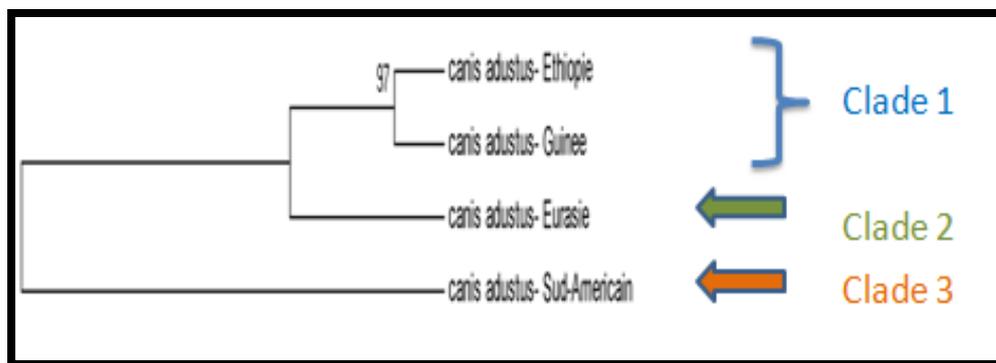


Figure 27 : L'arbre de l'espèce *Canis adustus* réalisé à l'aide du programme MEGA5.

Canis adustus : cet arbre est divisé en trois clades selon les continents du monde, le premier clade inclus une séquence de sud-américain, le deuxième clade comporte une séquence d'Eurasie (entre Asie et Europe), et le dernier clade comporte deux séquences : Ethiopie et Guinée (l'Afrique).



Figure 28 : L'arbre de l'espèce *Canis adustus* réalisé à l'aide du programme DAMBE.

Canis adustus : cet arbre est divisé en trois clades selon les continents du monde, le premier clade inclus une séquence de d'Eurasie (entre Asie et Europe), le deuxième clade comporte une séquence d'Ethiopie, et le dernier clade comporte deux séquences de Guinée et Sud-Américain.

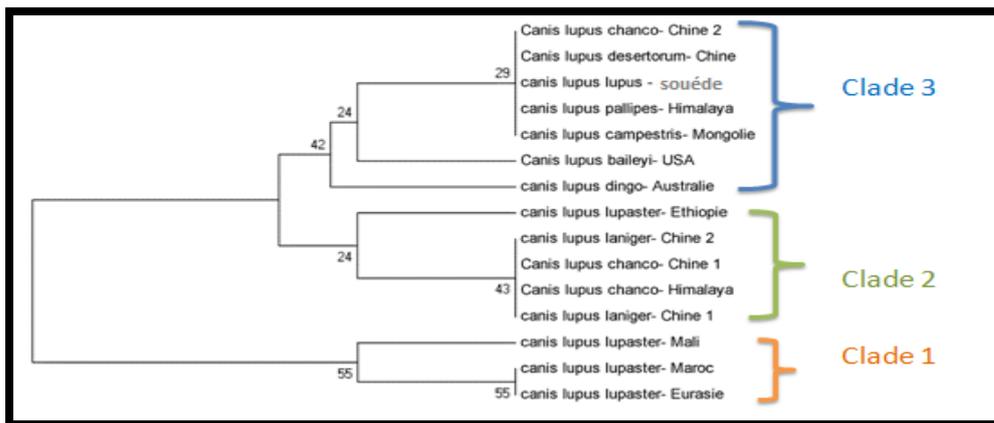


Figure 29 : L'arbre de différentes séquences de sous-espèce de l'espèce *Canis lupus* réalisé par le programme MEGA5.

Canis lupus : dans cet arbre nous avons trouvé 3 clades : le premier inclus les séquences de *C. l. lupaster* de Maroc, l'Eurasie et Mali, le deuxième lui-même subdivise en deux sous-clades, l'un comporte la séquence de *C. l. lupaster* (Maroc), et l'autre inclus les séquences suivantes : *C. l. laniger* (Chine 1 et 2), *C. l. chanco* (Chine 1 et Himalaya), et le dernier clade divisé en 2 sous-clades, le premier comprend la séquence de *C. l. dingo* d'Australie, et le second contient les séquences suivantes : *C. l. baileyi* (USA), *C. l. campestris* (Mongolie), *C. l. pallipes* (Himalaya), *C. l. upus* (Suède), *C. l. desertorum* et *C. l. chanco* (Chine).

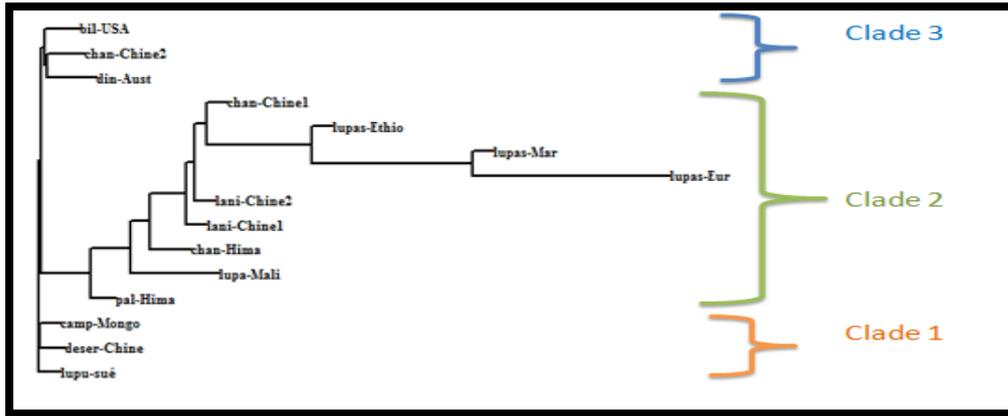


Figure 30 : L’arbre de différentes séquences de sous-espèce de l’espèce *Canis lupus* réalisé à l'aide du programme DAMBE.

Canis lupus : dans cet arbre nous avons 3 clades : le premier clade inclus les séquences de Mongolie, Suède et Chine, le deuxième clade lui-même subdivise en plusieurs sous clade qui inclus les séquences de Himalaya, Mali, Chine, Eurasie, Maroc, Ethiopie. Le troisième clade contient les séquences d’Australie, USA, et la Chine.

III.2.2. D-loop :

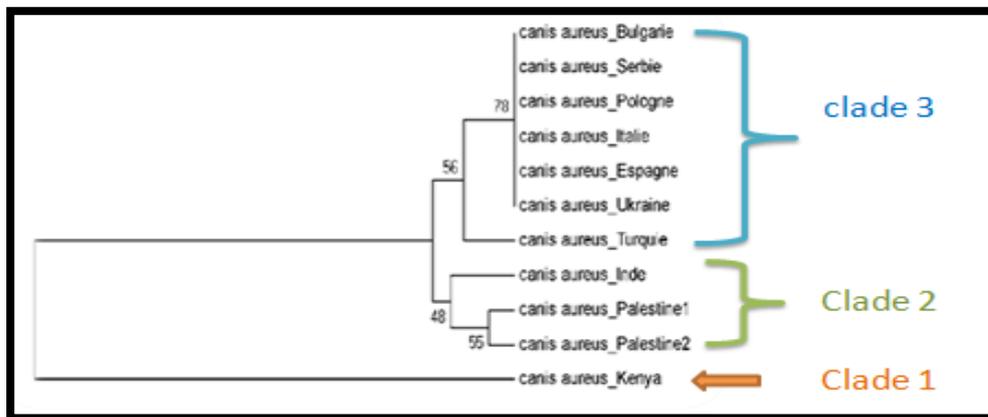


Figure 31 : L’arbre de l’espèce *Canis aureus* réalisé à l'aide du programme MEGA5.

Canis aureus : cet arbre est divisé en trois clades selon les continents du monde, le clade 1 comprend la séquence de Kenya (l’Afrique), le clade 2 comprend la séquence de l’Inde et Palestine (1 et 2) (l’Asie), le clade 3 est lui-même se subdivise en sous-clades chacun contient une séquences du : Bulgarie, Serbie, Pologne, Italie, Espagne, Ukraine et Turquie (l’Europe).

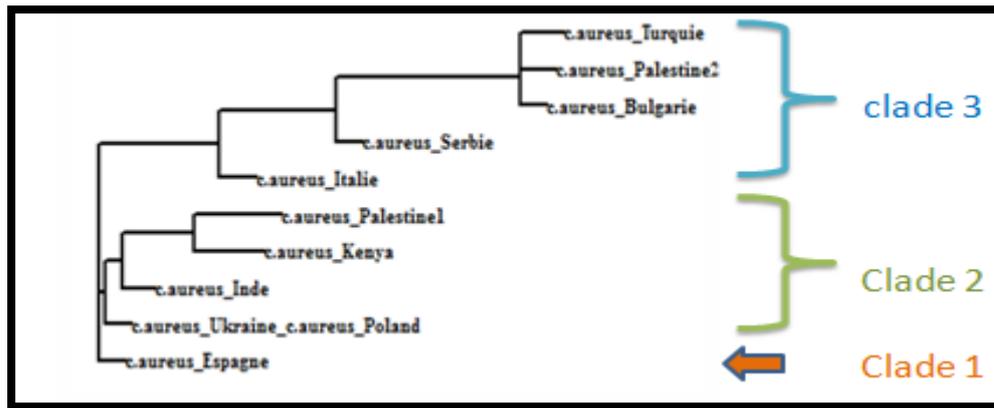


Figure 32 : L'arbre de l'espèce *canis aureus* réalisé à l'aide du programme DAMBE.

Dans cet arbre, nous avons trouvé trois clades, le clade 1 comprend la séquence d'Espagne, le clade 2 est divisé en trois sous-clades qui comportent les séquences suivantes : Palestine 1, Kenya, Inde, Ukraine et Pologne, le clade 3 est subdivisé en trois sous-clades, le premier inclus les séquences de Turquie, Palestine 2 et Bulgarie, le second inclus la séquence de Serbie, et le dernier comprend la séquence d'Italie.

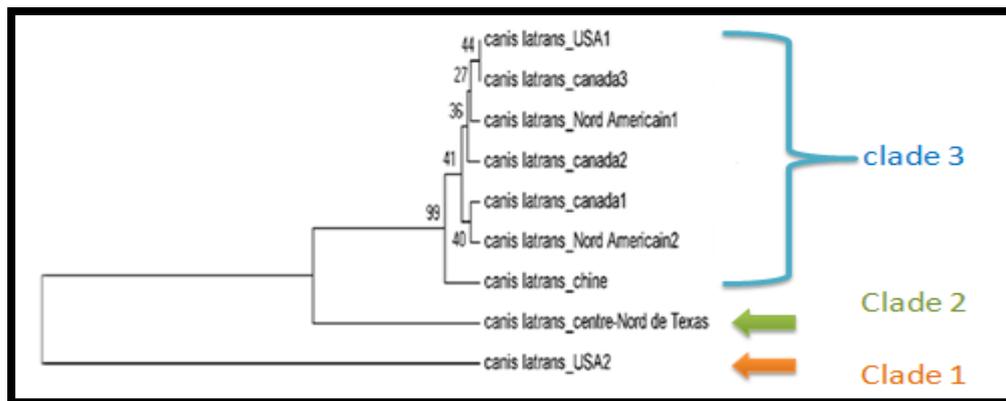


Figure 33 : L'arbre de l'espèce *Canis latrans* réalisé à l'aide du programme MEGA5.

Canis latrans : nous avons remarqué que cet arbre est divisé en trois clades, tous ces derniers appartiennent à l'Amérique du Nord, donc le premier inclus une séquence de l'USA 2, le deuxième inclus une séquence du centre-nord de Texas, et le troisième clade subdivisé en six sous-clades qui comportent les séquences d'USA 1, Canada 3, Nord-Américain 1, Canada 2, Canada 1, Nord-Américain 2 et Chine.

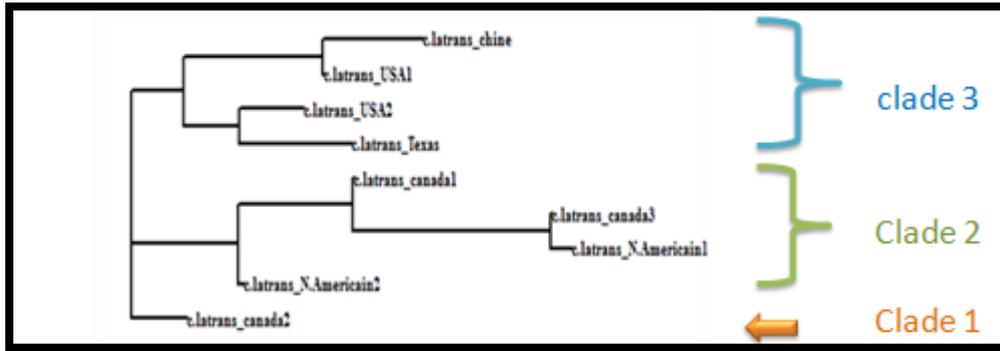


Figure 34 : L’arbre de l’espèce *Canis latrans* réalisé à l’aide du programme DAMBE. Cet arbre est divisé en trois clades, le premier inclus la séquence de Canada 2, le deuxième est divisé en deux sous-clades qui comportent les séquences de Canada 1 et 3, Nord-Américain 1 et 2, et le troisième est subdivisé en deux sous-clades, l’un comprend les séquences de Chine, USA1, et l’autre contient les séquences d’USA2 et Texas.

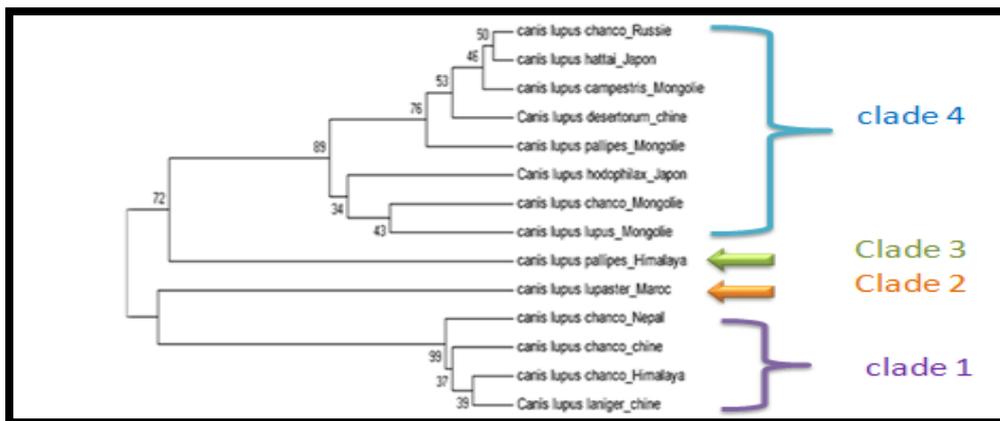


Figure 35 : L’arbre de différentes séquences de sous-espèce de l’espèce *Canis lupus* réalisé à l’aide du programme MEGA5.

Canis lupus : dans cet arbre, nous avons trouvé quatre clades, le premier clade comporte les séquences de : Népal, Chine, Himalaya, le clade 2 inclus une séquence de Maroc, le clade 3 inclus une séquence d’Himalaya, et le dernier clade lui-même se subdivise en deux 2 clade chacun contient un sous-clade qui contient les séquences du : Russie, Japon, Mongolie, Chine. Et Donc nous avons remarqué que toutes ces séquences appartiennent à l’Asie et l’Afrique du nord.

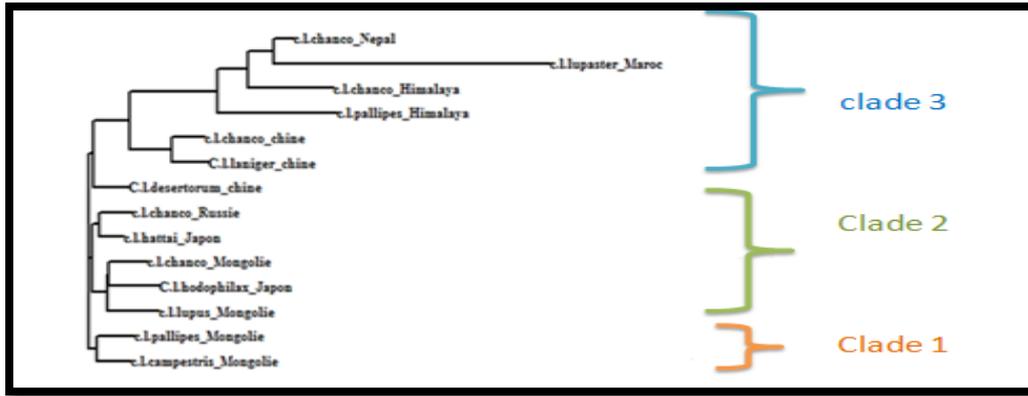


Figure 36 : L’arbre de différentes séquences de sous-espèce de l’espèce *Canis lupus* réalisé à l'aide du programme DAMBE.

Cet arbre est divisé en trois clades, le clade 1 inclus les séquences de Mongolie (*Canis lupus pallipes* et *Canis lupus campestris*), le clade 2 est subdivisé en deux sous-clades, l’un contient les séquences de Russie (*Canis lupus chanco*), Japon (*Canis lupus hattai*), et l’autre comprend deux séquences de Mongolie (*Canis lupus chanco* et *Canis lupus lupus*) et une séquence de Japon (*Canis lupus hadophilax*), et le dernier clade est divisé en trois sous-clades, le premier inclus les séquences de Népal (*Canis lupus chanco*), Maroc (*Canis lupus lupaster*), Himalaya (*Canis lupus chanco* et *Canis lupus pallipes*), le second comporte les séquences de Chine (*Canis lupus chanco* et *Canis lupus laniger*), et le troisième sous-clade inclus la séquence de Chine pour l’espèce *Canis lupus desertorum*.

III.2.3. Misc-feature :

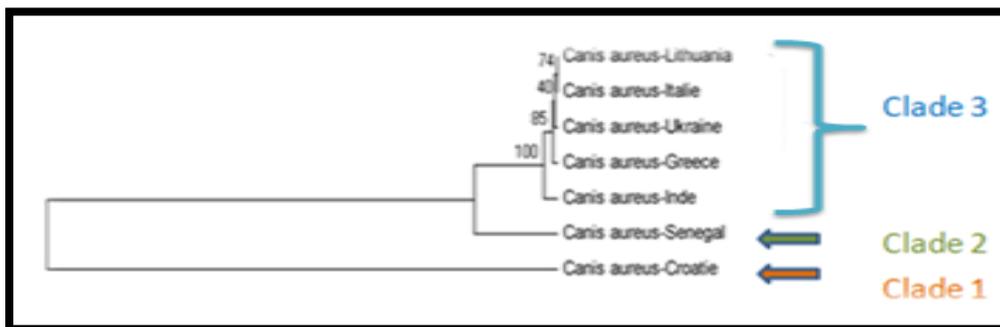


Figure 37 : L’arbre de l’espèce *Canis aureus* réalisé à l'aide du programme MEGA5.

Canis aureus : cet arbre est divisé en trois clades selon les continents du monde, le clade 1 comprend la séquence de Croatie (l’Europe), le clade 2 comprend la séquence de senegal (l’Afrique) et le clade 3 qui lui-même se subdivise en 2 clades

l'un contient les séquences du Inde (Asie) et l'autre contient les séquences de l'Europe (Ukraine, Grèce, Lituanie, Italie).

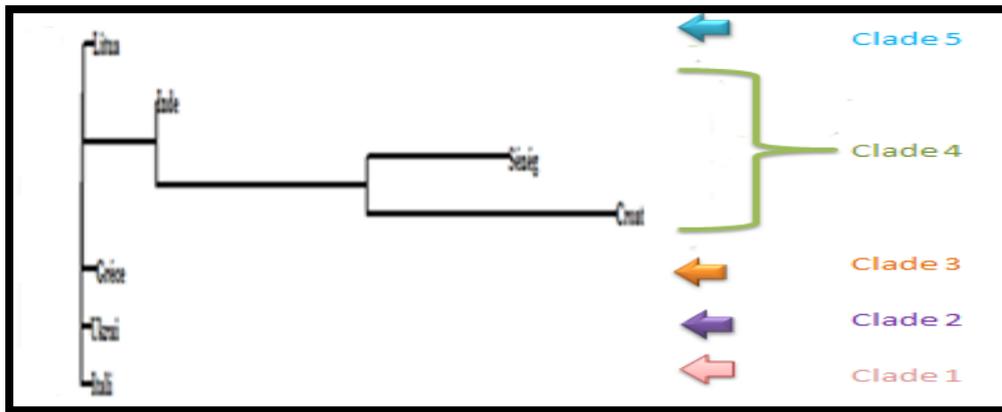


Figure 38 : L'arbre de l'espèce *Canis aureus* réalisé à l'aide du programme DAMBE.

Canis aureus : cet arbre est divisé en cinq clades, le clade 1 comprend la séquence de Italie, le clade 2 comprend la séquence de Ukraine, le clade 3 comprend la séquence de Grèce, le clade 4 inclus les séquences de Inde, Sénégal et Croatie et le dernier clade 5 comprend la séquence de Lituanie .

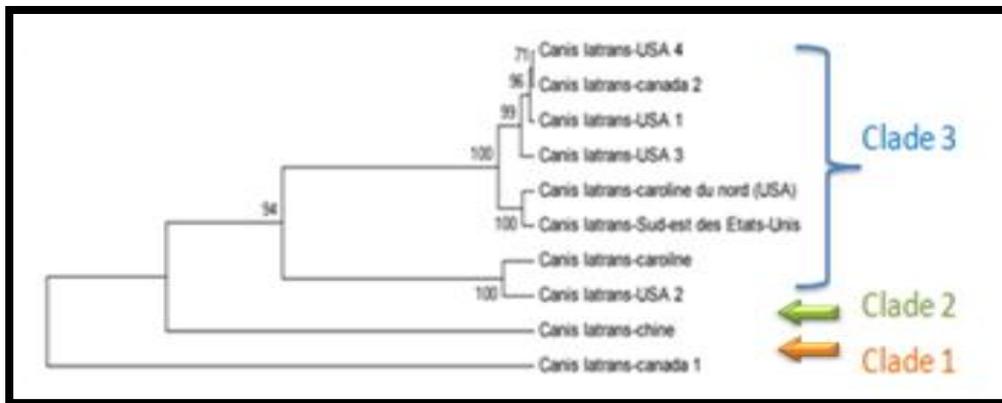


Figure 39 : L'arbre de l'espèce *Canis latrans* réalisé à l'aide du programme MEGA5.

Canis latrans : dans cet arbre nous avons trois clade : clade 1 qui inclus une séquence du Caroline et USA 2, clade 2 qui inclus une séquence de Canada 1, clade 3 inclus deux sous clade contenant 5 séquences.

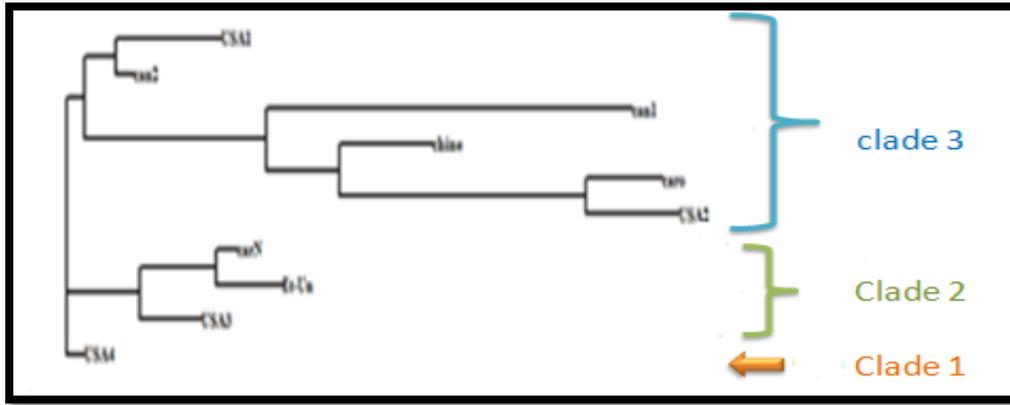


Figure 40 : L'arbre de l'espèce *Canis latrans* réalisé à l'aide du programme DAMBE. Cet arbre est divisé en 3 clades, le premier inclus une séquence de l'Amérique, le deuxième comporte les séquences de l'Amérique sud, Etats Unis, Caroline du nord et le dernier clade subdivise en deux autres clades l'un divisé en sous-clade comporte les séquences de l'Amérique, Canada, Caroline du nord et Chine, et l'autre comprend deux séquences de l'Amérique et Canada.

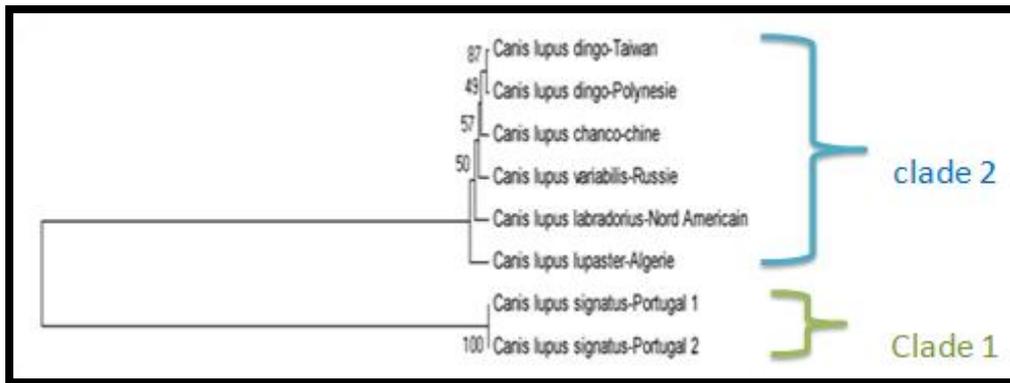


Figure 41 : L'arbre de différentes séquences de sous-espèce de l'espèce *Canis lupus* réalisé à l'aide du programme MEGA5.

Canis lupus : dans cet arbre nous avons trouvées trois clades : clade 1 qui inclus une séquence du Portugal 1 et 2, clade 2 qui inclus une séquence du l'Algérie, clade 3 inclus deux sous clade contenant cinq séquences.

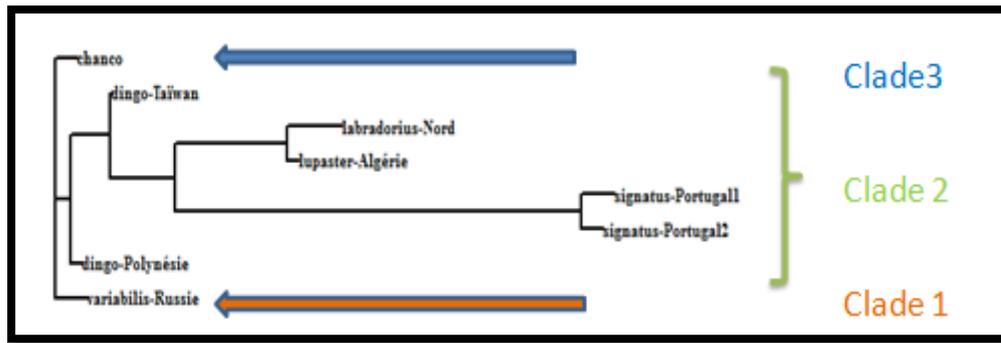


Figure 42 : L'arbre de différentes séquences de sous-espèce de l'espèce *Canis lupus* réalisé à l'aide du programme DAMBE.

Dans cet arbre, nous avons trouvé trois clades, le premier clade comporte une seule séquences de Russie, le deuxième clade subdivise en deux autres clades l'un comporte les séquences de Taiwan, Nord-Américain, l'Algérie et deux séquences de Portugal et l'autre comprend une seul séquences de Polynésie, le troisième clade comprend une seul séquence de Chine.

D'après nos résultats de MEGA et DAMBE de ces espèces, nous remarquons qu'il y a une différence entre les clades et une apparition de nouvelle similarité des séquences dans le logiciel DAMBE, par contre les clades de logiciel MEGA sont bien classés selon les continents du monde.

Nous concluons que le logiciel MEGA est plus spécialisé pour la reconstitution des arbres phylogénétique.

III.3. Résultat des haplotypes :

On parle d'un haplotype lorsque les allèles de plusieurs polymorphismes sont en position cis, à savoir sur le même chromosome. Le polymorphisme d'ADN et la variation entre espèces est sensiblement une approche efficace pour comprendre l'évolution processus et pour obtenir des informations sur la signification fonctionnelle de régions génomiques.

Pour cela le DnaSP compte les statistiques couramment utilisées quantifiant les niveaux du polymorphisme d'ADN, nous avons étudié :

1. Le nombre de sites de ségrégation : c'est juste le nombre de variables positions (c'est-à-dire polymorphes) dans un échantillon de séquences d'ADN.

2. Le nombre de haplotypes c'est-à-dire différentes séquences d'ADN.
3. La diversité des haplotypes (Hd) également connue sous le nom de diversité de gènes / allèles ou hétérozygotie attendue (la probabilité que deux séquences aléatoires sont différentes).

Remarque :

Ces séquences sont alignés à partir de logiciel FaBox ; cette alignement est appelée alignement trimer.

III.3.1 Cytochrome b :

```

Input Data File: G:\... \alignement trimer d canis aureus2.txt
Number of sequences: 11   Number of sequences used: 11
Selected region: 1-332   Number of sites: 332
Total number of sites (excluding sites with gaps / missing data): 332

Sites with alignment gaps: not considered
Number of variable sites: 310

----- Haplotype Distribution -----
Number of haplotypes, h: 8
Haplotype diversity, Hd: 0.9273

Hap_1: 1 [1]
Hap_2: 1 [2]
Hap_3: 3 [3 5 10]
Hap_4: 1 [4]
Hap_5: 2 [6-7]
Hap_6: 1 [8]
Hap_7: 1 [9]
Hap_8: 1 [11]

Hap_1: 1 [c.aureus-USA]
Hap_2: 1 [c.aureus-Sénégal]
Hap_3: 3 [c.aureus-palestine c.aureus-Serbie c.aureus-Egypte]
Hap_4: 1 [c.aureus-Inde]
Hap_5: 2 [c.aureus-Maroc c.aureus-Mauritanie]
Hap_6: 1 [c.aureus-Afghanistan]
Hap_7: 1 [c.aureus-Kenya]
Hap_8: 1 [c.aureus-Bulgarie]

```

Figure 43 : Le nombre des haplotypes de *Canis aureus* a été réalisé par le logiciel DnaSP.

Le nombre de séquences de l'espèce *Canis aureus* est : 11 séquences, le nombre de sites de ségrégation est : 332 sites, le nombre des haplotypes (h) est : 8 haplotypes, chaque haplotype comporte une seule séquence, sauf le haplotype (3) inclus trois séquences identiques (Palestine, Serbie et Egypte) et le haplotype (5) inclus deux séquences (Maroc et Mauritanie), et la diversité des haplotypes (Hd) est : 0,9273.

Les séquences identiques des haplotypes 3 et 5 correspondent au même clade dans l'arbre phylogénétique de MEGA.

```

Input Data File: G:\...\alignement traîner d caanis adustus2.txt
Number of sequences: 4   Number of sequences used: 4
Selected region: 1-385   Number of sites: 385
Total number of sites [excluding sites with gaps / missing data]: 385

Sites with alignment gaps: not considered
Number of variable sites: 357

===== Haplotype Distribution =====
Number of haplotypes, h: 4
Haplotype diversity, Hd: 1,0000

Hap_1: 1 [1]
Hap_2: 1 [2]
Hap_3: 1 [3]
Hap_4: 1 [4]

Hap_1: 1 [c.adustus-S.Américai]
Hap_2: 1 [c.adustus-Guinée]
Hap_3: 1 [c.adustus-Ethiopie]
Hap_4: 1 [c.adustus-Eurasie]

```

Figure 44 : Le nombre des haplotypes de *Canis adustus* réalisé par le logiciel DnaSP.

Le nombre de séquences de l'espèce *Canis adustus* est : 4 séquences, le nombre de sites de ségrégation est : 385 sites, le nombre de haplotypes (h) est : 4 haplotypes, chaque haplotype comporte une seule séquence, et la diversité des haplotypes (Hd) est : 1,0000.

```

Input Data File: G:\...\alignement traîner d canis lupus2.txt
Number of sequences: 15   Number of sequences used: 15
Selected region: 1-189   Number of sites: 189
Total number of sites [excluding sites with gaps / missing data]: 189

Sites with alignment gaps: not considered
Number of variable sites: 185

===== Haplotype Distribution =====
Number of haplotypes, h: 9
Haplotype diversity, Hd: 0,8762

Hap_1: 5 [1-2 4 6 14]
Hap_2: 3 [3 8-9]
Hap_3: 1 [5]
Hap_4: 1 [7]
Hap_5: 1 [10]
Hap_6: 1 [11]
Hap_7: 1 [12]
Hap_8: 1 [13]
Hap_9: 1 [15]

Hap_1: 5 [C.l.baileyi-USA c.l.campestris-Mongo c.l.chanco-Chine C.l.desertorum-Chine
c.l.lupus-suède]
Hap_2: 3 [C.l.chanco-Chine c.l.laniger-Chine c.l.laniger-Chine]
Hap_3: 1 [C.l.chanco-Himalaya]
Hap_4: 1 [c.l.dingo-Australie]
Hap_5: 1 [c.l.lupaster-Ethiopie]
Hap_6: 1 [c.l.lupaster-Maroc]
Hap_7: 1 [c.l.lupaster-Eurasie]
Hap_8: 1 [c.l.lupaster-Mali]
Hap_9: 1 [c.l.pallipes-Himalay]

```

Figure 45 : Le nombre des haplotypes de *Canis lupus* réalisé par le logiciel DnaSP.

Le nombre de séquences de l'espèce *Canis lupus* est : 15 séquences, le nombre de sites de ségrégation est : 189 sites, le nombre de haplotypes (h) est : 9 haplotypes, chaque haplotype comporte une seule séquence, sauf le haplotype 1 comporte cinq séquences identiques qui sont les suivantes : une séquence de l'USA (*Canis lupus baileyi*), une séquence de la Mongolie (*Canis lupus campestris*), une séquence de la Chine (*Canis lupus desertorum* et *Canis lupus chanco*) et une séquence de la Suède (*Canis lupus lupus*), et le haplotype 2 inclus trois séquences de la Chine qui sont : (2 *Canis lupus laniger* et 1 *Canis lupus chanco*), et la diversité des haplotypes (Hd) est : 0,8762.

D'après la comparaison de ces résultats avec l'arbre phylogénétique, nous avons remarquées que les séquences identiques des haplotypes 1 et 2 appartiennent au même clade dans l'arbre phylogénétique.

III.3.2 D-loop :

```

Input Data File: C:\...Ntrimer_alignment aureus_2.txt
Number of sequences: 11 Number of sequences used: 11
Selected region: 1-305 Number of sites: 305
Total number of sites [excluding sites with gaps / missing data]: 260
Sites with alignment gaps: not considered
Number of variable sites: 260

===== Haplotype Distribution =====
Number of haplotypes, h: 9
Haplotype diversity, Hd: 0.9455

Hap_1: 1 [1]
Hap_2: 1 [2]
Hap_3: 1 [3]
Hap_4: 1 [4]
Hap_5: 1 [5]
Hap_6: 1 [6]
Hap_7: 3 [7-9]
Hap_8: 1 [10]
Hap_9: 1 [11]

Hap_1: 1 [c.aureus_Turquie]
Hap_2: 1 [c.aureus_Bulgarie]
Hap_3: 1 [c.aureus_Palestine1]
Hap_4: 1 [c.aureus_Palestine2]
Hap_5: 1 [c.aureus_Inde]
Hap_6: 1 [c.aureus_Serbie]
Hap_7: 3 [c.aureus_Ukraine c.aureus_Poland c.aureus_Espagne]
Hap_8: 1 [c.aureus_Italie]
Hap_9: 1 [c.aureus_Kenya]
    
```

Figure 46 : Le nombre des haplotypes de *canis aureus* réalisé par le logiciel DnaSP.

Le nombre de séquences de l'espèce *Canis aureus* est : 11 séquences, le nombre de sites de ségrégation est : 305 sites, le nombre de haplotypes (h) est : 9 haplotypes, tous les haplotypes comportent une seule séquence, sauf l'haplotype numéro 7 comporte trois séquences identiques : une séquence : de l'Ukraine, de la Pologne et de l'Espagne, et ces derniers appartient à un même clade dans l'arbre phylogénétique, la diversité des haplotypes (Hd) est : 0,9455.

```

Input Data File: D:\...Ntrimer_alignment latrans.txt
Number of sequences: 9 Number of sequences used: 9
Selected region: 1-229 Number of sites: 229
Total number of sites [excluding sites with gaps / missing data]: 228
Sites with alignment gaps: not considered
Number of variable sites: 228

===== Haplotype Distribution =====
Number of haplotypes, h: 9
Haplotype diversity, Hd: 1.0000

Hap_1: 1 [1]
Hap_2: 1 [2]
Hap_3: 1 [3]
Hap_4: 1 [4]
Hap_5: 1 [5]
Hap_6: 1 [6]
Hap_7: 1 [7]
Hap_8: 1 [8]
Hap_9: 1 [9]

Hap_1: 1 [c.latrans_chine]
Hap_2: 1 [c.latrans_USA1]
Hap_3: 1 [c.latrans_USA2]
Hap_4: 1 [c.latrans_canada1]
Hap_5: 1 [c.latrans_canada2]
Hap_6: 1 [c.latrans_canada3]
Hap_7: 1 [c.latrans_N.Americai]
Hap_8: 1 [c.latrans_N.Americai]
Hap_9: 1 [c.latrans_c.N.Texas]
    
```

Figure 47 : Le nombre des haplotypes de *Canis latrans* réalisé par le logiciel DnaSP.

Le nombre de séquences de l'espèce *Canis latrans* est : 10 séquences, le nombre de sites de ségrégation est : 229 sites, le nombre de haplotypes (h) est : 10 haplotypes, chaque haplotype comporte une seule séquence, et la diversité des haplotypes (Hd) est : 1,0000.

```

Input Data File: C:\... \trimer_alignment lupus.txt
Number of sequences: 14   Number of sequences used: 14
Selected region: 1-220   Number of sites: 220
Total number of sites [excluding sites with gaps / missing data]: 219
Sites with alignment gaps: not considered
Number of variable sites: 219

===== Haplotype Distribution =====
Number of haplotypes, h: 14
Haplotype diversity, Hd: 1.0000

Hap_1: 1 [1]
Hap_2: 1 [2]
Hap_3: 1 [3]
Hap_4: 1 [4]
Hap_5: 1 [5]
Hap_6: 1 [6]
Hap_7: 1 [7]
Hap_8: 1 [8]
Hap_9: 1 [9]
Hap_10: 1 [10]

Hap_1: 1 [C.l.chanco_Nepal]
Hap_2: 1 [C.l.chanco_Chine]
Hap_3: 1 [C.l.chanco_Russie]
Hap_4: 1 [C.l.chanco_Himalaya]
Hap_5: 1 [C.l.chanco_Mongolie]
Hap_6: 1 [C.l.hattai_Japon]
Hap_7: 1 [C.l.hubasti_Maroc]
Hap_8: 1 [C.l.hubasti_Mongolie]
Hap_9: 1 [C.l.pallipes_Himalay]
Hap_10: 1 [C.l.pallipes_Mongolie]
Hap_11: 1 [C.l.campestris_Mongol]
Hap_12: 1 [C.l.desertorum_Chine]
Hap_13: 1 [C.l.hodophilax_Japon]
Hap_14: 1 [C.l.laniger_Chine]

```

Figure 48 : Le nombre des haplotypes de *Canis lupus* réalisé par DnaSP.

Le nombre de séquences de l'espèce *Canis lupus* est : 14 séquences, le nombre de sites de ségrégation est : 220 sites, le nombre de haplotypes (h) est : 14 haplotypes, chaque haplotype comporte une seule séquence, et la diversité des haplotypes (Hd) est : 1,0000.

III.3.3 Misc-feature :

```

Input Data File: G:\... \alignement trimer canis aureus.fasta
Number of sequences: 7   Number of sequences used: 7
Selected region: 1-289   Number of sites: 289
Total number of sites [excluding sites with gaps / missing data]: 289
Sites with alignment gaps: not considered
Number of variable sites: 289

===== Haplotype Distribution =====
Number of haplotypes, h: 7
Haplotype diversity, Hd: 1.0000

Hap_1: 1 [1]
Hap_2: 1 [2]
Hap_3: 1 [3]
Hap_4: 1 [4]
Hap_5: 1 [5]
Hap_6: 1 [6]
Hap_7: 1 [7]

Hap_1: 1 [Litua]
Hap_2: 1 [Grèce]
Hap_3: 1 [Ukrai]
Hap_4: 1 [Itali]
Hap_5: 1 [Inde]
Hap_6: 1 [Sénégal]
Hap_7: 1 [Croat]

```

Figure 49 : Le nombre des haplotypes de *Canis aureus* réalisé par le logiciel DnaSP.

Le nombre de séquences de l'espèce *Canis aureus* est : 7 séquences, le nombre de sites de ségrégation est : 289 sites, le nombre de haplotypes (h) est : 7 haplotypes, tous

les haplotypes comportent une seule séquence, et la diversité des haplotypes (Hd) est : 1,0000.

```

Input Data File: G:\...\Alignement traimer d canis latrans.fasta
Number of sequences: 10 Number of sequences used: 10
Selected region: 1-223 Number of sites: 223
Total number of sites (excluding sites with gaps / missing data): 223
Sites with alignment gaps: not considered
Number of variable sites: 223

----- Haplotype Distribution -----
Number of haplotypes, h: 10
Haplotype diversity, Hd: 1,0000

Hap_1: 1 [1]
Hap_2: 1 [2]
Hap_3: 1 [3]
Hap_4: 1 [4]
Hap_5: 1 [5]
Hap_6: 1 [6]
Hap_7: 1 [7]
Hap_8: 1 [8]
Hap_9: 1 [9]
Hap_10: 1 [10]

Hap_1: 1 [USA1]
Hap_2: 1 [can1]
Hap_3: 1 [chine]
Hap_4: 1 [caro]
Hap_5: 1 [can2]
Hap_6: 1 [Et-Un]
Hap_7: 1 [USA2]
Hap_8: 1 [USA4]
Hap_9: 1 [USA4]
Hap_10: 1 [can2]

```

Figure 50 : Le nombre des haplotypes de *Canis latrans* réalisé par le logiciel DnaSP.

Le nombre de séquences de l'espèce *Canis latrans* est : 10 séquences, le nombre de sites de ségrégation est : 223 sites, le nombre de haplotypes (h) est : 10 haplotypes, chaque haplotype comporte une seule séquence, et la diversité des haplotypes (Hd) est : 1,0000.

```

Input Data File: G:\...\Alignement traimer d canis lupus.txt
Number of sequences: 8 Number of sequences used: 8
Selected region: 1-277 Number of sites: 277
Total number of sites (excluding sites with gaps / missing data): 277
Sites with alignment gaps: not considered
Number of variable sites: 277

----- Haplotype Distribution -----
Number of haplotypes, h: 8
Haplotype diversity, Hd: 1,0000

Hap_1: 1 [1]
Hap_2: 1 [2]
Hap_3: 1 [3]
Hap_4: 1 [4]
Hap_5: 1 [5]
Hap_6: 1 [6]
Hap_7: 1 [7]
Hap_8: 1 [8]

Hap_1: 1 [chanco-chine]
Hap_2: 1 [dingo-Taiwan]
Hap_3: 1 [dingo-Polynésie]
Hap_4: 1 [labradorius-Nord]
Hap_5: 1 [lupaster-Algérie]
Hap_6: 1 [signatus-Portugal1]
Hap_7: 1 [signatus-Portugal2]
Hap_8: 1 [variabilis-Russie]

```

Figure 51 : Le nombre des haplotypes de *Canis lupus* réalisé par le logiciel DnaSP.

Le nombre de séquences de l'espèce *Canis lupus* est : 8 séquences, le nombre de sites de ségrégation est : 277 sites, le nombre de haplotypes (h) est : 8 haplotypes, chaque haplotype comporte une seule séquence, et la diversité des haplotypes (Hd) est : 1,0000.

D'après la comparaison des résultats de DnaSP avec les arbres phylogénétiques de MEGA, nous avons remarquées que les séquences identiques de chaque haplotype

des différentes espèces pour les trois gènes correspondent au même clade de l'arbre phylogénétique.

III.4. Résultats statistiques :

Nous avons utilisé le logiciel DAMBE pour traiter les séquences de chaque gène (D-loop, cytochrome b, misc-feature) par la méthode nucleotide skew.

III.4.1. GC/AT skew :

Nous avons travaillé sur les séquences de genre *Canis* ; les 3 gènes téléchargés de la GenBank sous format FASTA ; dans le but d'obtenir le graphique GC skew qui correspond à $G-C / G+C$ et le graphique AT skew qui correspond à $A-T / A+T$.

III.4.1.1. Cytochrome b :

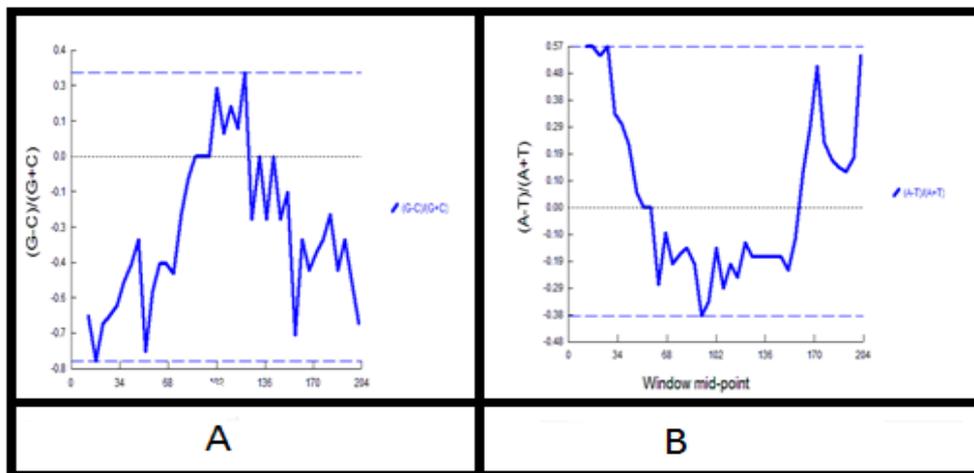


Figure 52 : La variation du GC skew (A) et AT skew (B) du genre *Canis* avec 204 bases réalisée par le logiciel DAMBE.

D'après ces 2 variations (A et B) nous remarquons que le spectre de dérivation de la composition nucléotidique (A) varie de -1 à +1, par conséquent l'écart de GC est négative ce qui représente la richesse de C sur G. Et nous remarquons aussi que le spectre de dérivation de la composition nucléotidique (B) varie de +1 à -1, par conséquent l'écart de AT est positive ce qui représente la richesse de A sur T.

III.4.1.2. D-loop :

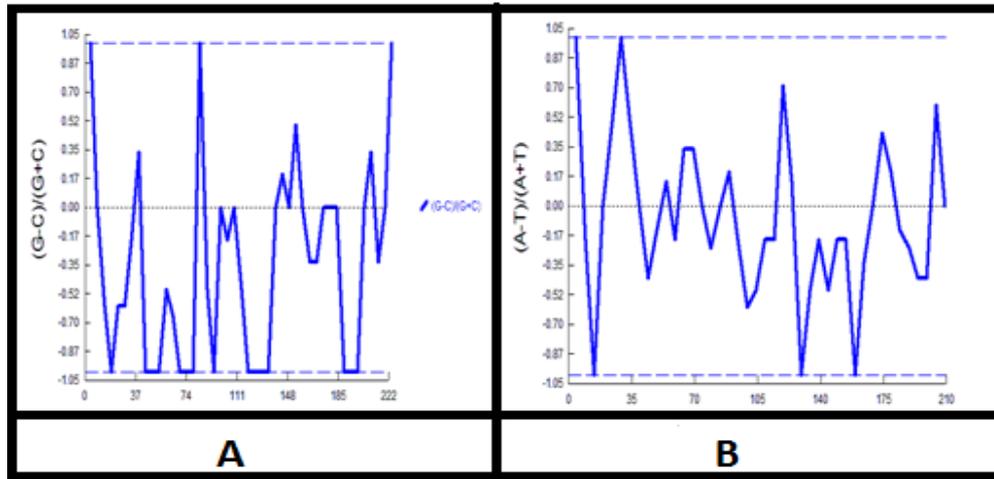


Figure 53 : La variation du GC skew (A) et AT skew (B) du genre *Canis* avec 222 bases réalisée par le logiciel DAMBE.

D’après ces 2 variations (A et B) nous remarquons que les deux spectres de dérivation de la composition nucléotidique (A et B) varient de -1 à +1, par conséquent l’écart de GC et AT sont négatives ce qui représente la richesse de C sur G et la richesse de T sur A.

III.4.1.3. Misc-feature :

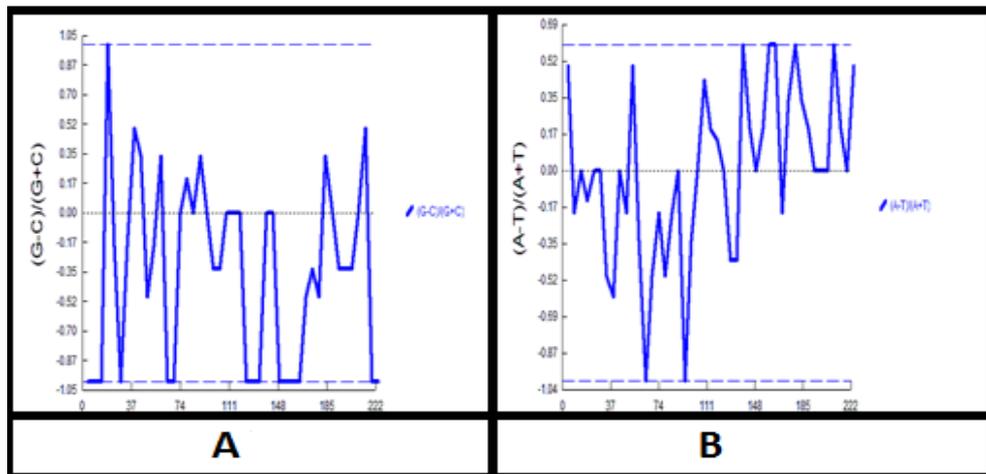


Figure 54 : La variation du GC skew pour le gène misc-feature du genre *Canis* avec 222 bases réalisée par le logiciel DAMBE.

D’après ces 2 variations (A et B) nous remarquons que les deux spectres de dérivation de la composition nucléotidique (A et B) varient de -1 à +1, par

conséquent l'écart de GC et AT sont négatives ce qui représente la richesse de C sur G et la richesse de T sur A.

Nucleotide skews for sequence: canis Global GC Skew = -0,29870 Global AT Skew = 0,01786 Sliding window size optimized with GC skew. Window size = 22; stepsize = 5 Circular genome: No	cytochrome B
Nucleotide skews for sequence: canis Global GC Skew = -0,38462 Global AT Skew = -0,02326 Sliding window size optimized with GC skew. Window size = 9; stepsize = 5 Circular genome: No	D-loop
Nucleotide skews for sequence: Canis Global GC Skew = -0,35632 Global AT Skew = -0,00763 Sliding window size optimized with GC skew. Window size = 9; stepsize = 5 Circular genome: No	misc feature

Figure 55 : Comparaison des fréquences entre les séquences de cytb, D-loop et misc-feature réalisée par logiciel DAMBE.

D'après cette comparaison nous remarquons que la fréquence globale de GC skew de gène cytochrome b est plus élevée par rapport aux D-loop et misc-feature, et nous remarquons aussi que ces derniers sont proches.

Selon **Tigran (1999)** la stabilité de la molécule d'ADN bicaténaires de même longueur croît avec le nombre de paires GC qu'elles contiennent.

Par conséquent le gène cytochrome b c'est le gène idéal pour étudié les espèces car il présente une forte stabilité par rapport au D-loop et misc-feature, ces derniers sont utilisés pour classer les sous-espèces.

III.5. La construction des réseaux :

Les relations phylogénétiques entre les haplotypes ont donc été mises en évidence grâce à la construction d'un réseau avec le logiciel Network (**Bandelt et al., 1999**). Nous avons choisi d'intégrer l'ensemble des données obtenues au cours de ce mémoire afin d'obtenir le réseau le plus informatif et le mieux résolu possible. Ces réseaux ont été construits sur la base d'un jeu de données représentant les espèces de genre *Canis*. Bien que le réseau tracé représente l'ensemble des haplotypes de notre jeu de données.

III.5.1. Cytochrome b :

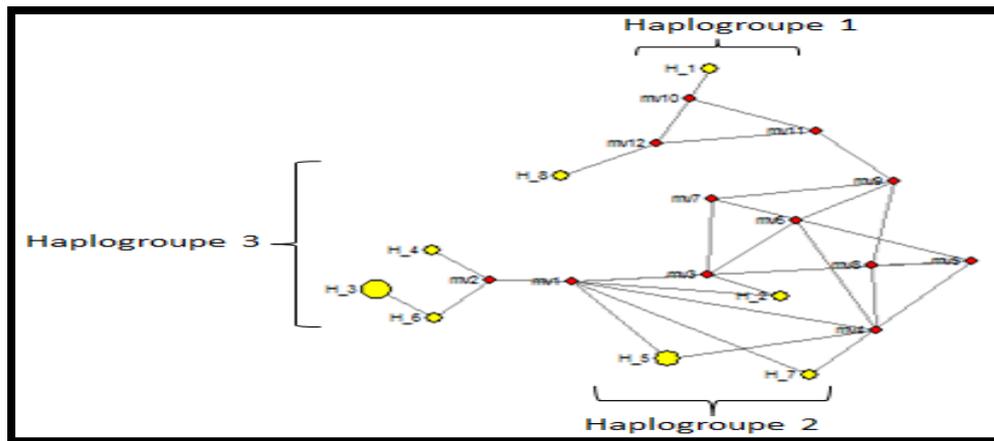


Figure 56 : Réseau des haplotypes de l’espèce *Canis aureus* construit grâce au logiciel Network.

Ce réseau comprend trois haplogroupes, le premier inclus H1, le deuxième comprend H2, H5 et H7 où le haplotype 5 comporte deux séquences de l’Afrique (Maroc et Mauritanie), et le dernier haplogroupe inclus H3, H4, H6 et H8 où le haplotype 3 comporte trois séquences identiques des régions suivants : Palestine, Serbie et Egypte. Nous avons remarquées que chaque haplogroupe correspond à un clade dans l’arbre phylogénétique de MEGA.

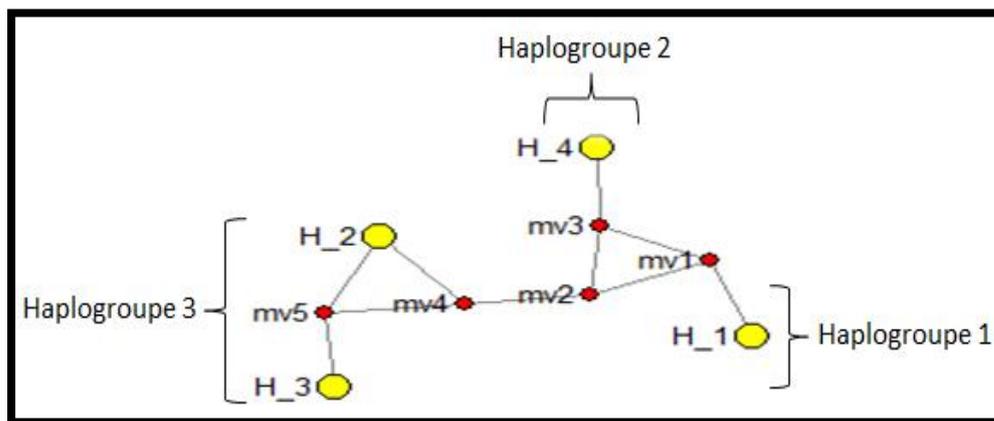


Figure 57 : Réseau des haplotypes de l’espèce *Canis adustus* construit grâce au logiciel Network.

Ce réseau comprend trois haplogroupes selon la distribution géographique de cette espèce, nous avons remarquées que ces haplogroupes correspondent aux clades de l’arbre phylogénétique de MEGA.

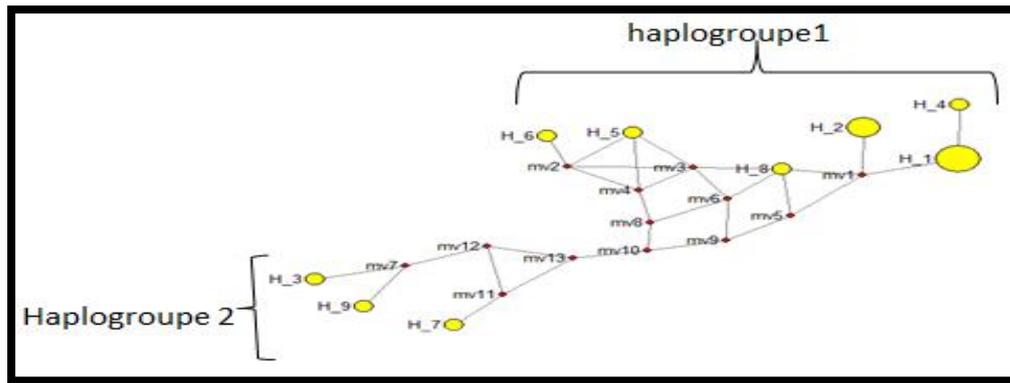


Figure 58 : Réseau des haplotypes de l'espèce *Canis lupus* construit grâce au logiciel Network.

Nous avons remarquées dans ce réseau deux haplogroupes, le premier inclus H3, H7 et H9, qui représente *C. l. chanco* (Himalaya), *C. l. lupaster* (Ethiopie) et *C. l. pallipes* (Himalaya), et le second haplogroupe comprend le reste des haplotypes H1, H2, H4, H5, H6 et H8. Les haplotypes 1 et 2 apparaissent plus grands par rapport les autres haplotypes, le haplotype 1 représente les séquences suivantes : *C. l. baileyi* (USA), *C. l. campestris* (Mongolie), *C. l. chanco* (Chine), *C. l. desertorum* (Chine), *C. l. lupus* (Suède), et le haplotype 2 représente 3 séquences de la Chine (1 de *C. l. chanco* et 2 de *C. l. laniger*), nous avons remarquées que les haplotypes 1 et 4 sont plus proches et cela est en accord avec l'arbre phylogénétique où ces haplotypes appartient au même clade.

III.5.2. D-loop :

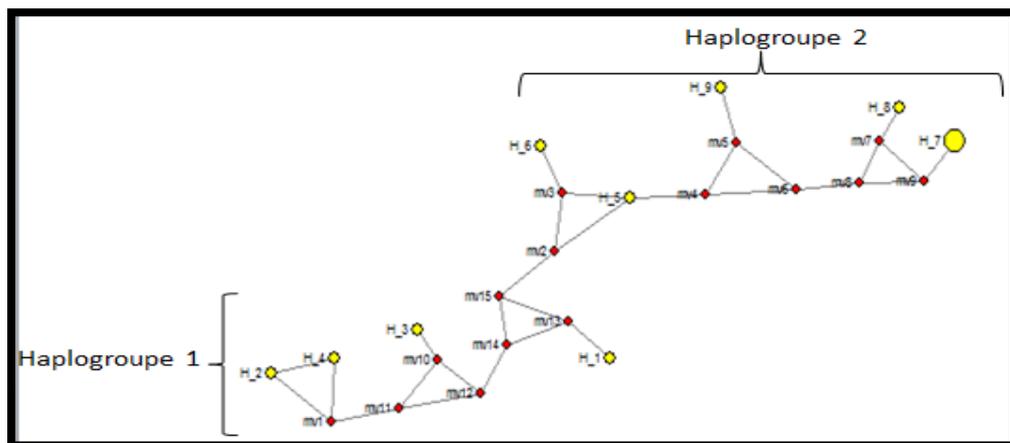


Figure 59 : Réseau présentant les relations phylogénétiques entre les haplotypes de l'espèce *Canis aureus* réalisé par le logiciel NETWORK.

Nous avons dans ce réseau deux haplogroupes, la taille des cercles est proportionnelle à la fréquence des haplotypes dans la population étudiée, en effet le H 7 comporte trois séquences identiques qui sont les suivantes une séquence : de l’Ukraine, de la Pologne et de l’Espagne, les haplotypes 7 et 8 sont plus proches selon la distribution géographique et cela est en accord avec l’arbre phylogénétique de MEGA, la même chose pour les haplotypes 3 et 4.

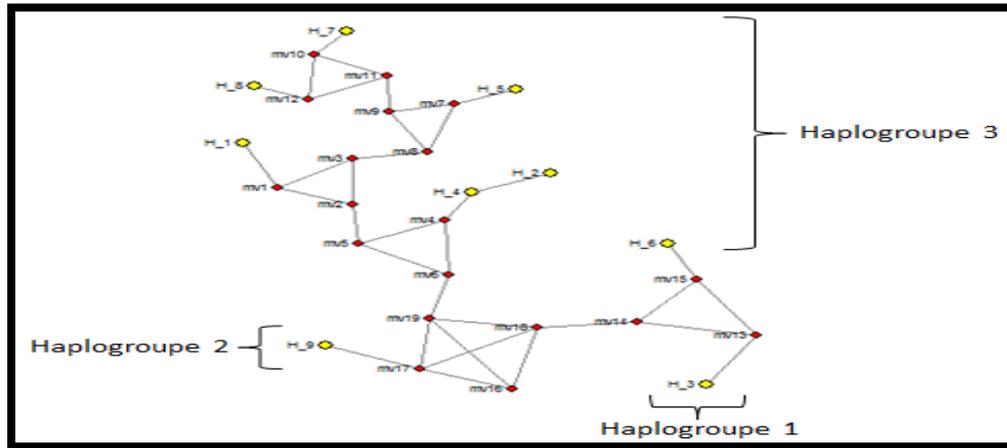


Figure 60 : Réseau présentant les relations phylogénétiques entre les haplotypes de l’espèce *Canis latrans* réalisé par le logiciel NETWORK.

Ce réseau comprend 3 haplogroupes, le premier inclus H3, le deuxième inclus H9, et le dernier haplogroupe inclus H1, H2, H4, H5, H6, H7 et H8. Chaque haplogroupe correspond à un clade dans l’arbre phylogénétique.

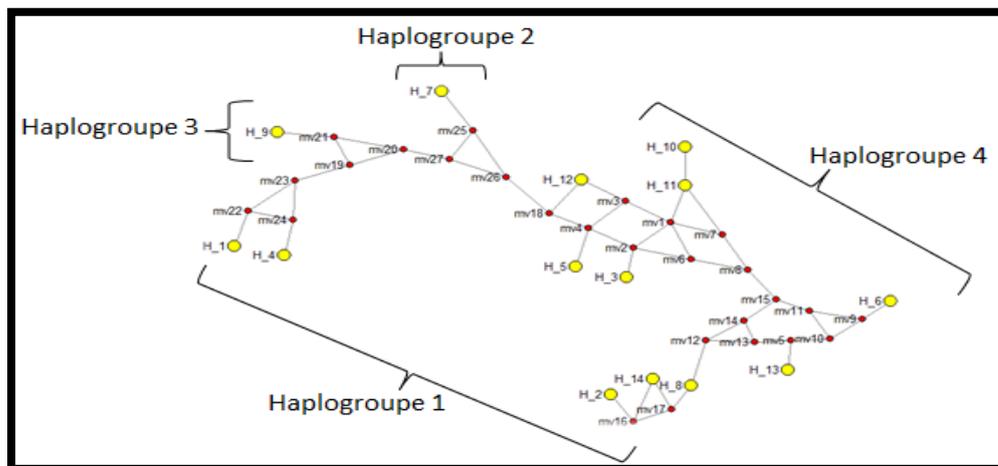


Figure 61 : Réseau présentant les relations phylogénétiques entre les haplotypes de l’espèce *Canis lupus* réalisé par le logiciel NETWORK.

Ce réseau a une diversité génétique très élevée (14 haplotypes), nous avons remarquées quatre haplogroupes, le premier haplogroupe comprend H1, H2, H4 et H14, le deuxième comprend H7, le troisième inclus H9, et le dernier haplogroupe inclus H3, H5, H6, H8, H10, H11, H12, H13 chacun de ces haplogroupes a une distribution géographique très proche et correspond aux quatre clades de l'arbre phylogénétique.

III.5.3. Misc-feature :

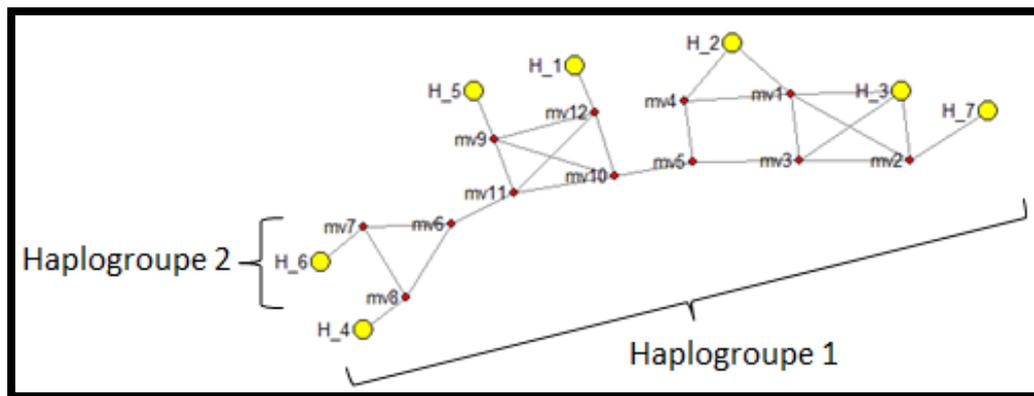


Figure 62 : Réseau des haplotypes de l'espèce *Canis aureus* construit grâce au logiciel NETWORK.

D'après les résultats de DnaSP, nous avons 7 haplotypes, ce réseau comporte trois haplogroupes, le premier inclus H1, H2, H3, H4, H5 et H7 (les séquences de Lituanie, Grèce, Ukraine, Italie, Inde et Croatie), le deuxième inclus H6 (la séquences de Sénégal), les haplotypes sont réparties selon la distribution géographique et cela est en accord avec les clades de l'arbre phylogénétique.

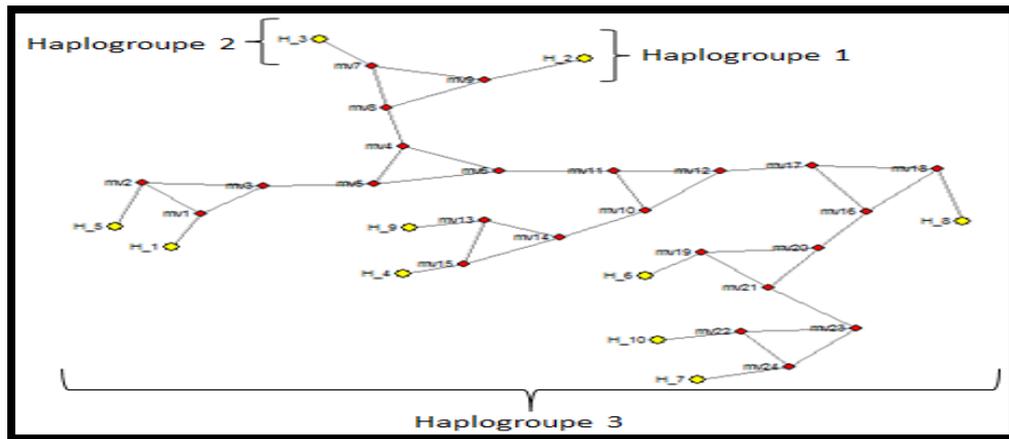


Figure 63 : Réseau présentant les relations phylogénétiques entre les haplotypes de l'espèce *Canis latrans* réalisé par le logiciel NETWORK.

Ce réseau est composé de quatre haplogroupes, le premier inclus H2, le deuxième inclus H3, et le dernier haplogroupe inclus H1, H4, H5, H6, H7, H8, H9 et H10. D'après la comparaison de ce réseau avec l'arbre phylogénétique nous avons remarquées que chaque haplogroupe correspond à un clade dans l'arbre phylogénétique de MEGA.

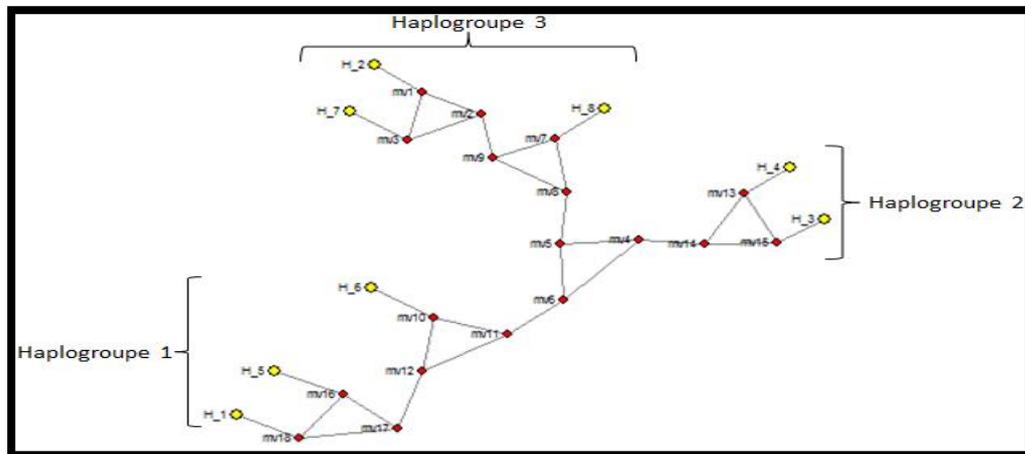


Figure 64 : Réseau présentant les relations phylogénétiques entre les haplotypes de l'espèce *Canis lupus* réalisé par le logiciel NETWORK.

Ce réseau comprend 3 haplogroupes, le premier contient H1, H5 et H6 (*C. l. chanco*, *C. l. lupaster* et *C. l. signatus*), le deuxième inclus H3 et H4 (*C. l. dingo* et *C. l. laboratorius*), et le troisième inclus H2, H7 et H8 (*C. l. dingo*, *C. l. signatus* et *C. l. variabilis*).

Nous avons basé dans ce travail sur l'histoire évolutive des sous espèce *Canis lupus lupaster*, d'après la comparaison des résultats de l'arbre de MEGA et le réseau de NETWORK, nous avons conclu que le *C. l. lupaster* de l'Algérie est plus proche au sous espèce *C. l. chanco* (Chine) et *C. l. singnatus* (Portugal).

C. a. lupaster, le loup africain, habite à au moins deux localités différentes en Ethiopie et a environ 2500 km au sud-est de l'Égypte (**Rueness, 2011**). Les résultats de la comparaison des séquences d'ADN prouvent que : *C. a. lupaster* à plus de similitude avec *C. lupus* qu'avec *C. aureus* et les espèces de loup gris complexe, y compris *C. a. lupaster*, forme un groupe monophylétique avec soutien statistique fort.

L'apparition de nouveau loup d'Afrique ouvre une étude d'actualité basée sur le genre *Canis*, ce dernier est un genre de mammifères carnivores de la famille des canidés qui regroupe les chiens, les dingos, les loups, les coyotes et les chacals. Ce genre est réparti sur les continents eurasiatique et africain et sur le continent Nord-américain. Il est représenté par 33 espèces, parmi ces derniers les espèces (*Canis aureus*) et (*Canis lupus*) qui est représenté par le sous espèce (*Canis lupus lupaster*) seraient les seules espèces de genre *Canis* qui existent en Algérie.

Après le traitement de nos séquences des trois gènes mitochondriaux par les cinq logiciels : FaBox, MEGA, DnaSP, DAMBE et NETWORK, nous avons mis en évidence les résultats suivants :

➤ Pour MEGA et DAMBE :

Il y a une différence entre les clades et une apparition de plusieurs similarités entre les séquences dans le logiciel DAMBE, par contre les clades de logiciel MEGA sont bien classés selon les continents du monde. Donc le logiciel MEGA est plus spécialisé pour la reconstitution des arbres phylogénétique.

➤ Pour DnaSP et MEGA :

Les séquences de chaque haplotype des différentes espèces pour les trois gènes correspondent au même clade de l'arbre phylogénétique.

➤ Pour NETWORK et MEGA :

Les séquences de chaque haplogroupe de NETWORK correspondent au même clade dans l'arbre phylogénétique de MEGA et ont la même distribution géographique.

➤ Pour DAMBE (GC skew) :

Le gène de cytochrome b c'est le gène idéal pour étudié les espèces car il présente une forte stabilité par rapport D-loop et misc-feature, ces derniers sont utilisés pour classer les sous espèces.

La fréquence globale de GC skew de gene cytochrome b (-0.29870) est plus élevée par rapport aux D-loop (-0.38462) et misc-feature (-0.35632), et ces derniers sont proches.

Dans ce travail notre attention s'est donc portée sur le positionnement et l'histoire évolutive de sous espèce *Canis lupus lupaster* par rapport aux autres espèces du monde de genre *Canis*, d'après la comparaison de nos résultats entre l'arbre de MEGA et le réseau de NETWORK, nous avons remarquées que le *C. l. lupaster* de l'Algérie de l'origine Egypte est plus proche aux sous espèces *Canis lupus chanco* et *Canis lupus singnatus*.

- **Agrawal R et al .**, 2007. Encyclopedia of Mammals, Volume Four, pgs. 107-114.
- **Alauzet C.**, 2009. Taxonomie des bactéries anaérobies : de la reclassification a la découverte de nouveaux pathogènes. Thèse doctorat, université de Nancy.
- **Anderson S., Bankier T., Barrell G., 1981.** Sequence and organization of the human mitochondrial genome. Nature 290, 457 – 465.
- **Baker R., 1983.** Michigan Mammals. Michigan State University Press, Detroit, pg: 390-399.
- **Banfield A., 1974.** Mammals of Canada. University of Toronto Press, Toronto and Buffalo, pg: 286-289.
- **Barabasi A., Gulbahce N., Loscalzo J., 2011.** Network medicine: a network-based approach to human disease. 12(1):56-68.
- **Beaulieu M. E., 2007.** Caractérisation moléculaire des champignons ophiostomatoï des 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.
- **Blanchet C., 1999.** Logiciel MPSA et ressources bioinformatiques client-serveur Web dédiés à l'analyse de séquences de protéine. Thèse de doctorat en Bioinformatique. Dir. Thèse : Deleage G. Univ. Claude Bernard-Lyon I : 23-36. 158p.
- **Bonfield, J., Smith K., Staden R., 1995.** A new DNA sequence assembly program. Nucleic Acids Res 23, 4992-4999.
- **Broder S., Venter J., 2000.** Whole genomes : the foundation of new biology and medicine. Curr Opin Biotechnol 11, 581-585.
- **Colby C., 1965.** Wild Dogs. New York, New York: Duell, Sloan and Pearce.
- **Corbet G., Hill J., 1991.** A World List of Mammalian Species. Third Edition. Oxford University Press, NY.
- **Corbett L., 1995.** The Dingo in Australia and Asia.. Ithaca, New York: Cornell University Press.
- **Darlu P., Tassy P., 1993.** La reconstruction phylogénétique : concepts et méthodes. Ed. Masson, Paris. P1et 2 et 3.
- **Dear S., Staden R., 1991.** A sequence assembly and editing program for efficient management of large projects. Nucleic Acids Res 19, 3907- 3911.

- **Diallo A. B., 2009.** Bioinformatique avancée .Université du Québec à Montréal. Canada.
- **Dorst et al ., 1976.** The ecology and evolution of the golden jackel (canis aureus) : investigating a cryptic species. Leeds,UK : University of Leeds.
- **Downs C., Bowland A., Bowland M., 1991.** Thermal parameters of serval felis-serval felidae and black backed jackal canidae. Journal of Thermal Biology, 16(5): 277-280.
- **Esther et al., 2015.** Encyclopedia of mammals, pgs. 63-67. Facts on file publications,NY.
- **Ferguson W., 1981.** The systematic position of canis aureus lupaster (Carnivora : Canidae) and the occurrence of Canis lupus in North africain. Egypt and Sinai.
- **Fleischmann R., Adams M ., White O., Clayton R., Kirkness E., Kerlavage A., Bult C., Tomb J., Dougherty B.,Merrick., 1995.** Whole-genome random sequencing and assembly of Haemophilus influenzae Rd. Science 269, 496-512.
- **Fournel M., Tardif C., 1834.** Traitâe du voisinage: considâerâe dans l'ordre judiciaire et dans ses rapports avec le code civil, Volume 1.Ed . Videcoq , Paris .52p.
- **Fox M., 1971.** Behaviour of Wolves, Dogs, and Related Canids. Malabar, Florida: Robert E. Krieger Publishing Company, Inc.
- **Fox M., 1975.** The Wild Canids: Their Systematics, Behavioral Ecology and Evolution. Van Nostrand Reinhold Co., New York, London and Melbourne, pg: 247-262.
- **Ginsberg J., 1990.** Foxes, Wolves, Jackals, and Dogs. Gland, Switzerland: International Union for Conservation of Nature and Natural Resources (IUCN).
- **Ginsberg J., Macdonald D., 1990.** Foxes, Wolves, Jackals, and Dogs: An action plan for the conservation of canids. Canid Specialist Group and Wolf Specialist Group, the World Conservation Union, Gland, Switzerland.
- **Ginsberg J., Macdonald D., 1990.** Foxes, Wolves, Jackals, and Dogs: An action plan for the conservation of canids. Canid Specialist Group and Wolf Specialist Group, the World Conservation Union, Gland, Switzerland.
- **Golding B., Morton D., 2003.** Elementary Sequence Analysis. Department of Biology. McMaster. Univ. Hamilton, Ontario. 190 p.

- **Gottelli D., Sillero-Zubiri C., Applebaum D., Roy M. S., Girman D. J., Garcia-moreno J., Ostrander E. A., R. K. Wayne., 1994.** Molecular genetics of the most endangered canid: the Ethiopian wolf *Canis simensis*. *Molecular Ecology* 3: 301-312.
- **Grzimek B., 1990.** Encyclopedia of Mammals, Volume Four, pgs. 107-114.
- **Guindon S., Gascuel O., 2003.** A Simple, Fast, and Accurate Algorithm to Estimate Large Phylogenies by Maximum Likelihood (PHYML). *Syst. Biol.* 52(5): P 696–704.
- **Honghai Z., 2013.** *canis lupus*. *Mammals Species*.
- **Kawalski K., Rzebick-kowalska B., 1991.** *Mammals of Algeria*.
- **Kingdon J., 1977.** *East African Mammals*. New York, NY: University of Chicago.
- **Lander E. S., 2001.** Initial sequencing and analysis of the human genome. *Nature* 409, 860-921.
- **Lander E., Waterman M., 1988.** Genomic mapping by fingerprinting random clones: a mathematical analysis. *Genomics* 2, 231-239.
- **Le T. H., Blair D., M. C. and Manus D. P., 2000.** Mitochondrial genomes of human helminths and their use as markers in population genetics and phylogeny. *Acta. Trop.*, 77: 243-256.
- **Librado p., Rozas J ., 2009.** DnaSP v5: a software for comprehensive analysis of DNA polymorphism data, *Genetics and population analysis*, Vol. 25 no. 11, pages 1451–1452.
- **Linnaeus C., 1758.** *Systema naturae per regna tria naturae, secundum classes, ordines , genera , species, cum characteribus, diffentis, synonymis, locis*. Tomus .Editio decimal, reformata.-pp. (1-4), 1-824.Holmiae.(Salvius).
- **MacDonald D., 1984.** *Encyclopedia of Mammals*, pgs. 63-67. Facts on File Publications, NY.
- **Macdonald D., 1984.** *Encyclopedia of Mammals*. Facts on File Publications, NY.
- **Malcolm J., 1997.** The diet of the Ethiopian wolf (*Canis simensis* Ruppell) from a grassland area of the Bale Mountains, Ethiopia. *African Journal of Ecology* 35: 162-164.
- **Mochtman P., Rowe D., 1998.** "International Union for the Conservation of Nature and Natural Resources" (On-line). Accessed November 20, 1999.
- **Mullis K., Saiki R., Scharf S., Faloona F., Horn G., Erlich H ., 1985.** Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science* 230: 1350-54.
- **Myers E., 2000.** A whole-genome assembly of *Drosophila*. *Science* 287, 2196-2204.

- **Nijman I. J., Otsen M., Verkaar E. L., Ruijter C., Hanekamp E., 2003.** Hybridization of banteng (*Bos javanicus*) and zebu (*Bos indicus*) revealed by mitochondrial DNA, satellite DNA, AFLP and microsatellites. *Heredity*, **90**: 10–16.
- **Nowak J., Ronald M., 1991.** Walker's Mammals of the World, pgs. 1065-1068. Fifth Edition. Johns Hopkins University Press, Baltimore and London.
- **Nowak L., Ronald M., John L., 1983.** Walkers Mammals of the World. 4th Ed. The Johns Hopkins University Press, Baltimore, MD, pg: 949-951.
- **Nowak R., 1999.** Walker's Mammals of the World, 6th Ed.. Baltimore, MD: The John Hopkins University Press.
- **Nowak R., Paradiso ., 1991.** *Canis rufus*: Mammalian Species No. 22. The American Society of Mammologists.
- **Oakman B., 2004.** «The Australian Dingo Conservation Association Inc." (On-line).
- **O'Neill T., 1997.** Traveling the Australian Dog Fence. *National Geographic*, 191(4): 18-38.
- **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, pp. 681–693.
- **Perière 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.
- **Philip W., Hedrick, Rhonda N., Karen M., 2000.** Major histocompatibility complex (MHC) variation in the endangered Mexican wolf and related canids.
- **Philippe G., 2012.** Reviving the African Wolf *Canis lupus lupaster* in North and West Africa: A Mitochondrial Lineage Ranging More than 6,000 km Wide. *PLoS ONE* 7(8): e42740.
- **Philippe G., 2012.** Reviving the African Wolf *Canis lupus lupaster* in North and West Africa: A Mitochondrial Lineage Ranging More than 6,000 km Wide. *PLoS ONE* 7(8): e42740.
- **Pillet V., 2000.** Méthodologie d'extraction automatique d'information à partir de la littérature scientifique en vue d'alimenter un nouveau système d'information. Thèse de doctorat en Sciences de l'Information et de la communication. Dir. Thèse : L. Quoniam et B. Jac. Univ. Aix-Marseille III. 20-43. 134p.
- **Pocock R ., 1935.** The Races of *canis lupus*.

Références bibliographiques

- **Rannala B., Huelsenbeck J., Yang Z., Nielsen R., 1998.** Taxon Sampling and the Accuracy of Large Phylogenies. *Syst. Biol.* 47(4): 702–710.
- **Ranwez V., 2013.** Méthodes efficaces pour reconstruire de grandes phylogénies suivant le principe du maximum de vraisemblance. Thèse de doctorat en Bioinformatique. Université Montpellier II - Sciences et Techniques du Languedoc, 2002. Français : 49-50.122p.
- **Ranwez V., Gascuel O., 2003.** "Phylogenetic Reconstruction Algorithms Based on Weighted 4-Trees." *Lecture Notes in Computer Science* 2066: 337-348.
- **Regnier J., 1960.** Les mammifères du Hoggar. *Bulletin de Liaison Sahara* 40:300-20.
- **Riddle M., 1979.** *The Wild Dogs in Life and Legend.* New York, NY: Howell Book House, Inc.
- **Rischkowsky B., Pilling D., 2008.** L'Etat des Ressources Zoogénétiques pour l'Alimentation et l'Agriculture dans le Monde, pp393-410.
- **Sanger., 1977.** Frederick Sanger, S. Nicklen, et A. R. Coulson (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Science, USA (PNAS)*, 74(12): 5463{7. | Cite page(s) 28.
- **Sarker N., 1990.** *Biological Abstracts*, Volume 92. Philadelphia, PA.
- **Sillero-Zubiri C., Gottelli D., 1994.** *Canis simensis.* *Mammalian Species* 485: 1-6.
- **Sillero-Zubiri C., Gottelli D., 1995.** Spatial organization in the Ethiopian wolf *Canis simensis*: large packs and small stable home ranges. *Journal of Zoology, London* 237: 65-81.
- **Sillero-Zubiri C., Gottelli D., Macdonald D., 1996.** Male philopatry, extra-pack copulations and inbreeding avoidance in Ethiopian wolves (*Canis simensis*). *Behavioral Ecology and Sociobiology* 38: 331-340.
- **Sillero-Zubiri C., Gottelli, 1994.** *Canis simensis.* *Mammalian Species* 485: 1-6.
- **Smithers R., 1983.** *The Mammals of the Southern African Subregion.* Pretoria, Transvaal-Republic of South Africa: University of Pretoria.
- **Straham R., 1983.** *The Australian Museum's Complete Book of Australian Mammals.* Sydney, Australia: Angus and Robertson Publishers.
- **Tamura K., Peterson D., Peterson N., Stecher G., Nei M., Kumar S., 2011.** MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood,

Références bibliographiques

- Evolutionary Distance, and Maximum Parsimony Methods. *Mol. Biol. Evol.* 28:2731-2739.
- **Tedford H., 2009.** Phylogenetic Systematics of the north american fossil caninae . Department of Vertebrate Paleontology Natural History Museum of Los Angeles County, California Number 325, 218 pp.
 - **Thinkquest T., 1998.** "Wildlife" (On-line). Accessed November 20,1999
 - **Van Valkenburgh B., Wayne R., 1994.** Shape divergence associated with size convergence in sympatric East African jackals. *Ecology*, 75(6): 1567-1581.
 - **Venter J., 2001.** The sequence of the human genome. *Science* 291, 1304-1351.
 - **Venter J., Adams M., Sutton G., Kerlavage A., Smith H., Hunkapiller M., 1998.** Shotgun sequencing of the human genome. *Science* 280, 1540-1542.
 - **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.
 - **Wayne B., 1995.** "Red wolves: to conserve or not to conserve" (On-line). Accessed November, 2001
 - **Wilson D., Ruff S., 1999.** The Smithsonian Book of North American Mammals. Washington: Smithsonian Institution Press.
 - **Xia X., 2013.** DAMBE5: a comprehensive software package for data analysis in molecular biology and evolution. *Mol Biol Evol* . Jul; 30(7):1720-8 .

Biblio-Web

Bioedit:<http://www.mbio.ncsu.edu/BioEdit/bioedit.html>

Blast:<https://blast.ncbi.nlm.nih.gov/Blast.cgi>

Clustalw:<http://www.ebi.ac.uk/Tools/msa/clustalw2/>

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

DNAdist:<http://mycor.nancy.inra.fr/egglib/v2/installation.html>

DnaMLK:<http://fr.freedownloadmanager.org/Windows-PC/T-Rex-GRATUIT.html>

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

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

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

PAUP:<http://fr.freedownloadmanager.org/Windows-PC/PAUP-GRATUIT.html>

PHYML:<http://www.atgc-montpellier.fr/phyml/versions.php>

Sequencher:<http://www.commentcamarche.net/download/telecharger-34084043-sequencher>

Références bibliographiques

Liste des figures

N° de figure :	Titre :	Page :
01	Carte de distribution du chacal doré (<i>Canis aureus</i>) et du loup africain (<i>Canis lupus lupaster</i>) en Algérie.	18
02	Le chacal doré <i>Canis aureus</i> (le parc national de Ben Aknoun 2017).	26
03	Le coyote <i>Canis latrans</i> (le parc national de Ben Aknoun 2017).	26
04	Le loup gris <i>Canis lupus</i> (le parc national de Ben Aknoun 2017).	27
05	Les étapes initiales pour récupérer les séquences de la GenBank.	29
06	Les 34 espèces de genre « <i>Canis</i> » obtenues de la GenBank.	30
07	Exemple des séquences nucléotidiques trouvées pour l'espèce « <i>Canis adustus</i> ».	30
08	Exemple d'un fichier sous format Fasta de « <i>Canis latrans</i> ».	31
09	Menu principal de logiciel FaBox.	32
10	L'outil « format convertir » de FaBox utilisé pour convertir les séquences Fasta à partir d'un bloc note.	32
11	L'outil « DNA to haplotype collapse and convert » pour l'obtention des haplotypes.	33
12	L'outil « alignement trimmer » utilisé pour l'alignement de la séquence la plus courte, elle supprime les zones qui sont pleines de gaps.	33
13	Menu principal de logiciel MEGA version 5.	34
14	Exemple des 3 séquences ajoutées dans le logiciel MEGA.	35
15	L'alignement multiple des séquences du genre « <i>Canis</i> » en utilisant le programme ClustalW.	35
16	La construction de l'arbre phylogénétique en utilisant la méthode UPGMA.	36
17	Les paramètres utilisés pour la reconstruction phylogénétique.	36
18	Menu principal de logiciel DnaSP 5.10.	37
19	Les paramètres « Haplotype Data File » et « Roehl Data File » ont été utilisés pour avoir les Haplotypes, et enregistrer le fichier sous format *.rdf.	38
20	Menu principal de logiciel DAMBE.	39
21	Menu principal de logiciel NETWORK 5.0.	39

Liste des figures

22	L'ouverture de fichier sous format rdf produit par logiciel DnaSp.	40
23	L'enregistrement de notre fichier sous format *out pour la reconstruction de rés de NETWORK.	40
24	Organisation de l'ADN mitochondrial humain.	42
25	L'arbre de l'espèce <i>Canis aureus</i> réalisé à l'aide du programme MEGA5.	47
26	L'arbre de l'espèce <i>Canis aureus</i> réalisé à l'aide du programme DAMBE.	48
27	L'arbre de l'espèce <i>Canis adustus</i> réalisé à l'aide du programme MEGA5.	48
28	L'arbre de l'espèce <i>Ccanis adustus</i> réalisé à l'aide du programme DAMBE.	49
29	L'arbre de différentes séquences de sous-espèce de l'espèce <i>Canis lupus</i> réalisé à l'aide du programme MEGA5.	49
30	L'arbre de différentes séquences de sous-espèce de l'espèce <i>Canis lupus</i> réalisé à l'aide du programme DAMBE.	50
31	L'arbre de l'espèce <i>Canis aureus</i> réalisé à l'aide du programme MEGA5.	50
32	L'arbre de l'espèce <i>Canis aureus</i> réalisé à l'aide du programme DAMBE.	51
33	L'arbre de l'espèce <i>Canis latrans</i> réalisé à l'aide du programme MEGA5.	51
34	L'arbre de l'espèce <i>Canis latrans</i> réalisé à l'aide du programme DAMBE.	52
35	L'arbre de différentes séquences de sous-espèce de l'espèce <i>Canis lupus</i> réalisé à l'aide du programme MEGA5.	52
36	L'arbre de différentes séquences de sous-espèce de l'espèce <i>Canis lupus</i> réalisé à l'aide du programme DAMBE.	53
37	L'arbre de l'espèce <i>Canis aureus</i> réalisé à l'aide du programme MEGA5.	53
38	L'arbre de l'espèce <i>Canis aureus</i> réalisé à l'aide du programme DAMBE.	54
39	L'arbre de l'espèce <i>Canis latrans</i> réalisé à l'aide du programme MEGA5.	54
40	L'arbre de l'espèce <i>Canis latrans</i> réalisé à l'aide du programme DAMBE.	55
41	L'arbre de différentes séquences de sous-espèce de l'espèce <i>Canis lupus</i> réalisé à l'aide du programme MEGA5.	55
42	L'arbre de différentes séquences de sous-espèce de l'espèce <i>Canis lupus</i> réalisé à l'aide du programme DAMBE.	56
43	Le nombre des haplotypes de <i>Canis aureus</i> réalisé par le logiciel DnaSP.	57
44	Le nombre des haplotypes de <i>Canis adustus</i> réalisé par le logiciel DnaSP.	58
45	Le nombre des haplotypes de <i>Canis lupus</i> réalisé par le logiciel DnaSP.	58
46	Le nombre des haplotypes de <i>Canis aureus</i> réalisé par le logiciel DnaSP.	59

Liste des figures

47	Le nombre des haplotypes de <i>Canis latrans</i> réalisé par le logiciel DnaSP.	59
48	Le nombre des haplotypes de <i>Canis lupus</i> réalisé par le logiciel DnaSP.	60
49	Le nombre des haplotypes de <i>Canis aureus</i> réalisé par le logiciel DnaSP.	60
50	Le nombre des haplotypes de <i>Canis latrans</i> réalisé par le logiciel DnaSP.	61
51	Le nombre des haplotypes de <i>Canis lupus</i> réalisé par le logiciel DnaSP.	61
52	La Variation du GC skew (A) et AT skew (B) du genre <i>Canis</i> avec 204 bases réalisée par le logiciel DAMBE.	62
53	La Variation du GC skew (A) et AT skew (B) du genre <i>Canis</i> avec 222 bases réalisée par le logiciel DAMBE.	63
54	La Variation du GC skew pour le gène misc-feature du genre <i>Canis</i> avec 222 bases réalisée par le logiciel DAMBE.	63
55	Comparaison des fréquences entre les séquences de cytb et D-loop et misc-feature réalisée par logiciel DAMBE.	64
56	Réseau des haplotypes de l'espèce <i>Canis aureus</i> a été construit grâce au logiciel Network.	65
57	Réseau des haplotypes de l'espèce <i>Canis adustus</i> a été construit grâce au logiciel Network.	65
58	Réseau des haplotypes de l'espèce <i>Canis lupus</i> a été construit grâce au logiciel Network.	66
59	Réseau présentant les relations phylogénétiques entre les haplotypes de l'espèce <i>Canis aureus</i> réalisé par le logiciel NETWORK.	66
60	Réseau présentant les relations phylogénétiques entre les haplotypes de l'espèce <i>Canis latrans</i> réalisé par le logiciel NETWORK.	67
61	Réseau présentant les relations phylogénétiques entre les haplotypes de l'espèce <i>Canis lupus</i> réalisé par le logiciel NETWORK.	67
62	Réseau des haplotypes de l'espèce <i>Canis aureus</i> construit grâce au logiciel Network.	68
63	Réseau présentant les relations phylogénétiques entre les haplotypes de l'espèce <i>Canis latrans</i> réalisé par le logiciel NETWORK.	69
64	Réseau présentant les relations phylogénétiques entre les haplotypes de l'espèce <i>Canis lupus</i> réalisé par le logiciel NETWORK.	69

Liste des tableaux

N° de tableau	Titre	Page
I	La classification de toutes les espèces de genre <i>Canis</i> .	14
II	La répartition et la morphologie des deux espèces dans l'Algérie et dans le monde.	16
III	La répartition géographique dans le monde et l'étude morphologique des 31 espèces de genre <i>Canis</i> .	19
IV	Les séquences d'ADN mitochondrial (cyt b) extraites de la GenBank pour le genre <i>Canis</i> .	43
V	Les séquences d'ADN mitochondrial (D-loop) extraites de la GenBank pour le genre <i>Canis</i> .	44
VI	Les séquences d'ADN mitochondrial (misc-feature) extraites de la GenBank pour le genre <i>Canis</i> .	46
VII	Liste des séquences d'ADN trouvées dans la GenBank pour l'étude phylogénétique du genre « <i>Canis</i> ».	Annexe

Tableau VII : Liste des séquences d'ADN trouvées dans la GenBank pour l'étude phylogénétique du genre « *Canis* ».*Canis adustus* :

Région	Gène	Taille	Locus
Maroc	tRNA, CDS, D-loop	16721	NC_027956
	tRNA, CDS, D-loop	16734	KT378606
	ND6, trnE, CYTB, trnp, trnT, D-loop	2789	KT378607

Canis himalayensis :

Région	Gène	Taille	Locus
Inde	cytochrome b (19)	1140	AY291431
	Misc-feature, D-loop, repeat-region (18)	1316	AY289977
	rRNA (18)	554	AY289962

Canis indica :

Région	Gène	Taille	Locus
Inde	cytochrome b (4)	1140	AY291432
	misc-feature, D-loop, repeat-region (9)	1323	AY289973
	rRNA	554	AY289968

Canis aureus :

Région	Gène	Taille	Locus
Thaïlande	ZFX	960	KX065058
	ZFY	1258	KX065055
Kenya	ZFX(4)	838	KT448250
	ZFY(3)	923	KT448270
	GHR(2)	855	KT447925
	APOB(2)	910	KT447815
	cytochrome b (18)	1140	KT447733
	TTR (2)	1060	KT448201
	ND1, ND2, COX1, COX2, ATP8, ATP6, COX3, ND3, ND4L, ND4, ND5, ND6, CYTB, D-loop	16727	KT448273
	RASA2 (2)	484	KT448178
	RABEP1 (2)	781	KT448160
	PRKDC (2)	743	KT448137
	NUP (2)	775	KT448063
	NTRK3 (2)	783	KT448040
	PNOC (2)	259	KT448114
	PDE6B (2)	512	KT448091
	RBP3 (2)	770	KT447948
	MDH1(2)	601	KT448017
	MACF1 (2)	716	KT447976
MC4R (2)	857	KT447994	
FBN1 (2)	675	KT447884	

	CHRNA1 (2)	359	KT447861
	APOE (2)	637	KT447833
	ACTN4 (2)	503	KT447792
	AAMP (2)	513	KT447769
	FGG(2)	664	KT447907
Palestine	ZFX(7)	838	KT448243
	ZFY(7)	1146	KT448261
	GHR(2)	855	KT447929
	APOB(2)	910	KT447811
	Cytochrome b (21)	1140	KT447731
	D-loop	438	KT343802
	TTR (2)	1060	KT448202
	ND1, ND2, COX1, COX2, ATP8, ATP6, COX3, ND3, ND4L, ND4, ND5, ND6, CYTB, D-loop	16730	KT448274
	RASA2 (2)	481	KT448182
	RABEP1 (2)	781	KT448156
	PRKDC (2)	743	KT448136
	NUP (2)	821	KT448067
	NTRK3 (2)	792	KT448044
	PONC (2)	259	KT448113
	PDE6B (2)	512	KT448090
	RBP3 (2)	770	KT447952
	FGG	664	KT447903
	MDH1 (2)	601	KT448018
	MACF1 (2)	716	KT447972
	MC4R (2)	857	KT447995
	FBN1 (2)	679	KT447880
	CHRNA1 (2)	359	KT447860
	APOE (2)	637	KT447837
	ACTN4 (2)	503	KT447791
	AAMP (2)	513	KT447768
	tRNA, rRNA	404	KP191627
	Maroc	ZFX(2)	838
ZFY		923	KT448267
GHR		855	KT447931
APOB(2)		910	KT447816
Cytochrome b (2)		1140	KT447759
TTR		1060	KT448207
RASA2		484	KT448184
RABEP1		781	KT448161
NTRK3		792	KT448046
PDE6B		512	KT448092
RBP3		770	KT447954
FGG		654	KT447908
MDH1		601	KT448023

	MACF1	716	KT447977
	MC4R	857	KT448000
	FBN1	626	KT447885
	CHRNA1	359	KT447862
	APOE	637	KT447839
	ACTN4	417	KT447793
	AAMP	510	KT447770
	PNOC	259	KT448115
	NUP	765	KT448069
Afghanistan	ZFX	838	KT448231
	GHR	830	KT447928
	APOB	910	KT447813
	Cytochrome b	1140	KT447748
	TTR	1057	KT448204
	RASA2	481	KT448181
	RABEP1	781	KT448158
	PRKDC	743	KT448135
	NUP	821	KT448066
	NTRK3	393	KT448043
	PNOC	259	KT448112
	PDE6B	512	KT448089
	RBP3	770	KT447951
	FGG	664	KT447905
	MDH1	601	KT448020
	MACF1	716	KT447974
	MC4R	857	KT447997
	FBN1	679	KT447882
	CHRNA1	359	KT447859
	APOE	637	KT447836
ACTN4	503	KT447790	
AAMP	513	KT447767	
Serbie	ZFX (5)	838	KT448236
	ZFY (3)	1146	KT448259
	GHR	830	KT447927
	APOB	910	KT447812
	Cytochrome b (10)	1140	KT447758
	D-loop	305	HQ845260
	TTR	1060	KT448203
	Misc-feature	392	GU936680
	RASA2	481	KT448180
	RABEP1	781	KT448157
	PRKDC	743	KT448134
	NUP	818	KT448065
	NTRK3	792	KT448042
	PNOC	259	KT448111

	PDE6B	510	KT448088
	RBP3	770	KT447950
	FGG	664	KT447904
	MDH1	601	KT448019
	MACF1	716	KT447973
	MC4R	857	KT447996
	FBN1	679	KT447881
	CHRNA1	359	KT447858
	APOE	637	KT447835
	ACTN4	503	KT447789
	AAMP	513	KT447766
	ZFX (2)	838	KT448245
	GHR	855	KT447932
	APOB(2)	910	KT447817
	Cytochrome b (2)	1140	KT447761
	TTR	1060	KT448208
	RASA2	484	KT448185
	RABEP1	781	KT448162
	PRKDC	743	KT448139
	NTRK3	792	KT448047
	PNOC (2)	512	KT448093
	NUP	775	KT448070
	RBP3	768	KT447955
	FGG	664	KT447909
	MDH1	601	KT448024
	MACF1	716	KT447978
	CHRNA1	359	KT447863
	APOE	637	KT447840
	ACTN4	417	KT447794
	AAMP	513	KT447771
Mauritanie	MC4R	857	KT448001
	FBN1	661	KT447886
Japon	ZFX	838	AB622134
	ZFY	1146	AB622145
	ZFY	566	KF021269
	DLA-DQA1	246	KT159767
	DLA-DQB1(2)	267	KT182925
	DLA-DRB1(3)	270	KT182926
	Misc-feature	648	JX220172
	Misc-feature	586	JX220170
	Misc-feature	664	JX220168
	BRCA1(2)	741	DQ240427
	VWF	732	DQ240791
USA	VANGL2	546	DQ240757
	TMEM20	615	DQ240724

	RAG1	741	DQ240690
	GHR(2)	811	AY885383
	CMKOR1	735	DQ240590
	CHST12	705	DQ240557
	APOB (2)	702	DQ240292
	BDNF	489	DQ240359
	mRNA	449	AY885359
	mRNA	359	AY885312
	vitronectin gene, exon 4, intron 4, exon 5 and partia cds	481	AY885407
	cytochrome b (2)	385	AF028138
	cytochrome P450	598	AY885336
	COII (2)	684	AY609143
	COI	588	AF028186
	RPPH1 (2)	671	AY609127
	growth factor receptor 3	492	DQ240624
	TRSP	710	AY609091
	PRDM9	1296	JF750659
	1.75	714	DQ240528
	7.92	583	DQ240495
	12.149	917	DQ240461
	Cytochrome b (9)	1140	AY291433
	COI	580	KJ887398
	D-loop (15)	468	KT343797
	misc-feature, D-loop, repeat-region	1333	AY289996
	rRNA	554	AY289969
	rRNA	436	DQ102371
Inde	rRNA	389	EU084670
Égypte	Cytochrome b	1140	KT447732
	Cytochrome b	332	KT343786
Bulgarie	D-loop	442	KT343803
	COI (3)	1506	KX156586
Turquie	D-loop(4)	439	KT988009
	D-loop	640	KT268319
Ukraine	Misc-feature	406	KT362176
Pologne	D-loop	640	KT268318
Espagne	D-loop	659	KJ490946
Lituanie	Misc-feature	460	KT123040
Grèce	Misc-feature	406	KT362174
Italie	D-loop	595	AF184048
	Misc-feature	450	KF588364
Algérie, Mali, Sénégal	Misc-feature	289	JQ088675
Canada	EPM2B	276	AY664373

Canis mesomelas elongae :

Région	Gène	Taille	Locus
USA	COI (2)	588	AF028191
	COII (2)	684	AF028214
	cyb (4)	396	AF028143

Canis rufus :

Région	Gène	Taille	Locus
USA	D-loop	341	HM755703
	bêta-fibrinogen (2)	786	AY882045
	TR	1060	AY750580
	TBG	442	AY750654
	ND2	1044	AY882058
	RNASE1 (2)	456	KC885046
	UGT1A1	642	EF440368
	UGT1A6	664	EF440405
	Cyb (11)	235	CRU47042
	Misc-feature	420	GU903017
	tRNA;CDS	1067	HM755690
Nord Carolina (USA)	Caru-DRB (7)	280	AF516916
	Misc-feature	325	AY280913
Canada	DLA-DQA1	246	DQ777755
	EPM2B	273	AY664371
	Tas1r2	3011	JN130360

Canis simensis :

Région	Gène	Taille	Locus
USA	CYPIA1	533	DQ239447
	CHRNA1	359	DQ239439
	COI	588	AF028192
	growth factor receptor 3	492	DQ240628
	TRSP	710	DQ239471
	BRCA1(2)	741	DQ240431
	VWF	732	DQ240795
	VANGL2	546	DQ240761
	TMEM20	615	DQ240728
	RAG1	741	DQ240694
	CMKOR1	735	DQ240594
	CHST12	705	DQ240561
	VTN	481	DQ239479

	COII	684	AF028216
	GHR (2)	810	DQ239463
	fes	448	DQ239455
	BDNF	489	DQ240363
	APOB (2)	702	DQ240296
	Misc-feature	233	AY562100
	1.75	714	DQ240532
	7.92	591	DQ240499
	12.149	917	DQ240465
Éthiopie	COII	684	DOGCOII
	Cyt b (10)	1140	KT447693
	TTR	1060	KT448211
	RASA2	509	KT448188
	RABEP1	785	KT448165
	PRKDC	743	KT448142
	PNOC	259	KT448119
	PDE6B	513	KT448096
	NUP	776	KT448073
	NTRK3	792	KT448050
	MDH1	601	KT448027
	MC4R	857	KT448004
	MACF1	710	KT447981
	RBP3	770	KT447958
	GHR	855	KT447935
	FGG	664	KT447912
	FBN1	675	KT447889
	ACTN4	503	KT447797
	AAMP	513	KT447774
	DLA-DQB1(3)	267	FM246844
	DLA-DQA1	246	AM182469
	DLA-DRB1(4)	270	AM182468
	CHRNA1	359	KT447866
	COI	588	DOGCOI
	APOE	637	KT447843
	APOB	910	KT447820
	mitochondrion, partial genome	9419	KT448281
	D-loop	234	AY551930
	D-loop	310	HQ845261
	D-loop	234	AY551930

Canis latrans :

Région	Gène	Taille	Locus
Chine	ND1, ND2, COX1, COX2, ATP8, ATP6, COX3, ND3, ND4L, ND4, ND5, ND6, CYTB, D-loop	16196	EU789789
	Misc-feature	606	HQ389435
	Misc-feature	545	HQ389371
	Misc-feature	1241	HQ389383
	Misc-feature	1489	HQ389389
	Misc-feature	1459	HQ389392
	Misc-feature	1376	HQ389396
	Misc-feature	839	HQ389400
	Misc-feature	396	HQ389406
	Misc-feature	761	HQ389415
	Misc-feature	734	HQ389422
	Misc-feature	1147	HQ389431
	Misc-feature	368	HQ389374
USA	Cala-DRB (14)	280	AF516924
	Cala-DRB1 (8)	335	EU400582
	ND1, ND2, COX1, COX2, ATP8, ATP6, COX3, ND3, ND4L, ND4, ND5, ND6, CYTB, D-loop (8)	16732	DQ480509
	BRCA1 (2)	741	DQ240428
	VWF	732	DQ240792
	VANGL2	546	DQ240758
	TMEM20	615	DQ240725
	RAG1	741	DQ240691
	GHR (4)	855	KT447933
	CMKOR1	735	DQ240591
	CHST12	705	DQ240558
	APOB (3)	910	KT447818
	BDNF	489	DQ240360
	mRNA, CDS	448	AY885361
	mRNA, CDS	359	AY885314
	mRNA, CDS	481	AY885409
	Cytochrome P450	598	AY885338
	Cytochrome b (59)	1140	KT447695
	HLA-DOA	807	AY745979
	COII (2)	684	AY609145
	ZFY (4)	1037	AY145848
	ZFX (3)	921	AY145847
	DLA-DRB1 (2)	270	AF343743
	DLA-DQA	246	AY126647
	DLA-DQA1 (2)	246	DQ777756
	Misc-feature (123)	461	KT946980
	TTR (2)	1060	KT448222

	RASA2 (2)	484	KT448199
	RABEP1 (2)	781	KT448176
	PRKDC (2)	743	KT448153
	PNOC (2)	259	KT448117
	PDE6B (2)	512	KT448107
	NUP (2)	775	KT448071
	NTRK3 (2)	792	KT448061
	MDH1 (2)	601	KT448025
	MC4R (2)	857	KT448015
	MACF1 (2)	716	KT447992
	RBP3 (2)	770	KT447969
	FGG (2)	664	KT447910
	FBN1 (2)	675	KT447900
	CHRNA1 (2)	359	KT447877
	APOE (2)	637	KT447854
	AAMP (2)	513	KT447785
	ACTN4 (2)	503	KT447808
	D-loop (64)	538	EU400578
	TFRC	2313	JN887450
	CDS	581	CLU86661
	tRNA, rRNA	1062	CLU78329
	STS (10124)	736	BV397186
	PRNP	910	EU341496
	rRNA	1040	DQ334812
	Misc-feature	223	GU062764
	Misc--feature	1268	AY240094
	Misc-feature(44)	444	EF508155
	growth factor receptor 3	492	DQ240625
	7.92	590	DQ240496
	12.149	917	DQ240462
	1.75	714	DQ240529
	rRNA	375	FJ797662
	Misc-feature (21)	402	GQ849379
	prdm9 (2)	1886	HE590880
Sud-Américain	TRSP (2)	710	AY609094
Nord de Caroline	Misc-feature (27)	329	AY280925
Centre-nord du Texas (USA)	D-loop (24)	229	GQ981348
Nord Américain	D-loop (61)	421	FM209368
	D-loop(2)	708	AF008158
Sud-est des États-Unis	Misc-feature	423	AF541876
	tRNA, CDS (14)	1067	HM755696
	ASIP (2)	489	DQ238596

Canada	PPARGC1A, Misc-feature	590	JN036836
	COI (8)	657	JF443196
	SLC45A2	1658	KM222820
	MC1R	1215	KJ363162
	D-loop (6)	702	AF098154
	Misc-feature (42)	225	GU014250
	EPM2B	286	AY664372
	tRNA-Pro, D-loop (22)	361	GQ863734
	Misc-feature(5)	225	FJ889993
	D-loop	398	AF020700
Corée	D-loop(2)	1266	AY172674
Norvégien, Scotland	PRNP (3)	774	JX218953

Canis lupus actros :

Région	Gène	Taille	Locus
Eurasie	Mcr1 (2)	942	JX273636

Canis lupus baileyi :

Région	Gène	Taille	Locus
USA	Calu-DRB(5)	207	AY009509
	Cytb	1140	HM222711
Eurasie	ZFX	837	KT448226
	TTR	1060	KT448220
	RASA2	484	KT448197
	RABEP1	781	KT448174
	PRKDC	730	KT448151
	PNOC	259	KT448128
	PDE6B	512	KT448105
	NUP	772	KT448082
	NTRK3	792	KT448059
	MDH1	601	KT448036
	MC4R	857	KT448013
	MACF1	716	KT447990
	RBP3	770	KT447967
	GHR	855	KT447944
	FGG	664	KT447921
	FBN1	675	KT447898
	CHRNA1	359	KT447875
	APOE	637	KT447852
	APOB	910	KT447829
	ACTN4	503	KT447806
	AAMP	513	KT447783

Canis lupus campestris :

Région	Gène	Taille	Locus
Mongolie	ND2, ND3, ND4, ND5, ND6, ND4L, COX1, COX2, COX3, ATP6, ATP8, cytb, D-loop	16719	KC896375

Canis lupus chanco :

Région	Gène	Taille	Locus
Chine	ND2, ND3, ND4, ND5, ND6, ND4L, COX1, COX3, ATP6, ATP8, cytb, D-loop	16774	EU442884
	NADH1, NADH2, NADH3, NADH4, NADH5, NADH6, NADH4L, COX1, COX2, COX3, ATP6, ATP8, cytb, D-loop	16709	GQ374438
	D-loop (11)	588	JX415348
	CDS	582	KC545621
	Misc-feature (16)	591	KC414568
Himalaya	Cytb	385	AY333748
	D-loop (4)	335	AY333741
Népal	D-loop	220	KT321360
Russie	D-loop (2)	598	AB480743
Mongolie	D-loop (5)	673	AB007378

Canis lupus desertorum :

Région	Gène	Taille	Locus
Chine	ND1, ND2, ND3, ND4, ND5, ND6, ND4L, COX1, COX2, COX3, ATP6, ATP8 CYTB, D-loop	16770	KC461238

Canis lupus dingo :

Région	Gène	Taille	Locus
Australie	COX2, ATP8, ATP6 (7)	677	JX090194
	ND6, CYTB (7)	1029	JX094432
Taiwan	Misc-feature (8)	610	KC106733
	EPM2B	278	AY664369
Polynésie	Misc-feature (20)	583	AY660656

Canis lupus hattai :

Région	Gène	Taille	Locus
Japon	D-loop	599	AB500701
	D-loop (6)	598	LC064093

Canis lupus hodophilax :

Région	Gène	Taille	Locus
Japon	D-loop (8)	598	AB480736
	Amlex (2)	351	AB080689

Canis lupus labradorius :

Région	Gène	Taille	Locus
Nord Américain	Misc-feature	425	GQ849370

Canis lupus laniger :

Région	Gène	Taille	Locus
Chine	mRNA, CDS	708	FJ861077
	ND1, ND2, ND3, ND4, ND5, ND6, ND4L, COX1, COX2, COX3, ATP6, ATP8, CYTB, D-loop	16729	KF573616
	ND1, ND2, ND3, ND4, ND5, ND6, ND4L, COX1, COX2, COX3, ATP6, ATP8, CYTB, D-loop	16723	FJ032363
	ND1, ND2, ND3, ND4, ND5, ND6, ND4L, COX1, COX2, COX3, ATP6, ATP8, CYTB, D-loop	16723	NC_011218

Canis lupus lupaster :

Région	Gène	Taille	Locus
Éthiopie	CYTB, CDS	1140	HQ845258
	rRNA (3)	354	HQ845256
Maroc	CYTB, CDS	474	KM670014
	D-loop (2)	306	KM670012
Eurasie	CYTB, CDS	189	KT447711
Algérie, Mali et Sénégal	Misc-feature (7)	290	JQ088679
	CYTB, CDS (7)	402	JQ088665

Canis lupus lupus :

Région	Gène	Taille	Locus
Sibérie	Mc1r (3)	942	JX273639
	ncRNA (2)	1076	KC970150
Mongolie	D-loop	673	AB007372
	tRNA-Phe, tRNA-Ile, tRNA-Val, tRNA-Leu, tRNA-Gln, tRNA-Gln, tRNA Met, tRNA-Trp, tRNA-Ala, tRNA-Asn, tRNA-Cys, tRNA-Tyr, tRNA-Asp, tRNA-Lys, tRNA-Gly, tRNA-Arg, tRNA-His, tRNA-Ser-tRNA-Thr tRNA-pro, 12S r RNA, 16S r RNA, tRNA-Glu, tRNA-Coi, tRNA-Coiii, tRNA- Atpase8et 6, NADH1, NADH2, NADH3, NADH4, NADH5, NADH6, NADH4L, CYTB	16757	AM711902
Suède	tRNA-Phe, tRNA-Ile, tRNA-Val, tRNA-Leu	16757	NC_009686
	tRNA-Gln, tRNA-Gln, tRNA-Met, tRNA-Trp, tRNA-Ala, tRNA-Asn, tRNA-Cys, tRNA-Tyr, tRNA-Asp, tRNA-Lys, tRNA-Gly, tRNA-Arg, tRNA-His, tRNA-Ser- tRNA-Thr, tRNA-pro, 12S r RNA ,16S r RNA, ND1, ND2, ND3, ND4, ND5, ND6, ND4L, COX1, COX2, COX3, ATP6, ATP8,CYTB		

Canis lupus pallipes:

Région	Gène	Taille	Locus
Himalaya	CYTB, CDS	385	AY333749
	D-loop (4)	336	AY333746
Mongolie	D-loop	672	AB007374

Canis lupus signatus:

Région	Gène	Taille	Locus
Portugal	Misc-feature (5)	142	KT967969
	Misc-feature (11)	423	KT967960
	ZFY (2)	194	GQ870640
	ZFX (5)	194	GQ870637
	tRNA;D-loop (5)	417	JX845625
	Misc-feature (26)	399	GQ366772
Espagne	D-loop (4)	647	EF380229

Canis lupus variabilis :

Région	Gène	Taille	Locus
Russie	Misc-feature	277	KJ909853

Canis lupus x Canis lupus familiaris :

Région	Gène	Taille	Locus
Estonie	Misc-feature (17)	1402	JN182137

Canis lycaon :

Région	Gène	Taille	Locus
Nord	tRNA, CDS (2)	1067	HM755686
Américain	ND1, ND2, ND3, ND4, ND5, ND6, ND4L, COX1, COX2, COX3, ATP6, ATP8	16753	JF342907
	CYTB, D -loop		
	Zfy	463	FJ687620
	D-loop (5)	341	HM755706
	Misc-feature	407	GQ849369
Canada	Misc-feature (3)	224	AY267720

Canis adustus :

Région	Gène	Taille	Locus
USA	BRCA1(2)	741	DQ240689
	VWF	732	DQ240790
	VANGL2	546	DQ240756
	TMEM20	603	DQ240723
	RAG1	741	DQ240689
	GHR(2)	736	DQ240655
	CMKOR1	735	DQ240589
	CHST12	705	DQ240556
	BDNF	489	DQ240358
	APOB(2)	702	DQ240291
	growth factor receptor 3	492	DQ240623
	1,75	720	DQ240527
	7,92	590	DQ240494
	12,149	917	DQ240460
Sud- Américain	mRNA; CDS	359	AY885310
	mRNA; CDS	481	AY885405
	mRNA; CDS	599	AY885334
	COII	684	AF028208
	COII	684	AY609141
	Cytb (2)	385	AF028136
	TRSP	708	AY609088
	RPPH1	584	AY609128
	GHR(2)	807	KT447924
	APOB(2)	910	KT447809
	CYTB	15263	KT448271

Eurasie	TTR	1046	KT448200
	RASA2	459	KT448177
	RABEP1	781	KT448154
	PRKDC	746	KT448131
	PNOC	259	KT448108
	PDE6B	510	KT448085
	NUP	775	KT448062
	NTRK3	792	KT448039
	MDH1	600	KT448016
	MC4R	857	KT447993
	MACF1	716	KT447970
	RBP3	770	KT447947
	FGG	664	KT447901
	FBN1	675	KT447878
	CHRNA1	341	KT447855
	APOE	643	KT447832
	ACTN4	503	KT447786
	AAMP	513	KT447763
Guinée	Cytb (6)	402	JQ088650
	COI (5)	658	KJ192748
	rRNA (5)	525	KJ193015
	rRNA (5)	389	KJ192451
Algérie, Mali et Sénégal	Misc-feature	291	JQ088674
	Misc-feature (5)	287	JQ088669
Éthiopie	Cytb	1121	KP874955
	D-loop	303	KP874956
	COI	588	AF028184
	CDS	448	AY885357
	mRNA, CDS	807	AY885381

Canis sp. Belgium/36,000 :

Région	Gène	Taille	Locus
Belgique	ND1, ND2, COX1, COX2, ATP8, ATP6, COX3, ND3, ND4L, ND4, ND5, ND6, CYTB, D-Loop	16494	KF661079

Canis sp. CANInt1 :

Région	Gène	Taille	Locus
Canada	Zfy	463	FJ687618

Canis sp.KEB-2016 :

Région	Gène	Taille	Locus
USA	Misc-featurre	306	KU556391

Canis sp.Russia/33,500 :

Région	Gène	Taille	Locus
Russie	ND1, ND2, COX1, COX2, ATP8, ATP6, COX3, ND3, ND4L, ND4, ND5, ND6, CYTB, D-loop	16727	KF661092

Canis environmental sample :

Région	Gène	Taille	Locus
USA	Misc-feature (60)	218	JF298093

Canis sp :

Région	Gène	Taille	Locus
Japon	(1)	561	LF806145
	(1)	686	BD408531
	(8)	501	E28740
	Interleukin 1-bêta convertase	582	E32835
	(4)	501	E28566
	mRNA for interleukin 12 (2)	990	E15017
	cDNA encoding peptide fragment of dog immunoglobulin E	372	E13301
	CZP2	2216	E07830
	CZP-3	1320	E06068
	DNA sequence coding for unchangeable region of dog immunoglobulin gamma chain	1007	E03345
	(1)	686	BD411926
	mat-peptide (2)	1961	CQ878845
	(genomic DNA)	459	BD188078
	PPAR-alpha activity	2032	BD171828
	(1)	540	BD162194
	(28)	517	BD107636
	(32)	517	BD015478
	(32)	517	BD015418
	(4)	990	BD014663
	(16)	517	E59431
	Canine obsity gene (2)	501	E59784
	Interferon-gamma mutant (31)	517	E58991
	Sig-peptide;mat-peptide, CDS	567	DOGPPLA2
	Misc-feature	291	AH001620
(3)	71	HZ427142	
(3)	71	LG064780	

Allemagne	(5)	76	HW805256
	(2)	71	HW802911
	(2)	28	DI441345
	(2)	31	HW798564
	(4)	62	HW808777
	(2)	28	HW795892
	(1)	71	DI442315
	(1)	435	DI016822
	(1)	1291	HW287310
	Misc-feature (2)	3008	HV745889
	(2)	76	HV945897
	Srp	127	2GO5_A
	Canada	Cytb (3)	421
Protein-bind, CDS (2)		3199	DOGSNCE
CDS		832	DOGPLBA
USA	(2)	597	HW340988
	CMR1	3815	DI156457
	(1)	901	DI088996
	(2)	70	JC361111
	(1)	1291	HV581979
	CDS (2)	831	DOGPSAP
	CDS	1080	DOGGAR1
	Sig-peptide, mat-peptide, CDS	1442	DOGSP18A
	Misc-feature (3)	1814	JA256017
	(2)	31	JA084317
	MIR (2)	76	JA684745
	(5)	2154	HV447625
	RNAS (2)	31	HV444938
	BMP-7 (2)	1296	HV234435
	miR (5)	76	FW592362
	miR (3)	71	FW592397
	SMH (2)	597	FW397586
	BMP-7 (8)	1296	FW363035
	10q23.3	1291	DM146413
	TAA	343	GN093658
	U2 snRNA	125	X69482
	clox, CDS	3379	X69017
	(8)	26	DL074661
	(2)	582	DJ065604
	(2)	24	DL074648
	CDS	480	CS467597
	RHO (14)	1569	CS566362
	(6)	813	BD313797

Annexe

	(7)	31	BD283945
	(genomic DNA) (4)	1353	BD183703
	Novel melanocortin-4 receptor sequences	1985	BD181636
	Anti-IgE vaccines (14)	102	BD177577
	CDS	582	AX011256
	Sig-peptide, mat-peptide, IDUA	2175	DOGIDUA
Angleterre	mRNA, CDS (2)	869	DOGTRYPC
	mat-peptide, CDS	762	DOGNKBNT
	CDS	1440	DOGGSTRN
	mat-peptide,CDS, sig-peptide, mRNA	2575	DOGCAL
	Sequence 5 from Patent WO0238612	34	AX467375
	Sequence 1 from Patent WO0238612	1263	AX467371
	Sequence 3 from Patent EP1160323	901	AX316088
	Sequence 1 from Patent WO0046402	132	AX033072
Sequence 3 from Patent WO0029428	901	AX025013	
Israel	repeat-region,regulatory	784	DOGMYCRC
	misc-feature	281	DOGMYCC
Chine	(4)	990	HV194784
	CDR	594	FW397162
	mat-peptide,sig-peptide	498	DM380012
	(2)	41	FB985031
France	Sequence 66 from Patent WO0077043	435	AX067829
	(7)	1429	CS458969
	(7)	1429	DJ442813
Himalaya, Inde	rRNA	103	KJ155700
Turquie	Srp	128	2J37_A
Argentine	Sequence 4 from Patent WO0071576	1803	AX050405
Pays de Galles	Sequence 59 from Patent WO03018632, Sequence 43 from Patent WO03018632, Sequence 27 from Patent WO03018632, Sequence 11 from Patent WO03018632 (4)	351	AX772071
Espagne	(3)	1528	DJ430829
Italie	BNP	1803	DOGBNPA
Romania	Misc-feature	3008	FW392675
	CDS	3978	CS723370
	(3)	2898	CS799978
	(8)	701	BD261938
	(1)	582	BD205497
	(1)	12	BD242999
	(1)	901	BD249733
	Unassigned DNA	2925	E59801
	Unassigned DNA	459	E59007

Annexe

Cholecystokinin-B/gastrin receptor(2)	1338	E10611
Cholecystokinin-B/gastrin receptor(2)	1450	E10409
Cholecystokinin-B/gastrin receptor(2)	1450	E09597
Interferon-beta	819	E11229
(4)	1560	E32838