

N° D'ordre : 46 / 2017 - D/S.B.

RÉPUBLIQUE ALGÉRIENNE DÉMOCRATIQUE ET POPULAIRE
MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE SCIENTIFIQUE
UNIVERSITÉ DES SCIENCES ET DE LA TECHNOLOGIE HOUARI BOUMEDIENE
Faculté des Sciences Biologiques



THESE

de Doctorat en Science

Présentée pour l'obtention du grade de

DOCTEUR

En Sciences Biologiques

Option Eco-Biologie et Amélioration Végétale

par

MOUSSOUNI Souhila

SUJET

**Caractérisation moléculaire de cultivars de palmier
dattier (*Phoenix dactylifera* L.)**

Soutenue publiquement, le 14 Décembre 2017 devant le jury composé de :

Mme ABROUS-BELBACHIR O., Professeur à l'USTHB/Alger

Présidente

Mme BOUGUEDOURA N., Professeur à l'USTHB/Alger

Directrice de thèse

Mme BENNACEUR M., Professeur à l'U.A.B.B/Oran1

Examinatrice

Mr AMIROUCHE R., Professeur à l'USTHB/Alger

Examinateur

Mr BRAHIM-ERRAHMANI M., Maître de Conférences A à l'U.S.D/Blida

Examinateur

Mr DJEBBAR R., Professeur à l'USTHB/Alger

Examinateur

Résumé

Le palmier dattier (*Phoenix dactylifera* L.) est le pilier de l'agriculture oasienne. Nous avons échantillonné 10 oasis réparties sur tout le Sahara algérien pour comprendre la structure de la diversité du palmier dattier. Nous avons utilisé 18 marqueurs microsattellites et un minisatellite chloroplaste pour caractériser 414 individus de dattier correspondant à 114 variétés nommées. Nous avons trouvé un coefficient de consanguinité négatif significatif, ce qui suggère la sélection active des agriculteurs pour les individus hétérozygotes. Cette analyse nous a permis d'établir la clé de détermination moléculaire des variétés algériennes. Cependant la variabilité intra et inter-variétale fournit une lecture complète de l'agrobiodiversité du palmier dattier algérien.

Afin d'étudier la diversité et la structure génétique du palmier dattier dans le monde, nous avons évalué le polymorphisme des marqueurs SSR nucléaires sur un échantillonnage important et sur une large distribution géographique pour révéler les relations entre les populations de palmier des différents pays. Nous avons analysé également un minisatellite chloroplastique (CpSSR), qui est un marqueur haploïde avec un taux inférieur de mutation que les SSR nucléaires. En effet, étant hérité de la mère, le génome du chloroplaste nous fournit des informations sur l'origine de la plante mère.

Le séquençage de ce minisatellite de douze paires de bases dans une région non codante au niveau de l'espaceur intergénique *psbZ-trnfM* montre une variation entre trois et quatre répétitions chez le palmier dattier. L'étude de ce minisatellite, chez quatorze espèces de *Phoenix* a montré la présence de six haplotypes dont certains présents chez plusieurs espèces. Le séquençage de quatre locus chloroplastiques chez 24 variétés a permis d'identifier les deux chlorotypes chez le palmier dattier, caractérisés par un ensemble de polymorphismes liés.

Mots clés : *Phoenix dactylifera* L., diversité génétique, microsattellites, minisatellite chloroplastique.

Abstract

The date palm (*Phoenix dactylifera* L.) is the mainstay of Oasis agriculture. To understand the genetic structure of date palm, we have sampled 10 Oasis in Algeria. Eighteen molecular markers and one chloroplastic marker were used to characterize 414 individuals which correspond to 114 named varieties. We found that, the negative consanguinity coefficient is significant and involved active heterozygote selection by the farmers. This analysis allows us to identify the molecular key to characterize Algerian varieties. However, the intra and inter variability study furnishes the complete knowledge of agro-biodiversity of Algerian date palm.

In order to study the genetic structure and diversity of date palm in the world, we have evaluated the nuclear SSR marker polymorphism on a large number of samples covering important geographic distribution to make relations among date palm populations collected in different countries. We have also analyzed a chloroplast minisatellite marker (CpSSR), which is known as haploid with a mutation rate less than nuclear SSR markers and so evolves differently. This chloroplast minisatellite marker, as maternal heritage, gives information on the origin of the plant.

The sequencing of this minisatellite marker of 12 base pairs in the non-coding region in the level of *psbZ-trnfM* shows variability within three or four tandem repeats in date palm. The study of this minisatellite in 14 species of *Phoenix* has shown the presence of six haplotypes and some of them are present in many species.

The sequencing of four chloroplastic loci in 24 varieties allowed identifying the two chlorotypes in date palm characterized by linked polymorphism.

Keywords: *Phoenix dactylifera* L., genetic diversity, microsatellites, chloroplast minisatellite.

نخيل التمر هو أساس زراعة الواحات. استعملنا عينات من 10 واحات في الجزائر لفهم بنية التنوع الوراثي لنخيل التمر.

18 القصيرة المتكررة النووية و مؤشر التتابع الطويل المتكرر بلاستيدي لتوصيف 414

يمتلون 114 . لقد وجدنا معامل القرابة سلبي، مما يدل على التحديد النشط من المزارعين للأفراد المتخالفة.

هذا التحليل بإنشاء مفتاح التقرير الجزئي للأصناف الجزائرية. . ان التباين داخل وبين الأصناف يقدم قراءة كاملة للتنوع البيولوجي الزراعي للنخيل الجزائري.

لدراسة التنوع والهيكل الوراثي للنخيل في العالم، قمنا بتقييم تباينات المؤشرات النووية على عينة كبيرة من النخيل على توزيع جغرافي واسع للكشف عن العلاقة بين مجموعات النخيل في مختلف البلدان. قمنا أيضا بتحليل مؤشر التتابع الطويل رر بلاستيدي الذي هو مؤشر انفرادي مع انخفاض معدل الطفرة لديه مقارنة مع المؤشرات النووية، وهكذا تطوره . لأنه موروث من الأم، جينوم بلاستيديات الخضراء يوفر لنا معلومات حول أصل النبتة .

تسلسل مؤشر التتابع الطويل المتكرر الذي يحتوي على اثني عشر قاعدة ازوتية و الموجود في منطقة بين الجينات يظهر الاختلاف بين ثلاثة وأربعة تكرار في النخيل. وقد أظهرت دراسة هذا المؤشر التتابع المتكرر الطويل، عند أربعة عشر نوع فينيكس وجود ستة نسخ، مع وجود بعض النسخ في العديد من أنواع فينيكس.

نسختين جينوم بلاستيديات

ضع جيني للبلستيديات الخضراء في 24

في النخيل، وتتميز بمجموعة من الأشكال المرتبطة.

Remerciements

Voici sans doute la partie la plus libre de ce manuscrit, mais pas nécessairement la plus évidente à écrire.

Cette thèse est le fruit de longues années de travail au cours desquelles plusieurs personnes ont aidé de près ou de loin à la réalisation de cette thèse. Je vais donc essayer de n'oublier personne dans la longue liste que je dois dresser, et leur exprimer toute ma gratitude.

Je remercie très sincèrement Mme le professeur BOUGUEDOURA Nadia, ma directrice de thèse, pour son encadrement et son implication dans le travail. Cette thèse n'aurait pas vu le jour sans son soutien, ses encouragements et sans les efforts qu'elle a déployés pour avoir un cadre de travail confortable. Ses conseils, ses critiques, ses corrections et tous les échanges réguliers avec elle au cours de ma thèse ont certainement été la lumière qui a guidé et éclairé mes pas. Je n'oublie pas non plus que grâce à elle je suis partie au Sahara où j'ai pu visiter différentes oasis, des expériences qui ne s'oublient pas. De la science, de sa rigueur, de sa passion et son extraordinaire connaissance du palmier dattier, elle m'a donné du meilleur d'elle-même. Je suis devenue passionnée du palmier dattier, et je sais qu'il y a encore à apprendre. De tout cœur, professeur, je vous dis : MERCI.

Jean-Christophe PINTAUD...que lui dire ? Quels mots utiliser pour lui exprimer toute l'émotion de mon cœur, toute ma reconnaissance ? Il m'avait adopté comme étudiante et sœur, il m'avait tout appris, tout donné. Avec la plus incroyable des générosités. Je me souviendrais encore et toujours du début de l'histoire, son déroulement, et sa fin. Avec joie et bonheur, car des souvenirs, j'en ai eu avec lui et sa famille que des très bons. Sa disponibilité totale, sa patience et sa pédagogie m'ont permis de comprendre beaucoup de choses importantes. Jean-Christophe, aujourd'hui, tu n'es plus parmi nous. Je me souviens avec une grande nostalgie de notre première rencontre de travail à Montpellier, il y a huit ans, comme si c'était hier. Tu es devenu un précieux ami pas un collègue ni un encadrant. Tu étais un grand passionné de palmiers. J'ai beaucoup appris de toi et je ne t'oublierai jamais. Que ton âme repose en paix.

Mes premiers véritables contacts avec la génétique remontent à mon premier stage à l'IRD de Montpellier avec Bertha LUDENA-PINTAUD, à qui j'exprime toute ma reconnaissance pour son initiation aux techniques de biologie moléculaire. Sa patience et ses connaissances seront toujours pour moi une source d'inspiration. Son amitié et ses conseils ainsi que son soutien dans ma vie personnelle et de recherche ne m'ont jamais fait défaut. Bertha, je te dois beaucoup et j'ai eu la chance de t'avoir à mes côtés.

Yves Vigouroux, ma reconnaissance à votre égard est énormément large. Si j'ai pu finaliser mon article, c'est sans doute avec votre soutien, le plus ferme. Yves a toujours un nouveau test à faire, une nouvelle formule à redémontrer. Bon, même si je n'arrivais pas toujours à le suivre dans ses délires il m'a quand même fait découvrir les statistiques, les analyses de structure, le logiciel R, la génétique....voilà qu'il m'avait légué une base certainement durable. Je n'oublierais pas aussi toute l'aide et facilitation qu'il m'a apportée pour les démarches administratives.

Aux membres du jury :

Je remercie Mme le professeur ABROUS Ouzna d'avoir accepté de présider le jury de cette thèse.

Je remercie Mr le professeur AMIROUCHE Rachid, d'avoir bien voulu accepter d'évaluer ces travaux pour améliorer la qualité de ce manuscrit.

Je voudrais également remercier Mme le professeur BENNACEUR Malika, pour son soutien constant, ses précieux conseils et pour l'ensemble de toutes les discussions scientifiques et autres que nous avons pu échanger, qu'elle trouve ici mes sincères reconnaissances.

A Mr le professeur DJEBBAR Reda, j'exprime toute ma gratitude pour avoir accepté d'examiner cette thèse et de faire partie de mon jury. J'ai eu grand bénéfice à profiter de ses conseils, savoir et expérience.

Mes vifs remerciements s'adressent à Mr le professeur BRAHIM-ERRAHMANI, qui malgré ses multiples occupations s'est déplacé pour examiner ce travail.

J'exprime toute ma reconnaissance et mes remerciements à Mr Hassen GHERBI, chercheur à l'IRD Montpellier, pour ses conseils et suggestions ainsi que son grand soutien. Qu'il soit assuré de ma profonde et respectueuse gratitude. Son encouragement m'a grandement aidé tout durant ces années.

Je tiens particulièrement à remercier et exprimer également toute ma gratitude à mes amis chercheurs Abdou-Aziz SAIDOU et Issaka SALIA OUSSEINI (IRD) pour leur soutien, leur disponibilité sans pareil et leur aide pour les analyses de bioinformatique.

SNV L1, l'équipe était pour moi une source de motivation pédagogique et de convivialité, Sabrina BOUDJEMLINE, Kerima BOULAHIA, Hassina MEGUELATI, Salima BAHAR, Ouahiba BELOUI, Souad SMAI, Saliha OUALI, Karima LOTMANI, Dalila BOUBETRA, Assia CHAKER, Hamida SADJI, Baya BOUCENNA, Nassima BAHA, Tassadit AIT KETTOUT, Malha AMGOUD, Nawel MORZOUGLAL, Akila KOROGHLI, Toufik MEFTAH, DERRICHE, Smain TRABSI, Samir DJOUADI, Amine CHAFAI, Faiza BELAID, Dalila TOUAM, Soraya SADAoui, Siham OSMANE, Zakia SELMANI, Radia BOUSSOUAR, Djahida BOUAOUNE, Assia KAMEL, Reda AOUAMEUR, Hassiba HARANI, Kahina CHABANE, Fouzia BEDDOU, Fatima ABDELKRIM, Fatima GRIM, Yasmine BELLABES NABI, Lilia BOUCELEHA, Amel BENSLAMA, Saida OUAFI, Rabea GACEB, OMARI, KANDSI et Houria OUZNADJI. Trouvez tous, ici, mon expression amicale et ma joie d'avoir passé des moments agréables avec vous. Merci à tous mes collègues qui ont assuré mes heures d'enseignement pendant toutes les périodes de mes stages à l'IRD.

Je ne saurais oublier d'adresser mes vifs remerciements à mes amies particulièrement Sabrina B, Lilia B, Yasmine B, Houria B, Salima B, Faiza B, Fouzia B, Amel B, Hassina M, Radai B, Siham M, Kerima B, Samia K, Lila D, Lila K, qui en me témoignant leur profonde amitié et leur attention, m'ont soutenu à chaque instant.

J'adresse mes remerciements à mes enseignants professeurs de l'USTHB, notamment Mme OUAFI S, BOUHIREL L, CHABANE D, AMIROUCHE N et AID F, pour tout leur soutien et leurs encouragements.

Un merci à tous les membres du laboratoire de LRZA qui ont su créer une bonne ambiance de travail en particulier Karima MAFOUFI, Karima OUALI, Cherifa SELMANI, Amina LAHMISSI, Hafida AMEDJKOUH, Djamila YATTA, Fatiha LEBTAHI et Ammar.

Mes remerciements s'adressent à Mr SALAMANI, Riad BOUBENIA, Samir FOURAR pour leur aide.

J'exprime également tous mes remerciements à tous les membres de l'équipe DYNADIV et Palmiers de l'IRD, Yves VIGOUROUX, Cédric MARIAC, Leila ZERKAOUI, Muriel GROS-BALTHAZARD, Thoma COUVREUR, Nora SCARCELLI, Marie C, Adéline B, Jérémy C, Nathalie CHABRILLANGE, Emira CHERIF et Frédérique ABERLENC-BERTOSSI, pour la bonne entente et leur collaboration durant mes formations passées à l'IRD. Chaque année et pendant mes stages à l'IRD, vous avez eu l'amabilité d'écouter mes exposés, apprécier mon travail et mes résultats, et poser les questions qui font avancer. Vos expériences m'ont très largement enrichi, orienté, et surtout canalisé. Merci pour toutes ces discussions, pour ces apports très significatifs dans ma formation, et pour les encouragements renouvelés à chaque fois.

Mes remerciements s'adressent également aux collègues et amis du projet MeRCI : Claudio LITTARDI et Marco BALLARDINI du Centro Studi e Ricerche perle Palme – Sanremo (Italie), Selwa ZEHDİ et Emira CHERIF de l'Université de Tunis, pour leur soutien et l'intérêt qu'ils ont porté à ce travail.

Merci aux collègues doctorants et stagiaires à l'IRD, pour les moments passés ensemble et le soutien mutuel : Kheira ROUANE, Julissa RONCAL, Summar NAQVI, Boubaker, Abdou-Aziz SAIDOU, Chouaib MEZIADI, Meriem, Julie ORJUELA, Andres GUTIERREZ, Karina CASTILLO, Hicham HAOUANE, Maria José SANIN, Maria Fernandez ALARCON, Ai My LUONG, Kim FOUYONTPHANICH...et tous les autres.

Je me dois de rendre hommage aux phoeniciculteurs algériens. Tous ont gentiment accepté de collaborer avec moi, passant parfois des heures à répondre à mes questions et n'hésitant jamais à me fournir le matériel végétal nécessaire à ce travail.

Je ne dois pas oublier M. Hamid CHEMLAL qui a été de toutes mes missions dans tout le Sud du pays et qui par ses nombreuses connaissances du terrain et des hommes m'a grandement facilité mes récoltes.

Bien sûr je n'oublie pas ma famille, toute ma fierté et ma joie, qui m'a toujours beaucoup soutenue et m'a donné le courage de continuer quand je n'avais pas le moral, en particulier ma mère et ma sœur qui m'ont donné beaucoup d'amour et surtout leurs prières. Je dédie cette thèse et tout le travail qu'elle représente à mon père ALLAH Yarahmou.

LISTE DES ABREVIATIONS

% : Pourcentage

°C : Degrés Celsius

ABI : Applied Biosystems

ADN : Acide Désoxyribo Nucléique

AFLP : Amplified Fragment Length Polymorphism, Polymorphisme de Longueur des Fragments Amplifiés

FAO : Food and Agriculture Organization, organisation des Nations Unies pour l'alimentation et l'agriculture

h : heure

IRD : Institut de Recherche pour le Développement

ISSR : Inter Simple Sequence Repeats

l : litre

min : minute

ml : millilitre

ng : nanogramme

pb : Paire de base

PCR : Polymerase Chain Reaction, Technique de polymérisation en chaîne

qsp : Quantité suffisante pour

RAPD: Random amplified polymorphic DNA, Amplification Aleatoire d'ADN polymorphe

sec : seconde

SNP : Single Nucleotide Polymorphism

SSRs : Simple Sequence Repeat

µl : microlitre

LISTE DES FIGURES

Figure 1 : Evolution spatio-temporelle et diversification des palmiers	6
Figure 2 : La carte du génome chloroplastique du palmier dattier.	38
Figure 3 : Distribution géographique du palmier dattier en Algérie.	40
Figure 4 : Position de cinq gènes sur la carte du génome chloroplastique de <i>Phoenix dactylifera</i> L.	52
Figure 5 : Séparation sur gel d'agarose (2.5%) des produits d'amplification par PCR de l'ADN des 24 variétés de <i>P. dactylifera</i> en utilisant les <i>psbA-trnH</i> et <i>rps15-ycf1</i>	54
Figure 6 : Carte de distribution du genre <i>Phoenix</i>	57
Figure 7 : Localisation du minisatellite (12pb) sur l'espaceur <i>psbZ-trnfM</i>	60
Figure 8 : Localisation du minisatellite (20pb) sur le locus <i>rpl16-rps3</i>	61
Figure 9 : Séparation et détection des fragments d'ADN chloroplastique (<i>trnG-trnfM</i>) à l'aide du système QIAxcel. Alignement de la séquence <i>trnG-trnfM</i> des espèces de <i>Phoenix</i> étudiées.	62
Figure 10 : Phylogénie du réseau d'haplotypes pour 301 échantillons représentant 14 espèces du genre <i>Phoenix</i> basées sur la combinaison des données de génotypage des locus <i>trnG-trnfM</i> et <i>rpl16-rps3</i>	65
Figure 11 : Histogramme illustrant les fréquences alléliques obtenues pour chaque locus microsatellite.	70
Figure 12 : Arbre phylogénétique des espèces basé sur la distance des allèles partagés (DAS) de 18 loci microsatellites pour l'ensemble des espèces du genre <i>Phoenix</i>	73
Figure 13 : Séparation sur gel d'agarose (2.5%) des produits d'amplification par PCR de l'ADN des 24 variétés de <i>P. dactylifera</i> en utilisant les <i>rps15-ycf1</i> et <i>matk</i> exon. Le locus <i>matk</i> exon n'a pas été amplifié pour toutes les variétés.	79
Figure 14 : Locus <i>atpH-atpI</i> : Indel/Substitution/Homopolymère.....	82
Figure 15 : Locus <i>rps15-ycf1</i> : Inversion (TTC/GAA).	82
Figure 16 : <i>rpoB</i> locus monomorphe à la position 1881.....	83
Figure 17 : Evolution des chloroplastes des espèces du genre <i>Phoenix</i>	85
Figure 18 : Bayesian cluster analysis using the STRUCTURE program: results for K = 3. (A) Plots of (a) maximum log likelihood over the 10 runs and (b) delta k from the structure analysis was calculated according to the method of Evanno et al.	

(B) Estimated population structure inferred from all the individual date palms for $K = 3$. Each individual is represented by a thin vertical line divided into K colored segments representing the fraction of the individual's estimated membership of the K clusters. Pie charts show the frequencies of the haplotypes belonging to the three structure groups.	
(C) Geographic distributions of the 10 oases and two haplotypes. The pie chart shows the proportions of haplotypes in each oasis.	
(D) Sampling location of the date palms. Pie charts show the proportion of membership of each sampled population inferred by structure analysis ($K = 3$).....	97
Figure 19 : Neighbor-joining tree with microsatellite genotypes using shared allele distance. The neighbor-joining tree shows the genetic relationships among the date palm genotypes included in the study, where each branch represents a single individual. The individuals in the population STRUCTURE1 are in red, those in population 2 in green, and those in population 3 in blue.	99
Figure 20 : Fonctionnement du système d'identification du dattier à 7 marqueurs.	106
Figure 21 : Clé d'identification de 155 génotypes de palmier dattier (<i>Phoenix dactylifera</i>), basée sur six marqueurs microsatellites nucléaires et un marqueur minisatellite chloroplastique.	117
Figure 22 : Résultat de l'approche bayésienne de regroupement avec l'identification de la valeur de K la plus vraisemblable suivant la méthode du Delta K en fonction du nombre possible de sous-populations.	131
Figure 23 : Résultat de l'approche bayésienne de regroupement avec l'assignement de 481 individus suivant leur localisation géographique pour un nombre de sous-population de $K = 2$ à $K = 6$. Chaque couleur représente une sous-population génétique, les lignes verticales représentent les individus. Un individu défini par plusieurs couleurs, est considérée comme ayant des ancêtres issus de plusieurs sous-populations génétiques. Les groupes de palmier dattier sont séparés par une ligne pointillée blanche. K mesure l'indice de stabilité d'Evanno <i>et al.</i> (2005).	133
Figure 24 : Arbre Neighbor Joining des groupes géographiques basé sur les valeurs de distance génétique D_{AS} , ainsi que la répartition des deux clusters génétiques définis par l'analyse STRUCTURE à $K = 2$ au sein de chaque pays. Le cluster occidental en bleu et le cluster oriental en rouge.	137
Figure 25 : Analyse discriminante des composantes principales (DAPC) sur les données SSR des génotypes mondiaux.	139

Figure 26 : Analyse discriminante des composantes principales (DAPC) sur les données SSR des géotypes occidentaux.....	140
Figure 27 : Distribution des deux chlorotypes du palmier dattier au sein des 481 accessions selon leur origine géographique.....	142

LISTE DES ENCADRES

Encadré 1 : Diversité génétique : définition et intérêt	15
Encadré 2 : Les principaux modèles mutationnels des microsatellites.....	20
Encadré 3 : Différents types de mutations de séquence	31
Encadré 4 : Polymerase chain reaction (PCR).....	46
Encadré 5 : Statistiques et paramètres de diversité génétique	48
Encadré 6 : Définitions des termes ‘cultivar’ et ‘landrace’	120

LISTE DES TABLEAUX

Tableau 1 : Génomes chloroplastiques entièrement séquencés des monocotylédones.....	37
Tableau 2 : Caractéristiques des 18 loci microsatellites de palmier dattier (<i>Phoenix dactylifera</i> L.)	43
Tableau 3 : Caractéristiques des locus utilisés pour l'étude du polymorphisme chloroplastique	53
Tableau 4 : Classification de la sous-famille <i>Coryphoideae</i>	58
Tableau 5 : Répartition des haplotypes observés dans les régions <i>trnG-trnfM</i> et <i>rpl16-rps3</i> pour les espèces du genre <i>Phoenix</i>	63
Tableau 6 : Paramètres descriptifs de la diversité génétique révélés par les 18 loci microsatellites analysés sur 301 génotypes de palmiers.	67
Tableau 7 : Diversité génétique moyenne sur tous les loci, par espèce.	72
Tableau 8 : Sept polymorphismes au niveau des quatre loci polymorphes.	81
Tableau 9 : Substitutions au niveau des loci <i>rpoB</i> et <i>atpF</i> intron.	83
Tableau 10 : Les états plésiomorphes et apomorphes du polymorphisme chloroplastique.	86
Table 11 : Descriptive genetic parameters for 18 microsatellite loci analyzed on 192 individual date palms.	93
Table 12 : Summary statistics for 18 microsatellite loci in 10 oases.	94
Table 13 : Genetic distance between oases (F_{ST}).	95
Table 14 : Description and variability parameters of the 18 polymorphic microsatellite loci in the three STRUCTURE populations.	98
Table 15 : Genetic distance between Structure populations (F_{ST}).....	98
Tableau 16 : Synonyme dénomination variétale des pays.	107
Tableau 17 : Synonyme dénomination variétale des oasis algériennes.	108
Tableau 18 : Diversité génétique des 481 accessions de dattier au niveau de 18 loci SSR.....	130
Tableau 19 : La diversité génétique du palmier dattier dans 13 pays pour les 18 SSR microsatellites.....	135

SOMMAIRE

<i>Résumé</i>	<i>i</i>
<i>Remerciements</i>	<i>ii</i>
<i>Liste des abréviations</i>	<i>iii</i>
<i>Liste des figures, des tableaux et encadrés</i>	<i>iv</i>

INTRODUCTION GENERALE	1
------------------------------------	----------

SYNTHESE BIBLIOGRAPHIQUE

I. Présentation du palmier dattier	5
1. Position systématique	5
2. Le genre Phoenix.....	7
3. Origine du palmier dattier	8
4. Répartition géographique du palmier dattier.....	9
5. Travaux réalisés sur le palmier dattier	10
II. La diversité génétique	12
1. Le polymorphisme génétique	12
2. Les forces évolutives	12
2.1. Les mutations	12
2.2. La dérive génétique	13
2.3. La migration	14
2.4. La sélection naturelle	14
3. Les marqueurs génétiques d'estimation de la structuration génétique.....	16
3.1. Les marqueurs phénotypiques	16
3.2. Les marqueurs biochimiques.....	16
3.3. Les marqueurs moléculaires.....	16
3.3.1. Les marqueurs RFLP.....	17
3.3.2. Les marqueurs RAPD.....	17
3.3.3. Les marqueurs AFLP	18
3.3.4. Les marqueurs minisatellites	19
3.3.5. Les marqueurs microsatellites, des témoins de l'histoire génétique récente.....	19
- Définition et modèles mutationnels	19
- Applications	21
3.3.6. Les marqueurs SNP (Polymorphisme de simple nucléotide).....	21
4. Caractérisation et mesures de la diversité génétique.....	21
4.1. Diversité intra population.....	22
4.1.1. Nombre moyen d'allèles par locus (A)	22
4.1.2. Taux de polymorphisme (PIC).....	22
4.1.3. Taux d'hétérozygote observé (Ho).....	22

4.1.4. Taux d'hétérozygote attendu (H_e).....	22
4.1.5. Indice de Fixation F_{IS}	23
4.1.6. L'équilibre de Hardy-Weinberg	23
4.2. Diversité inter population.....	24
4.3. Structuration spatiale de la diversité génétique.....	25
5. Détection de la structure génétique	26
5.1. Analyse factorielle des correspondances.....	26
5.2. Analyse en composantes principales.....	26
5.3. Arbre phylogénétique.....	26
5.4. Analyse Bayésienne pour inférence de la structure génétique des populations.....	27
III. Généralités sur le génome chloroplastique	28
1. Fonctions biologiques des chloroplastes	28
2. Origine et caractéristiques des chloroplastes	29
3. Hérité des chloroplastes chez les Angiospermes.....	29
4. Caractéristiques de l'ADN chloroplastique.....	30
5. Taille du génome chloroplastique	32
6. Les gènes chloroplastiques.....	32
6.1. Les gènes impliqués dans la photosynthèse	33
6.2. Les gènes impliqués dans l'expression génétique.....	34
7. Intérêt et développement des marqueurs chloroplastiques.....	34
8. Le génome chloroplastique des monocotylédones et du palmier dattier en particulier	36

MATERIEL & METHODES

1. Matériel végétal et zones d'étude	39
2. Extraction et quantification d'ADN.....	41
3. Génotypage microsatellite.....	41
3.1. Choix des marqueurs génétiques moléculaires	41
3.2. Marqueurs microsatellites utilisés.....	41
3.3. Réaction d'amplification en chaîne de l'ADN microsatellite nucléaire et génotypage	45
3.4. Diversité et analyse statistique	47
3.4.1. Paramètres de diversité génétique	47
3.4.2. Isolement par la distance	49
3.4.3. Analyse discriminante en composante principale (DAPC).....	49
3.4.4. Arbre phylogénétique et analyse Bayésienne de regroupement.....	49
4. Polymorphisme des séquences de loci chloroplastiques.....	50
4.1. Génotypage du minisatellite <i>psbZ-trnfM</i>	50
4.2. Séquençage des amorces chloroplastiques	51
4.2.1. Définition des locus chloroplastiques	51
4.2.2. Réaction d'amplification en chaîne (PCR).....	54
4.2.3. Séquençage des produits PCR.....	55
4.2.4. Alignement et analyse des séquences.....	55

RESULTATS ET DISCUSSIONS

Chapitre 1 : Caractérisation génétique et phylogénie du genre *Phoenix*

Introduction.....	56
Résultats.....	60
1. Polymorphisme chloroplastique des locus <i>trnG-trnfM</i> et <i>rpl16-rps3</i> chez les espèces du genre <i>Phoenix</i>	60
2. Phylogénie du chloroplaste... ..	64
3. Diversité génétique globale et puissance des microsatellites.....	66
4. Relations génétiques entre les espèces de <i>Phoenix</i>	72
Discussion	74
a. Polymorphisme chloroplastique des locus <i>trnG-trnfM</i> et <i>rpl16-rps3</i>	74
b. Phylogénie du chloroplaste	75
c. Structuration génétique des espèces de <i>Phoenix</i>	76

Chapitre 2 : Apport des séquences chloroplastiques à la recherche de chlorotypes de palmier dattier

Introduction.....	78
Résultats.....	79
1. Mise au point des loci chloroplastiques et séquençage	79
2. Caractère ancestral et dérivé du dattier cultivé	84
Discussion	87
a. Polymorphisme au sein du génome chloroplastique de <i>Phoenix dactylifera</i>	87
b. Caractère ancestral et dérivés du dattier cultivé.....	89

Chapitre 3 : Diversity of Algerian Oases date palm (*Phoenix dactylifera* L., *Arecaceae*): heterozygote excess and cryptic structure suggest farmer management had a major impact on diversity

Introduction.....	91
Results	92
1. Chloroplast minisatellite and nuclear SSR genotyping.....	92
2. Genetic differentiation between populations.....	94
3. Variety name and diversity	95
4. Population structure and differentiation.....	96
5. Genetic distance and dendrogram construction	98
Discussion	100
a. Genetic diversity and population structure of date palm.....	100
b. Explanation for the genetic history of diffusion of date palm cultivars in Algeria.....	101

Chapitre 4 : Constitution d'une clé d'identification de variétés de palmier dattier

Introduction.....	103
Résultats.....	105
Discussion	118
a. Concepts de variété, de cultivar.....	118
b. Etablissement de la clé d'identification variétale.....	121
c. L'efficacité de la clé d'identification.....	124

Chapitre 5 : Diversité génétique du palmier dattier dans l'ancien monde : Origine et processus de domestication

Introduction.....	126
Résultats.....	129
1. Polymorphisme des marqueurs microsatellites	129
2. Analyse Bayésienne de regroupement	131
3. Diversité et structure génétique entre les groupes géographiques du palmier dattier	134
4. Diversité génétique du minisatellite chloroplastique	141
Discussion	143
a. Structure et diversité génétique de l'ADN nucléaire et chloroplastique du palmier dattier.....	143
b. Origines géographiques et Histoire génétique du palmier dattier cultivé indiquent deux centres de domestication indépendants	145

CONCLUSION GENERALE.....	148
---------------------------------	------------

REFERENCES BIBLIOGRAPHIQUES	151
--	------------

ANNEXES

INTRODUCTION GENERALE

INTRODUCTION GENERALE

Le palmier dattier (*Phoenix dactylifera* L.) est l'un des plus anciens arbres fruitiers cultivés, et est intimement lié à l'histoire de la migration humaine. C'est une espèce monocotylédone de la famille des Arecaceae qui comprend également d'autres espèces telles que le cocotier et le palmier à huile commercialement importantes.

Le genre *Phoenix* comprend 14 espèces (Barrow 1998) qui sont endémiques dans diverses régions s'étendant des îles de l'Afrique de l'Ouest à l'Asie de l'Est. La dioécie et l'hybridation entre les espèces de *Phoenix* dans les zones de sympatrie (Gros-Balthazard *et al.* 2013) offre la possibilité d'une grande diversité génétique au sein du genre.

Le palmier dattier est actuellement la principale culture des zones arides et semi-arides d'Afrique du Nord et du Moyen-Orient, avec une estimation de plus de 3000 cultivars importants pour la production de fruits (dattes) dans le monde, dont environ 60 sont largement cultivés et ont d'importants marchés nationaux et internationaux (Chao & Krueger 2008).

La production mondiale du fruit (la datte) atteint près de 8 millions de tonnes, générant chaque année des millions de dollars américains en faveur des économies locales et nationales (FAOstat). En effet, la production du palmier dattier est passée de la culture traditionnelle dans des agrosystèmes riches et diversifiés à une monoculture intensive (Jain *et al.* 2011 ; Bouguedoura 2012). Ce développement a entraîné une érosion génétique sévère, avec la perte de cultivars et l'appauvrissement de l'agrobiodiversité du palmier dattier (Jain *et al.* 2011). En outre, le palmier dattier subit un stress biotique et abiotique, ce qui a permis de mettre en place des programmes de sélection des variétés tolérantes. De ce fait, la conservation des ressources génétiques du palmier dattier est devenue une question cruciale pour la stabilité des zones désertiques et semi-désertiques.

En Algérie, le palmier dattier conserve une diversité importante, plus d'un millier de cultivars traditionnels sont recensés (Benkhalifa *et al.* 2004). Cependant la reconnaissance des cultivars est toujours traditionnellement basée sur des paramètres morphologiques de l'appareil végétatif et les caractères organoleptiques du fruit. De plus, sur l'ensemble des cultivars dénombrés en Algérie, peu d'entre eux présentent une importance économique réelle, tel que la variété Deglet-Nour. Les autres sont rares et produisent des dattes de faible valeur marchande. A cet effet, beaucoup de variétés inventoriées sont menacées de disparition car elles sont âgées et ne produisent plus de rejets. Cette approche a des limites car d'une part les jeunes agriculteurs ne maîtrisent plus les données morphologiques et d'autres part ces

données varient avec l'environnement ce qui se traduit par de nombreuses synonymies et homonymies. C'est pourquoi les scientifiques ont orienté leurs travaux sur d'autres marqueurs plus stables et plus performants.

L'analyse de la diversité génétique du palmier dattier est aujourd'hui abondamment étudiée. Ainsi, de nombreux marqueurs biochimiques tels que les flavonoïdes, et moléculaires tels que les RFLP (Restriction Fragment Length Polymorphism), les RAPD (Random Amplified Polymorphic DNA), les AFLP (Amplified Fragment Length Polymorphism) ou encore les SSR (Simple Sequence Repeat) sont exploités pour discriminer d'une façon fiable les variants génotypiques chez le palmier dattier. Parmi ces marqueurs, les microsatellites ou SSR sont réputés être les plus performants pour les analyses de la diversité génétique et la structuration géographique des espèces.

Bien que l'histoire de la domestication de la palmeraie au Moyen-Orient soit maintenant connue, grâce à la récolte de fruits de type sauvage datés du 6ème millénaire av. J.-C., suivie de l'établissement de la culture dans le sud de la Mésopotamie au 5ème millénaire av. J.-C. et, par la suite la propagation de la technologie de l'agriculture d'oasis autour du Golfe Persique vers la fin du 4ème millénaire av. J.-C. (Tengberg 2012). Cependant, l'origine et le développement de la culture du palmier dattier dans le Sahara restent peu connus. Car, contrairement au bilan archéologique de la palmeraie au Moyen-Orient (Tengberg 2012), il n'y a pratiquement aucune preuve d'une culture significativement ancienne de la datte au Sahara. Les peintures murales et les pétroglyphes des montagnes sud-centrales du Sahara comme le Tassili n'Ajjers et Ahaggar (Algérie), Oua n-Moulin (Libye) représentent rarement les palmiers dattiers, et toutes ces représentations appartiennent à la période "Cabaline", au 1er millénaire av. J.-C. (Lhote 1964 ; Camps 1989). Étant donné qu'il existe un enregistrement archéologique très ancien, bien que fragmentaire, au Sahara, il est probable que le développement précoce des établissements humains dans cette région ne dépendait pas de la culture du palmier dattier. En effet, la preuve de la technologie de la céramique datant du 10ème millénaire av. J.-C. au site Tagalagal (Aïr, nord du Niger) suggère l'utilisation d'autres ressources, en particulier des céréales, au moment où les conditions climatiques sont particulièrement favorables dans le sud du Sahara (Roset 1982). Même en Égypte, qui est de loin la zone la mieux documentée du Sahara, il n'y a aucune preuve de palmier dattier domestique dans un contexte d'oasis avant le 14ème siècle av. J.-C. (Terral *et al.* 2012). Cet écart chronologique de près de deux millénaires avec le Moyen-Orient est très intrigant, compte tenu du développement historique et de la propagation de l'agriculture dans le

Croissant Fertile (Zeder 2011 ; Zohary *et al.* 2012). On ne peut pas écarter que ce diachronisme dans le développement de l'agriculture oasisienne du palmier dattier soit un artefact d'un bilan archéologique incomplet en Afrique. Cependant, il existe de fortes preuves que les antécédents génétiques de la domestication du palmier dattier en Afrique du Nord et au Moyen-Orient soient distincts. En particulier, plusieurs études basées sur le génotypage des SSR nucléaires (Arabnezhad *et al.* 2012 ; Zehdi *et al.* 2012 ; 2015 ; Pintaud *et al.* 2013) ont montré qu'il existe des pools de gènes distincts dans les régions de l'Ouest et de l'Est de la ceinture de culture traditionnelle du palmier dattier. Il a été démontré que le pool de gènes de l'Ouest existe dans son état pur dans la Méditerranée occidentale et la région macaronésienne (Murcia en Espagne, les îles Gabès en Tunisie, les îles du Cap-Vert) et le pool de gènes Oriental typiquement en Oman et aux Émirats arabes unis (Gros-Balthazard *et al.* 2013 ; Pintaud *et al.* 2013). Les deux principaux pools de gènes (Ouest et Est) sont les plus facilement suivis par le génome du chloroplaste à hérédité maternelle. Les génomes plastidiaux correspondant aux pools de gènes primaires de l'Ouest et de l'Est sont nommés "Occidental" et "Oriental", respectivement. Ces génomes peuvent être facilement identifiés par chlorotypage à un locus de minisatellites dodécanucléotidiques situé dans l'espaceur intergénique *trnG-trnfM*. Le chlorotype Occidental est caractérisé par trois répétitions du motif de 12 pb (chlorotype 3), tandis que le chlorotype Oriental est de type 4 (Moussouni *et al.* 2017). Ce polymorphisme est l'accès le plus simple à la différenciation des génomes des chloroplastes Occidentaux et Orientaux, avec de grandes conséquences sur la compréhension de l'histoire de la domestication de cette espèce.

Il a été intéressant d'étudier la diversité génétique nucléaire et chloroplastique du palmier dattier en Algérie, au carrefour des influences occidentales et orientales, avec un échantillonnage génétique très important et une bonne couverture géographique. Ainsi, l'Algérie est considéré comme l'un des pays de la ceinture aride Saharo-Sindienne ayant une structure génétique de l'agrobiodiversité du dattier qui montre un mélange significatif entre les pools de gènes de l'Ouest et de l'Est, ainsi qu'une diversité génétique exceptionnellement élevée.

Le premier objectif de cette thèse a été de caractériser génétiquement l'agrobiodiversité du palmier dattier en Algérie, à partir des données de génotypage et séquençage des marqueurs nucléaires et chloroplastiques ; en comparant la structuration génétique obtenue avec les marqueurs nucléaires (à transmission biparentale) avec ceux obtenus avec les marqueurs chloroplastiques uniparentaux.

Dans le cadre d'un projet AUF impliquant plusieurs pays phoenicicoles, il s'est avéré important d'appréhender le premier objectif dans celui global de l'étude de la domestication du palmier dattier, de comprendre la diffusion de la phoeniciculture, de distinguer le palmier dattier par rapport aux autres espèces de *Phoenix* ce qui permettra d'identifier les plus proches parents du dattier par le biais d'analyses phylogénétiques.

Pour répondre à ces objectifs, notre travail a été organisé en trois parties :

- La première partie est consacrée à une synthèse bibliographique portant sur l'état des connaissances du palmier dattier et la diversité génétique nucléaire et chloroplastique.
- Dans une deuxième partie, nous décrirons le matériel ainsi que les différentes techniques et approches utilisées au cours de ce travail.
- La troisième partie qui sera consacrée à la présentation des résultats obtenus, est articulée en cinq chapitres discutés et conclus :

- ❖ Le **chapitre 1** vise à caractériser phylogénétiquement le palmier dattier au sein du genre *Phoenix* afin d'identifier ses plus proches parents.
- ❖ Dans le **chapitre 2**, nous présenterons les apports de l'analyse du génome chloroplastique. L'étude du polymorphisme des séquences chloroplastiques a pour but de trouver les chlorotypes présents chez le palmier dattier.
- ❖ Le **chapitre 3** concernera la diversité génétique du palmier dattier en Algérie. Ce chapitre s'appuie sur la combinaison de données nucléaires et chloroplastiques et permet de cerner le rôle de la sélection humaine dans le maintien de la diversité variétale du palmier dattier en Algérie.
- ❖ Dans le **chapitre 4** nous exposerons les résultats de la clé d'identification des variétés algériennes du palmier dattier.
- ❖ Le **chapitre 5** consiste en une étude génétique détaillée de la diversité du palmier dattier à travers sa zone de culture traditionnelle dans l'Ancien Monde. L'objectif de cette étude est de fournir de nouvelles idées sur l'influence des origines géographiques et de l'activité humaine sur la structure génétique du palmier dattier.

Une conclusion générale et des perspectives de travail permettent de clore ce manuscrit en ouvrant à de nouvelles recherches.

SYNTHESE BIBLIOGRAPHIQUE

I. Présentation du palmier dattier

Le palmier dattier (*Phoenix dactylifera* L.) est une plante monocotylédone dioïque cultivée dans les zones arides et semi-arides. La description de la plante a été réalisée par Linné (1734) qui lui donne son appellation actuelle. Le nom *Phœnix* viendrait de l'antiquité qui le considérait comme l'arbre des Phéniciens. *Dactylifera* vient du latin *dactylus* (doigts); qui fait référence aux fruits du palmier sous forme de doigts et du latin *fero* signifiant « je porte ». En arabe, il est appelé *nakhil*, *date palm* en anglais. La culture du palmier dattier est associée aux premières civilisations urbaines et agricoles. Il a joué un rôle capital dans l'épanouissement des civilisations des régions arides car il constitue l'armature de ces îlots de verdure que sont les oasis. Son expansion est actuellement quasi planétaire et son utilisation universelle.

1. Position systématique

Le palmier dattier appartient à la famille des *Arecaceae* (anciennement *Palmaceae*) comprenant 183 genres et plus de 2500 espèces, ce qui en fait l'une des plus importantes des monocotylédones et une des plus anciennes des Angiospermes (Dransfield *et al.* 2008 ; Couvreur *et al.* 2011).

Cette famille comporte 15 groupes taxonomiques définis selon des critères morphologiques (foliaires et floraux) et de répartition géographique (Moore 1973). Le genre *Phœnix* fait partie du groupe des palmiers *Phoenicoïdeae*. Ce groupe partage certaines caractéristiques florales, comme le dimorphisme entre les fleurs mâles et femelles des espèces dioïques, avec le groupe des *Coryphoïdeae*. Dans la nouvelle classification de la famille des *Arecaceae* (Asmussen *et al.* 2006), basée aussi bien sur les caractères morphologiques que sur les séquences d'ADN nucléaires, nous retrouvons 5 sous-familles : *Coryphoïdeae*, *Calamoïdeae*, *Nypoïdeae*, *Ceroxyloïdeae* et *Arecoïdeae*. La phylogénie des palmiers selon la théorie de l'horloge moléculaire à l'aide de fossiles a été effectuée par Couvreur *et al.* (2011) sur 183 genres de palmiers actuellement décrits. L'origine et le début de la diversification des palmiers remonte à environ 100 millions d'années du milieu du Crétacé (Figure 1).

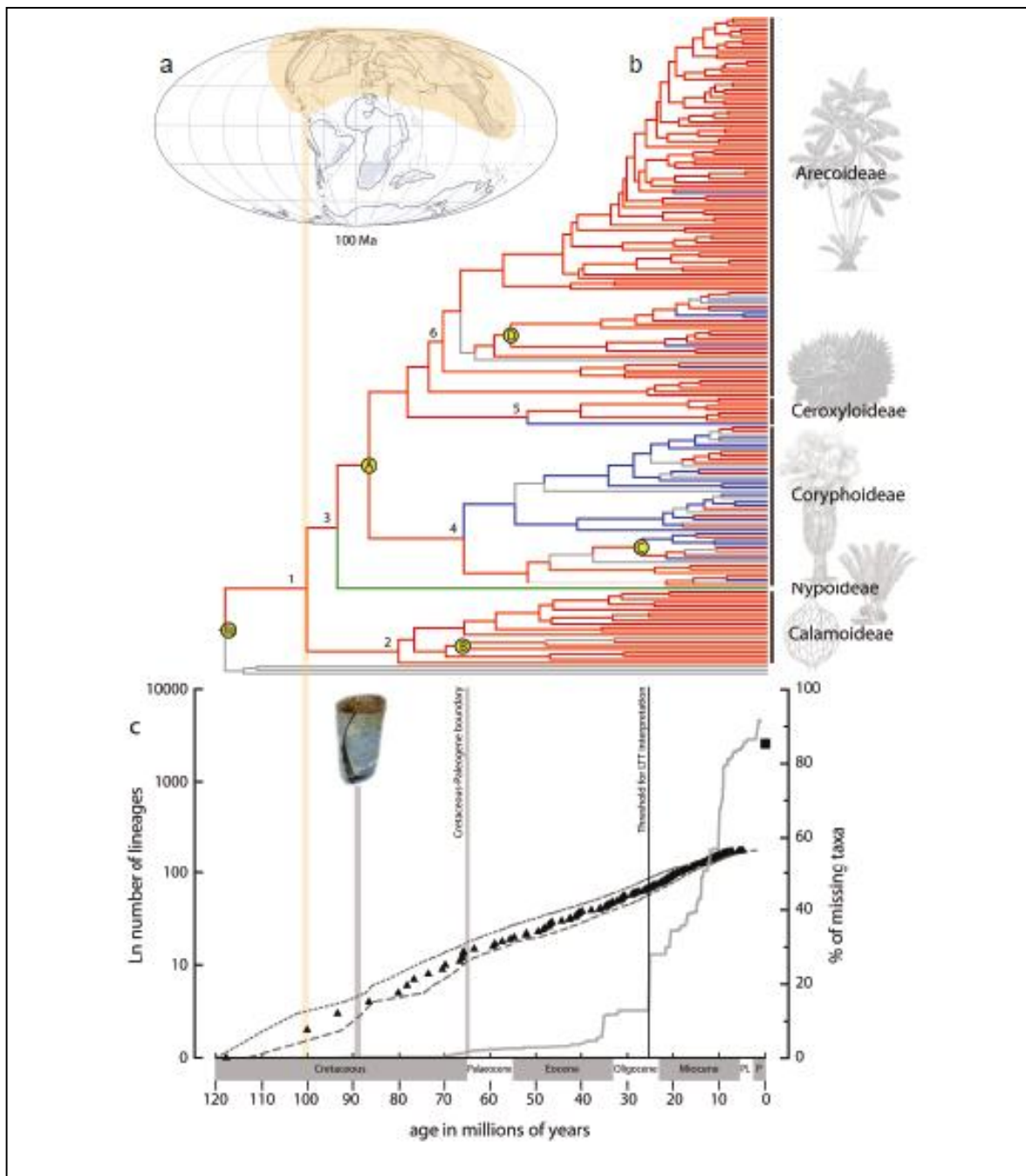


Figure 1 : Evolution spatio-temporelle et diversification des palmiers (Couvreur *et al.* 2011).

Les espèces de genre *Phoenix* qui est le seul genre de la tribu des *Phoenixeae* sont actuellement classées dans la sous-famille des *Coryphoideae* Griffith. Elles sont caractérisées par des longues feuilles pennées dont les folioles de la base sont modifiées en épines. La divergence des *Phoenix* remonterait, selon l'analyse de phylogénie moléculaire de Couvreur *et al.* (2011), au début du Tertiaire.

Les palmiers présentent une très faible diversité dans leur équipement chromosomique. En effet, les valeurs $2n$ les plus représentées sont 26, 28, 32 et 36 chromosomes, mais il existe

une relation entre les nombres chromosomiques et la tendance évolutive des palmiers, puisque le nombre haploïde 18 a été trouvé chez les palmiers les plus primitifs, tels les *Phoenix*, alors que les nombres 17, 16, 15, 14 voire 13 ont été comptés chez les palmiers les plus évolués (Hussein *et al.* 1984). L'adaptation évolutive des palmiers serait donc accompagnée d'une diminution du nombre de chromosomes.

Classification systématique du palmier dattier (Asmussen *et al.* 2006 ; Dransfield *et al.* 2008) :

Ordre : *Arecales*

Famille : *Arecaceae*

Sous-famille : *Coryphoideae*

Tribu : *Phæniceae*

Genre : *Phoenix*

Espèce : *P. dactylifera* L.

L'ordre des *Arecales* comprend cinq autres espèces économiquement intéressantes, le palmier à huile (*Elaeis guineensis* Jacq.), le cocotier (*Cocos nucifera* L.), l'aréquier (*Areca catechu*), le raphia (*Raphia farinifera*) et le palmier pêche (*Bactris gasipaes*).

2. Le genre *Phoenix*

Le palmier dattier appartient au genre *Phoenix*, monotypique de la tribu des *Phæniceae* qui représente un groupe à part au sein des palmiers.

Le genre *Phoenix* se distingue des autres genres de la famille des *Arecaceae* par les caractéristiques suivantes (Ozenda 1958) :

- Les feuilles sont pennatisséquées et la spathe est unique,
- L'ovaire à trois carpelles libres.

Le genre *Phænix* comprend 14 espèces botaniques (Govaerts & Dransfield 2005) avec une large aire de répartition géographique, depuis l'Asie du Sud (Indonésie, Philippines) jusqu'aux îles Canaries en passant par la Chine, l'Inde, le Moyen-Orient et l'Afrique (Moore 1973 ; Dransfield *et al.* 2008).

Toutefois, l'espèce *Phoenix dactylifera* L. se distingue des autres espèces du même genre par un stipe long et grêle et par des feuilles glauques. La majeure partie des espèces de ce genre

est utilisée comme plantes ornementales et/ou comme source de fibres textiles, d'amidon, de sucre ainsi que de chaume.

3. Origine du palmier dattier

L'origine du palmier dattier fait l'objet de plusieurs hypothèses depuis la fin du XIX^{ème} siècle. Les travaux les plus récents font état d'une identification d'un ancêtre sauvage de l'espèce sur la frange méridionale chaude et sèche du Proche-Orient, sur le bord nord-est du Sahara et le nord du désert d'Arabie (Zohary & Hopf 1988). La morphologie et les exigences climatiques de cet ancêtre sont les mêmes que celles du dattier cultivé ; la seule différence réside dans la taille des fruits qui est plus petite avec une pulpe très réduite et indigeste. L'analogie des formes sauvages avec les arbres cultivés les a fait classer par les botanistes avec *Phoenix dactylifera*. L'idée d'un ancêtre sauvage avait déjà été émise par Werth (1933) qui refusait la pluralité des ancêtres en raison de la stabilité des formes des organes floraux du palmier dattier cultivé dans toutes les régions du monde. La domestication du palmier dattier sauvage remonterait vers 6000 ans avant J.C. dans la région de Dalma aux Emirats (Newton *et al.* 2013). Cette domestication serait associée à la naissance des premières civilisations agricoles du croissant fertile, entre la Mésopotamie et l'Égypte, vers 4000 ans avant J.C. il a fait partie de la période Chalcolithique avant l'âge de bronze.

Des vestiges de palmiers dattiers ont été mis en évidence dans le site archéologique de Hili, dans la péninsule d'Oman, qui révèle l'existence d'oasis depuis plus de 3000 ans avant J.C. (Cleuzio & Constantini 1982). Son origine remonterait aux dernières périodes pluviales du début de l'ère tertiaire et se situerait probablement en Europe Méridionale. Les auteurs ont surtout montré, compte tenu de ses exigences hydriques (1200 à 1500 mm par an), que le palmier dattier n'est pas originaire de zones désertiques mais de régions à climat tropical humide (Jahiel 1996). La domestication et la sélection naturelle ont amélioré la qualité et la taille du fruit.

L'histoire de l'origine géographique et de la domestication du palmier dattier *Phoenix dactylifera* est restée longtemps obscure, dont le principal problème est qu'il est difficile de démontrer si l'espèce existe vraiment à l'état sauvage.

Trois hypothèses principales sont proposées pour expliquer l'origine du *Phoenix dactylifera* (Munier 1973 ; Zohary & Spiegel-Roy 1975 ; Barrow 1998 ; Rossignol-Strick 2003) :

1- Le palmier dattier, une plante strictement cultivée, provient d'une ou de plusieurs formes sauvages, de la même espèce, plus ou moins passées dans les formes cultivées.

2- Le palmier dattier, provient de l'hybridation de plusieurs *Phoenix* existant encore dans le voisinage de son aire de dispersion, en particulier *Phoenix sylvestris*, *P. atlantica*, *P. canariensis* et *P. reclinata*.

3- Le palmier dattier a existé à l'état sauvage en tant qu'une espèce distincte, il provient d'une hybridation entre des dattiers sauvages et une autre espèce du genre *Phoenix*.

Une analyse génétique basée sur des marqueurs microsatellites nucléaires et un minisatellite chloroplastique a récemment réfuté les deux dernières hypothèses (Pintaud *et al.* 2010). En effet, le profil allélique du dattier apparaît fortement divergeant des autres *Phoenix* indiquant que *P. dactylifera* est une espèce distincte qui a été domestiquée indépendamment des autres (Pintaud *et al.* 2010).

4. Répartition géographique du palmier dattier

Le palmier dattier est cultivé historiquement dans les zones arides et semi-arides chaudes du Moyen et du Proche Orient, de l'Afrique du Nord et d'une partie du Nord- Ouest de l'Inde et du Pakistan.

Au Moyen-Orient, les plus importants peuplements de palmiers dattiers se rencontrent en Irak, en Iran, en Arabie Saoudite, au Yémen et aux Emirats Arabes Unis. Le palmier dattier se rencontre aussi, au Pakistan, au Liban, en Palestine.

En Afrique, cette espèce est très anciennement cultivée dans la région périméditerranéenne, depuis l'Atlantique à l'Ouest jusqu'en Egypte à l'Est. Les principales régions productrices se situent au sud du Maroc, au sud de l'Algérie, la Tunisie, la Libye, l'Egypte, au nord et au centre de la Mauritanie (Munier 1973). Aussi, dans les aires sahariennes du Mali, du Cameroun, du Niger, du Tchad, du Soudan, en Somalie, en Ethiopie et à Djibouti. La culture du palmier dattier est aussi pratiquée en Tanzanie et à Madagascar dans des aires analogues de l'hémisphère sud.

Il existe aussi des sites de culture du palmier dattier dispersés dans le monde. Ces sites sont caractérisés le plus souvent par des sujets isolés ou de petits peuplements.

En Europe, le dattier est cultivé dans les rivages européens de la Méditerranée et ceux du secteur méridional de la Péninsule ibérique. En Espagne, on rencontre les plus importants peuplements européens de palmier dattier. L'espèce est surtout cultivée comme plante ornementale notamment sur la Côte d'Azur en France, en Italie, au Portugal et en Grèce.

En Amérique, le palmier dattier est essentiellement cultivé aux Etats-Unis d'Amérique. Les principaux centres de production sont situés en Californie, en Amazonie et au Texas où le palmier fut introduit au 18^{ème} siècle mais sa culture débute réellement vers les années 1900 (Hilgeman 1972). Le palmier dattier se rencontre également, mais en nombre réduit, au Mexique, aux Antilles, en Colombie, au Brésil, en Equateur et en Argentine.

En Australie, les principaux centres de production de dattes se localisent en Queensland et en Australie du nord.

Il a aussi été importé en Nouvelle Calédonie et on le rencontre sporadiquement comme plante d'ornement à la Réunion (Ouennoughi & Dubost 2005).

La présence du palmier dattier n'est pas toujours liée à la pratique de la phoeniculture. L'existence de cette pratique s'est faite progressivement au gré des déplacements de populations et d'échanges commerciaux.

5. Travaux réalisés sur le palmier dattier

Depuis une époque récente, le palmier dattier est l'objet d'importants travaux scientifiques basés sur l'analyse moléculaire, afin de mieux connaître sa génétique et d'élucider l'organisation de son génome. Ainsi, plusieurs marqueurs moléculaires ont été employés avec succès pour étudier sa diversification variétale (Cornicquel & Mercier 1996 ; Billotte *et al.* 2004 ; Zehdi *et al.* 2004, 2006 ; Elshibli & Korpelainen 2009 ; Pintaud *et al.* 2010).

En Algérie, les travaux de Bennaceur *et al.* (1991) basés sur des systèmes enzymatiques ont été d'un grand apport dans la discrimination des cultivars de palmier dattier. De même, en se basant sur la transférabilité des marqueurs microsatellites du cocotier (*Cocos nucifera* L) et du palmier à huile (*Elaeis guineensis* Jac.) chez le palmier dattier, Bennaceur *et al.* (2006) ont pu analyser la diversité génétique de quelques cultivars de palmier dattier.

Au Maroc, Sedra *et al.* (1998) ont rapporté l'importance de l'utilisation de la technique RAPD chez le palmier dattier. Ils ont montré l'efficacité de cette technique par rapport à la RFLP en termes de mise en évidence du polymorphisme chez cette espèce.

En Tunisie, les travaux de Ould Mohamed Salem *et al.* (2001), sur les génotypes multilocus de quatre systèmes enzymatiques, ont pu caractériser 27 cultivars parmi 29 étudiés. Sakka *et al.* (2004) ont exploité l'analyse de la diversité moléculaire du génome chloroplastique pour générer des marqueurs moléculaires susceptibles de différencier les cultivars tunisiens de palmier dattier. Ces auteurs ont démontré que 38 écotypes de palmier dattier se répartissent sur 29 haplotypes mis en évidence, qui ont permis d'identifier 28 génotypes. Cependant, l'utilisation des marqueurs microsatellites a été rapportée pour établir une clé d'identification variétale. Zehdi *et al.* (2006) ont pu déterminer 48 cultivars tunisiens parmi les 200 répertoriés dans le pays. L'apport du marquage moléculaire a permis de répondre à quelques questions relatives à la compréhension de l'évolution du palmier dattier

En Algérie, le patrimoine génétique phoénicicole joue un rôle très important dans la stabilité des régions désertiques situées au sud du pays. En dépit de l'importance économique de cette espèce, beaucoup de travaux sont réalisés actuellement afin de mieux apprécier la diversité génétique de ce patrimoine palmicole algérien. Le palmier dattier est issu soit de graine soit de rejets. Les phoéniculteurs ont toujours essayé de discriminer traditionnellement les variétés de palmier dattier, en se basant sur l'utilisation des caractères morphologiques, notamment les caractères du fruit qui sont nécessaires pour toute reconnaissance variétale des plantes issues de rejets ou de graines. Sur le millier de variétés dénombrées en Algérie, peu de variétés présentent une importance économique réelle, telle que la variété Deglet-Nour, et qui ont été sélectionnées avec le temps pour leur qualité, les autres sont rares et produisent des dattes de faible valeur marchande. Parmi ces variétés certaines sont menacées de disparition car très âgées elles ne produisent plus de rejets indispensables à leur multiplication.

II. La diversité génétique

1. Le polymorphisme génétique

Le polymorphisme génétique est à la base de la diversité génétique, il correspond à des variations de séquences d'ADN au sein d'un groupe d'individus. Ces variations naturelles sont dues à des mutations successives au cours de l'évolution qui permettent de caractériser la diversité génétique entre individus et populations. En général, la majorité des polymorphismes sont neutres (Kimura 1968, 1983), mais une partie de ces variations peut influencer les différences phénotypiques observées entre individus qui pourront leur permettre d'avoir un meilleur fitness dans leur environnement. Le polymorphisme peut nous renseigner sur les différents processus qui façonnent la variabilité génétique. Ils peuvent être de nature démographique (taille des populations, migrations) et sélectif (lié à l'environnement). L'étude de l'importance relative de ces différentes composantes façonnant le polymorphisme permet de mieux comprendre l'histoire évolutive des populations et des espèces.

2. Les forces évolutives

Le niveau de diversité génétique des populations et des variations de fréquences alléliques dépend de l'action respective de quatre forces évolutives pouvant interagir les unes avec les autres : la mutation, la sélection, la migration et la dérive. Elles sont à l'origine de la structure de la diversité génétique et de son évolution.

2.1. Les mutations

Les mutations sont des modifications de l'information génétique. Elles peuvent survenir aléatoirement au cours de la réplication de l'ADN mais peuvent aussi être dues à des agents mutagènes physiques (par exemple, les rayonnements ultra-violets), chimiques ou biotiques (virus et transposons). La recombinaison est un processus qui entraîne un brassage de matériel génétique intra- ou inter-chromosomique et dont la conséquence est la création de nouvelles combinaisons alléliques, c'est-à-dire d'haplotypes. Les mutations (et recombinaisons) qui augmentent ainsi la diversité génétique globale en générant de nouveaux allèles et de nouveaux haplotypes peuvent être de différentes classes. Le polymorphisme chromosomique peut être dû soit à une variation du nombre des chromosomes soit à un changement de leur structure (délétion, duplication, inversion, translocation). Le polymorphisme d'insertion

caractérise des éléments d'ADN capables de s'insérer dans différentes régions du génome. Inversement, le polymorphisme de délétion supprime certaines régions du génome, alors que le polymorphisme d'inversion comme son nom l'indique provoque un changement dans l'ordre de certaines séquences dans le génome. Le polymorphisme ponctuel signifie des changements à l'échelle du nucléotide (SNP pour « Single Nucleotide Polymorphism »). Enfin, le polymorphisme de taille correspond à une variation du nombre de motifs répétés sur un brin d'ADN, comme nous pouvons l'observer chez les microsatellites (SSR pour « Simple Sequence Repeat »). Différents marqueurs moléculaires existent pour travailler sur ces différents types de polymorphismes.

Les mutations affectent l'ensemble du génome. Celles-ci sont silencieuses lorsqu'elles n'ont pas d'impact évolutif. Par contre, des changements fonctionnels peuvent avoir lieu d'une part, lorsqu'elles entraînent des changements de nucléotides (mutation non-synonyme) dans les régions codantes et donc des modifications structurales et fonctionnelles des protéines.

D'autre part, les mutations affectant les régions non-codantes peuvent aussi avoir un impact évolutif. En effet, elles peuvent entraîner des modifications dans les régions de l'ADN ayant un rôle important dans la régulation de la transcription et dans l'organisation du génome et qui sont donc soumises à des pressions de sélection indiquant qu'elles ont un rôle évolutif et fonctionnel majeur (Wang *et al.* 1999).

Le phénomène de mutation est la source fondamentale de variation génétique et correspond à un changement héréditaire dans le matériel génétique de l'organisme. Les mutations sont importantes car elles apportent de la richesse par la création de nouveaux allèles.

2.2. La dérive génétique

La dérive génétique, qui agit sur tout le génome, représente un changement aléatoire des fréquences alléliques au cours des générations dans une population de taille finie (Wright 1931). En effet, dans une population de taille finie, les fréquences alléliques varient sous l'effet du hasard d'une génération à l'autre. Plus la population sera petite, plus les effets de la dérive seront grands car les écarts de fréquences alléliques y seront d'autant plus visibles d'une génération à l'autre. La dérive peut aboutir à l'élimination ou à la fixation d'un allèle pour tous les individus de la population. Les chances de fixation ou d'élimination d'un allèle dépendent évidemment de sa fréquence dans la population. Plus il est fréquent, plus il aura de chances d'être fixé tandis que plus l'allèle est rare, plus il sera sujet à l'élimination (s'il ne confère bien entendu aucun avantage sélectif). Contrairement aux mutations, la dérive génétique aura tendance à réduire la diversité génétique.

2.3. La migration

En génétique des populations, la migration se réfère au flux de gènes entre populations. La migration limite, voire empêche (selon son intensité), la divergence génétique entre populations engendrée par l'effet de la dérive génétique. Le système de dispersion du pollen et des graines est bien évidemment directement lié à ce processus (Hamrick *et al.* 1993).

Les flux géniques sont le résultat de la contribution génétique d'une population migrante sur la future génération. Ils ont pour effet l'augmentation de la richesse allélique pour les populations qui reçoivent ces migrants. A long terme, lorsque les flux de gènes sont réciproques d'une population à une autre, les fréquences alléliques entre populations s'homogénéisent (Slatkin 1985) induisant une réduction de la différenciation génétique. L'homogénéisation est d'autant plus forte que les flux de gènes sont importants. Il est utile de noter que les flux de gènes ont généralement lieu au sein d'une même espèce, mais il existe aussi des flux de gènes interspécifiques.

La dérive génétique tend à augmenter la différenciation entre les populations alors que la migration provoque l'effet contraire.

2.4. La sélection naturelle

La sélection naturelle est un mécanisme qui contribue à l'évolution des espèces. Elle fait le tri entre les individus montrant un différentiel de survie et de reproduction dans un environnement donné. Elle augmente ainsi la fréquence des phénotypes les plus favorables à la survie et à la reproduction tandis que les phénotypes les moins adaptés tendront à disparaître. Pour que la sélection agisse, il est nécessaire que les individus d'une population ne soient pas tous identiques (variation de traits), qu'une part de cette variation affecte les traits liés à la capacité de survivre et/ou de se reproduire, et que les différences de survie/reproduction soient dues à des variations génétiques héritables (par conséquent, ces différences doivent être contrôlées par des gènes transmissibles de génération en génération). Les porteurs des mutations avantageuses ont plus de descendants qui à leur tour auront plus de descendants et ainsi de suite, aboutissant alors à une population dont les individus sont adaptés à leur milieu. Contrairement à la migration et à la dérive génétique qui affectent l'ensemble du génome, la sélection n'agit que sur certains loci.

Encadré 1 : *Diversité génétique : définition et intérêt*

La diversité génétique est une composante essentielle de la biodiversité. Elle décrit la variabilité des gènes entre ou à l'intérieur des espèces et de leurs populations. En d'autres termes, la diversité génétique est définie par le niveau de similarité ou de différence dans la composition génétique des individus, des populations et des espèces ou des communautés. Elle peut donc se mesurer à différents niveaux : de l'individu à la communauté.

La diversité génétique est générée au niveau du génome. Une étape préliminaire à sa compréhension est donc la compréhension des mécanismes d'évolution moléculaire de ces génomes, c'est à dire l'apparition des mutations et leur devenir en fonction de la dérive, de la sélection et de la migration.

L'intérêt actuel porté à la biodiversité montre à quel point il est nécessaire de veiller au maintien d'une ressource génétique suffisamment large pour garantir l'adaptation des organismes face aux changements environnementaux directs (les exploitations forestières et les changements d'usage des terres par exemple) et indirects sur le long terme (changements climatiques globaux). En effet, cette diversité génétique est très importante car elle représente le support de base sur lequel peut agir la sélection. Elle peut représenter un avantage direct pour une population, la valeur adaptative d'un caractère étant généralement supérieure pour un gène présentant plusieurs états alléliques (hétérozygotie) ou pour une population formée d'individus différents et complémentaires. Elle peut de plus se concevoir comme une assurance permettant l'adaptation à de nouvelles conditions environnementales. Il est admis que plus la diversité génétique est grande dans un groupe d'individus (sous-population, population, espèce) plus il sera facile pour certains individus de s'adapter à de nouvelles conditions environnementales. En plus de permettre une plus grande adaptabilité des individus, elle permettra également de réduire la dépression de consanguinité et diminuera ainsi le risque d'extinction (Frankham 2003, 2005). L'étude de la diversité génétique à différents niveaux hiérarchiques d'organisation peut apporter une aide précieuse à la biologie des populations et à la biologie évolutive ; deux disciplines importantes pour la biologie de la conservation.

3. Les marqueurs génétiques d'estimation de la structuration génétique

L'étude de la diversité génétique passe par le choix du type de marqueur susceptible de la traduire le plus fidèlement possible. Les marqueurs génétiques sont des caractéristiques héréditaires qui renseignent sur le génotype de l'individu qui les porte (de Vienne 1998). Trois types sont largement utilisés pour l'évaluation de la variabilité génétique, à savoir les marqueurs phénotypiques, les marqueurs biochimiques et les marqueurs moléculaires.

3.1. Les marqueurs phénotypiques

Les marqueurs morphologiques ont été retenus dans un premier temps parce qu'ils présentaient l'avantage d'être immédiatement disponibles et nécessitent seulement un équipement simple. Ils représentent des variations de type qualitatif (couleur..), morphologiques (forme), des résistances à des maladies ou des ravageurs. Ces critères sont utilisés pour décrire et identifier les lignées et les variétés chez les végétaux. Cependant, ces marqueurs sont peu polymorphes et en général dominants, ils ont fréquemment une base génétique complexe, sont limités en nombre et leur expression est souvent fortement influencée par l'environnement.

3.2. Les marqueurs biochimiques

Les marqueurs biochimiques, généralement des isozymes ou des protéines de réserve, ont un pouvoir de discrimination plus élevé que les marqueurs morphologiques, car ils traduisent directement l'activité des gènes. C'est grâce à la technique de l'électrophorèse sur gel qu'il a été possible de mettre en évidence les variantes protéiques (allozymes) (Neel & Ellstrand 2001). Ces marqueurs ont été et sont encore largement utilisés pour des études de génétique des populations.

3.3. Les marqueurs moléculaires

Ils se développent rapidement et les plus utilisés en raison de leur fiabilité et leur précision de réponse sont les SSR (Simple Sequence Repeats ou microsatellites). Les autres marqueurs utilisés sont les ISSR (Inter Simple Sequence Repeats), les RAPD (Random Amplified Polymorphic DNA), les AFLP (Amplified Fragment Length Polymorphism), les RFLP (Restriction Fragment Length Polymorphism) et plus récemment les SNP (Single Nucleotide Polymorphism) (Santoni *et al.* 2000).

3.3.1. Les marqueurs RFLP

La technique RFLP repose sur la mise en évidence de la variabilité de la séquence nucléotidique de l'ADN génomique après digestion par des enzymes de restriction. Les fragments générés par digestion enzymatique de l'ADN de deux individus sont séparés selon leur taille sur gel d'agarose, dénaturés et transférés par capillarité sur une membrane de nylon. La position relative des fragments d'ADN est conservée durant le transfert. La dernière étape consiste en l'hybridation moléculaire avec une sonde d'ADN marquée préalablement. La sonde employée permet de détecter une région du génome où la digestion donne des fragments de restriction identiques pour les deux individus. Les deux profils révélés après électrophorèse ne permettent donc pas de les distinguer. Avec le couple enzyme/sonde, aucun polymorphisme n'est mis en évidence entre ces deux individus. En revanche, pour l'enzyme, un individu présente une mutation au niveau d'un site de restriction, entraînant la perte de ce site. Ainsi, dans un second cas de figure, la digestion avec cette enzyme et l'emploi de la même sonde met en évidence un fragment plus petit chez un individu que chez l'autre. Dans ce cas, le couple enzyme/sonde permet ainsi de mettre à jour le polymorphisme existant au niveau des sites de restriction dans cette région du génome. Les marqueurs RFLP présentent plusieurs des qualités espérées d'un bon marqueur moléculaire. Ils peuvent être mono-locus ou multi-locus, codominants, bi- ou multialléliques. Ils permettent une analyse génétique complète, présentent un très bon potentiel de transfert entre espèces, même assez éloignées, grâce à certaines sondes de séquence très conservée (Helentjaris *et al.* 1985) et sont généralement très reproductibles.

Les RFLP ne sont cependant pas des marqueurs parfaits pour un travail à grande échelle. En effet, cette technique prend du temps (plusieurs jours de travail en laboratoire), nécessite beaucoup d'ADN pour chaque analyse (quelques microgrammes).

3.3.2. Les marqueurs RAPD

La technique RAPD développée par Williams *et al.* (1990) consiste en l'amplification par PCR de fragments de l'ADN génomique en utilisant des amorces arbitraires de taille courte (10 pb). Une amorce RAPD permet généralement l'amplification d'une dizaine de fragments. Les fragments produits par cette amplification sont séparés par électrophorèse sur un gel d'agarose. Pour deux individus, si deux sites d'hybridation sont proches et sur deux brins complémentaires d'ADN, il y aura amplification. En revanche, si ces deux sites sont trop éloignés, il ne peut y avoir amplification. Le polymorphisme décelé est dû à des mutations soit dans les régions amplifiées soit au niveau des sites de fixation des amorces. Le

polymorphisme révélé est un polymorphisme de sites d'hybridation d'amorce traduit par la présence ou l'absence de la bande.

Cette technique est rapide et d'une faible technicité, elle génère un grand nombre de fragments à partir d'une faible quantité d'ADN et ne nécessite pas des informations génétiques ou génomiques préalables. Les marqueurs RAPD sont des marqueurs multilocus, bialléliques, dominants.

Bien que cette technique soit simple, le désavantage est que les marqueurs sont dominants et ne sont pas reproductibles (Williams *et al.* 1990 ; Welsh & McClelland 1990). Les marqueurs RAPD ont été utilisés chez plusieurs espèces pour l'analyse de la diversité génétique, la caractérisation de ressources génétiques, l'identification de cultivars et la cartographie des génomes.

3.3.3. Les marqueurs AFLP

Développée par Vos *et al.* (1995), cette technique utilise à la fois les enzymes de restriction et l'amplification PCR. Après la digestion de l'ADN génomique par deux enzymes de restriction, des adaptateurs de séquence connue et spécifiques des deux sites de restriction sont ajoutés aux extrémités des fragments de restriction. Une fois que les fragments de restriction sont ainsi bordés de séquences connues, il devient possible de les amplifier par PCR. Grâce à l'emploi d'amorces sélectives qui ne peuvent amplifier qu'une petite portion des milliers de fragments de restriction obtenus, seule une centaine de fragments sont amplifiés. Ces fragments sont séparés par électrophorèse sur gel de polyacrylamide dénaturant. Le polymorphisme révélé (de type présence ou absence de bandes) est provoqué par des changements dans les sites de restriction ou les régions bordantes.

Les marqueurs AFLP sont des marqueurs bialléliques dominants (Vos *et al.* 1995) et génèrent un grand nombre de bandes polymorphes. Ils sont aussi extrêmement reproductibles et aucune information préalable de séquence ou génération de sonde ne sont nécessaires ce qui les rend très utiles pour créer rapidement des cartes génétiques. Toutefois, la dominance et les difficultés techniques liées au marquage par AFLP limitent son utilisation à grande échelle pour des applications comme la sélection assistée par marqueurs.

La technique AFLP est largement utilisée dans le domaine végétal en tant que technique d'empreintes génétiques hautement informatives et permet de comparer des génotypes génétiquement très proches. Elle connaît aussi une large application dans l'identification des cultivars et la détermination de leurs relations phylogénétiques, la cartographie des génomes et la saturation d'une région du génome au voisinage d'un gène en vue de son clonage.

3.3.4. Les marqueurs minisatellites

Jeffreys *et al.* (1985) ont découvert au niveau du génome un type de marqueurs moléculaires polymorphes connus sous le nom de minisatellites hypervariables. Ces marqueurs sont constitués des répétitions en chaîne d'un motif formé de 15 à 70 nucléotides. Les minisatellites appartiennent à la classe des VNTR (Variable Number Tandem Repeat), et présentent un polymorphisme de taille dû à la variation d'unités de répétition qui les constituent.

3.3.5. Les marqueurs microsatellites, des témoins de l'histoire génétique récente

- Définition et modèles mutationnels

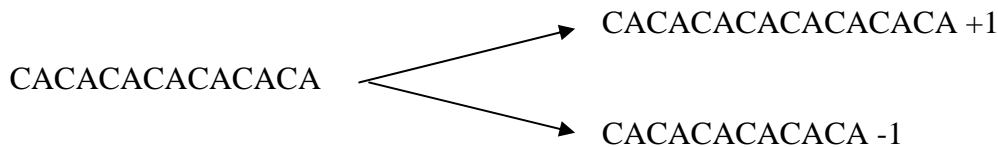
Les microsatellites (SSR, pour Simple Sequence Repeats) sont de petites portions d'ADN, constituées d'un motif variant d'un à quatre nucléotides, répété plusieurs fois. Ces régions sont particulièrement intéressantes car elles évoluent plus rapidement que d'autres portions du génome (10^{-2} à 10^{-6} mutation par locus par génération) (Li *et al.* 2002 ; Selkoe & Toonen 2006). Les SSR présentent un taux de polymorphisme extrêmement élevé qui repose sur la variation du nombre d'unités de répétition (n) constituant le microsatellite. On les trouve dans le génome nucléaire ainsi que dans le génome chloroplastique (Goldstein & Schlötterer 1999 ; Provan *et al.* 2001). Les microsatellites nucléaires sont codominants et sont transmis de manière biparentale.

Avec leur densité importante dans les génomes eucaryotes, et la proportion de nucléotides qu'ils peuvent représenter, les microsatellites sont des acteurs importants de l'évolution moléculaire au sens large (Encadré 2).

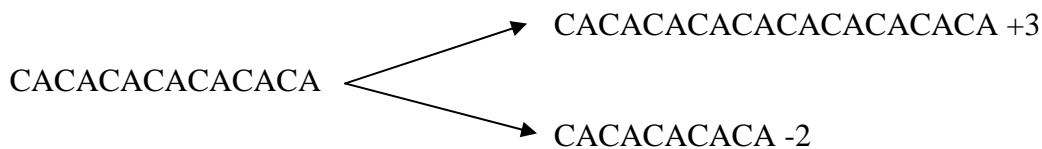
Encadré 2 : Les principaux modèles mutationnels des microsatellites

Différents modèles mutationnels ont été proposés pour les microsatellites (Jarne & Lagoda 1996 ; Estoup *et al.* 2002 ; Li *et al.* 2002). Les principaux modèles de mutations des microsatellites sont :

- le modèle de mutation strictement pas à pas (Stepwise Mutation Model ou SMM) : chaque mutation ajoute ou retranche un motif au microsatellite.

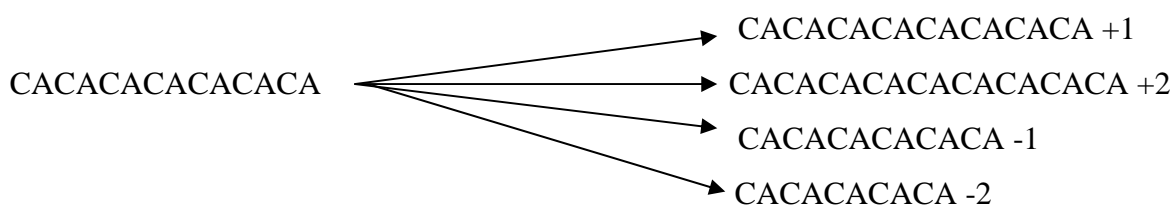


- le modèle de mutation généralisé (Generalized Stepwise Model) : ajoute ou retranche plus d'un motif à chaque évènement mutationnel.



- le modèle à deux phases (Two Phases Model, TPM), qui est une généralisation du GSM : chaque évènement mutationnel ajoute ou retranche un motif (comme dans le SMM) avec une probabilité p ou bien ajoute ou retranche X motifs (comme dans le GSM) avec une probabilité $(1-p)$. X est tirée dans une loi géométrique.

- le modèle à K allèles (KAM) : exactement K états équiprobables sont possibles (Exemple avec $K = 4$).



- le modèle à allèles infinis (IAM) : généralise le modèle KAM ($K \rightarrow \infty$) ; chaque mutation crée un nouvel allèle.

On décrit généralement un microsatellite sous la forme $(X)_n$, avec X le motif, et n le nombre de répétitions.

Les microsatellites sont des marqueurs codominants, extrêmement polymorphes, faciles à manipuler et permettant le multiplexage. Ils requièrent très peu d'ADN (quelques nanogrammes) et sont bien répartis dans le génome. Ces marqueurs sont très reproductibles et moins lourds d'utilisation que les RFLP ou AFLP avec une bonne résolution analytique.

- Applications

De par leur taux de mutation élevé, ces marqueurs sont en général très polymorphes. On peut donc les utiliser pour détecter des événements relativement récents à l'échelle historique, telles des expansions ou des réductions de taille de population, mais également pour effectuer des assignations d'individus (Zhang & Hewitt 2003 ; Selkoe & Toonen 2006), ainsi que des analyses de parenté, grâce à leur caractère codominant (Jones *et al.* 2010). Les microsatellites sont également fréquemment utilisés pour déterminer l'histoire démographique des populations ; de nombreuses méthodes d'inférence ont été développées pour détecter à partir de la variation génétique actuelle observée dans les populations naturelles, des événements historiques tels que des contractions, effets fondateurs ou expansions à l'aide de marqueurs microsatellites. Enfin les marqueurs microsatellites peuvent également être utilisés dans un but similaire à celui des séquences, afin de déterminer la structure génétique des populations d'une espèce, et de reconstituer les mécanismes à l'origine de la distribution génétique et géographique de l'espèce (Jarne & Lagoda 1996 ; Ellegren 2004).

3.3.6. Les marqueurs SNP (Polymorphisme de simple nucléotide)

Les SNP sont des marqueurs moléculaires constituant la forme la plus abondante de variations génétiques dans les génomes. C'est un type de polymorphisme de l'ADN dans lequel deux chromosomes diffèrent sur un segment donné par une seule paire de bases. L'inconvénient des SNP est leur faible degré de polymorphisme par rapport aux microsatellites, ce qui implique de tester plus de SNP que de microsatellites pour avoir la même information.

4. Caractérisation et mesures de la diversité génétique

L'étude de la diversité génétique est une des applications les plus répandues pour les marqueurs moléculaires. Le choix du marqueur approprié dépendra en partie du matériel au sein duquel on cherche à mesurer la diversité génétique. Selon le but de l'étude, on peut être intéressé à mesurer la diversité génétique entre des populations (niveau inter population) ou encore entre divers individus ou au sein d'une même population (niveau intra population).

4.1. Diversité intra population

La diversité intra-population est la diversité à l'intérieur d'une même population. Plusieurs marqueurs moléculaires sont utilisés pour l'évaluation de cette diversité au sein d'un matériel génétique. En termes de variabilité génétique intra-population, la comparaison directe des fréquences alléliques n'est pas facile à réaliser. Toutefois, il existe des paramètres susceptibles de synthétiser, moyennant une valeur globale, les informations les plus importantes comme le nombre moyen d'allèles par locus (A), taux de polymorphisme (PIC), le taux moyen d'hétérozygotie observé (H_o), le taux moyen d'hétérozygotie attendu (H_e) et indice de fixation F_{IS} .

4.1.1. Nombre moyen d'allèles par locus (A)

Ce paramètre traduit la richesse en allèles d'une population (Petit *et al.*, 1998), il est calculé selon la formule:

$$A = \frac{\sum a}{l}$$

Où a représente le nombre d'allèles à un locus et l le nombre des loci étudiés.

4.1.2. Taux de polymorphisme (PIC)

C'est le pourcentage des loci polymorphes dans l'échantillon étudié. Autrement dit, c'est la probabilité d'observer au moins deux allèles au même locus, cette probabilité dépend des fréquences respectives des allèles et aussi de la taille de l'échantillon.

4.1.3. Taux d'hétérozygote observé (H_o)

C'est la proportion d'individus hétérozygotes au locus K comme dans la formule :

$$H_o = \frac{\sum_{k=1}^l H_{ok}}{L}$$

4.1.4. Taux d'hétérozygote attendu (H_e)

L'hétérozygotie attendue (H_e) correspond à la diversité génétique de Nei (1973), et caractérise le niveau de variabilité des gènes au sein d'une même population, c'est-à-dire qu'elle exprime la probabilité pour que deux allèles tirés au hasard dans une population soient

différents à un même locus. Ce paramètre fournit une bonne estimation de la variabilité génétique de la population :

$$H_e = 1 - \sum_{i=1}^n p_i^2$$

Avec p_i la fréquence allélique pour chaque allèle et n le nombre d'allèle au locus considéré.

4.1.5. Indice de Fixation F_{IS}

Ce paramètre mesure l'écart entre le taux d'hétérozygote observée (H_o) et d'hétérozygote attendu (H_e) d'une population d'individus trouvés à l'écart de l'équilibre de Hardy Weinberg (Guo & Thompson 1992), il est appelé aussi l'écart à la panmixie et se calcule selon la formule :

$$F = 1 - \frac{H_o}{H_e}$$

A un locus donné, dans une population d'individus diploïdes l'association de deux gamètes pour former les individus se fait au hasard par rapport aux génotypes de ces gamètes, la structure génétique de cette population est appelée structure de Hardy Weinberg. L'indice de fixation est l'écart à la structure de Hardy Weinberg, il varie de -1 à +1 et permet de connaître le déficit en hétérozygote par population, par locus et pour l'ensemble de locus. F est positif quand la population présente un déficit en hétérozygote par rapport à l'équilibre panmictique et négatif dans le cas contraire. Un certain nombre de facteurs contribue à cet écart comme la consanguinité, la dérive, la sélection et la différenciation génétique.

4.1.6. L'équilibre de Hardy-Weinberg

La variabilité génétique d'une population peut être difficile à prévoir au fil des générations. De nombreux facteurs peuvent modifier la fréquence des allèles au sein des populations (les forces évolutives).

Les fréquences alléliques et génotypiques peuvent être prédites facilement pour un modèle théorique simple d'une population « idéalisée », dont les caractéristiques sont les suivantes :

- Population de taille infinie (pour éviter la dérive génétique) pour des individus diploïdes
- Croisements par reproduction sexuée aléatoires (population panmictique) pour des générations non-chevauchantes
- Absence de migration

- Absence de sélection
- Absence de mutation

Si ces conditions sont respectées, la loi de Hardy-Weinberg (Hardy 1908 ; Weinberg 1908) permet de prédire exactement les fréquences génotypiques à partir des fréquences alléliques de la population idéale. Les implications du principe de Hardy-Weinberg sont qu'une seule génération de panmixie suffit pour atteindre les fréquences génotypiques de Hardy-Weinberg et que les fréquences alléliques et les fréquences génotypiques (c'est à dire la structure génétique de la population) restent stables de génération en génération.

4.2. Diversité inter population

Pour décrire la diversité génétique des populations, Wright (1951) a développé l'indice F_{ST} qui est la corrélation entre deux gamètes tirés au hasard et provenant de deux sous-populations différentes.

Les F-statistiques (Wright 1969 ; Weir & Cockerham 1984) permettent de calculer la différenciation génétique entre populations (F_{ST}), aussi appelée indice de fixation (Nei & Chesser 1983). Il permet de mesurer la différenciation des populations à partir du polymorphisme génétique et est calculé sous l'équilibre migration/dérive (Wright 1951). Cet indice s'applique aux locus bialléliques et en l'absence de sélection. Lorsque les locus sont multialléliques, on utilise une généralisation de cet indice, dénommé G_{ST} de Nei (1973). Les valeurs de cette mesure sont comprises entre 0 et 1 ; si les populations présentent les mêmes fréquences alléliques, alors $F_{ST} = 0$. Le F_{ST} peut être interprété comme la part de variance des fréquences alléliques due aux différences entre populations :

$$F_{ST} = \frac{H_T - H_S^*}{H_T}$$

Avec

H_S l'hétérozygotie attendue pour une sous-population supposée à l'équilibre de Hardy-Weinberg

H_S^* la diversité moyenne des H_S sur les sous-populations

H_T l'hétérozygotie attendue pour une population globale supposée à l'équilibre de Hardy-Weinberg.

4.3. Structuration spatiale de la diversité génétique

Pour comprendre l'origine de la structuration de la diversité génétique entre populations, il est parfois utile de déterminer des distances génétiques entre paires de populations (par exemple avec les F_{ST}). Les matrices de distances génétiques peuvent alors être comparées à des matrices de distances (géographiques ou écologiques) entre sites afin d'évaluer leur corrélation.

Le test de Mantel (Mantel 1967 ; Smouse *et al.* 1986) permet d'évaluer cette corrélation par régression entre matrices à expliquer (distance génétique) et matrices explicatives (distance géographique ou écologique).

Le test de Mantel est la méthode la plus couramment utilisée pour étudier les relations entre variables environnementales et structure génétique. Il permet de calculer la corrélation entre deux matrices et d'évaluer si cette corrélation est significative en la comparant à la distribution de valeurs obtenues à la suite de permutations au sein des matrices (Mantel 1967). Cette méthode, tout à fait générale, a rapidement été utilisée pour tester la corrélation entre une matrice contenant les distances génétiques pour chaque paire d'individus et une matrice contenant les distances géographiques entre ces mêmes individus.

L'autocorrélation génétique spatiale permet une estimation du patron de répartition de la variable génétique. Elle va permettre de renseigner si les individus les plus proches géographiquement sont aussi les plus proches génétiquement. L'indice générique I de Moran (1950) est le plus souvent utilisé comme estimateur de l'autocorrélation des fréquences alléliques (Smouse & Peakall 1999):

$$I_k = \frac{n \sum_{i=1}^n \sum_{j=1}^n w_{ij} (x_i - \bar{x})(x_j - \bar{x})}{W \sum_{i=1}^n (x_i - \bar{x})^2} W$$

Avec n , nombre total d'individus ; $w_{ij} = 1$ si les individus i et j appartiennent à la classe de distance k , $w_{ij} = 0$ dans les autres cas ; $x_i = 1, 0.5$ ou 0 selon que l'individu i a 2, 1 ou 0 fois l'allèle considéré, représente la moyenne sur les n individus des fréquences alléliques individuelles.

D'autres méthodes pour analyser l'autocorrélation génétique spatiale s'avèrent plus spécifiques, notamment en utilisant des coefficients de parenté (Hardy & Vekemans 1999). Certaines analyses sont même conçues spécifiquement pour les marqueurs dominant en supposant des coefficients de consanguinité (Hardy 2003).

5. Détection de la structure génétique

Lorsque l'on connaît les données génétiques d'un ensemble d'individus pour un ensemble de marqueurs, un des objectifs est de détecter si les individus sont structurés en populations, et si c'est le cas, d'identifier le nombre de groupes (*clusters*) isolés ou partiellement isolés et les individus composant chacun de ces groupes. Plusieurs méthodes permettent d'effectuer ces recherches.

5.1. Analyse factorielle des correspondances

L'analyse factorielle des correspondances (AFC) vise à rassembler en un nombre réduit de dimensions la plus grande partie de l'information initiale en s'attachant aux valeurs relatives, soit les correspondances entre les variables. L'AFC fournit un espace de représentation commun aux variables et aux individus. Les individus analysés sont vus comme un nuage de points dans un hyperespace qui a autant de dimensions qu'il y a de modalités (les allèles) sur toutes les variables (les allèles aux différents loci).

5.2. Analyse en composantes principales

L'analyse en composantes principales (ACP) est une méthode classique d'analyse de données multidimensionnelles. Elle permet de projeter des données sur un nombre réduit d'axes orthogonaux, tout en maximisant la variance des données projetées sur chacun des axes.

5.3. Arbre phylogénétique

Les méthodes historiques pour étudier la proximité génétique d'individus sont basées sur la construction d'arbres phylogénétiques.

L'algorithme fréquemment appliqué pour la construction d'arbres phylogénétiques est l'algorithme de *neighbor-joining*, proposé par Saitou & Nei (1987), implémenté dans divers logiciels (Felsenstein 1989 ; Kumar *et al.* 2004), et utilisé dans de nombreuses études. La construction de l'arbre est faite à partir d'une matrice contenant les distances génétiques pour chaque paire d'individus. Ces distances génétiques sont classiquement calculées à l'aide des distances de Nei *et al.* (1983) ou de Reynolds *et al.* (1983). Une fois l'arbre construit, on peut définir les différents groupes génétiques.

5.4. Analyse Bayésienne pour inférence de la structure génétique des populations

L'analyse Bayésienne de regroupement vise à estimer la structure génétique d'un ensemble d'individus en rassemblant ces derniers dans des groupes, sur la base de leurs profils génétiques, elle permet l'identification de groupes génétiques différents. On obtient après analyse des groupes distincts, génétiquement différents. Cette approche est implémentée dans le programme STRUCTURE (Pritchard *et al.* 2000). Ce programme travaille sous deux hypothèses : les populations sont à l'équilibre de Hardy-Weinberg (panmixie et pas de pression évolutive), et il y a existence d'un équilibre de liaison entre les différents locus.

La méthode STRUCTURE repose sur les modèles de mélange entre populations (*admixture*) (le genotype d'un individu peut provenir de plusieurs populations) et la corrélation entre fréquences alléliques. L'algorithme décrit par Pritchard *et al.* (2000) place les individus dans K populations (ou *clusters*). En général, les runs de STRUCTURE sont réalisés avec 10^6 itérations (runs) précédés d'une période de burning de 10.000 itérations (Pritchard *et al.* 2000). Dix simulations indépendantes sont effectuées pour chaque valeur de K . L'intervalle des valeurs testées de K varie de 1 à 10. La procédure définie par Evanno *et al.* (2005) détermine le nombre optimal K de populations. La méthode STRUCTURE attribue à chaque individu sa probabilité d'appartenance à chaque population. Les résultats de STRUCTURE sont visualisés sous forme d'histogrammes (chaque barre correspond à un individu, le coefficient d'appartenance à chaque population étant donné en ordonnée).

III. Généralités sur le génome chloroplastique

Les végétaux sont caractérisés par trois types de génomes localisés dans des compartiments cellulaires distincts : le noyau, les mitochondries et les chloroplastes. Ces génomes présentent différentes caractéristiques selon leur nature et selon l'organisme au sein duquel ils évoluent. Le séquençage complet de ces différents génomes a été réalisé chez *Arabidopsis thaliana* (The Arabidopsis Genome Initiative 1999).

Le génome nucléaire, bien que constituant une source importante de marqueurs, est diploïde, et à ce titre présente diverses caractéristiques pouvant compliquer les analyses : hétérozygotie, recombinaisons inter-alléliques, ou encore duplications des locus. Les marqueurs cytoplasmiques, bien que constituant une réserve plus limitée de gènes et étant assimilés à un seul locus, semblent mieux adaptés pour révéler une structuration génétique. Leur hérédité non mendélienne, généralement uniparentale maternelle chez les angiospermes, entraîne une réduction de la taille effective des populations de gènes. Cependant ces génomes montrent des taux de substitution nucléotidique plus faible (Wolfe *et al.* 1987), ce qui constitue une limite dans la recherche de variabilité intraspécifique. Chez les végétaux, c'est l'ADN chloroplastique (ADNcp) qui est le plus utilisé pour étudier les relations phylogénétiques.

1. Fonctions biologiques des chloroplastes

Les chloroplastes sont des structures spécialisées présentes dans le cytoplasme des cellules eucaryotes végétales. Ils sont séparés de l'environnement cellulaire par deux membranes; la cellule eucaryote est donc compartimentée. Les chloroplastes possèdent leurs propres génomes, et leurs fonctions sont vitales pour la cellule.

Les chloroplastes assurent la photosynthèse : c'est un processus bioénergétique, composé de deux phases, permettant aux plantes de synthétiser de la matière carbonée organique ($C_6H_{12}O_6$) et du dioxygène (O_2) à partir des pigments, d'eau (H_2O), de dioxyde de carbone (CO_2) et d'énergie solaire. La photosynthèse se déroule en deux phases distinctes : la première phase, aussi appelée phase des réactions photochimiques, permet à partir de l'énergie solaire, d' H_2O et de NADP, de produire de l' O_2 , du NADPH et de l'ATP. Les réactions de la première phase nécessitent des complexes situés dans la membrane des thylacoïdes : le photosystème II (PSII), la plastoquinone, le cytochrome *b6f*, le photosystème I (PSI), un complexe ATP synthase ainsi que des molécules comme la plastocyanine. La seconde phase, ou phase de

fixation du carbone, est indépendante de la lumière et est assurée dans le stroma par le cycle de Calvin. L'ATP et le NADPH, synthétisés lors de la première phase, vont être utilisés lors de la seconde phase pour fixer le CO₂ et générer du triose phosphate exporté vers le cytosol ou stocké dans le chloroplaste sous forme d'amidon. L'enzyme permettant de fixer le CO₂ est la ribulose 1,5 bisphosphate carboxylase/oxygénase (Rubisco). La photosynthèse constitue l'activité principale des chloroplastes. Néanmoins, il existe également un métabolisme non photosynthétique du chloroplaste : certaines réactions impliquées dans la réduction de l'azote, du soufre et dans la biosynthèse d'acides gras et d'acides aminés sont localisées dans les chloroplastes.

2. Origine et caractéristiques des chloroplastes

Les chloroplastes sont des organites dérivés des cyanobactéries. Ils prennent origine d'un événement d'endosymbiose impliquant l'engouffrement d'une cyanobactérie photosynthétique par un eucaryote non-photosynthétique. Suite à cet événement, la taille et le contenu en gènes du génome de l'endosymbionte ont été largement réduits. C'est d'ailleurs cette réduction et la dépendance consécutive du chloroplaste envers la cellule-hôte qui distingue cet organite endosymbiotique d'un endosymbionte obligatoire (Delwiche 1999).

On trouve de 10 à 100 copies du génome par plaste, et de 10 à 100 plastes par cellule (Walsh 1993). Les chloroplastes sont le siège de la photosynthèse et leurs 120 gènes environ codent pour des ARNr, des ARNt et des protéines impliquées dans la photosynthèse.

Le génome chloroplastique présente une structure similaire à celle du génome des Procaryotes notamment au niveau de ses ribosomes (dans leur taille, leur sensibilité aux antibiotiques et leur réactivité immunologique croisée), mais contrairement à ce dernier, il possède des introns (Umesono & Ozeki 1987).

3. Hérité des chloroplastes chez les Angiospermes

Peu de temps après la redécouverte des lois de Mendel en 1900, Correns et Baur (1909) découvrent simultanément l'existence d'une hérédité non-Mendélienne chez des espèces appartenant à la famille des Angiospermes. Correns (1909) observa chez *Mirabilis jalapa* que suite à des croisements entre des plantes (ou des branches) ayant des feuilles de couleur jaune ou verte, le caractère « couleur des feuilles » était hérité seulement du parent maternel (hérédité maternelle). Les branches vertes produisent toujours des graines qui donnent des

plantes vertes et les branches jaunes produisent toujours des graines qui donnent des plantes jaunes. Le parent qui donne le pollen n'a pas d'influence sur le caractère « couleur des feuilles » de la descendance. Baur (1909) identifia un autre type d'hérédité non mendélienne : en réalisant des croisements réciproques avec des plantes de *Pelargonium zonale* ayant des feuilles de couleurs différentes, il mit en évidence que le caractère « couleur des feuilles » était hérité des deux parents (Hérédité biparentale) de façon non mendélienne. La découverte de l'existence des chloroplastes puis l'identification de la présence d'ADN dans les chloroplastes ont pu matérialiser cette hérédité non mendélienne. Pour *Mirabilis jalapa* et *Pelargonium zonale*, les chloroplastes, qui portent les traits associés à la couleur des feuilles, sont hérités de façon maternelle et biparentale respectivement.

A la différence de l'ADN nucléaire, transmis par les deux parents, l'ADN chloroplastique a généralement un mode de transmission uniparental : chez la grande majorité des espèces comme chez les Angiospermes, la transmission du chloroplaste se fait par la mère (Corriveau & Coleman 1988). Cependant, cette transmission peut également être paternelle (par exemple chez les gymnospermes (Corriveau et Coleman 1988 ; Mogensen 1996) ou très occasionnellement biparentale (Hansen *et al.* 2007). Chez les Areaceae, le mode de transmission est maternel.

4. Caractéristiques de l'ADN chloroplastique

L'ADN chloroplastique se présente sous la forme d'une molécule circulaire double brin, contenue dans le chloroplaste. Il est de taille généralement petite varie de 40 et 220 kilobases, soit environ 20 à 30 fois inférieure à celle du génome de son ancêtre bactérien. Il est constitué de trois régions chez la plupart des espèces : régions répétées inversées (Inverted Repeat : IR) de 25 kb en moyenne (Palmer 1991) séparent le reste du génome en deux autres régions, la grande région simple copie (Large Single Copy : LSC) et la petite région simple copie (Small Single Copy : SSC). La LSC et la SSC semblent évoluer plus rapidement en séquence que la région IR : leur taux de mutation serait de quatre à six fois plus élevé (Wolfe *et al.* 1987 ; Clegg *et al.* 1991). La présence des IR semble stabiliser structurellement le génome chloroplastique en diminuant la fréquence des réarrangements (Palmer & Thompson 1982).

Le caractère haploïde du génome chloroplastique augmente l'effet de la dérive sur l'ADN chloroplastique et la taille efficace des populations, mesurée par des marqueurs chloroplastiques, est la moitié de celle mesurée à partir de marqueurs nucléaires (Birky 1995).

De ce fait, la structure génétique a tendance à être plus importante pour l'ADN chloroplastique que pour l'ADN nucléaire.

Le taux de mutation de l'ADN chloroplastique est plus faible que celui de l'ADN nucléaire. Chez les palmiers, il est particulièrement faible : la vitesse d'évolution du génome chloroplastique, mesurée sur le gène *rbcL*, est en moyenne 5 fois moins importante que celle du génome chloroplastique des graminées (Gaut *et al.* 1992), et est estimée à $4,2 \times 10^{-11}$ substitutions par site et par an (Wilson *et al.* 1990). Ce faible taux de mutation permet donc d'inférer l'histoire évolutive des espèces en remontant jusqu'à des périodes très anciennes, de l'ordre du million d'années selon les espèces et les régions chloroplastiques considérées. Ce taux de mutation varie entre les portions codantes (les gènes) et les portions non-codantes. Différents types de mutations sont susceptibles d'affecter les séquences chloroplastiques (et nucléaires). Ces différents types de mutations de séquence sont des substitutions, des InDel et des Inversions (Encadré 3).

Encadré 3 : Différents types de mutations de séquence

Substitution : Un nucléotide est remplacé par un autre.

CTAGTCCAGGACTCTTT

CTAGTCCAGTACTCTTT

Insertion-délétion : Un nucléotide ou un groupe de nucléotides est inséré ou supprimé dans une séquence.

CTAGTCCAGG ACTCTTT

CTAGTCCAGGTCATACTCTTT

Inversion : Les motifs de part et d'autre de la zone d'inversion sont complémentaires, ce qui permet le repliement de la molécule d'ADN. De ce fait, au moment de la réplication, il arrive que la séquence se décroche et se raccroche dans l'autre sens, entraînant une inversion des paires de bases. Ce motif rare, dit « hairpin », est répertorié chez plusieurs espèces végétales, au niveau de l'ADN chloroplastique (Kelchner & Wendel 1996).

CTAGTCCAGGCTAATACGATAGCTACTAATTAGCCTGGACTAG

CTAGTCCAGGCTAATTAGTAGCTATCGTATTAGCCTGGACTAG

5. Taille du génome chloroplastique

Les premiers génomes de plantes à avoir été séquencés sont ceux de génomes chloroplastiques appartenant à *Marchantia polymorpha* et *Nicotiana tabacum*, (Ohyama *et al.* 1986 ; Shinozaki *et al.* 1986). En juillet 2011, 213 génomes chloroplastiques étaient séquencés et disponibles dans les bases de données.

Au sein des Spermatophytes, ou plantes à graines, 139 génomes chloroplastiques ont été séquencés dont 14 pour des plantes appartenant aux Conifères et 121 pour des plantes appartenant aux Magnoliophytes. Parmi les génomes chloroplastiques de conifères séquencés, on peut citer ceux de *Pinus contorta* et *Pinus thunbergii* qui font 120 kb et 119kb respectivement. Au sein des plantes à fleurs, le génome chloroplastique fait 134 kb chez *Oriza sativa*, 140 kb chez *Zea mays*, 150 kb chez *Helianthus annuus* et *Phaseolus vulgaris*, et 153 kb chez *Brassica rapa*. Le génome chloroplastique d'*Arabidopsis thaliana* disponible a été séquencé en 1999 sur l'accession Columbia (Col-0) et a une taille de 154 kb (Sato *et al.* 1999). D'après la classification phylogénétique, les organismes les plus proches d'*Arabidopsis* et dont les génomes chloroplastiques ont été séquencés, sont *Capsella bursa-pastoris* (154 kb), *Crucihimalaya wallichii* (155 kb), et *Olimarabidopsis pumila* (154 kb) : les tailles de ces génomes sont presque identiques à celle du génome d'*Arabidopsis thaliana*. En règle générale, la taille des génomes chloroplastiques varie peu d'une plante terrestre à l'autre.

6. Les gènes chloroplastiques

Chez les végétaux supérieurs, les gènes chloroplastiques peuvent être classés en trois catégories. La première regroupe ceux qui codent des protéines de l'appareil photosynthétique : le PSI, le PSII, le cytochrome *b6f*, l'ATP synthase, la grande sous-unité de la Rubisco et des protéines du complexe NAD(P)H déshydrogénase. La seconde catégorie de gènes chloroplastiques code des ARN et des protéines impliquées dans l'expression des gènes chloroplastiques : ARNt, ARNr, protéines de l'ARN polymérase, protéines ribosomiques et d'autres protéines comme la maturase K, qui intervient dans l'épissage des introns chloroplastiques. La troisième catégorie contient des gènes qui codent des protéines n'appartenant pas aux deux premières catégories : c'est le cas de Cema, une protéine présente dans la membrane interne du chloroplaste (Weng *et al.* 2017).

6.1. Les gènes impliqués dans la photosynthèse

Des gènes de l'ADNcp codent pour les sous-unités de quatre systèmes importants intégrés dans la membrane des thylakoïdes et impliqués dans la photosynthèse : le photosystème I, le photosystème II, le cytochrome *b6f* et l'ATP synthétase (Campbell & Mathieu 1995). Les gènes codant pour des sous-unités des protéines Rubisco et protochlorophyllide réductase sont aussi présents dans le génome des chloroplastes.

Parmi les gènes photosynthétiques, ceux des photosystèmes sont les plus nombreux. Les photosystèmes sont les unités photoréceptrices; ce sont eux qui captent l'énergie lumineuse. Leur différenciation s'est faite au cours de l'évolution. Les gènes *psa* et *psb* codent respectivement pour les composantes des photosystèmes I et II. Chez les plantes vertes, l'ADNcp contient six gènes codant pour des sous-unités du photosystème I et 14 gènes codant pour des composantes du photosystème II. Les gènes *pet* codent pour les six sous-unités qui composent le cytochrome *b6f*. Ce système sert à faire la navette entre les deux photosystèmes. L'ATP synthétase produit l'ATP nécessaire à la deuxième phase de la photosynthèse (cycle de Calvin) à l'aide des protons produits par les réactions photochimiques (première phase de la photosynthèse) (Campbell & Mathieu 1995).

Deux parties, le CF0 et le CF1 composent l'ATP synthétase. Le CF0 comprend lui-même cinq différentes sous-unités et le CF1 en comprend quatre. Les six gènes *atp* du génome chloroplastique des plantes vertes codent pour des sous-unités du CF0 et du CF1 de l'ATP synthétase (Sugiura 1992).

La ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) est la protéine la plus abondante dans le stroma des chloroplastes (Sugiura 1992). Elle est composée de huit grandes sous-unités identiques (LS) de 55 kDa et de huit petites sous-unités identiques (SS) de 12 kDa. Chez les plantes vertes, les sous-unités LS sont codées dans l'ADN chloroplastique par le gène *rbcL* mais le gène *rbcS*, codant pour les petites sous-unités, se retrouve dans l'ADN nucléaire.

Onze gènes chloroplastiques (*ndh*) codent pour des protéines similaires aux composantes de la chaîne respiratoire NADH déhydrogénase des mitochondries. Ils ont été trouvés dans l'ADNcp de plusieurs plantes terrestres (Sugiura 1992). Il a été montré récemment que les protéines Ndh des chloroplastes du tabac font partie d'un complexe fonctionnel indispensable à la croissance de la plante dans des conditions optimales (Burrows *et al.* 1998 ; Shikanai *et al.* 1998). Ce complexe semble participer à la chaîne cyclique d'électrons autour du photosystème I lors des deux phases de la photosynthèse (Turmel *et al.* 1999).

6.2. Les gènes impliqués dans l'expression génétique

La traduction des gènes chloroplastiques fait appel à des ribosomes 70S composés de deux sous-unités, une de 30S et une de 50S (Sugiura 1992). Les gènes *rps* codent pour les protéines de la petite sous-unité (30S) du ribosome tandis que les gènes *rpl* codent pour les protéines de la grande sous-unité (50S). Il est à noter que l'ADNcp code seulement pour le tiers des 60 protéines ribosomales (Sugiura 1992). Il y a trois principaux types d'ARNs ribosomiques dans le génome chloroplastique. Le 23S ou grande sous-unité de l'ARNr (LSU) et le 5S sont associés à la sous-unité 50S du ribosome. L'ARNr associé à la sous-unité 30S du ribosome est le 16S ou petite sous-unité de l'ARNr (SSU). Le gène *rri* code pour l'ARN ribosomique 23S le gène *rri* code pour l'ARN ribosomique 16S et le gène *rri* code pour l'ARN ribosomique 5S. Un ARNr 4,5S, aussi associé à la sous-unité 50S du ribosome, a été trouvé dans l'ADNcp des plantes terrestres évoluées.

Il y a entre 20 et 40 gènes *trn* codant pour les 30-35 ARNs de transfert présents dans le génome chloroplastique. Normalement, ces ARNs de transfert sont suffisants pour lire les 61 codons du code génétique universel, en assumant des cas de dégénérescence entre la première base des anticodons commençant par un U et la troisième base des codons associés (Sugiura 1992). Tous les ARNs chloroplastiques se replient en forme de trèfle à quatre feuilles, selon la conformation standard.

Les gènes chloroplastiques *rpoA*, *B*, *C1* et *C2*, codent pour des sous-unités similaires aux sous-unités (*rpoA*), (*rpoB*) et (*rpoC*) de l'ARN polymérase de *E. coli*. L'homologue du gène *rpoC* dans l'ADNcp est divisé en deux sous-unités, soit *rpoC1* et *rpoC2*. La présence de ces gènes indique que les chloroplastes synthétisent plusieurs si ce n'est toutes les sous-unités de l'ARN polymérase.

7. Intérêt et développement des marqueurs chloroplastiques

La petite taille et l'abondance du génome chloroplastique lui donnent de nombreux avantages, et expliquent pourquoi il est le génome le mieux caractérisé.

Les recherches sur l'ADNcp peuvent être divisées en deux catégories : l'une s'intéresse à l'étude des mécanismes qui gouvernent l'évolution de l'ADNcp, l'autre traite de l'utilisation de la variation de l'ADNcp comme un outil permettant d'étudier la diversité génétique des espèces végétales.

Les marqueurs chloroplastiques peuvent servir à de multiples applications. Les génomes chloroplastiques permettent de révéler une structuration génétique plus forte, et constituent donc une source idéale de marqueurs dans les études de diversité génétique.

Une des applications pour laquelle le génome chloroplastique s'avère être performant est celle de la phylogéographie. En effet, l'analyse du polymorphisme dans les génomes transmis uniquement par les graines autorise à retracer les voies de recolonisation suivies par les "lignées" issues des différentes zones refuges. L'hérédité maternelle et l'absence de recombinaison de l'ADN cytoplasmique au moment du brassage génétique lors de la fécondation facilitent la reconstruction de la généalogie des différents variants génétiques. Les génomes cytoplasmiques des plantes présentent également la caractéristique d'évoluer plus lentement en séquence que le génome nucléaire (le taux de substitution de l'ADN mitochondrial est inférieur d'un tiers à celui du génome chloroplastique, lui-même évoluant moitié moins rapidement que le génome nucléaire, (Wolfe *et al.* 1987). Le génome chloroplastique apparaît comme une entité dynamique évoluant lentement, tant dans sa taille, dans sa structure, que dans son répertoire de gènes et dans leur ordonnancement (Downie & Palmer 1992). Malgré sa nature très conservatrice, le taux d'évolution du génome chloroplastique est parfois suffisant pour effectuer des études de phylogéographie intraspécifique.

D'autres applications peuvent être envisagées concernant les marqueurs chloroplastiques. Les amorces chloroplastiques, lorsqu'elles sont chevauchantes, peuvent être utilisées dans des projets de séquençage de génomes chloroplastiques. De plus, le génome chloroplastique, tout comme le génome mitochondrial, constitue un outil pour l'étude de l'hérédité des chloroplastes. Il peut également être utilisé dans le cadre d'études traitant des interactions entre les espèces au cours de l'évolution : phénomènes d'introgession et d'hybridation, et origine de la polyploidie (Clegg *et al.* 1991 ; Rieserberg & Soltis 1991).

La combinaison de ces différentes caractéristiques fait des génomes chloroplastiques des outils pertinents pour la biologie évolutive. Obtenir un nombre important de marqueurs constitue donc un objectif essentiel pour de nombreuses études. De plus, du fait de sa nature conservée (séquence et ordre des gènes), les outils développés pour le génome chloroplastique sont transférables chez de nombreuses espèces. Les efforts de mise au point de marqueurs se sont concentrés sur la région LSC qui évolue plus rapidement que l'IR et contient plus de séquences non codantes que la SSC, augmentant ainsi la probabilité d'obtenir du polymorphisme. En effet, la recherche de polymorphismes s'est focalisée sur les introns et les espaces inter-géniques qui constituent de bons candidats. Ces régions non codantes subissent

une pression de sélection moins forte que celle exercée sur les exons, et par conséquent les mutations s'y accumulent plus rapidement (Clegg & Zurawski 1992). Depuis plusieurs années, de nombreuses paires d'amorces chloroplastiques ont été mises au point.

8. Le génome chloroplastique des monocotylédones et du palmier dattier en particulier

Depuis l'achèvement du premier génome chloroplastique de tabac (*Nicotiana tabacum*) (Shinozaki *et al.* 1986), il y a eu de nombreuses découvertes sur les réarrangements et les expansions IR dans les génomes chloroplastiques, la plus remarquable est celle de *Pelargonium x hortorum*, dans le génome duquel de nombreux réarrangements et des grandes expansions ont été trouvés (Chumley 2006). Des pertes de gènes ont été trouvées fréquemment dans les génomes des angiospermes (Mardanov *et al.* 2008 ; Hansen *et al.* 2007 ; Chang *et al.* 2006). Les plantes entièrement séquencées pour leur génome chloroplastique, représentent une source d'information qui sera de plus en plus importante.

Le domaine de la génomique végétale connaît une nouvelle accélération, depuis 2000. 132 génomes chloroplastiques sont à ce jour entièrement séquencés et disponibles dans GenBank, dont 25 sont des génomes chloroplastiques des Monocotylédones (Tableau1, Figure2).

Le génome chloroplastique du palmier dattier a été séquencé et annoté par Yang *et al.* (2010), en utilisant la technique de pyroséquençage (technologie de séquençage à haut débit). Yang *et al.* (2010) ont décrit le génome chloroplastique du *Phoenix dactylifera* comme contenant 112 gènes, beaucoup de séquences intergéniques ainsi que de nombreuses substitutions d'un nucléotide (SNP). Le génome chloroplastique de *Phoenix dactylifera* se présente comme une molécule circulaire de 158462 pb. Il présente une structure quadripartite avec une IR de 27276 pb et deux régions simples copies de 86198 pb et 17712 pb (respectivement LSC et SSC) (Figure 2).

Similaire aux autres génomes chloroplastiques (Gao *et al.* 2009 ; Raubeson *et al.* 2007), le génome chloroplastique du palmier dattier est également riche en A et T (62.77%). La valeur décroît dans les régions non-codantes, codantes pour les protéines, l'ARNt et l'ARNr, qui est 66.60%, 61.03%, 57.94% et 52.19%, respectivement. Contrairement à la plupart des monocotylédones, le palmier dattier a un génome chloroplastique typique et similaire à celui du tabac.

Tableau 1 : Génomes chloroplastiques entièrement séquencés des monocotylédones.

Famille	Genres/espèces	Numéro d'Accession	Longueur (Pb)
Acoraceae	<i>Acorus americanus</i>	NC_010093	153,819
Acoraceae	<i>A. calamus</i>	NC_007407	153,821
Araceae	<i>Lemna minor</i>	NC_010109	165,955
Arecaceae	<i>Phoenix dactylifera</i> L.	GU811709	158,462
Dioscoreaceae	<i>Dioscorea elephantipes</i>	NC_009601	152,609
Orchidaceae	<i>Phalaenopsis aphrodite</i>	NC_007499	148,964
Orchidaceae	<i>Oncidium Gower Ramsey</i>	NC_014056	146,484
Poaceae	<i>Agrostis stolonifera</i>	NC_008591	136,584
Poaceae	<i>Anomochloa marantoidea</i>	NC_014062	138,412
Poaceae	<i>Bambusa oldhamii</i>	NC_012927	139,350
Poaceae	<i>Brachypodium distachyon</i>	NC_011032	135,199
Poaceae	<i>Dendrocalamus latiflorus</i>	NC_013088	139,394
Poaceae	<i>Festuca arundinacea</i>	NC_011713	136,040
Poaceae	<i>Hordeum vulgare subsp. vulgare</i>	NC_008590	136,462
Poaceae	<i>Lolium perenne</i>	NC_009950	135,282
Poaceae	<i>Oryza nivara</i>	NC_005973	134,494
Poaceae	<i>O. sativa Indica Group</i>	NC_008155	134,496
Poaceae	<i>O. sativa Japonica Group</i>	NC_001320	134,525
Poaceae	<i>Saccharum hybrid</i>	NC_005878	141,182
Poaceae	<i>S. officinarum</i>	NC_006084	141,182
Poaceae	<i>Sorghum bicolor</i>	NC_008602	140,754
Poaceae	<i>Triticum aestivum</i>	NC_002762	134,545
Poaceae	<i>Coix lacryma-jobi</i>	NC_013273	140,745
Poaceae	<i>Zea mays</i>	NC_001666	140,384
Typhaceae	<i>Typha latifolia</i>	NC_013823	161,572

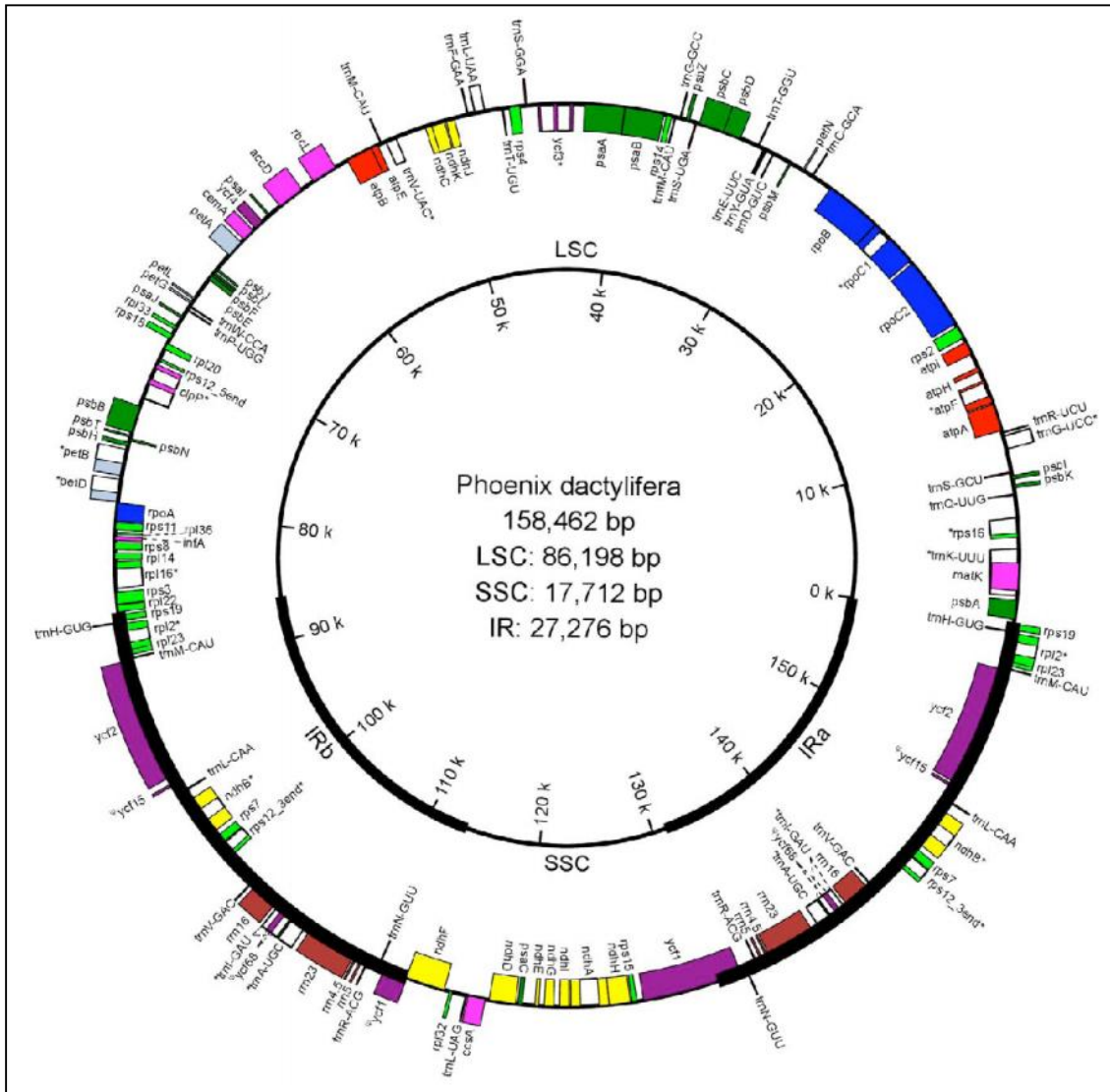


Figure 2 : La carte du génome chloroplastique du palmier dattier. Les gènes représentés à l'extérieur du cercle sont transcrits dans le sens des aiguilles d'une montre, tandis que ceux représentés à l'intérieur sont transcrits dans le sens inverse des aiguilles d'une montre. La ligne épaisse indique IRs. La coordonnée du génome est représentée dans le cercle intérieur. Les gènes appartenant à différents groupes sont codés par couleur (Y, pseudogène). Doi: 10.1371 / journal.pone.0012762.g001 (Yang *et al.* 2010).

MATERIEL & METHODES

1. Matériel végétal et zones d'étude

Le matériel végétal est constitué de jeunes folioles collectées à partir de palmiers adultes âgés de 12 à 15 ans, lorsque cela était possible. Pour les cultivars rares l'âge n'a pu être déterminé. La collection de variétés algériennes était choisie sur les bases de la qualité des dattes, la résistance aux maladies et certaines sur leur rareté.

Nous avons récolté 414 échantillons (individus) de palmier dattier (*Phoenix dactylifera* L.) à partir de 10 oasis de palmier dattier réparties dans le nord-est (4 oasis), le nord-centre (2 oasis), le nord-ouest (3 oasis), et le sud (1 oasis), du Sahara Algérien (Figure 3).

Ces 10 oasis sont : Biskra dans les Zibans (13 variétés / 72 individus), Oued Souf (21 variétés / 52 individus), Touggourt (30 variétés / 54 individus), Ouargla (6 variétés / 29 individus), Ghardaia dans le M'Zab (16 variétés / 65 individus), El Ménia ou Golea (10 variétés / 68 individus), Timimoun dans le Gourara (24 variétés / 24 individus), Adrar dans le Touat (21 variétés / 23 individus), Béni Abbès dans la Saoura (20 variétés / 20 individus) et Tamanrasset (4 variétés / 7 individus) (Figure 3, Annexe 1).

Pour couvrir la plus grande diversité génétique possible de *Phoenix dactylifera*, l'échantillonnage inclus dans cette étude 481 génotypes, provenant de 13 pays appartenant aux zones de culture traditionnelles occidentales (de la Mauritanie à l'Égypte) et orientales (du Djibouti au Pakistan) du dattier (Annexe 2).

De plus, cent treize échantillons appartenant à 13 espèces du genre *Phoenix*, ont été analysés aussi dans notre travail : *Phoenix andamanensis* (1 individu), *Phoenix acaulis* (1 individu), *Phoenix canariensis* (19 individus), *Phoenix atlantica* (4 individus), *Phoenix caespitosa* (1 individu), *Phoenix loureiroi* (30 individus), *Phoenix paludosa* (2 individus), *Phoenix pusilla* (2 individus), *Phoenix reclinata* (19 individus), *Phoenix roebelenii* (20 individus), *Phoenix rupicola* (1 individu), *Phoenix sylvestris* (6 individus), *Phoenix theophrasti* (7 individus).

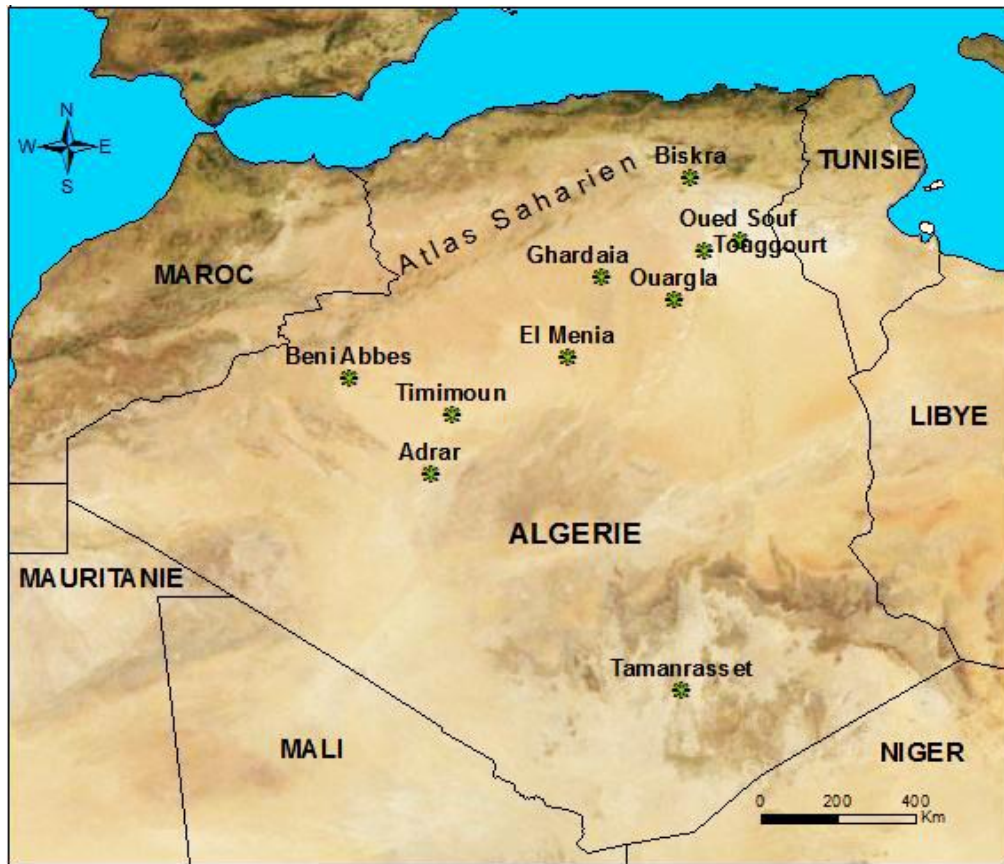


Figure 3 : Distribution géographique du palmier dattier en Algérie. Biskra (72 individus), Oued Souf (52 individus), Touggourt (54 individus), Ouargla (29 individus), Ghardaia (65 individus), El Ménia (68 individus), Timimoun (24 individus), Adrar (23 individus), Béni Abbès (20 individus) et Tamanrasset (7 individus).

2. Extraction et quantification d'ADN

La matière fraîche a été déshydratée et séchée avec du silicagel juste après la collecte et conservée à température ambiante.

Les échantillons ont été broyés en poudre fine à l'aide d'un homogénéisateur TissueLyser (Qiagen), en présence de billes d'acier. L'ADN est extrait à partir de 45 mg de jeune tissu foliaire séché en utilisant le Kit DNeasy Plant MiniKit (Qiagen) suivant les instructions du fabricant, dont le déroulement exact est fourni en annexe (Annexe 3). L'intégrité et la qualité de l'ADN a été vérifiée par une électrophorèse sur gel d'agarose à 1% en présence de bromure d'éthidium (1µg/ml) afin de tester son aspect qualificatif sous rayons ultraviolets (Sambrook *et al.* 1989).

La concentration en ADN des différents échantillons a été déterminée selon la méthode spectrophotométrique par le NanoDrop et NanoQuant.

L'ADN total extrait est conservé sous forme de solution (100µl) à -20°C avant analyse.

3. Génotypage microsatellite

3.1. Choix des marqueurs génétiques moléculaires

Les principales sources de marqueurs moléculaires proviennent soit d'un polymorphisme de séquence (par exemple substitution, insertion, délétion), soit d'un polymorphisme de nombre d'unités de répétitions (microsatellites et minisatellites).

Nous avons fait le choix d'utiliser les marqueurs microsatellites pour l'étude de la diversité du palmier dattier car ils constituent d'excellents marqueurs génétiques, les autres marqueurs testés par nos prédécesseurs n'ayant pas donné les résultats escomptés.

3.2. Marqueurs microsatellites utilisés

Dix-huit marqueurs microsatellites nucléaires spécifiques au palmier dattier ont été utilisés pour cette étude afin de caractériser sa diversité moléculaire (Tableau 2).

Nous avons utilisé dix microsatellites dinucléotidiques (GA)_n issus des travaux de Billote (2004). Ces marqueurs dinucléotidiques appartiennent à des régions inconnues non codantes, et ils ont été choisis pour leur grande variabilité (Henderson *et al.* 2006 ; Pintaud *et al.* 2010). Deux marqueurs microsatellites dinucléotidiques décrits par Adam *et al.* (2010) ont été retenus. Quatre marqueurs microsatellites situés dans des régions codantes développés à l'IRD de Montpellier ont été utilisés au cours de cette étude (trois trinuéotidiques et un

microsatellite hexanucléotidiques) (Aberlenc-Bertossi *et al.* 2014). Les marqueurs AG1 et AP3 (Ludeña *et al.* 2011, Zehdi *et al.* 2015) ont complété le jeu de microsatellites utilisés pour le génotypage de *P. dactylifera* (Tableau 2).

Ces loci ont été analysés sur 192 échantillons (variétés algériennes) dont 178 clones femelles représentant 114 variétés identifiées, 11 clones femelles non identifiés, un génotype femelle issu de semence et 2 palmiers mâles. Les individus échantillonnés ont été répartis comme suit: Biskra (13 variétés / 20 individus), Oued Souf (22 variétés / 30 individus), Touggourt (30 variétés / 32 individus), Ouargla (5 variétés / 5 individus), Ghardaia (17 variétés / 29 individus), El Méria (10 variétés / 17 individus), Timimoun (16 variétés / 17 individus), Adrar (18 variétés / 20 individus), Béni Abbès (14 variétés / 15 individus) et Tamanrasset (4 variétés / 7 individus) (Annexe 1).

481 génotypes du dattier de différentes origines géographiques ont été analysés. Ces génotypes sont issus de différentes collections internationales en vue des analyses de diversité génétique mondiale (Annexe 2).

Ainsi, ces marqueurs microsatellites ont été testés sur cent treize échantillons appartenant à 13 espèces du genre *Phoenix*. L'analyse de ces microsatellites sur les espèces du genre *Phoenix* a permis d'étudier la phylogénie des 14 espèces.

Tableau 2 : Caractéristiques des 18 loci microsatellites de palmier dattier (*Phoenix dactylifera* L.).

Locus SSR	Nom court	Séquence de l'amorce Forward (5' → 3')	Séquence de l'amorce Reverse (5' → 3')	Motifs SSR	Taille (pb)	Tm (C°)	Référence	Accession GeneBank
mPdCIR010	Pd10	ACCCCGGACGTGAGGTG	CGTCGATCTCCTCCTTTGTCTC	(GA) ₂₂	164	56	Billotte <i>et al.</i> 2004	AJ571673
mPdCIR015	Pd15	AGCTGGCTCCTCCCTTCTTA	GCTCGGTTGGACTTGTCT	(GA) ₁₅	135	52	Billotte <i>et al.</i> 2004	AJ571674
mPdCIR016	Pd16	AGCGGGAAATGAAAAGGTAT	ATGAAAACGTGCCAAATGTC	(GA) ₁₄	134	52	Billotte <i>et al.</i> 2004	AJ571675
mPdCIR025	Pd25	GCACGAGAAGGCTTATAGT	CCCCTCATTAGGATTCTAC	(GA) ₂₂	230	49	Billotte <i>et al.</i> 2004	AJ571676
mPdCIR032	Pd32	CAAATCTTTGCCGTGAG	GGTGTGGAGTAATCATGTAGTAG	(GA) ₁₉	300	52	Billotte <i>et al.</i> 2004	AJ571677
mPdCIR035	Pd35	ACAAACGGCGATGGGATTAC	CCGCAGCTCACCTCTTCTAT	(GA) ₁₅	188	54	Billotte <i>et al.</i> 2004	AJ571678
mPdCIR057	Pd57	AAGCAGCAGCCCTTCCGTAG	GTTCTCACTCGCCCAAAAATAC	(GA) ₂₀	282	55	Billotte <i>et al.</i> 2004	AJ571682
mPdCIR063	Pd63	CTTTTATGTGGTCTGAGAGA	TCTCTGATCTTGGGTTCTGT	(GA) ₁₇	167	50	Billotte <i>et al.</i> 2004	AJ571683
mPdCIR078	Pd78	TGGATTTCCATTGTGAG	CCCGAAGAGACGCTATT	(GA) ₁₃	121	50	Billotte <i>et al.</i> 2004	AJ571685
mPdCIR085	Pd85	GAGAGAGGGTGGTGTATT	TTCATCCAGAACCACAGTA	(GA) ₂₉	179	50	Billotte <i>et al.</i> 2004	AJ571686

Locus SSR	Nom court	Séquence de l'amorce Forward (5'→ 3')	Séquence de l'amorce Reverse (5'→ 3')	Motifs SSR	Taille (pb)	Tm (C°)	Référence	Accession GeneBank
PdAG1-ssr	AG1	TCTGATTTTCGTTTACTTCTTAGGA	TTCATATTCAGTTGTCGGGTGTA	(GA)	260	52	Ludeña <i>et al.</i> 2011	
PdCUC3-ssr1	CUC3-1	CGTGGACTCATGACTCGCATGTCC	GGTCCTTGCCGGTGGCCTTC	(GT) ₁₄	330	60	Adam <i>et al.</i> 2010	HM622273
PdCUC3-ssr2	CUC3-2	ACATTGCTCTTTTGCCATGGGCT	CGAGCAGGTGGGGTTCGGGT	(GA) ₂₂	350	59	Adam <i>et al.</i> 2010	HM622273
PdAP3-ssr	AP3	GAGAAATAGAGAGCTGTGCAAG	CTGCAGTACTCGGAGAACTTG	(GA) ₂₅	331	57	Zehdi <i>et al.</i> 2015	KC188337
mPdIRD013	P13	GCGGAGACAGGAGATGGTAA	CTTGACTGCTTCTGCTGCTG	(CAC) ₆	204	60	Aberlenc-Bertossi <i>et al.</i> 2014	PDK_20s1496731g002
mPdIRD031	P31	GCAGGTGGACTGCAAAATCT	CTATTGGGGTGCTGATCCAT	(CCA) ₇	198	60	Aberlenc-Bertossi <i>et al.</i> 2014	PDK_20s1419261g003
mPdIRD033	P33	GGAGCATAACAGTGGGTTTGC	CAGCCTGGGAATGAGGATAG	(CAG) ₇	199	60	Aberlenc-Bertossi <i>et al.</i> 2014	PDK_20s1569281g001
mPdIRD040	P40	GAGAGATGCGTCAGGGAATC	CCAGAATCTTCCAAGCAAGC	(CCAGTG) ₄	193	60	Aberlenc-Bertossi <i>et al.</i> 2014	PDK_20s1327401g002

3.3. Réaction d'amplification en chaîne de l'ADN microsatellite nucléaire et génotypage

Pour tous les extraits d'ADN, la concentration a été homogénéisée à 10 ng/μl. Les réactions d'amplification ont été réalisées dans un volume final de 20 μl contenant 10 ng d'ADN génomique, 2 μl tampon réactionnel 10X, 5 pmol de chaque amorce sens (forward) et anti-sens (reverse), 0.2 mM de chaque désoxynucléotide (dNTPs), 2 mM de MgCl₂, et 1 unité de Taq polymérase (Sigma). Les amorces sens (forward) ont été marquées 5' avec l'un de trois composés fluorescents (6-FAM, NED ou HEX) pour permettre une analyse avec des séquenceurs automatisés. La PCR a été réalisée en utilisant un thermocycleur *Eppendorf Mastercycler pro vapo.protect* équipé d'une technologie de protection contre les vapeurs (AG, Hamburg, Allemagne). Pour chacune de nos deux plaques PCR, nous avons introduit un témoin négatif (tous les composants de la réaction sans l'ADN matrice) pour vérifier qu'il n'y a pas eu de contamination externe.

Les conditions de PCR étaient une dénaturation initiale à 95°C pendant 5 minutes, suivie par 35 cycles composés chacun d'une dénaturation à 95 °C pendant 30 secondes, une hybridation à 57 °C pendant 60 secondes et d'une élongation à 72 °C pendant 30 secondes ; une dernière étape d'élongation à 60 °C pendant 30 minutes met fin au programme (Encadré 4).

Les produits amplifiés ont été analysés sur un analyseur génétique ABI 3130XL (Applied Biosystems, USA) et les allèles ont été évalués à l'aide du GeneMapper V3.7 logiciel (Applied Biosystems). Grâce au logiciel GeneMapper, nous avons pu déterminer la taille des 2 allèles de chaque locus microsatellite pour chaque échantillon analysé.

Le génotypage microsatellite a été réalisé sur la plateforme de génotypage ADNid.

Encadré 4 : Polymerase chain reaction (PCR)

La PCR : La technique de polymérisation en chaîne (en anglais « polymerase chain reaction ») ou PCR, est une méthode d'amplification in vitro, qui permet d'amplifier des millions de fois un unique fragment d'ADN. L'invention de la PCR revient à Kary Mullis dans les années 80. Il obtint pour cette découverte le Prix Nobel de Chimie en 1993.

Une réaction de PCR nécessite :

- de l'ADN à amplifier
- des amorces ou primers
- des desoxynucléotides triphosphates (dATP, dCTP, dGTP, dTTP)
- une ADN polymérase thermorésistante (la Taq polymérase)
- du magnésium (Mg^{++}) indispensable au fonctionnement de l'ADN polymérase

Ces 5 ingrédients sont mélangés dans un tube à essai et soumis à différents cycles de températures. Chaque cycle de PCR comprend trois étapes :

1^{ère} étape : **la dénaturation** (94 °C).

L'ADN est chauffé à 94°C. Les deux brins de l'ADN se séparent. On parle de dénaturation de l'ADN.

2^{ème} étape : **l'hybridation** (40-65 °C).

En descendant la température, les amorces s'apparient (s'hybrident) par complémentarité à leurs séquences cibles sur l'ADN.

3^{ème} étape : **l'élongation** (72 °C).

A cette température, la Taq synthétise les brins d'ADN complémentaires en ajoutant des desoxynucléotides triphosphates à la suite des amorces. Le temps d'élongation dépend de la taille du fragment à amplifier et de la vitesse de polymérisation de l'enzyme (la Taq ajoute environ 1000 nucléotides par minute).

Environ une trentaine de cycle sont effectués dans une expérience de PCR classique.

3.4. Diversité et analyse statistique

Une fois les allèles présents pour chaque marqueur identifié grâce à GeneMapper, on établit une matrice de données des allèles pour l'ensemble des accessions.

Pour cette étude, nous avons utilisé les logiciels Excel, PowerMarker Version 3.25 (Liu & Muse 2005), Genetix 4.05 (Belkhir *et al.* 2004), Fstat (Goudet 2002), R (<http://www.R-project.org>), Mega version 5.05 (Kumar *et al.* 2004) et STRUCTURE version 2.3 (Pritchard *et al.* 2000).

Excel est le logiciel de base. Il nous a servi pour la construction des matrices. Les matrices comportent en colonnes l'ensemble des allèles existant pour un même locus et en lignes les individus.

3.4.1. Paramètres de diversité génétique

La première étape des analyses est celle des statistiques descriptives. Les paramètres de la diversité génétique, à savoir le nombre d'allèles par locus (N_a), les fréquences alléliques, l'hétérozygotie et la valeur de contenu d'information sur le polymorphisme (valeur PIC) ont été estimés avec le logiciel PowerMarker Version 3.25 (Liu & Muse 2005).

Les indices F_{IS} et F_{ST} ont été calculés selon la méthode de Weir & Cockram (1984) en utilisant le logiciel Fstat (Goudet 2002) et Genetix 4.05 (Belkhir *et al.* 2004). La richesse allélique a été calculée avec le logiciel Fstat (Goudet 2002).

Un test non paramétrique de Wilcoxon (statistique inférentielle) a été utilisé pour déterminer l'importance des différences significatives de diversité entre les groupes à l'aide du logiciel R (<http://www.R-project.org>). Les valeurs de probabilité (valeur P ou P value) du test sont données pour chaque comparaison par paires de populations. Le niveau de significativité utilisé est de $P < 0,05$. La correction séquentielle de Bonferroni a été utilisée pour les comparaisons multiples (Encadré 5).

Encadré 5 : Statistiques et paramètres de diversité génétique

Statistique descriptive : Ensemble de méthodes de dénombrement, de classement, de synthèse et de présentation de données quantitatives relatives à un ensemble d'individus.

Statistique inférentielle : Ensemble des méthodes et des théories permettant de généraliser à une population de référence des conclusions obtenues à partir de l'étude d'un échantillon extrait de cette population.

Correction de Bonferroni : C'est une correction apportée au seuil de significativité en cas de comparaison multiple pour compenser le fait que lors de tests multiples, un test peut apparaître significatif par simple fait au hasard. La correction donne un seuil $\alpha' = \alpha/n$, avec n le nombre de tests fait.

Paramètres de diversité génétique :

- Le nombre moyen d'allèles et de génotypes par locus : est le premier indicateur de la diversité génétique d'une population.
- Le taux de polymorphisme, ou Polymorphic Information Content (PIC) : il met en évidence la capacité d'un marqueur à détecter un polymorphisme dans une population (Botstein *et al.* 1980).
- L'hétérozygotie observée (H_o) : c'est la moyenne des fréquences des hétérozygotes observés à chacun des locus étudié. C'est le pourcentage réel d'individus hétérozygotes dans un échantillon.
- L'hétérozygotie attendue (H_e) : c'est la probabilité qu'à un locus donné, deux allèles choisis aléatoirement soient différents.
- F_{IS} : Déficit en hétérozygote à l'intérieur des populations. L'indice F_{IS} décrit l'écart à la panmixie à l'échelle intra-populationnelle
- F_{ST} : Déficit en hétérozygote dû à la subdivision de la population totale en sous-populations ; correspond à la différenciation des populations par rapport au total. L'indice F_{ST} décrit la différenciation allélique entre les sous-populations.

3.4.2. Isolement par la distance

L'existence d'isolement par la distance a été testée en tenant compte des paires de populations. Pour cela, des tests de Mantel de corrélations entre les matrices des distances géographiques et des distances génétiques représentées par la valeur de F_{ST} , entre paires de populations ont été effectués sous R (<http://www.R-project.org>), et la significativité du coefficient de Pearson (r) a été établie en effectuant 1000 permutations.

3.4.3. Analyse discriminante en composante principale (DAPC)

L'analyse discriminante en composante principale (DAPC) (Jombart *et al.* 2010) permet l'identification du nombre de clusters génétiques via l'étude de la variation génétique des individus échantillonnés. Cette méthode multivariée définit un modèle qui sépare la variation entre groupes (variabilité génétique structurée) de celle à l'intérieur du groupe (variabilité génétique aléatoire), estimant ainsi une variable maximisant la première et minimisant la deuxième. Le principal avantage de cette approche multivariée est l'absence d'hypothèse préalable sur la génétique des populations (équilibre HW, déséquilibre de liaison). Le nombre de clusters génétiques est identifié à l'aide d'un modèle statistique et de sa vraisemblance associée (fonction *find.cluster*) qui correspond à la plus faible valeur du critère bayésien d'information (BIC). Le nombre de clusters génétiques permet ensuite d'effectuer la DAPC. Les analyses ont été réalisées avec le logiciel R à l'aide du package Adegenet.

3.4.4. Arbre phylogénétique et analyse Bayésienne de regroupement

L'analyse des clusters (arbre phylogénétique) a été réalisée en générant une matrice de distance génétique en calculant les fréquences alléliques partagées (Shared Alleles Distance) et utilisant l'algorithme de neighbor-joining à l'aide du logiciel Mega version 5.05 (Kumar *et al.* 2004). Nous avons également évalué la classification basée sur l'approche bayésienne développée dans STRUCTURE version 2.3 (Pritchard *et al.* 2000), sous les modèles de mixité ou « admixture » et de fréquences alléliques corrélées. Il permet d'affecter les individus dans des sous-populations liées à des informations multilocus. Pour ce faire, nous avons mené une série d'analyses pour différents nombres de groupes (K) allant de 1 à 10, avec 10 répétitions pour chaque valeur donnée de K. Pour chaque essai (valeur de K), nous avons utilisé 10000 chaînes de Markov Monte Carlo (MC répétitions) et 100000 itérations pour la période de chauffe (Burn-in period). Le nombre le plus probable de clusters a été évalué à l'aide de likelihood ainsi que de la quantité *ad hoc* basée sur le taux de changement du second ordre

dans la probabilité logarithmique des données entre différentes valeurs de K (Evanno *et al.* 2005). L'estimation du nombre de populations (K) est calculée suivant la méthode donnée pour Evanno *et al.* (2005). La diversité (nombre d'allèles, hétérozygotie) a été évaluée dans les différents groupes et comparée en utilisant le test de Wilcoxon. Nous avons assigné un individu à un groupe si son ascendance dans ce groupe était > 75% (Vigouroux *et al.* 2008). Pour vérifier la différence de fréquence des chlorotypes entre les groupes, nous avons effectué un test de Chi-carré, en utilisant SPSS Statistics 20.

4. Polymorphisme des séquences de loci chloroplastiques

4.1. Génotypage du minisatellite *psbZ-trnfM*

Cette analyse a été faite sur un échantillonnage de 414 échantillons algériens pour déterminer les chlorotypes présents.

Pour le génotypage avec les amorces *trnG-trnfM*, le mélange réactionnel était le suivant : 1 µl d'ADN, 5 µl de tampon FailSafe PCR PreMixE, 0.1 µl de Taq ADN polymérase, 0.1 µl d'amorce sens, 0.1 µl d'amorce anti-sens, et 3.6 µl qsp d'eau promega pour atteindre un volume final de 10 µl.

Le protocole d'amplification comprend une phase de dénaturation à 95 °C durant 3 minutes, suivie par 30 cycles, comprenant chacun une étape de dénaturation à 94 °C pendant 30 secondes, une étape d'hybridation à 56 °C pendant 1 minute 30 secondes, et une étape d'élongation à 72 °C pendant 1 minute 30 secondes.

Le génotypage a été effectué grâce à l'analyseur QIAXCEL. Ce système Qiaxcel utilise l'électrophorèse sur gel capillaire pour permettre une séparation rapide des acides nucléiques basée sur leur taille, et permet le traitement de 96 échantillons en une seule opération. Nous avons utilisé le kit de haute résolution QX DNA et le marqueur de taille d'ADN QX 25-500 pb. La collecte et l'analyse des données (taille des bandes) ont été réalisées avec le logiciel QIAxcel ScreenGel.

Une liste d'individus avec génotype algérien et d'endroits de prélèvement est donnée dans un fichier de données supplémentaire (Annexe 4).

Un génotypage ABI de ce minisatellite a été effectué sur 481 génotypes du *Phoenix dactylifera* provenant de différentes origines géographique (13 pays), et 113 échantillons du genre *Phoenix*.

4.2. Séquençage des amorces chloroplastiques

4.2.1. Définition des locus chloroplastiques

Dans cette étude nous avons analysé le polymorphisme chloroplastique d'un échantillon représentatif composé de 24 variétés différentes, prélevées des palmeraies algériennes. Cette expérimentation renferme aussi une variété 'Iberica' originaire d'Elche (Espagne), et une variété 'Khalass' d'Arabie Saoudite.

Les analyses ont portés sur six locus chloroplastiques (Figure 4):

atpH-atpI, *psbZ-trnfM*, *rps15-ycf1*, *psbA-trnH*, *rpoB*, et *matk* exon (Tableau 3).

Il est admis que ces amorces sont des marqueurs d'espèces intéressants pour les palmiers compte tenu de leur variabilité (Scarcelli *et al.* 2011).

Tableau 3 : Caractéristiques des locus utilisés pour l'étude du polymorphisme chloroplastique.

Région des locus	Type des locus	Nom des locus	Nom des amorces	Séquences (5'→3')
LSC	Espaceur intergénique (IGS)	atpH/atpI	atpH-IGS-F	CCAGCAGCAATAACRGAAGC
			atpI-IGS-R	TTCAAGCTCTTATTTTTGCAACKT
LSC	Espaceur intergénique (IGS)	psbZ/trnfM	psbZ-IGS-F	GCTACMTCATTATGGATTGG
			trnfM-IGC-R	GCGGAGTAGAGCAGTTTGGT
LSC	Espaceur intergénique (IGS)	psbA/trnH	psbA-IGS-F	GTTATGCATGAACGTAATGCTC
			trnH-IGS-R	CGCGCATGGTGGATTCACAAATC
SSC	Espaceur intergénique (IGS)	rps15/ycf1	rps15-IGS-F	CGTTGACGTTTTCCYAGRATT
			rps15-IGS-R	GCTTGTATGAATCGBTATTGGTTT
LSC	Intron	rpoB intron	rpoB-In-F	ATGCAACGTCAAGCAGTTCC
			rpoB-In-R	GATCCCAGCATCACAATTCC
LSC	Exon	matk exon	matk-Ex-F	RAKAATATCCAATACCAA
			matk-Ex-R	ACCATATYGYACTATGTATC

4.2.2. Réaction d'amplification en chaîne (PCR)

Les concentrations d'ADN extraits ont été homogénéisées par dilution et ramenées à 30 ng/μl. Les réactions d'amplification en chaîne (PCR) ont été réalisées à l'aide d'un thermocycleur Biometra, dans un volume final de 25 μl, et contenant le mélange réactionnel suivant : 2 μl d'ADN cellulaire total, 12.5 μl de tampon FailSafe PCR PreMixE, 0.25 μl de Enzyme Mix PCR (Taq DNA polymérase), et 0.4 μl de chaque amorce sens (forward) et anti-sens (reverse) à 25 μM. Pour atteindre le volume réactionnel, 9.45 μl d'eau pure (Nuclease Free Water) ont été ajoutés. Un témoin négatif correspondant au mélange réactionnel contenant tous les composants excepté l'ADN cellulaire a été utilisé.

Le cycle d'amplification consiste en une première étape de dénaturation totale d'ADN à 96 °C pendant 3 minutes, suivi d'un cycle de dénaturation de 30 secondes à 94 °C, d'une hybridation à 60 °C pendant 1 minute 30 secondes (cette température est identique pour tous les locus utilisés pour le séquençage dans cette étude) et d'une élongation à 72 °C pendant 1 minute 30 secondes. Ce cycle est répété 34 fois. Une étape d'élongation finale de 10 minutes à 72 °C achève l'amplification finale par PCR.

Les produits PCR à séquencer ont été séparés sur gel d'agarose à 2.5% afin de vérifier la concentration de l'amplimère (Figure 5).

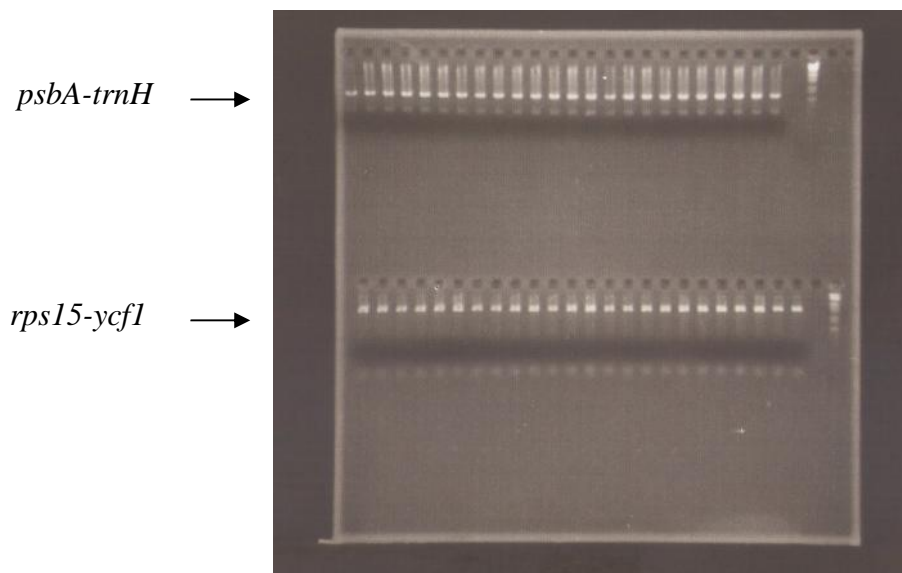


Figure 5 : Séparation sur gel d'agarose (2.5%) des produits d'amplification par PCR de l'ADN des 24 variétés de *P. dactylifera* en utilisant les *psbA-trnH* et *rps15-ycf1*.

4.2.3. Séquençage des produits PCR

25 µl de la réaction PCR sont purifiés grâce au kit Agencourt Ampure. L'estimation de la concentration des amplimères purifiés a été faite selon la méthode spectrophotométrique par le NanoDrop.

Les réactions de séquençage ont été effectuées grâce à l'utilisation du kit BigDye Terminator v3.1 Cycle Sequencing, de produits PCR purifiés et dilués à 5 ng/µl, et des amorces ayant servi à la PCR à une concentration de 10 µM.

Le milieu réactionnel contenant un volume final de 10 µl est le suivant : 1 µl de BigDye Terminator, 2 µl de 5xSequencing Buffer, 1 µl d'amorce reverse ou forward à 10 µM, 1 µl d'eau Promega, et 5 µl d'ADN.

La réaction a été effectuée dans un thermocycleur durant 39 cycles et comprenait les étapes suivantes : 10 secondes à 96 °C, 20 secondes à 50 °C et 3 minutes 30 secondes à 60 °C. L'étape à 96 °C sert à dénaturer l'ADN, l'étape à 50 °C à l'hybridation et l'étape à 60 °C à l'élongation des molécules d'ADN.

Les produits de la réaction ont été purifiés par le kit Agencourt Clean SEQ.

L'électrophorèse a été effectuée sur un séquenceur 16 capillaires AIB (Applied Biosystems, DNA Sequencer).

4.2.4. Alignement et analyse des séquences

Les séquences obtenues ont été alignées et assemblées à l'aide du logiciel GENEIOUS version 5.3.6 (Drummond *et al.* 2010). L'utilisation de ce logiciel permet l'alignement et l'obtention de séquences consensus selon des critères menant à la création de contigs de séquences. Ce logiciel nous a permis d'identifier les séquences intéressantes polymorphes chez les différentes variétés du palmier dattier. Les alignements ont été étudiés par le logiciel MESQUITE version 2.75 (Maddison & Madisson 2007), afin de déterminer les polymorphismes.

Les polymorphismes observés ont été positionnés en référence à la séquence complète du génome chloroplastique de *P. dactylifera* cv. khalass, disponible dans GenBank (accession NC_013991.2).

RESULTATS ET DISCUSSIONS

Chapitre 1
Caractérisation génétique et phylogénie du
genre *Phoenix*

Introduction

Les recherches sur les palmiers sont un défi pour les chercheurs en raison de leur diversité phénologique et de leur long cycle de reproduction (plusieurs années avant de fleurir). La connaissance de leur biologie et de leurs caractéristiques offre des perspectives de recherche très intéressantes aussi bien au niveau fondamental qu'appliqué. Les originalités de la famille, au niveau de leurs formes et de leurs fonctions, donnent à ce groupe d'espèces un intérêt tout particulier qui n'a fait que s'amplifier au fur et à mesure du développement des outils d'analyse, et plus spécifiquement ceux liés à la génomique.

La famille des palmiers, ou *Arecaceae*, regroupe environ 2500 espèces distribuées dans toutes les régions chaudes du Monde.

L'origine et le début de la diversification des palmiers remonte à environ 100 millions d'années durant le milieu du Crétacé, les fossiles les plus anciens connus datant de 95 Millions d'années (Harley 2006). De nombreuses espèces à intérêt économique et commercial font partie de cette famille : on peut citer évidemment le dattier (*Phoenix dactylifera* L.), le cocotier (*Cocos nucifera* L.), le palmier à huile (*Elaeis guineensis* Jacq.), ou encore le coeur de palmier (*Euterpe edulis* Mart.). Cette famille joue un rôle écologique important tant au niveau du nombre d'espèces que par sa répartition mondiale.

D'un point de vue systématique, la famille des *Arecaceae* est divisée en cinq sous familles d'inégale importance en terme de nombre d'espèces (Asmussen *et al.* 2006 ; Baker *et al.* 2009 ; Couvreur *et al.* 2011): *Arecoideae* (106 genres), *Ceroxyloideae* (8 genres), *Coryphoideae* (46 genres), *Nypoideae* (1 genre) et *Calamoideae* (23 genres), la plus importante étant celle des *Coryphoideae*, composée d'espèces à feuilles pennées et fleurs unisexuées.

Selon Govaerts & Dransfield (2005), le genre *Phoenix* est constitué de 14 espèces distribuées des îles de l'Atlantique à travers l'Afrique, le Sud de l'Europe, le Moyen-Orient, et l'Asie du Sud jusqu'aux Philippines : *Phoenix acaulis*, *Phoenix canariensis*, *Phoenix dactylifera*, *Phoenix atlantica*, *Phoenix caespitosa*, *Phoenix adamanensis*, *Phoenix loureiroi*, *Phoenix paludosa*, *Phoenix pusilla*, *Phoenix reclinata*, *Phoenix roebelenii*, *Phoenix rupicola*, *Phoenix sylvestris*, *Phoenix theophrasti* (Figure 6).

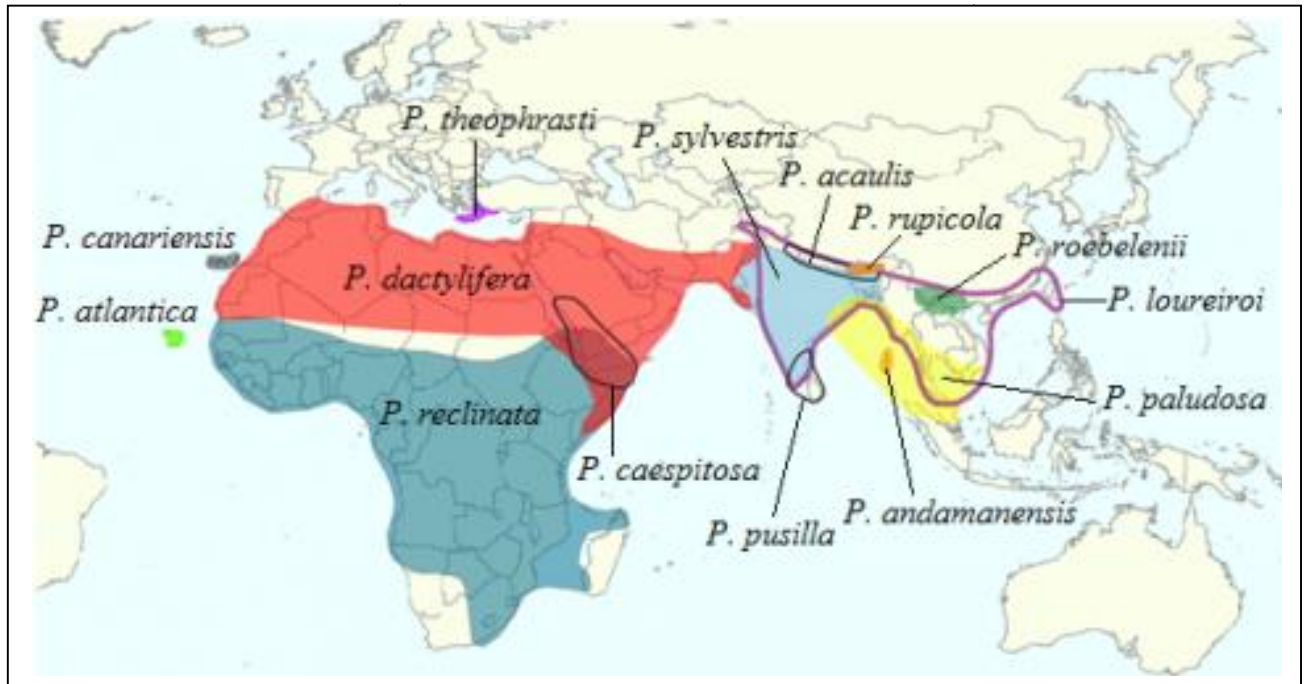


Figure 6 : Carte de distribution du genre *Phoenix* (Gros-Balthazard *et al.* 2013 ; d'après Munier 1973 ; Barrow 1998 ; Henderson 2009).

Le genre *Phoenix* et ses 14 espèces est le seul genre de la tribu *Phoenixaceae* dans la sous-famille des *Coryphoideae* (Asmusen *et al.* 2006 ; Dransfield *et al.* 2008). La position du genre *Phoenix* dans la sous-famille *Coryphoideae* a été confirmée par une analyse phylogénétique de toute la famille des *Arecaceae*, qui comprenait des données de séquences d'ADN chloroplastique et nucléaire, des données de la RFLP d'ADN chloroplastique et des données morphologiques (Baker *et al.* 2009). Au niveau moléculaire, le genre *Phoenix* est fortement divergent de tous les autres genres des palmiers (Asmussen & Chase 2001 ; Hahn 2002) (Tableau 4).

Tableau 4 : Classification de la sous-famille *Coryphoideae*.

Sous Famille	Tribu	Genre
<i>Coryphoideae</i>	<i>Corypheeae</i>	<i>Corypha</i>
	<i>Borasseae</i>	<i>Latania, Borassus, Borassodendron, Lodoicea, Hyphaene, Bismarckia, Medemia, Satranala.</i>
	<i>Caryoteae</i>	<i>Caryota, Arenga, Wallichia.</i>
	<i>Chuniophoeniceae</i>	<i>Chuniophoenix, Kerriodoxa, Nannorrhops, Tahina.</i>
	<i>Cryosophileae</i>	<i>Cryosophila, Chelyocarpus, Coccothrinax, Itaya, Schippia, Thrinax, Hemithrinax, Trithrinax, Zombia.</i>
	<i>Livistoneae</i>	<i>Livistona, Johannesteijsmannia, Lanonia, Licuala, Pholidocarpus, Rhapsis, Chamaerops, Guihaia, Maxburretia, Rhapsidophyllum, Trachycarpus, Acoelorrhapha, Brahea, Colpothrinax, Copernicia, Pritchardia, Saribus, Serenoa, Washingtonia.</i>
	<i>Phoeniceae</i>	<i>Phoenix</i>
	<i>Sabaleae</i>	<i>Sabal</i>

La taxonomie, la phylogénie et l'évolution du genre *Phoenix* ont été évalués par des études morphologiques et moléculaires. Selon Barrow (1998), la combinaison des données morphologiques et moléculaires sur les séquences d'ADN de l'espaceur intergénique 5S (ADN ribosomale nucléaire) a montré l'existence de deux clades d'espèces étroitement liées. Le premier clade inclut *P. dactylifera*, *P. sylvestris*, *P. canariensis*, et *P. atlantica*. Le second groupe comprend les espèces sœurs *P. paludosa* et *P. roebelenni*. Ces données morphologiques et moléculaires ont montrés une faible variation, ce que signifie la nécessité d'employer des marqueurs plus polymorphes pour démêler les modèles de différenciations génétiques dans le genre *Phoenix*.

Henderson *et al.* (2006) ont utilisé des microsatellites pour distinguer entre *P. dactylifera*, *P. sylvestris*, *P. canariensis* et *P. atlantica*. Chacune des quatre espèces a été clairement séparée dans cette étude.

Les données de Pintaud *et al.* (2010) sur le génotypage de 17 SSR, à savoir 16 microsatellites nucléaires dinucléotidiques et un minisatellite chloroplastique dodecanucléotidique, sur 308 accessions de *Phoenix* représentant 12 espèces, ont démontré que toutes les espèces de *Phoenix* sont distinctes, et suggèrent que *P. atlantica* et *P. dactylifera* sont des espèces sœurs.

Le *P. dactylifera* est l'espèce la plus importante cultivée dans les pays du Moyen-Orient et d'Afrique du Nord. Cette espèce a probablement été domestiquée autour de 4000 avant JC, dans la Mésopotamie et la région du Golfe Persique (Zohary & Hopf 2000 ; Tengberg 2012).

Au sein du genre *Phœnix*, les espèces sont proches morphologiquement, et parfois difficiles à distinguer (Pintaud *et al.* 2010).

Le nombre d'espèces reconnu par les chercheurs a changé au cours des dernières décennies. D'abord Munier (1973) a considéré que le genre *Phoenix* est composé de 12 espèces, Uhl & Dransfield (1987) ont ramené le nombre à 17, puis Barrow (1998) a proposé le nombre de 13 espèces compte tenu des fortes similitudes constatées entre plusieurs espèces d'appellation différente. Enfin, Govaerts & Dransfield (2005) et Henderson *et al.* (2006) ont proposé de ramener le nombre d'espèce à 14.

Cette confusion sur le nombre exact d'espèces du genre *Phœnix* résulte de la difficulté de distinction des espèces entre elles. En effet, l'étude de la distribution et de l'écologie des espèces du genre *Phœnix* révèle que la structure génétique du genre est mal connue, du fait des hybridations interspécifiques car tous les *Poenix* possèdent 36 chromosomes. Les espèces de *Phoenix* étant susceptibles de se polliniser mutuellement, des hybrides peuvent apparaître dans les zones de contact. L'intervention humaine joue aussi un rôle important dans ce processus.

En raison de cette situation confuse, l'utilisation des marqueurs SSR nucléaires et des marqueurs chloroplastiques fiables à discrimination entre les espèces de *Phoenix* serait extrêmement nécessaire. L'objectif principal de cette partie de travail est de distinguer les différentes espèces du genre *Phoenix* et identifier les plus proches parents du dattier.

Résultats

1. Polymorphisme chloroplastique des locus *trnG-trnfM* et *rpl16-rps3* chez les espèces du genre *Phoenix*

Deux séquences chloroplastiques (*trnG-trnfM* et *rpl16-rps3*) ont été génotypées (Scarcelli *et al.* 2011), dans le but d'identifier les haplotypes et déterminer les espèces proches du palmier dattier.

Il s'agit d'un minisatellite chloroplastique dodécanucléotidique situé dans l'espaceur intergénique *trnG-trnfM* (F: *trnG*-IGS-, R: *trnfM*-IGS-) (Figure 7), et un minisatellite de 20 pb de la séquence *rpl16-rps3* (F: *rpl16*-IGS-, R: *rps3*-IGS-) (Figure 8).

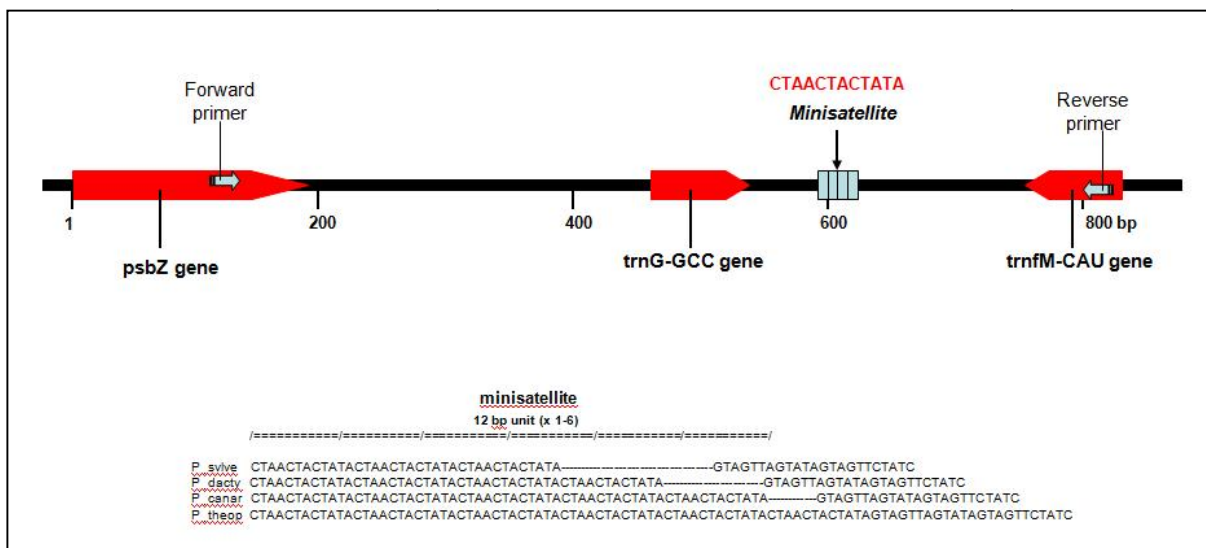


Figure 7 : Localisation du minisatellite (12pb) sur l'espaceur *psbZ-trnfM*.

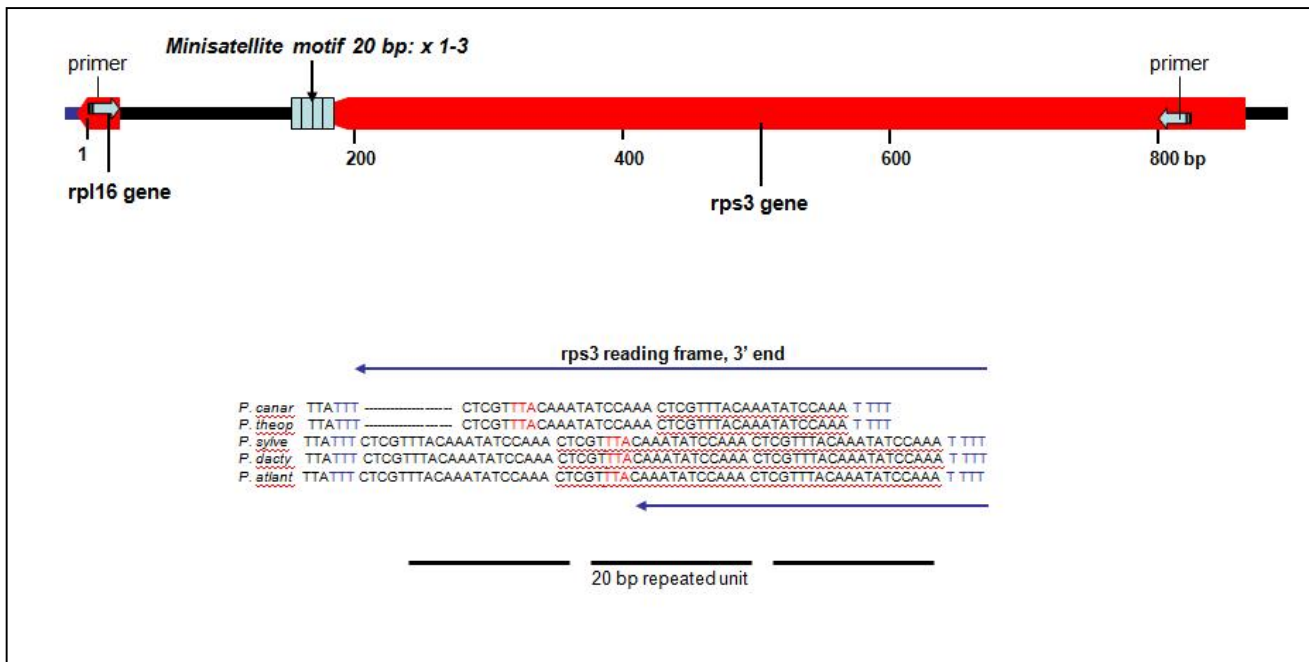


Figure 8 : Localisation du minisatellite (20pb) sur le locus *rpl16-rps3*.

Le polymorphisme des séquences d'un IGS *trnG-trnfM* (*psbZ-trnfM*) a été observé sur quatorze espèces du genre *Phoenix*.

Le premier objectif était de vérifier si l'amplification de cette amorce pourrait être adaptée aux espèces de *Phoenix*. La lecture des gels de génotypage de l'amorce utilisée sur le QIAxcel et ABI a montré des résultats chez toutes les espèces analysées (Figure 9). Le minisatellite de l'ADN chloroplastique situé dans l'espaceur intergénique *trnG* (GCC) -*trnfM* (CAU) a été amplifié avec un indice de réussite de 100%. Bien que ce minisatellite présente une seule mutation moins polymorphe que les microsatellites nucléaires, il montre un modèle de variation fortement lié aux espèces, y compris un haplotype privé fixé dans *Phoenix canariensis* (Tableau 5).

Dans cette présente étude le génotypage (QIAxcel et ABI) de l'espaceur intergénique a révélé six haplotypes. Ce dernier montre que *P. dactylifera* possède deux haplotypes. *Phoenix canariensis* est caractérisé par un haplotype privé avec cinq répétitions du motif de 12pb. *Phoenix theophrasti* et *P. rupicola* partagent le même haplotype, et sont séparés de *Phoenix canariensis* sur la base d'un seul minisatellite.

P. dactylifera et *P. atlantica* ont le même profil d'haplotype. *P. sylvestris* a deux haplotypes partagés avec *P. dactylifera*, *P. atlantica* et *P. roebelenii* ; et un haplotype partagé avec d'autres espèces (Tableau 5).

Tableau 5 : Répartition des haplotypes observés dans les régions *trnG-trnfM* et *rpl16-rps3* pour les espèces du genre *Phoenix*.

Espèces	Minisatellite 20pb	Minisatellite 12pb
	<i>rpl16-rps3</i>	<i>trnG-trnfM</i>
<i>P. canariensis</i>	454 (2rep)	266 (5rep)
<i>P. andamanensis</i>	454 (2rep)	230 (2rep)
<i>P. paludosa</i>	454 (2rep)	230 (2rep)
<i>P. pusilla</i>	454 (2rep)	230 (2rep)
<i>P. caespitosa</i>	454 (2rep)	242 (3rep)
<i>P. rupicola</i>	454 (2rep)	278 (6rep)
<i>P. acaulis</i>	454 (2rep)	242 (3rep)
<i>P. roebelinii</i>	454 (2rep)	242 (3rep) 254 (4rep)
<i>P. loureiroi</i>	454 (2rep)	230 (2rep) 242 (3rep)
<i>P. reclinata</i>	434 (1rep) 454 (2rep)	230 (2rep) 236 (2.5rep) 242 (3rep) 254 (4rep)
<i>P. theophrasti</i>	454 (2rep)	278 (6rep)
<i>P. dactylifera</i>	474 (3rep)	242 (3rep) 254 (4rep)
<i>P. sylvestris</i>	474 (3rep)	230 (2rep) 242 (3rep) 254 (4rep)
<i>P. atlantica</i>	474 (3rep)	242 (3rep) 254 (4rep)

La combinaison des données des deux loci *trnG-trnfM* et *rpl16-rps3* a révélé l'identification de treize haplotypes (Tableau 5). Huit espèces sont représentées par un seul haplotype (*Phoenix canariensis*, *P. andamanensis*, *P. rupicola*, *P. theophrasti*, *P. acaulis*, *P. pusilla*, *P. paludosa*, *P. caespitosa*). Cinq espèces présentent deux haplotypes ou plus (*P. dactylifera*, *P. sylvestris*, *P. loureiroi*, *P. roebelenii*, *P. reclinata*). La plupart des haplotypes sont spécifiques à l'espèce.

2. Phylogénie du chloroplaste

La phylogénie des haplotypes est constituée de 6 lignées phylogénétiques (Figure 10). La première regroupe 3 espèces réparties entre les îles de l'atlantique (*P. atlantica*), en Afrique et au Moyen-Orient (*P. dactylifera*) et en Asie (*P. sylvestris*). La seconde (*P. loureiroi*, *P. roebelenii*, *P. acaulis*) et la troisième lignée (*P. pusilla*, *P. paludosa*, *P. adamanensis*) sont strictement asiatiques. La quatrième regroupe 3 espèces réparties entre la corne de l'Afrique et la péninsule arabique (*P. caespitosa*), en Europe (*P. theophrasti*) et en Asie (*P. rupicola*). Les cinquième et sixième lignées comprenant une seule espèce sont limitées à l'Afrique (*P. reclinata*) et aux îles de l'atlantique (*P. canariensis*).

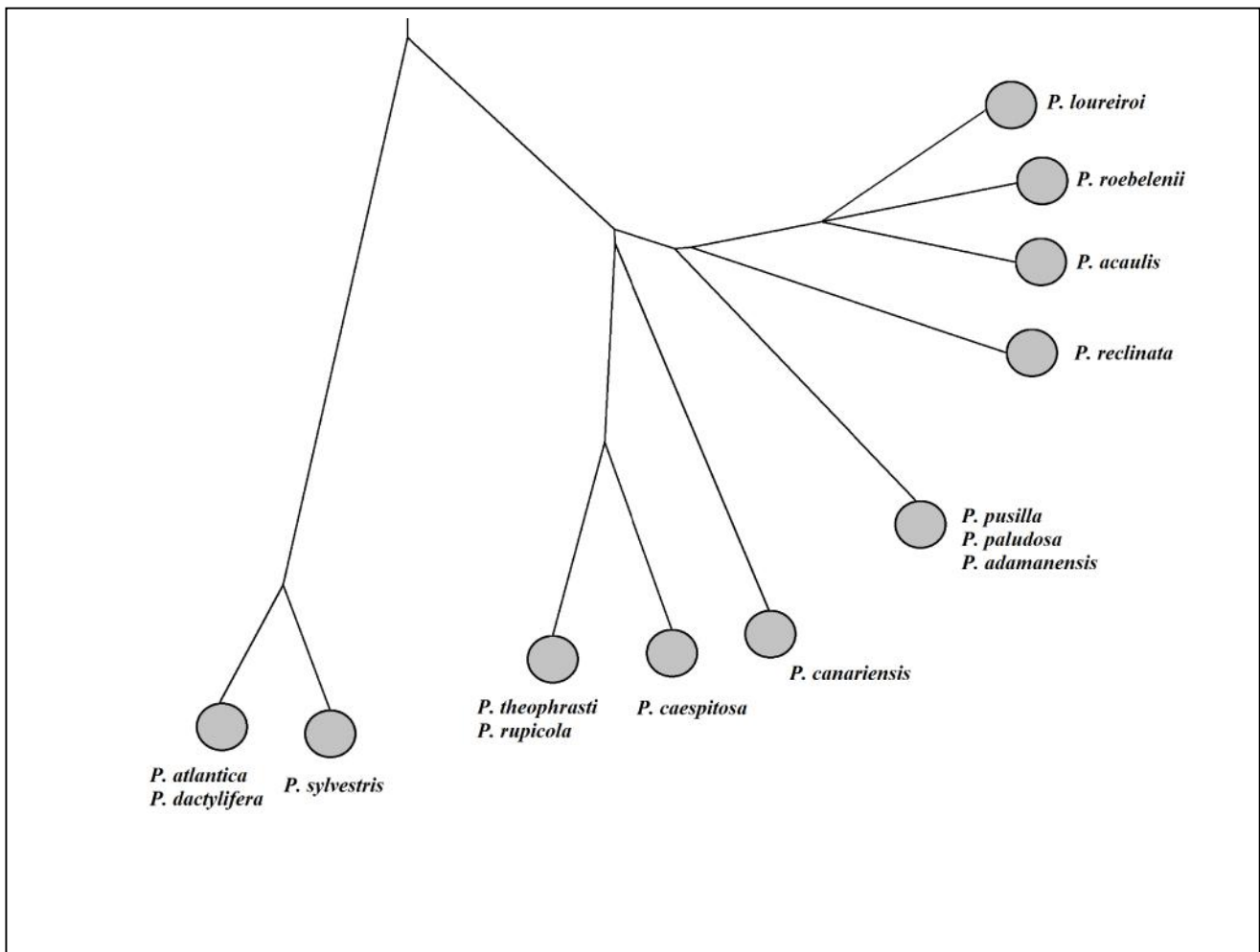


Figure 10 : Phylogénie du réseau d'haplotypes pour 301 échantillons représentant 14 espèces du genre *Phoenix* basées sur la combinaison des données de génotypage des locus *trnG-trnfM* et *rpl16-rps3*.

3. Diversité génétique globale et puissance des microsatellites

Le succès de l'amplification des 18 microsatellites nucléaires à travers le genre *Phoenix* a dépassé 98% (Tableau 6). Tous les loci ont amplifié entre 99% et 100%, à l'exception des loci mPdCIR035, mPdCIR63 et mPdCIR085 qui ont posé des problèmes d'amplification pour quelques espèces de *Phoenix*. Ces trois loci n'ont pas amplifié du tout dans *P. roebelenii*. Les loci mPdCIR035 et mPdCIR063 ont échoué avec les accessions de *P. theophrasti*. Les loci mPdIRD031, mPdIRD033, mPdCIR035, mPdCIR057, mPdCIR078, mPdCIR063, mPdCIR085, PdCUC3-ssr2 et PdAG1-ssr n'ont pas amplifié une accession dans *P. pusilla*, en raison de la mauvaise qualité de son ADN.

Les loci mPdIRD013, mPdIRD040, mPdCIR015, mPdCIR016, mPdCIR025, mPdCIR032, PdAP3-ssr et PdCUC3-ssr1 ont amplifié parfaitement tous les génotypes étudiés du genre *Phoenix*.

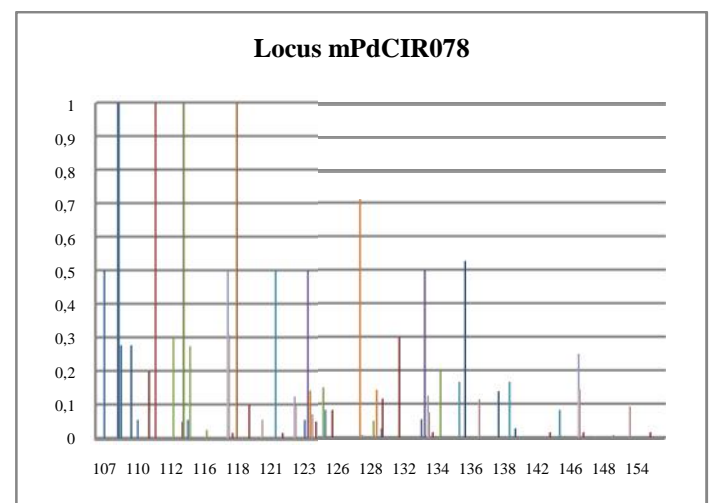
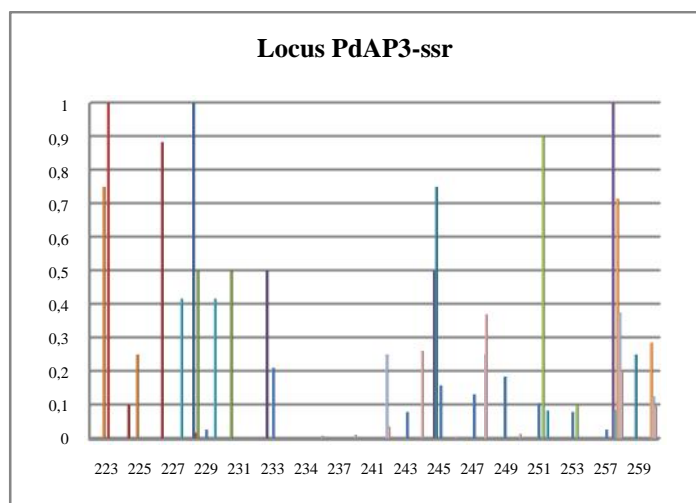
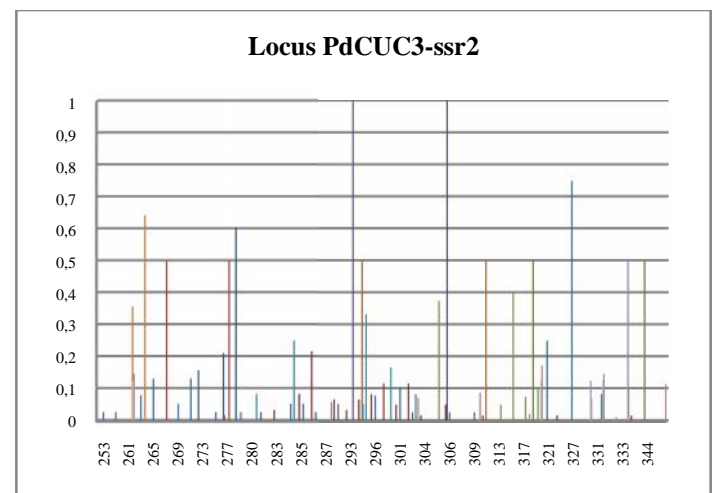
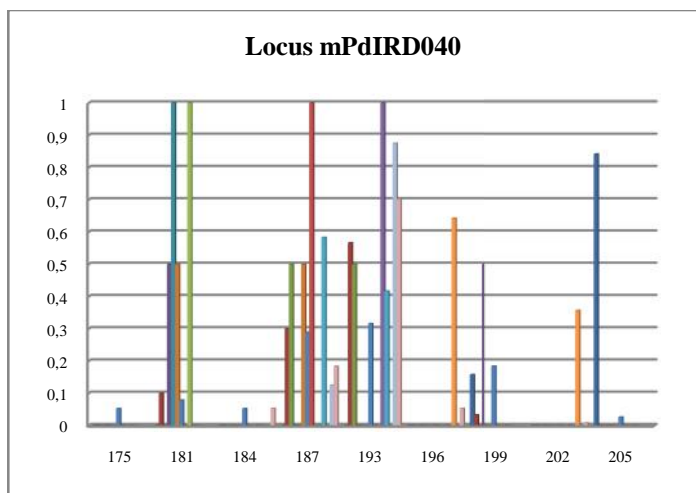
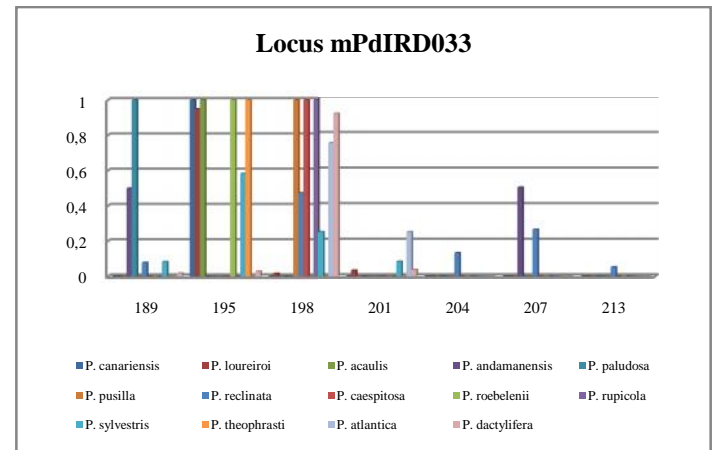
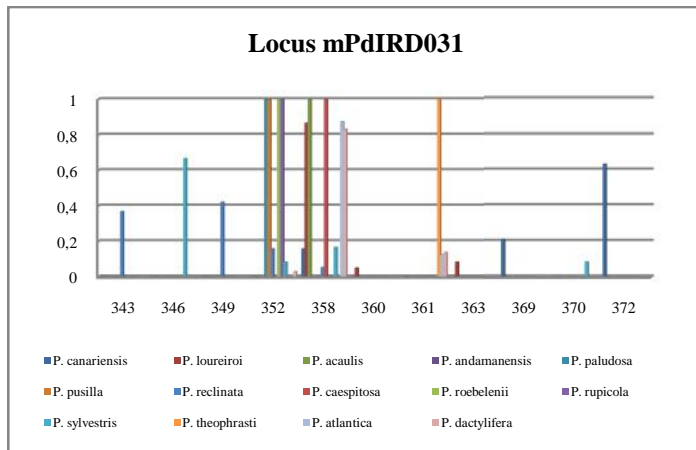
Au total 403 allèles ont été identifiés pour les 18 marqueurs SSR utilisés. Le nombre d'allèles par locus varie de 5 (PdCUC3-ssr1) à 46 (PdCUC3-ssr2), avec une moyenne de 22.4. Les valeurs de PIC varient de 0.16 (PdCUC3-ssr1) à 0.94 (PdCUC3-ssr2), avec une moyenne de 0.73. La fréquence de l'allèle le plus commun à chaque locus varie de 12% (PdCUC3-ssr2) à 91% (PdCUC3-ssr1). En moyenne, 38% des 301 accessions partagent un allèle majeur commun à chaque locus considéré (Tableau 6).

L'hétérozygotie attendue (H_e) pour chaque locus va de 0.05 à 0.74, avec une moyenne de 0.53. mPdCIR010, mPdCIR063, mPdCIR085, PdAG1-ssr et PdCUC3-ssr2 les locus les plus polymorphes, tandis que PdCUC3-ssr1 le moins polymorphe. La valeur moyenne de l'hétérozygotie observée est de 0.75. H_o varie de 0.16 (PdCUC3-ssr1) à 0.94 (PdCUC3-ssr2). La majorité des loci ont montré un haut niveau de diversité génétique.

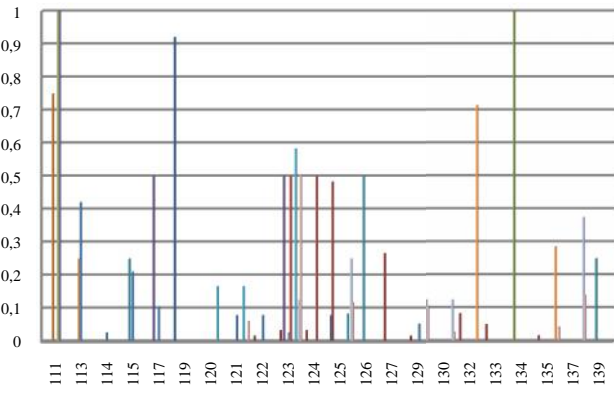
Tableau 6 : Paramètres descriptifs de la diversité génétique révélés par les 18 loci microsatellites analysés sur 301 génotypes de palmiers.

Marqueur	Nombre de génotype total	Nombre de génotype identifié	Amplification (%)	Fréquence de l'allèle le plus commun	Nombre de génotype identifié par locus	Nombre d'allèles par locus	Hétérozygotie observée (Ho)	Hétérozygotie attendue (He)	PIC
mPdIRD031	301,00	299,00	99,34	0,65	20,00	11,00	0,56	0,25	0,53
mPdIRD033	301,00	300,00	99,67	0,63	15,00	7,00	0,52	0,15	0,46
mPdIRD040	301,00	301,00	100	0,54	23,00	9,00	0,66	0,46	0,63
PdCUC3-ssr2	301,00	300,00	99,67	0,12	106,00	46,00	0,94	0,74	0,94
PdAP3-ssr	301,00	301,00	100	0,24	43,00	17,00	0,86	0,61	0,85
mPdCIR078	301,00	298,00	99	0,20	83,00	34,00	0,92	0,68	0,91
mPdCIR015	301,00	301,00	100	0,34	49,00	22,00	0,84	0,57	0,83
mPdCIR016	301,00	301,00	100	0,35	43,00	21,00	0,81	0,49	0,79
mPdCIR032	301,00	301,00	100	0,28	69,00	22,00	0,86	0,69	0,85
mPdCIR035	301,00	265,00	88	0,51	50,00	22,00	0,72	0,32	0,71
mPdCIR057	301,00	297,00	98,67	0,27	72,00	33,00	0,85	0,73	0,84
mPdCIR085	301,00	280,00	93	0,17	73,00	29,00	0,90	0,74	0,90
PdAG1-ssr	301,00	300,00	99,67	0,18	115,00	44,00	0,93	0,74	0,92
mPdCIR010	301,00	298,00	99	0,22	65,00	23,00	0,87	0,74	0,85
mPdCIR025	301,00	301,00	100	0,23	35,00	17,00	0,85	0,66	0,83
mPdCIR063	301,00	272,00	90,4	0,24	60,00	30,00	0,86	0,74	0,84
PdCUC3-ssr1	301,00	301,00	100	0,91	11,00	5,00	0,16	0,05	0,16
mPdIRD013	301,00	301,00	100	0,80	22,00	11,00	0,35	0,18	0,34
Moyenne	301,00	295,89	98,13	0,38	53,00	22,39	0,75	0,53	0,73

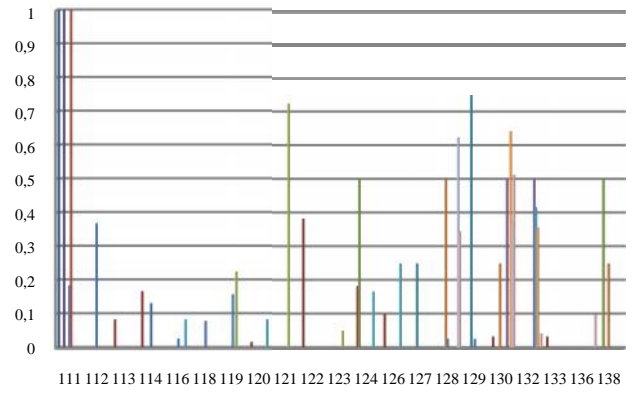
La distribution des fréquences alléliques au sein des espèces de *Phoenix* et aux différents loci est reportée sur la figure 11. Nos résultats montrent que certains allèles ont des fréquences basses. La fréquence de l'allèle la plus élevée de 1 a été enregistrée pour des allèles de plusieurs loci comme le montre la Figure 11.



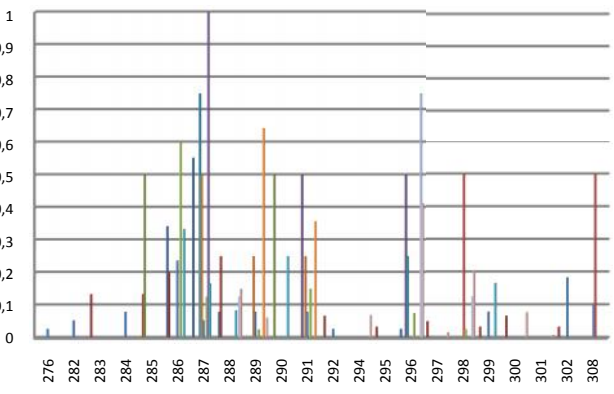
Locus mPdCIR015



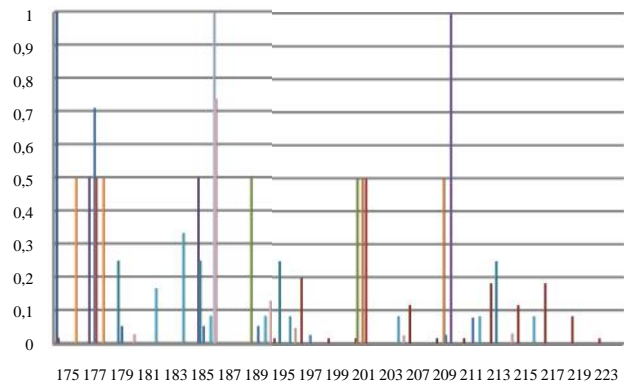
Locus mPdCIR016



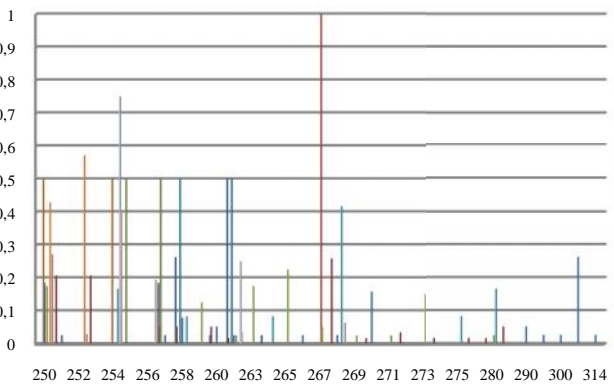
Locus mPdCIR032



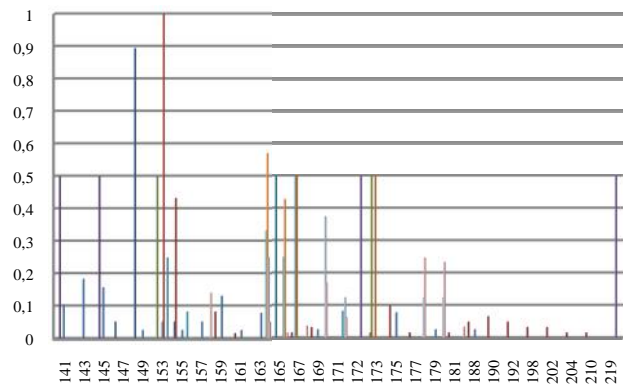
Locus mPdCIR035



Locus mPdCIR057



Locus mPdCIR085



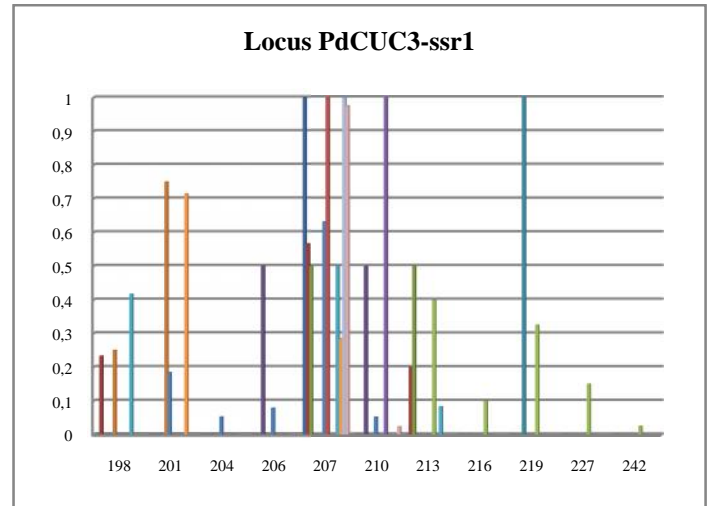
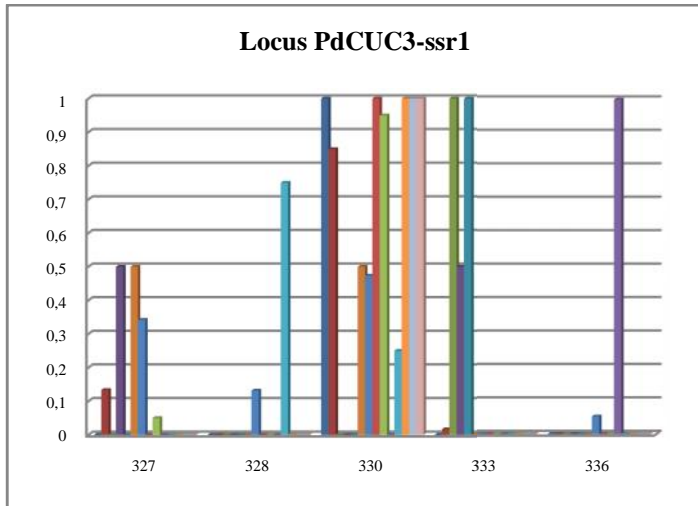
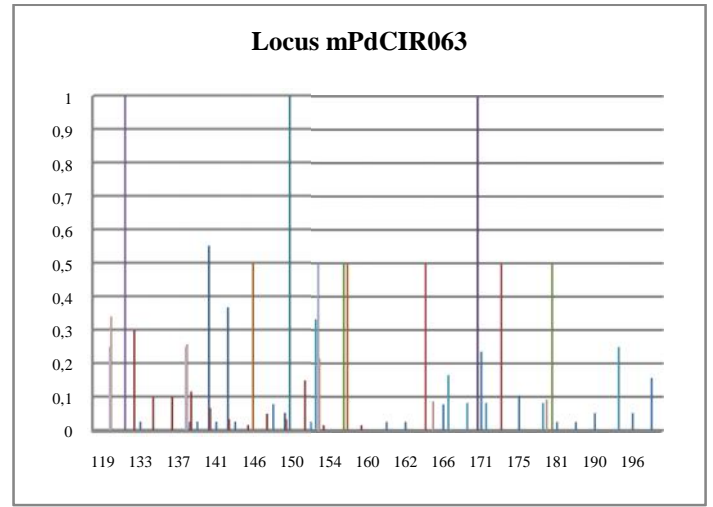
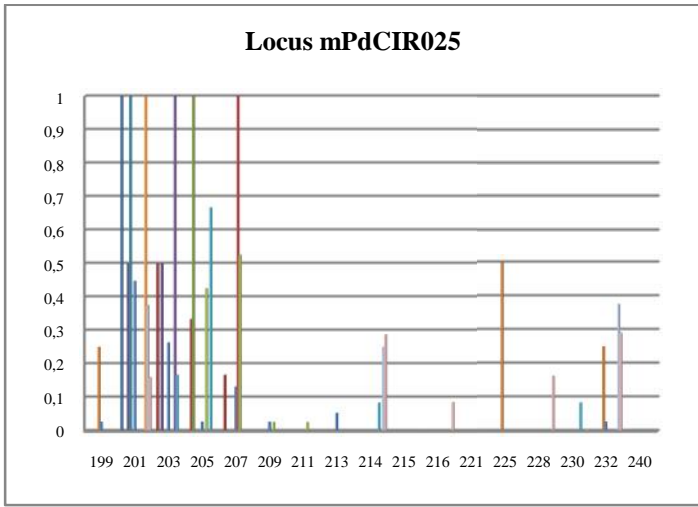
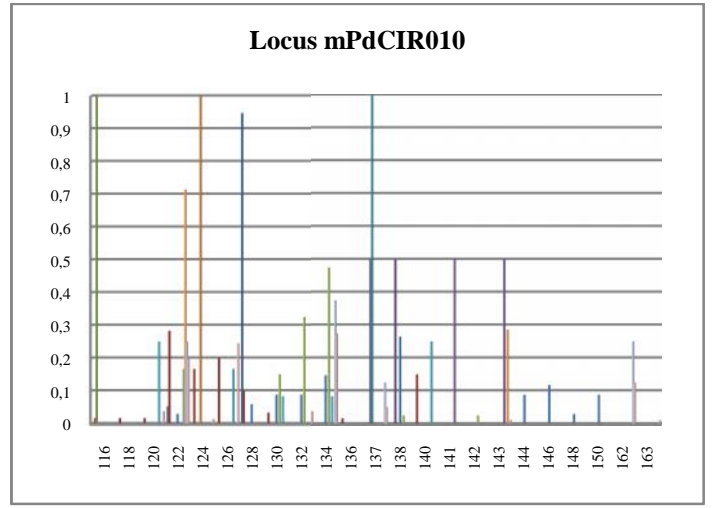
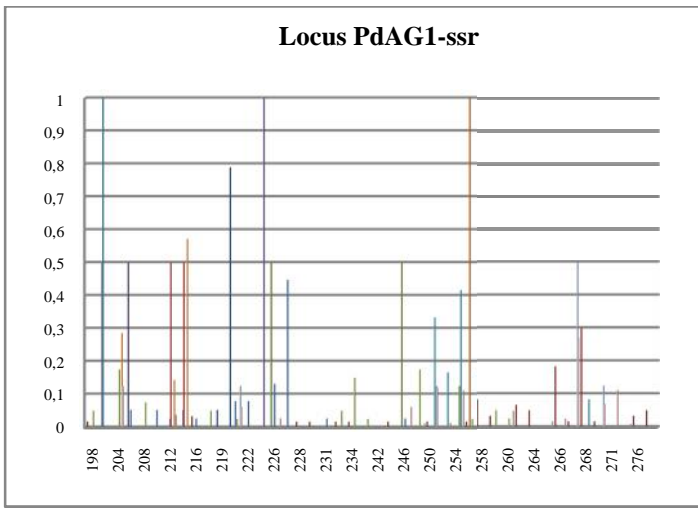


Figure 11 : Histogramme illustrant les fréquences alléliques obtenues pour chaque locus microsatellite.

Nous avons calculé pour chaque espèce différents indices de diversité génétique : l'hétérozygotie observée H_o , l'hétérozygotie attendue H_e , la richesse allélique et le nombre d'allèles grâce aux logiciels Genetix 4.05 (Belkhir *et al.* 2004) et Fstat (Goudet 2002).

Par espèce, le nombre d'allèles moyen par locus varie de 1.17 (pour *P. rupicola*) à 9.17 (pour *P. reclinata*). L'hétérozygotie attendue varie de 0.12 (*P. rupicola*) à 0.75 (*P. reclinata*), tandis que l'hétérozygotie observée varie de 0.17 (*P. theophrasti*) à 0.67 (*P. andamanensis*) (Tableau 7). Cette diversité génétique est donc variable d'une espèce à l'autre.

Tableau 7 : Diversité génétique moyenne sur tous les loci, par espèce.

Espèce	N°	Na	Ap	Ho	He
<i>P. canariensis</i>	19	2.44	8	0.19	0.25
<i>P. loureiroi</i>	30	8.94	42	0.51	0.64
<i>P. acaulis</i>	1	1.61	4	0.61	0.31
<i>P. andamanensis</i>	1	1.56	0	0.67	0.38
<i>P. paludosa</i>	2	1.61	5	0.25	0.22
<i>P. pusilla</i>	2	1.89	2	0.50	0.36
<i>P. reclinata</i>	19	9.17	45	0.54	0.75
<i>P. caespitosa</i>	1	1.28	2	0.33	0.18
<i>P. roebelenii</i>	20	3.61	21	0.30	0.42
<i>P. rupicola</i>	1	1.17	5	0.22	0.12
<i>P. sylvestris</i>	6	4.61	10	0.53	0.66
<i>P. theophrasti</i>	7	1.78	0	0.17	0.35
<i>P. atlantica</i>	4	3	0	0.54	0.46
<i>P. dactylifera</i>	188	7.89	29	0.60	0.59

N° : nombre d'individus échantillonnés par espèce, Na : nombre d'allèles par locus, Ap : Allèles privés, Ho : Hétérozygotie observée, He : Hétérozygotie attendue.

4. Relations génétiques entre les espèces de *Phoenix*

L'arbre de Neighbour Joining (Figure 12), basé sur la distance des allèles partagés (shared allele distance 'Das'), a clairement séparé toutes les espèces de *Phoenix*. Deux individus de *P. reclinata* ont été mal placés, ce qui pourrait s'expliquer par des données manquantes. Dans cette analyse, le cluster *Phoenix dactylifera* est parfaitement maintenu malgré l'échantillonnage plus large de cette espèce.

Sur cet arbre de Neighbour Joining, on observe que les individus de *P. atlantica* présentent une affinité avec les génotypes de *P. dactylifera*. Globalement, on observe un regroupement des autres espèces, ce qui montre une structuration très marquée de la diversité des 14 espèces.

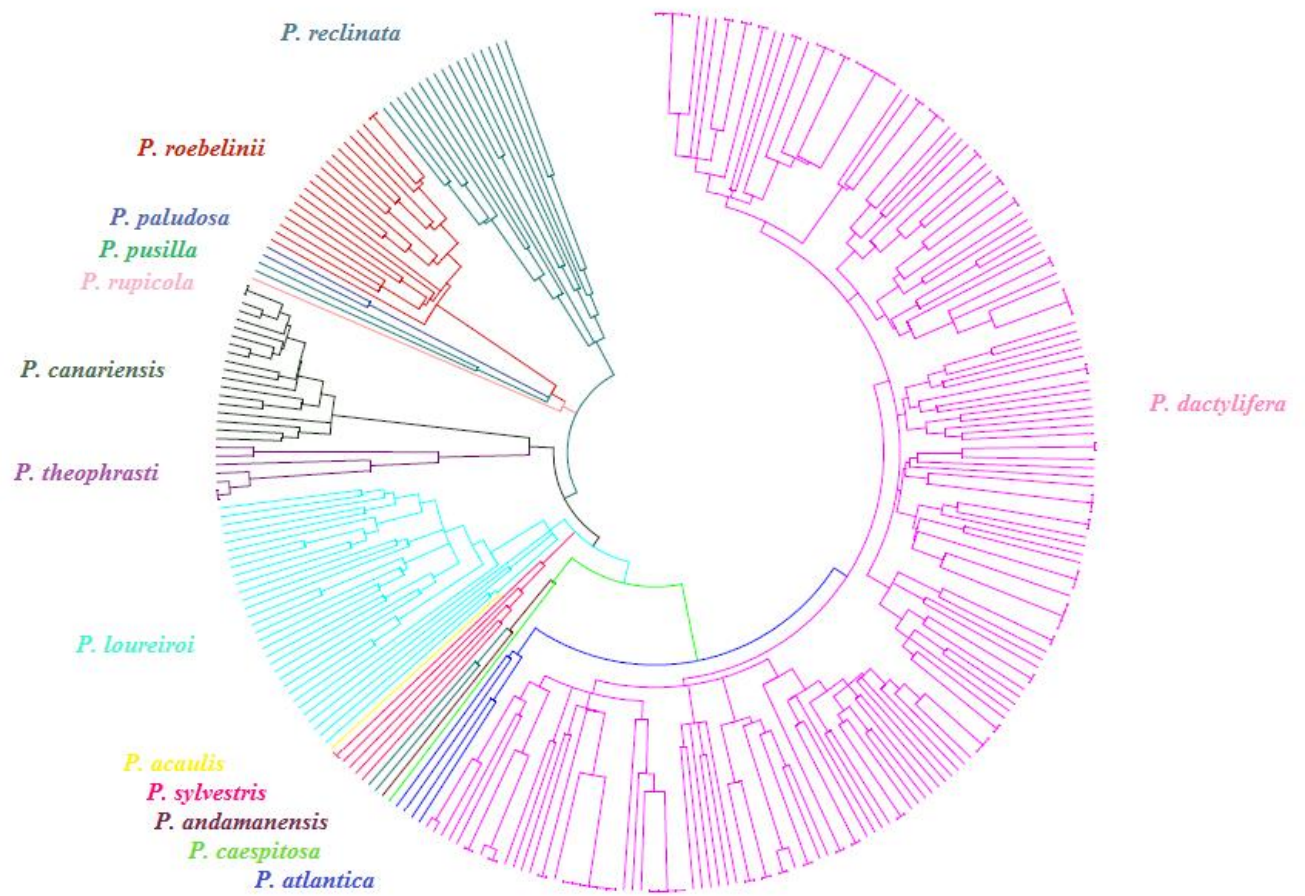


Figure 12 : Arbre phylogénétique des espèces basé sur la distance des allèles partagés (DAS) de 18 loci microsatellites pour l'ensemble des espèces du genre *Phoenix*.

Discussion

a. Polymorphisme chloroplastique des locus *trnG-trnfM* et *rpl16-rps3*

Les répétitions en tandem fournissent des informations précieuses sur la variabilité intra et interspécifique des locus de l'ADN chloroplastique. Les substitutions nucléotidiques et les indels ne sont pas très conservés dans le génome chloroplastique au sein de *Phoenix* (Asmussen *et al.* 2006). Cette constatation est conforme au modèle des variations des SSR chloroplastiques rapportés dans *Bactris gasipaes* (Pintaud *et al.* 2008).

Dans cette étude, nous avons testé l'utilité des deux régions chloroplastiques *trnG-trnfM* (CAU) et *rpl16-rps3* pour le genre *Phoenix*. La réussite de l'amplification, le séquençage et le génotypage de ces marqueurs dans toutes les espèces analysées confirme leurs efficacités. Leurs hautes performances permettent d'acquérir l'information de locus code-barres.

En ce qui concerne le minisatellite à 12 pb, nos résultats ont révélé beaucoup plus d'informations que ceux précédemment rapportées par Pintaud *et al.* (2010). Ceci pourrait s'expliquer par l'augmentation du nombre d'échantillons de la présente étude, et aussi par l'utilisation des deux méthodes : le séquençage et le génotypage. En effet, les données de génotypage de Pintaud *et al.* (2010) et de séquençage de Ballardini *et al.* (2013) n'ont pas détecté tous les haplotypes minisatellite de certaines espèces. Nous recommandons donc que les données de séquençage et de génotypage soient menés ensembles afin d'avoir une base solide pour interpréter les résultats. Au total, nos résultats ont permis d'identifier efficacement treize haplotypes. Ceci indique que les deux loci chloroplastiques ont un certain potentiel pour l'identification des espèces au sein du genre *Phoenix*.

Néanmoins, cette étude sur les polymorphismes au sein du genre *Phoenix* doit se poursuivre par l'utilisation du polymorphisme d'autres IGS dans l'ensemble des espèces. D'autres régions d'ADN devraient être examinées, à la recherche de caractères permettant l'identification de tous les taxons. Compte tenu de leur utilité prouvée dans les palmiers, le locus *psbA-trnH* (Al-Qurainy *et al.* 2011), l'ITS2 ribosomique (Jeanson *et al.* 2011) et *atpH-atpI* pourraient être étudiés en combinaison avec *psbZ-trnfM* et *rpl16-rps3* pour cet effet.

Une attention particulière devrait être accordée au groupe d'espèces *P. dactylifera*, *P. atlantica* et *P. sylvestris*, partageant les mêmes haplotypes (Tableau 7). Ce groupe est composé d'espèces étroitement apparentées et liées, de sorte que la difficulté de différencier ces espèces est attendue.

Enfin, dans certains cas, la divergence morphologique n'est pas associée à une divergence de séquence dans les régions chloroplastiques *psbZ-trnfM* et *rpl16-rps3*. Par exemple, *P. rupicola* et *P. theophrasti* partagent le même haplotype malgré une différenciation morphologique considérable et un isolement géographique, le premier étant originaire des forêts montagneuses de l'Inde et restreint aux contreforts himalayens (Asie), et le second est originaire d'Europe et il existe également en Turquie. Ces deux espèces montrent une convergence dans l'haplotype minisatellite. En revanche, le *P. theophrasti* est un palmier assez proche du *P. dactylifera*. Ils sont phénotypiquement très similaires, mais peuvent facilement être distingués par les loci *psbZ-trnfM* et *rpl16-rps3*.

La plupart des espèces de *Phoenix* peuvent être différenciées en fonction des séquences *psbZ-trnfM* et *rpl16-rps3*. Seule la distinction entre *Phoenix dactylifera*, *P. atlantica* et *P. sylvestris*, entre *P. pusilla*, *P. paludosa* et *P. andamanensis*, et entre *P. rupicola* et *P. theophrasti* est encore problématique. Nos travaux antérieurs sur le séquençage complet de ses deux loci ont pu séparer entre *P. pusilla*, *P. paludosa* sur la base des SNP (Pintaud *et al.* 2013).

b. Phylogénie du chloroplaste

La phylogénie d'haplotypes a permis d'identifier six lignées phylogénétiques dans le genre *Phoenix*. Deux lignées (*P. loureiroi-roebelenii -acaulis* et *P. pusilla -paludosa-adamanensis*) sont asiatiques ; une lignée est restreinte aux îles de l'atlantique (*P. canariensis*) et une à l'Afrique (*P. reclinata*) ; une lignée comprend trois espèces (*P. caespitosa-rupicola-theophrasti*) avec une vaste distribution de la corne de l'Afrique et la Péninsule Arabique vers l'Asie atteignant l'Europe. Avec les deux marqueurs chloroplastiques utilisés, *P. rupicola* est réellement indiscernable du *P. theophrasti*, morphologiquement et géographiquement lointain, probablement en raison d'une combinaison d'effet de plesiomorphie et d'homoplasie.

Dans cette phylogénie, les affinités directes de *Phoenix dactylifera* peuvent être retracées à deux espèces, *P. sylvestris* et *P. atlantica*. Ces trois dernières espèces

appelées «complexe du palmier dattier » (Pintaud *et al.* 2010), constituent une lignée à part et se distinguent de toutes les autres espèces de *Phoenix* au niveau moléculaire par la présence de trois répétitions du motif minisatellite de 20 pb de l'espaceur intergénique *rpl16-rps3*, au lieu de deux dans les autres espèces.

c. Structuration génétique des espèces de *Phoenix*

Les loci microsattellites conçus pour *Phoenix dactylifera* utilisés dans cette étude ont montré une bonne amplification et un polymorphisme élevé dans la plupart des espèces de *Phoenix*, confirmant ainsi la bonne transférabilité indiquée par Billotte *et al.* (2004). Les seuls loci problématiques étaient mPdCIR035 et mPdCIR063, avec un taux d'amplification de 88% et 90% respectivement. Dans d'autres travaux sur la diversité des espèces du genre *Phoenix*, les locus mPdCIR044 et mPdCIR048, non utilisés dans notre étude, ne pouvaient pas être amplifiés et ont produit des amplifications erratiques (Billotte *et al.* 2004 ; Zehdi *et al.* 2004 ; Henderson *et al.* 2006 ; Ahmed & Al-Qaradawi 2009 ; Pintaud *et al.* 2010). Les pertes de polymorphisme dans les expériences de transférabilité ont souvent été rapportées (Ellegren *et al.* 1997 ; Hutter *et al.* 1998 ; Butcher *et al.* 2000).

Les espèces les plus répandues (*P. reclinata*, *P. sylvestris*, *P. loureiroi* et *P. dactylifera*) présentaient une grande diversité génétique. Ces résultats sont comparables à d'autres espèces de palmiers tropicaux comme *Euterpe edulis* (Gaiotto *et al.* 2003), qui présentent également une grande diversité génétique basée sur des marqueurs microsattellites. *Phoenix dactylifera* possédait un nombre important d'allèles privés (29), ce qui l'isolait fortement d'autres espèces ; plus le nombre d'allèles privés dans une population est élevé plus cette population est isolée des autres. Par exemple, au niveau du locus mPdCIR025, six allèles sur dix indiqués chez *P. dactylifera* étaient privés.

Diverses espèces à distribution restreinte (*P. canariensis*, *P. acaulis*, *P. andamanensis*, *P. paludosa*, *P. rupicola*, *P. pusilla*, *P. caespitosa*, *P. theophrasti*) ont une faible diversité génétique et certains allèles privés, un modèle cohérent avec une évolution de petites populations isolées. Cependant, *Phoenix atlantica*, espèce endémique du Cap vert, a montré une diversité génétique considérable, proche à celle de *P. dactylifera* ou même plus élevée dans certains loci (Henderson *et al.* 2006).

La phylogénie établie par les microsatellites a séparé toutes les espèces de *Phoenix* étudiées (Figure 12). Tous les individus de la même espèce se regroupent. Ainsi, les allèles privés sont des signes de monophylie (synapomorphies) et d'isolement génétique des espèces correspondantes. Nos résultats confirment la classification morphologique et anatomique de Barrow (1998). Cependant, l'analyse phylogénétique de Barrow (1998) des données de séquençage de la région intergénique de l'espaceur ADN 5S, a identifié un clade composé de *P. dactylifera*, *P. theophrasti* et *P. sylvestris*. Ce regroupement n'est pas récupéré par l'analyse de données microsatellites. Ainsi, nos résultats concordent avec ceux de Pintaud *et al.* (2010), les trois espèces étant isolées dans leurs profils alléliques.

L'isolement génétique nucléaire caractérise également *P. atlantica* et *P. dactylifera*, malgré les polymorphismes chloroplastiques partagés. *Phoenix atlantica* est placé en tant que sœur de *P. dactylifera* (Pintaud *et al.* 2010), bien que l'isolement de *Phoenix atlantica* de *P. dactylifera* ait été montré quelque peu affaibli dans leur étude (Pintaud *et al.* 2010). Le fait que les deux espèces partagent le même profil d'haplotype chloroplastique, différent de toutes les autres espèces, peut aussi être indicatif d'une relation frère.

Phoenix theophrasti, qui est morphologiquement très similaire à *P. dactylifera*, a été suggéré par Barrow (1998) comme étant éventuellement une forme sauvage de *P. dactylifera*. Nos données montrent que ces deux espèces sont très distinctes et divergentes. Elles ont des allèles différents ou des allèles partagés à fréquences différentes dans les loci nucléaires, et elles ne partagent pas les mêmes haplotypes chloroplastiques.

En conclusion, l'analyse des marqueurs moléculaires a fourni des connaissances importantes sur la taxonomie du genre *Phoenix*.

Chapitre 2

**Apport des séquences chloroplastiques à la
recherche de chlorotypes de palmier dattier**

Introduction

Chez les plantes supérieures, l'information génétique est localisée dans les génomes chloroplastique, mitochondrial et nucléaire. Le génome chloroplastique, le moins complexe, est couramment utilisé en biologie moléculaire.

La petite taille du génome chloroplastique et l'abondance des chloroplastes dans les tissus chlorophylliens procurent de nombreux avantages pratiques, ce qui explique que ce soit le mieux caractérisé. 132 génomes chloroplastiques sont à ce jour entièrement séquencés et disponibles dans GenBank (Yang *et al.* 2010), dont 25 sont des génomes chloroplastiques de Monocotylédones.

L'étude du génome chloroplastique caractérisé par sa grande stabilité, a intéressé les scientifiques passionnés par les palmiers.

Malgré sa nature conservatrice, le génome chloroplastique des palmiers constitue une source idéale de marqueurs dans les études de diversité génétique (Pintaud *et al.* 2006, 2010, 2013), et pour la recherche de polymorphismes distribués dans les introns, les exons et les espaceurs intergéniques (Scarcelli *et al.* 2011 ; Ballardini *et al.* 2013). Deux haplotypes majeurs ont été trouvés dans le chloroplaste, un haplotype nommé Occidental et un Oriental (Pintaud *et al.* 2013)

Notre étude est réalisée sur 24 variétés algériennes. Nous avons étudié la structure du polymorphisme de l'ADN chloroplastique en séquençant des fragments situés dans les régions LSC et SSC, dans le but de déterminer les chlorotypes présents chez *Phoenix dactylifera* L. L'objet de l'étude sera donc d'approfondir l'étude du polymorphisme au sein de l'espèce *P. dactylifera* à l'aide d'une batterie de séquences chloroplastiques localisés dans les régions LSC et SSC, dans le but de déterminer les chlorotypes présents chez *Phoenix dactylifera*, et aussi d'apporter des éléments de réponses relatives à l'histoire du polymorphisme et de l'origine du palmier dattier.

Résultats

1. Mise au point des loci chloroplastiques et séquençage

Le succès d'amplification des locus choisis a dépassé 90%. La lecture des gels d'amplification des amorces utilisés a montré que seule l'amorce *matk* exon n'a donné aucun résultat pour tous les individus analysés (Figure 13).

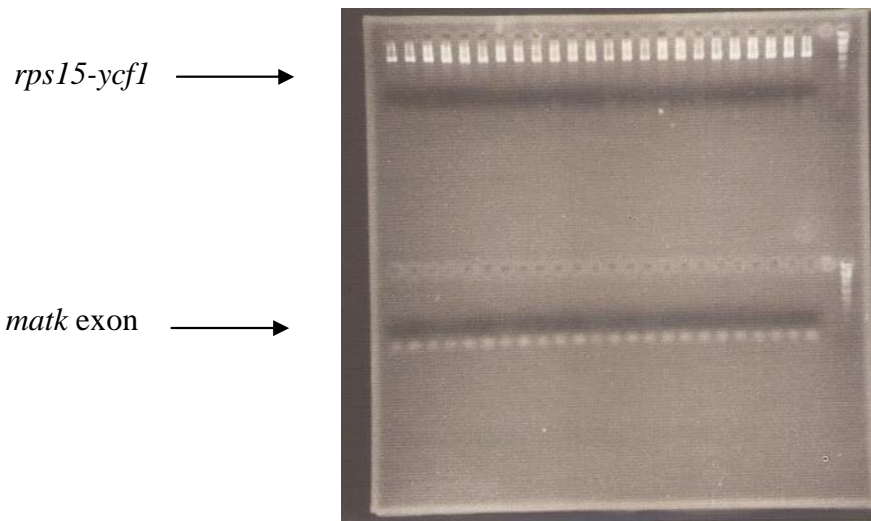


Figure 13 : Séparation sur gel d'agarose (2.5%) des produits d'amplification par PCR de l'ADN des 24 variétés de *P. dactylifera* en utilisant les *rps15-ycf1* et *matk* exon. Le locus *matk* exon n'a pas été amplifié pour toutes les variétés.

Les techniques de PCR et de séquençage ont permis de déterminer la variation de l'ADN chloroplastique. Un polymorphisme des séquences de quatre locus des régions LSC et SSC, a été observé sur 24 variétés algériennes de dattier.

Le séquençage a révélé que ce polymorphisme observé était dû à des mutations localisées dans les espaces intergéniques. Au total, deux microsattellites mononucléotidiques, un minisatellite, une petite inversion, une insertion ou délétion de 5pb (Indel), et deux substitutions (Tableau 8 ; Figure 14, 15). Le locus *rpoB* s'est avéré monomorphe.

Par séquençage, il a été montré que sept polymorphismes sont liés, avec absence de combinaison entre eux. Les sept mutations vont toujours ensemble, chaque individu possède soit : un minisatellite à quatre répétitions de motif, une insertion de 5pb, une inversion de 3pb

(TTC), deux homopolymère (T à 10pb et T à 11pb), deux SNP (A, T); soit un minisatellite à trois répétitions, une délétion de 5pb, une inversion de 3pb (GAA), deux homopolymère (T à 9pb et T à 10pb), deux SNP (C, A) (Tableau 8).

Tableau 8 : Sept polymorphismes au niveau des quatre loci polymorphes.

Locus	psbZ-trnfM	atpH-atpI	atpH-atpI	atpH-atpI	rps15-ycf1	psbA-trnH	psbA-trnH
Variété	<i>minisatellite</i>	<i>indel</i>	<i>substitution</i>	<i>homopolymer</i>	<i>Petite Inversion</i>	<i>substitution</i>	<i>homopolymer</i>
Khalass (GenBank)	4 rep	insert présent	A pos 60	T(10)	SI TTC	T pos 80	T(11)
Deglet Nour	4 rep	insert présent	A pos 60	T(10)	SI TTC	T pos 80	T(11)
Aghaliane	4 rep	insert présent	A pos 60	T(10)	SI TTC	T pos 80	T(11)
Horra	4 rep	insert présent	A pos 60	T(10)	SI TTC	T pos 80	T(11)
Taramount	4 rep	insert présent	A pos 60	T(10)	SI TTC	T pos 80	T(11)
Bacheir	4 rep	insert présent	A pos 60	T(10)	SI TTC	T pos 80	T(11)
Tati	4 rep	insert présent	A pos 60	T(10)	SI TTC	T pos 80	T(11)
Dfor Lgot	4 rep	insert présent	A pos 60	T(10)	SI TTC	T pos 80	T(11)
Zaghraya	4 rep	insert présent	A pos 60	T(10)	SI TTC	T pos 80	T(11)
Tacherwint	4 rep	insert présent	A pos 60	T(10)	SI TTC	T pos 80	T(11)
Lulu jaune	4 rep	insert présent	A pos 60	T(10)	SI TTC	T pos 80	T(11)
Kesba	4 rep	insert présent	A pos 60	T(10)	SI TTC	T pos 80	T(11)
Ghars	4 rep	insert présent	A pos 60	T(10)	SI TTC	T pos 80	T(11)
Iberica	3 rep	insert absent	C pos 60	T(9)	SI GAA	A pos 80	T(10)
Degla Baida	3 rep	insert absent	C pos 60	T(9)	SI GAA	A pos 80	T(10)
Arécthi	3 rep	insert absent	C pos 60	T(9)	SI GAA	A pos 80	T(10)
Mech Degla	3 rep	insert absent	C pos 60	T(9)	SI GAA	A pos 80	T(10)
Halwaya	3 rep	insert absent	C pos 60	T(9)	SI GAA	A pos 80	T(10)
Bayd Hmam	3 rep	insert absent	C pos 60	T(9)	SI GAA	A pos 80	T(10)
Deglet Gurara	3 rep	insert absent	C pos 60	T(9)	SI GAA	A pos 80	T(10)
Amari	3 rep	insert absent	C pos 60	T(9)	SI GAA	A pos 80	T(10)
Alig	3 rep	insert absent	C pos 60	T(9)	SI GAA	A pos 80	T(10)
Hamraya	3 rep	insert absent	C pos 60	T(9)	SI GAA	A pos 80	T(10)
Bu'Rus	3 rep	insert absent	C pos 60	T(9)	SI GAA	A pos 80	T(10)
Tinasser	3 rep	insert absent	C pos 60	T(9)	SI GAA	A pos 80	T(10)
Takerboucht	3 rep	insert absent	C pos 60	T(9)	SI GAA	A pos 80	T(10)

Le loci *rpoB* est monomorphe sur l'ensemble des variétés algériennes (Figure 16). Nous avons comparé nos séquences avec des séquences des variétés 'Khalass' et 'Aseel' extraites de Genbank (Yang *et al.* 2010 ; Khan *et al.* 2012). Nous avons trouvé une substitution à la position 1881 pour le locus *rpoB* avec la variété 'Aseel' (Tableau 9).

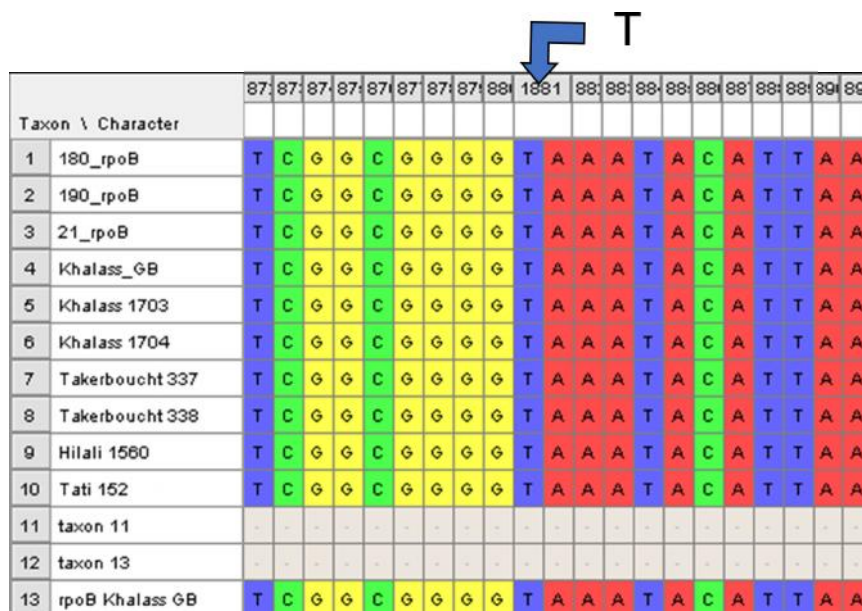


Figure 16 : *rpoB* locus monomorphe à la position 1881.

Tableau 9 : Substitutions au niveau du locus *rpoB*.

Locus	<i>rpoB</i>
Variété	Substitution
Khalass (GenBank)	T pos 1881
Aseel (GenBank)	G pos1881
Deglet Nour	T pos 1881
Ghars	T pos 1881
Iberica	T pos 1881
Degla Baida	T pos 1881
Takerboucht	T pos 1881

2. Caractère ancestral et dérivé du dattier cultivé

Les séquences d'ADN chloroplastique des quatre loci ont été obtenues pour chaque échantillon de *P. dactylifera* (24 variétés algériennes), *P. canariensis* (un individu) et *P. iberica* (un individu) avec des amorces et de protocole de Scarcelli *et al.* (2011), ou directement extraites de Genbank, en particulier dans les génomes chloroplastiques complètement séquencés de *P. dactylifera* variété 'Khalass' (Arabie Saoudite) et 'Aseel' (Pakistan) (Yang *et al.* 2010 ; Khan *et al.* 2011).

Phoenix Iberica est une nouvelle espèce, décrite récemment (Rivera *et al.* 2007). Endémique au sud de l'Espagne, elle serait trouvée naturellement dans les *barrancos* (ruisseaux intermittents) et les ravins du bassin de la rivière Chicamo (province de Murcie) et cultivée à Elche et à Fortuna (province d'Alicante, Espagne). Cette espèce est considérée comme synonyme de *Phoenix dactylifera* dont il pourrait s'agir de population naturelle ou de population ensauvagée.

Selon une étude sur un ensemble de 100 paires d'amorces d'ADN chloroplastique pour étudier la génétique et la phylogénie des monocotylédones (Scarcelli *et al.* 2011), *P. dactylifera*, *P. sylvestris* et *P. atlantica* ont une origine chloroplastique commune et ont divergé il y a 2.9 millions d'années. La divergence entre les chloroplastes de *P. dactylifera* et *P. canariensis* est placée à 4.6 millions d'années (Figure 17). Pour cela, nous avons choisi de comparer les mutations de *P. canariensis* avec celles de *P. dactylifera* pour détecter les caractères plésiomorphes et apomorphes.

Un résumé des résultats des polymorphismes est présenté dans le tableau 10. Les polymorphismes des variétés algériennes sont représentés par leur haplotypes Oriental ou Occidental.

Grace à ces données nous avons pu détecter les mutations plésiomorphes et apomorphes à partir de locus étudiés (Tableau 10).

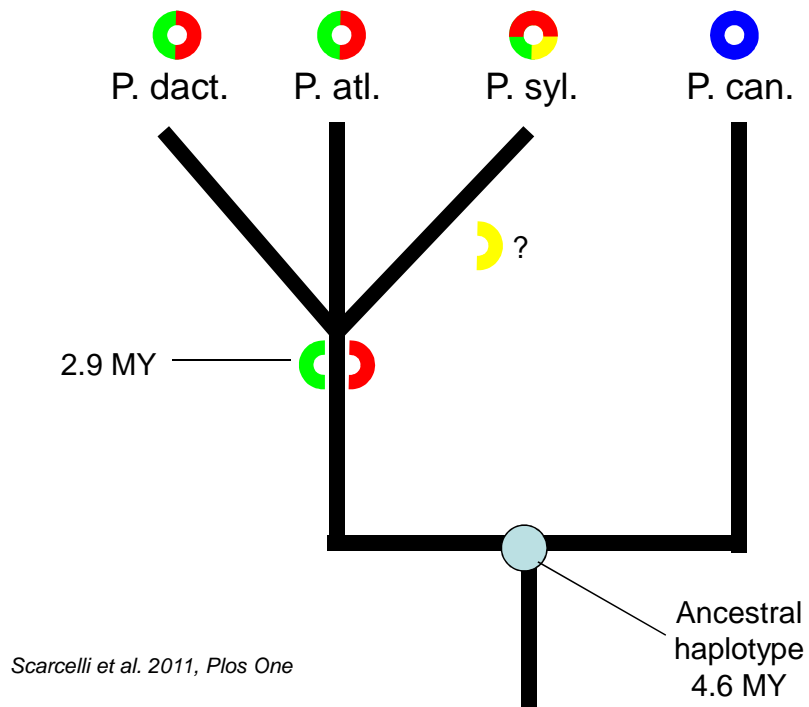


Figure 17 : Evolution des chloroplastes des espèces du genre *Phoenix*. (Réalisation : J.C. Pintaud, d'après Scarcelli *et al.* 2011).

Tableau 10 : Les états plésiomorphes et apomorphes du polymorphisme chloroplastique.

Locus	Longueur Totale	Position	Khalass	Aseel	Variétés à haplotype Oriental	Iberica (Elche)	Variétés à haplotype Occidental	canariensis
SNP								
psbA-trnH	770	94	T	T	T	A	A	A
atpH-atpI	860	57	A	A	A	C	C	A
Petite inversion								
rps15-ycf1	630	326-328	TTC	TTC	TTC	GAA	GAA	TTC
VNTR								
Minisatellites > 10 bp								
rpl16-rps3		157-196	3x	3x		3x		2x
psbZ-trnfM	700	599-646	4x	4x	4x	3x	3x	5x
Indels oligonucléotidique < 10 bp								
atpH-atpI	860	218-222	ATATA	ATATA	ATATA	0	0	0
Microsatellite mononucléotidique								
psbA-trnH	770	212-222	T(11)	T(11)	T(11)	T(10)	T(10)	T(10)
atpH-atpI	860	72-82	T(10)	T(10)	T(10)	T(9)	T(9)	T(11)

Etat Plésiomorphe : non coloré

Etat Apomorphe : coloré

Discussion

a. Polymorphisme au sein du génome chloroplastique de *Phoenix dactylifera*

Dans le cadre de l'étude du polymorphisme du génome chloroplastique de *Phoenix dactylifera*, nous avons examiné les variations de locus choisis pour y identifier des régions polymorphes et révéler la relation entre les différentes mutations. Ainsi, trois de la grande copie simple LSC, et un locus de la petite copie simple SSC, ont été analysés.

Les sept polymorphismes liés de quatre loci polymorphes ont pu séparer génétiquement les variétés de *Phoenix dactylifera* à l'aide de la PCR-séquençage, en deux chlorotypes distincts, un chlorotype spécifique des variétés à haplotype minisatellite Occidental, et un autre chlorotype spécifique des variétés à haplotype minisatellite Oriental.

Depuis plusieurs années, de nombreuses paires d'amorces chloroplastiques ont été mises au point. Les marqueurs chloroplastiques s'avèrent être performant et peuvent servir à de multiples applications. Parmi ces marqueurs, les microsatellites ou SSR (Simple Sequence Repeats) sont réputés être les plus performant pour des analyses de la diversité génétique (Provan *et al.* 2001), et les plus utilisés sont les microsatellites mono-nucléotidiques (Provan *et al.* 2004).

Plusieurs séquences simples répétées de minisatellites du génome chloroplastique ont été amplifiées à l'aide des amorces universelles chez les espèces de palmier (Demesure *et al.* 1995 ; Grivet *et al.* 2001 ; Shaw *et al.* 2005 ; Scarcelli *et al.* 2011). Les études faites sur des minisatellites chloroplastiques ont apporté des informations sur de nombreux genres appartenant aux Arecoïdeae, notamment pour le palmier pêche *Bactris gasipaes* et les différentes espèces du genre *Phoenix* (Pintaud *et al.* 2006 ; Couvreur *et al.* 2007 ; Pintaud *et al.* 2010). Ces travaux ont prouvé que le haut degré de conservation des séquences du génome chloroplastique des plantes pouvait répondre à plusieurs questions relatives à l'histoire de la domestication et de l'origine de ces espèces.

Le minisatellite du locus *psbZ-trnfM* est un outil pertinent pour l'étude de génétique des populations et la délimitation des populations des espèces de *Phoenix* (Pintaud *et al.* 2010 ; Scarcelli *et al.* 2011 ; Pintaud *et al.* 2013 ; Ballardini *et al.* 2013). Ce minisatellite a permis la séparation d'espèces proches du genre *Phoenix*.

L'espaceur inter-génique *rps15-ycf1*, est l'un des locus les plus variables de la région SSC chez les Areceae (Scarcelli *et al.* 2011). Ce locus contient une petite inversion de 3pb TTC vs GAA. Les inversions de taille sont fréquentes dans les régions non-codantes de l'ADN chloroplastique (Kelchner & Wendel 1996 ; Kim & Lee 2005).

Plus récemment, un nouveau type de marqueurs basés sur des changements ponctuels d'une base dans une séquence donnée et appelés SNP (Single Nucleotide Polymorphism) a été développé (Jehan *et al.* 2006 ; Angaji 2009). La recherche des SNP chez le dattier a été conduite à 62 substitutions identifiées sur 78 SNP, dans la variété Khalass (Yang *et al.* 2010). Deux SNP ont été détecté dans notre étude. Ces SNP sont présents cher la variété occidentale tel que Iberica qui montre les mêmes substitutions que les variétés algériennes à haplotype occidental. Néanmoins, ces SNP ont été identifiés dans les séquences des variétés orientales Khalass et Aseel de GenBank.

L'espaceur intergénique situé entre les gènes *atpH* et *atpI*, montre un polymorphisme important avec trois mutations : un microsatellite mononucléotidique, une substitution et un indel. La région *atpH-atpI* pourrait être considérée comme une bonne base pour l'établissement d'un système de code-barres d'ADN pour *Phoenix dactylifera*.

Le travail de Al-Qurainy *et al.* (2011) sur huit variétés de dattier d'Arabie Saoudite, a révélé un polymorphisme important pour les gènes *psbA-trnH* et *rpoB*. Le séquençage de ses deux locus sur nos variétés montre que l'espaceur intergénique *psbA-trnH* est polymorphe. Il présente deux mutations liées aux chlorotypes, une substitution (T vs A) et un microsatellite homopolymère (T)_n où n varie entre 10 et 11. Le gène *rpoB* s'est avéré monomorphe sur l'ensemble de nos variétés. Cependant, il présente une substitution avec la variété Aseel de GenBank (Khan *et al.* 2012). Nos variétés possèdent une substitution T, alors que la variété Aseel possède une substitution G.

D'une manière générale, l'échantillonnage important pour l'analyse du génotypage et le choix des variétés à haplotype 3 et 4 pour le séquençage, étaient suffisant pour conclure que chaque cultivar a un chlorotype unique.

Nos données de séquençage obtenues n'ont pas permis de révéler l'existence d'hétéroplasmie dans l'échantillonnage algérien, un phénomène suggéré par Yang *et al.* (2010).

Les travaux de Yang *et al.* (2010) révèlent la présence de plusieurs chlorotypes sur la base des SNP, sur le même individu de la variété Khalass. Selon ces auteurs l'hétéroplasmie existe chez le palmier dattier en présence d'un chlorotype majoritaire et d'un autre minoritaire. En revanche, de rares cas d'hétéroplasmie ont été rapportés chez les angiospermes.

Le séquençage des loci chloroplastiques a montré que chaque variété présente un unique chlorotype. Une nette différenciation génétique se dessine entre les deux chlorotypes et se traduit en deux chloroplastes différents.

b. Caractère ancestral et dérivés du dattier cultivé

Les espèces du genre *Phoenix* sont proches morphologiquement, parfois difficiles à distinguer. Par exemple, il existe peu de caractères permettant de différencier *P. dactylifera* de *P. theophrasti* (Barrow 1998) et de *P. atlantica* (Henderson *et al.* 2006).

Les populations sauvages de palmier dattier n'ont pas été caractérisées avec certitude jusqu'à présent. Plusieurs hypothèses ont longtemps coexisté concernant l'origine du dattier. Selon de nombreux auteurs, l'origine du dattier cultivé se trouverait au Moyen-Orient (Bonavia 1885 ; Beccari 1890 ; Popenoe 1973 ; Costantini 1985 ; Barrow 1998 ; Costantini & Audisio 2000 ; Zohary *et al.* 2012). Les données archéologiques confirment cette hypothèse (Tengberg 2012). Cependant les données archéologiques sont encore rares dans le nord de l'Afrique.

Le polymorphisme génique actuel est le résultat de mutations survenues dans le passé. Il existe de nombreux allèles pour un gène. Afin que ces nouveaux allèles jouent rôle d'un point de vue évolutif, il faut qu'ils soient transmis de générations en générations. Pour déterminer l'état ancestral ou l'état dérivé, pour une mutation donnée, il faut étudier l'état de cette mutation dans son groupe monophylétique.

Pour cela et dans le but de déterminer les caractères ancestraux et dérivés des mutations présentes chez les variétés de *P. dactylifera*, nous avons comparé les polymorphismes existant entre les variétés algériennes, la variété 'Khalass' d'Arabie Saoudite, la variété 'Aseel' du Pakistan, Iberica d'Espagne, et *P. canariensis*.

Les résultats montrent des autapomorphies liées à l'haplotype 3 (haplotype Occidental) et des autapomorphies confirmées pour l'haplotype 4 (haplotype Oriental) (Tableau 10).

La coexistence d'autapomorphie lié au deux chlorotypes suggère une introduction ancienne et le mélange de deux pools génétiques oriental et occidental.

Les résultats apportés d'une part par les deux haplotypes chloroplastiques et d'autre part par les polymorphismes liés (séquençage), nous amène à suggérer l'existence de deux chlorotypes qui se traduisent en deux chloroplastes différents.

Le palmier dattier présente deux chlorotypes ; un haplotype se trouve dans le palmier dattier oriental (variété Khalass) du Moyen-Orient, et un trouvé dans l'échantillon provenant d'Espagne. La palmeraie algérienne possède les deux chlorotypes. Cette découverte de deux génomes chloroplastiques, ouvre la question sur l'existence de deux centres de domestication

distincts, l'un au Moyen-Orient (déjà suggéré par Tengberg 2012), et un autre dans la partie Ouest de la zone de distribution du dattier.

Chapitre 3

Diversity of Algerian Oases date palm

**(*Phoenix dactylifera* L., Arecaceae):
heterozygote excess and cryptic structure
suggest farmer management had a major
impact on diversity**

Introduction

The date palm (*Phoenix dactylifera* L., Arecaceae) is a perennial monocotyledon ($2n = 36$). It is an ecologically, culturally and economically important crop, widely cultivated in arid and semi-arid Mediterranean regions, in the Sahara, and in the Middle East (Munier 1973 ; Barrow 1998 ; Dransfield *et al.* 2008). More than 3,000 cultivars are estimated to be used for date production worldwide, of which around 60 are widely grown and have important national and international markets (Chao & Krueger 2008).

Early cultivation of the date palm is recorded in the Eastern part of its cultivation area, in southern Mesopotamia in the 5th Millennium BC. In the Western part of its cultivation area, evidence for domesticated date palm has been found in Egypt in the 4th century BC (Terral *et al.* 2012). Genetic studies of date palm clearly separate the Eastern and Western group (Arabnezhad *et al.* 2012 ; Pintaud *et al.* 2013 ; Mathew *et al.* 2015 ; Zehdi *et al.* 2012 ; Zehdi *et al.* 2015). The two primary gene pools (Eastern and Western) are also observed in the maternally-inherited chloroplast genome (Pintaud *et al.* 2013) and the paternally-inherited Y chromosome (Cherif *et al.* 2013). Two major alleles were found in the chloroplast, an Occidental and Oriental haplotype (Pintaud *et al.* 2013).

In Algeria, chloroplast diversity includes roughly 70% of the Eastern chloroplast (Zehdi *et al.* 2015). In neighboring Egypt, Tunisia and Morocco, the proportion of Eastern haplotype only ranges from 11% to 42%, but Algerian nuclear diversity is similar to that found in its neighboring countries, Tunisia and Morocco (Zehdi *et al.* 2015). The contrast between nuclear and chloroplast diversity observed in Algeria thus remains largely unexplained, perhaps because of the only fragmentary analysis of the diversity of the 1,000 varieties described in Algeria.

Date palms are cultivated in Algerian oases in most of the regions south of the Saharan Atlas Mountains. In 2002, Algerian date palm groves contained 13.5 million trees occupying 120,830 ha, while in 2015, 18 million date palms occupied 169,380 ha (Bouguédoura *et al.* 2015). Nearly 1,000 cultivars clonally propagated from offshoots have been inventoried and their distribution shows a very marked breakdown into eastern, central and western parts of the country. Some cultivars are found in two or three regions but most are restricted to their area of origin (Bouguédoura *et al.* 2008). What is more, cultivars are not evenly distributed across oases, as they are adapted to slightly different types of soil, ranges of temperature and

humidity, and often do not produce a satisfactory yield when cultivated outside their place of origin (Bouguedoura 2012 ; Bouguedoura *et al.* 2015). Palms grown from seeds occur randomly in the oases and are called “khalts” or “dgouls”. Khalts represent up to 10% of a population of date palm sand area valuable resource for new selection by farmers.

The aims of the present study were to (1) genetically characterize date palm agrobiodiversity in Algeria using a set of representative female fruiting date palm cultivars and (2) investigate the structure of genetic diversity in the oases south of the Saharan Atlas based on nuclear and chloroplast genotyping.

Results

1. Chloroplast minisatellite and nuclear SSR genotyping

The two genotyping methods used (either QiAxel or the ABI genetic analyzer) were consistent and revealed two chlorotypes, previously reported as Occidental and Oriental (Pintaud *et al.* 2013) (Annex 4). The Occidental (Western) chlorotype was found in 31.3% and the Oriental (Eastern) in 68.7% of the total sample. The proportion varied among oases (Figure 18 C).

Linkage disequilibrium was very low between SSR markers (Annex 5). A total of 143 alleles were identified with the 18 nuclear markers. The average number of alleles was 7.9, average expected heterozygosity was 0.60, and the average PIC was 0.57 (Table 11). The average number of alleles per locus at the individual oases ranged from 3.83 (Ouargla) to 5.78 (Oued Souf). Allelic richness varied slightly between oases with values ranging from 3.36 (Touggourt) to 3.85 (Beni Abbes). The same was true of expected heterozygosity, which ranged from 0.53 (Ouargla) to 0.59 (Oued Souf and Beni Abbes), and of observed heterozygosity, which ranged from 0.56 (Tamanrasset) to 0.65 (Ghardaia).

No significant difference in diversity (allelic richness, heterozygosity observed or expected) between oases was observed using a Bonferroni corrected p-value for multiple testing (Annex 6, 7, 8). Significant negative fixation indices (F_{IS}) were obtained for two out of 10 oases (Table 12) (Annex 9). Only three isolated southern oases : Tamanrasset, Adrar and Beni Abbes, had a deficit of heterozygous individuals (Table 12).

Table 11 : Descriptive genetic parameters for 18 microsatellite loci analyzed on 192 individual date palms.

Marker	N° of genotypes	N° of alleles	Major allele frequency	Gene diversity	Expected heterozygosity He	PIC
mPdCIR010	31	10	0.28	0.80	0.86	0.78
mPdCIR015	19	8	0.50	0.70	0.68	0.67
mPdCIR016	8	4	0.51	0.61	0.56	0.53
mPdCIR025	19	10	0.29	0.77	0.87	0.74
mPdCIR032	25	9	0.41	0.75	0.77	0.72
mPdCIR035	15	7	0.70	0.48	0.27	0.46
mPdCIR057	18	7	0.40	0.72	0.77	0.68
mPdCIR063	14	6	0.34	0.75	0.74	0.71
mPdCIR078	41	16	0.31	0.84	0.83	0.82
mPdCIR085	30	10	0.25	0.83	0.84	0.80
PdAG1-ssr	54	19	0.27	0.87	0.86	0.86
PdCUC3-ssr1	1	1	1.00	0.00	0.00	0.00
PdCUC3-ssr2	42	12	0.17	0.88	0.85	0.87
PdAP3-ssr	20	9	0.37	0.74	0.83	0.70
mPdIRD013	2	2	0.98	0.05	0.05	0.05
mPdIRD031	5	3	0.83	0.29	0.31	0.26
mPdIRD033	5	4	0.91	0.16	0.16	0.16
mPdIRD040	9	5	0.70	0.47	0.49	0.43
Mean	19.89	7.94	0.51	0.59	0.60	0.57

The number of genotypes, the number of alleles, the frequency of the main alleles, gene diversity, heterozygosity and the polymorphism information content (PIC) of 18 loci in *Phoenix dactyifera* are given.

PIC = polymorphism information content.

Table 12 : Summary statistics for 18 microsatellite loci in 10 oases.

Oasis	N°	Na	Ar	Ho	He	F _{IS}	P value F _{IS}
Biskra	18	4.72	3.45	0.58	0.56	-0.02	0.3216
Touggourt	32	5.33	3.36	0.59	0.54	-0.06	0.0097
Oued Souf	30	5.78	3.73	0.60	0.59	-0.01	0.2457
Ouargla	5	3.83	3.83	0.61	0.53	-0.05	0.2563
Ghardaia	29	5.22	3.53	0.65	0.56	-0.13	0.0000
Tamanrasset	7	4.11	3.68	0.56	0.57	0.09	0.9603
El Menia	17	4.78	3.47	0.58	0.55	-0.03	0.1719
Timimoun	16	5.11	3.67	0.64	0.57	-0.04	0.1694
Adrar	20	5.06	3.53	0.58	0.56	0.004	0.5505
Beni Abbes	14	5.22	3.85	0.58	0.59	0.05	0.9013

The number of samples (N°), the number of alleles per locus (Na), allelic richness (Ar), observed heterozygosity (Ho), expected heterozygosity (He), the inbreeding coefficient (F_{IS}) and the p-value of F_{IS} are given for each population (averaged across the 18 loci).

2. Genetic differentiation between populations

Overall differentiation was weak between oases, with an overall average differentiation of 0.0119. Overall, there was isolation by geographical distance (Mantel test, p-value = 0.039). Geographically close populations were generally weakly differentiated while greater differentiation was observed for populations sampled further apart (Table 13).

Table 13 : Genetic distance between oases (F_{ST}).

	Biskra	Touggourt	Oued Souf	Ouargla	Ghardaia	Tamanrasset	El Menia	Timimoun	Adrar	Beni Abbes
Biskra		0.0288*	0.0288	0.0062	0.0701*	0.0205	0.0668*	0.0581*	0.0559*	0.0623*
Touggourt			0.0098	-0.0288	0.0379*	0.0473*	0.0420*	0.0479*	0.0492*	0.0555*
Oued Souf				-0.0065	0.0240*	0.0270	0.0289	0.0322*	0.0334*	0.0313*
Ouargla					0.0197	0.0049	0.0016	0.0063	0.0079	0.0132
Ghardaia						0.0322	0.0414*	0.0143	0.0365*	0.0372*
Tamanrasset							0.0245	0.0141	0.0129	0.0172
El Menia								0.0179	0.0072	0.0250
Timimoun									0.0119	0.0056
Adrar										0.0086
Beni Abbes										

The F_{ST} component was determined using a method that calculates the genetic distance between two oases, the differences being in pairs. A Bonferroni correction was applied, the significance level passing $p < 0.001$.

*: $P < 0.001$ (significant).

3. Variety name and diversity

Based on an individual phylogenetic tree, relationships between varieties with similar names are easy to see. Varietal names need to be compared because of the many synonyms, such as: Al Kayed and Tanteboucht; Tgaza and Takerboucht; Halimi and Masri; Aharthan and HarthanOumazer; Chikh, MhammedChikh and Hamuri; Bent Cherk and Cherka; Tati and Tacherwint; Kesba and Sokrya; Azizaou and Adam Zrak; Alig, Bu'Rus and Our'Rous. In a few cases, we also observed varieties that shared the same name but were relatively distant. For example, three samples of the variety Hamraya collected at three different oases (Biskra, Touggourt and Oued Souf) differed genetically.

4. Population structure and differentiation

Analysis of population structure led us to keep three sub-populations (Figure 18) based on Evanno statistics (Figure 18 b). Among the 192 accessions with distinct genotypes, 186 were assigned to a single group (Figure 18). Only six accessions remained "unclassified" and were considered "admixed". The first group included 51 varieties from the north-eastern and north-central part of the Algerian Sahara. The second group, which was the smallest with 48 varieties, was present in every oasis. The third group was the largest with 87 varieties, including varieties from the north-western, north-central and southern part of the Algerian Sahara, and 20% from the north-eastern part.

Allelic richness was lowest in the northern group (Table 14, Annex 10). No difference in observed and expected heterozygosity was found between the three groups (Table 14, Wilcoxon signed-rank test, p -value > 0.05) (Annex 11, 12). The inbreeding coefficient F_{IS} was negative and significant in the first and third group (Table 14, Annex 13). Only group 2 had no excess of heterozygotes. Pairwise F_{ST} values in the three groups ranged from 0.066 to 0.092 (Table 15). The highest differentiation was found between group 1 and group 2 ($F_{ST} = 0.0924$). The three groups are thus genetically different (significantly different at $p < 1\%$).

A relationship was found between the three genetic structure clusters and the two chlorotype frequencies (Figure 18, $\chi^2 = 28.1$, $dof = 2$, $p < 0.001$).

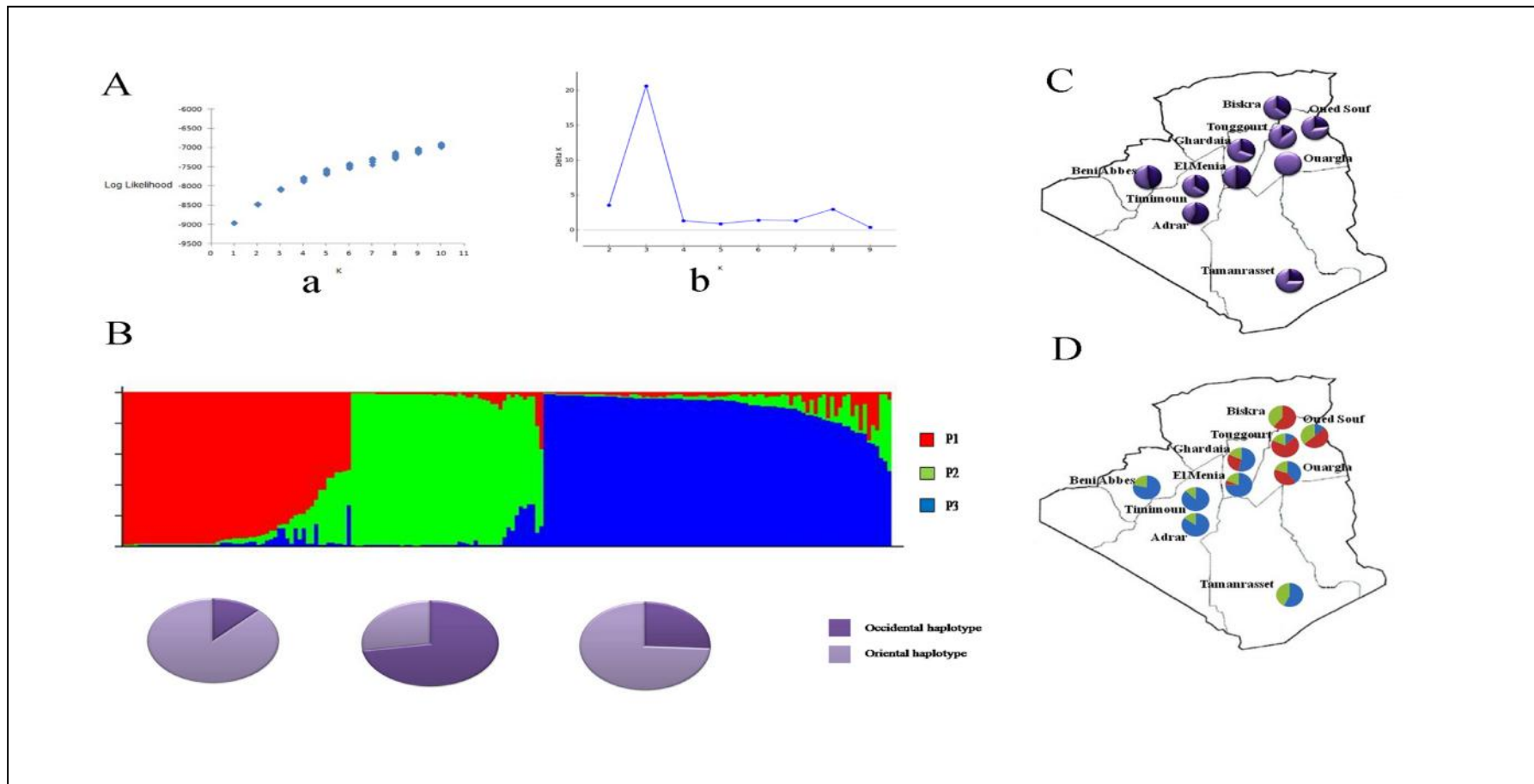


Figure 18 : Bayesian cluster analysis using the STRUCTURE program: results for K = 3.

(A) Plots of (a) maximum log likelihood over the 10 runs and (b) delta k from the structure analysis was calculated according to the method of Evanno et al.

(B) Estimated population structure inferred from all the individual date palms for K = 3. Each individual is represented by a thin vertical line divided into K colored segments representing the fraction of the individual's estimated membership of the K clusters. Pie charts show the frequencies of the haplotypes belonging to the three structure groups.

(C) Geographic distributions of the 10 oases and two haplotypes. The pie chart shows the proportions of haplotypes in each oasis.

(D) Sampling location of the date palms. Pie charts show the proportion of membership of each sampled population inferred by structure analysis (K = 3).

Table 14 : Description and variability parameters of the 18 polymorphic microsatellite loci in the three STRUCTURE populations.

Population	N°	Na	Ar	Ho	He	F _{IS}	P value F _{IS}
Population 1	51	5.28	5.15	0.60	0.53	-0.13	0.000002
Population 2	48	6	5.88	0.57	0.56	-0.004	0.426301
Population 3	87	6.94	6.32	0.61	0.58	-0.05	0.000101

The number of samples (N°), the number of alleles per locus (Na), allelic richness (Ar), expected heterozygosity (He), observed heterozygosity (Ho), the inbreeding coefficient (F_{IS}) and the p-value of F_{IS} are given for each population and each locus.

Table 15 : Genetic distance between Structure populations (F_{ST}).

	population 1	population 2	population 3
population 1		0.0924*	0.0742*
population 2			0.0659*
population 3			

The F_{ST} component was determined using a method that calculates the genetic distance between two assumed populations, the differences being in pairs. A Bonferroni correction was applied, the significance level passing $p < 0.01$.

*: $p < 0.01$ (significant)

5. Genetic distance and dendrogram construction

The matrix of estimated distances between individuals of the total population was used to construct a phylogenetic tree of individuals in the total population. The distances ranged from 0 to 0.75, underlining the wide genetic variability of our study population. Zero distance between two individuals suggests a clonal relationship (Figure 19).

Analysis of the phylogenetic tree showed that the individuals were grouped independent of their geographic origin and ethnic name ecotype. The distribution on the phylogenetic tree can be explained by the existence of a common genetic basis between different populations and ecotypes, despite geographic distance and phenotypic divergence.

Discussion

a. Genetic diversity and population structure of date palm

The diversity of Algerian date palm was shown to be high with an average number of alleles per locus of 7.94, similar to the number observed in Qatar, Iraq or Tunisia (Zehdi *et al.* 2004 ; Khierallah *et al.* 2011 ; Elmeur *et al.* 2011, 2012 ; Hamza *et al.* 2013). Seven populations had negative fixation indices (F_{IS}), suggesting an excess of heterozygosity at these loci. Two populations, Touggourt ($F_{IS} = -0.06$) and Ghardaia ($F_{IS} = -0.13$), had significant positive fixation indices (F_{IS}). A negative F_{IS} value suggests that different heterozygous genotypes have been maintained. Negative and significant F_{IS} value were also found when individuals were grouped based on Bayesian analysis of population structure. This particular signature is not very common in plants, and may indicate direct selection for individuals with the best performance, i.e. those with high heterozygosity. It suggests farmers had a direct impact on maintaining the most heterozygous plants (Pujol *et al.* 2005). As date palms are also propagated by stem (offshoots), it should be noted that if all the plants were at Hardy-Weinberg equilibrium, propagation by stem would not change or increase F_{IS} through the production of clones. So this excess of heterozygous plants could only have been maintained by choosing the fittest individual (Pujol *et al.* 2005).

The difference in F_{ST} between oases was relatively low, ranging from 0.0016 to 0.0701. However, slightly higher differentiation was observed between oases in the northern and southern Algerian Sahara. This could reflect the use of different varieties in different ecological conditions. In the Algerian Sahara, the climate is characterized by aridity and heat, and gradually becomes hotter and drier from north to south. The northern regions of the Sahara are characterized by a semi-arid climate, which is the case of Biskra located at the foot of the Saharan Atlas mountain range, the natural boundary between northern and southern Algeria. Similarly, Touggourt, Oued Souf, Ouargla and Ghardaia have a semi-arid to arid climate. Southern Saharan regions, i.e. Beni Abbas, Adrar, Timimoun and Tamanrasset are characterized by an extremely hot dry desert climate. The analyses of structure also support the northern and southern grouping of diversity.

b. Explanation for the genetic history of diffusion of date palm cultivars in Algeria

Recent studies based on the global diversity of date palm suggest a geographic differentiation between African and Middle Eastern date palms (Zehdi *et al.* 2015 ; Mathew *et al.* 2015). Chloroplastic diversity and Pairwise the non-recombinant portion of the segment of the Y chromosome both reveal differentiation between African and Middle Eastern cultivars (Cherif *et al.* 2013). A hypothesis of two origins, one in the east and the other in the west has been proposed (Zehdi *et al.* 2015 ; Mathew *et al.* 2015). However, it is likely that this differentiation was caused by successive bottlenecks during the diffusion of cultivated date palms. However that may be, the diversity of the chloroplast is structured with major Eastern and Western alleles. Chloroplast genomes help trace the path of the female lines (Pintaud *et al.* 2013), and our chloroplast data provided insight into this maternal structure.

Interestingly, we observed a high frequency (almost 75%) of the eastern (Oriental) chloroplast in a specific group of varieties. In this particular group, the frequency was similar to that found further east. Conversely, in the two other structure groups, chloroplastic allele frequency was similar to that observed further west.

The 18 microsatellite loci used in the study differed between Eastern and Western date palms (Zehdi *et al.* 2015). Algeria does not differ particularly from other Western countries and falls in the Western genetic group. While chloroplasts inherited maternally suggest some proximity to the east, the nuclear markers clearly show a proximity to the west. One hypothesis proposed to explain this peculiar pattern, is that seeds and offshoots were imported from the Middle East to North Africa, and, after generations of crossing with local males (Western nuclear genome), varieties acquired a Western nuclear genome although an Eastern chlorotype was maintained. Such varieties therefore carry the Eastern haplotype. Another possible explanation is that both Eastern and Western haplotypes were widespread in both the Eastern and Western Sahara, but because farmers selected a specific genetic group of plants, chloroplastic diversity varied considerably from one local genetic group to the other in the west. Whatever the scenario, this pattern is only compatible with human selection that maintained the maternal lineage separately. The fact that we also found that in Algeria, nuclear SSR is structured in three groups shows that the groups were also to a certain extent kept separate by controlled pollen flow.

In conclusion, humans shaped the diversity of date palms by selecting heterozygous individuals and maintaining them. Moreover, cryptic diversity was observed both at genome and chloroplast level. Taken together, our results suggest that human selection played a major role in maintaining the diversity and lineage of date palms in Algeria.

Chapitre 4
Constitution d'une clé d'identification de
variétés de palmier dattier

Introduction

Le palmier dattier (*Phoenix dactylifera* L.) représente le pilier de l'agriculture dans les zones oasiennes où il joue un rôle socioéconomique et écologique majeur. L'espèce *Phoenix dactylifera* L. conserve une diversité importante dans les oasis algériennes. Chaque palmeraie possède une composition variétale qui lui est propre, résultant d'une sélection locale établie selon les caractères du fruit, au sein d'une oasis déterminée (Benkhalifa 1999 ; Tirichine *et al.* 2011).

En dépit de son importance économique, les phoeniculteurs ont toujours essayé de discriminer traditionnellement les variétés de palmier dattier, en se basant sur des caractères morphologiques, notamment les caractères du fruit qui sont nécessaires pour toute reconnaissance variétale des individus issus de rejets ou de graines. Plus d'un millier de cultivars traditionnels sont recensés en Algérie (Benkhalifa *et al.* 2004). Cependant, des contraintes biotiques et abiotiques ainsi qu'une monoculture intensive ont entraîné un appauvrissement de l'agrobiodiversité du palmier dattier. Pour limiter l'érosion génétique, il est nécessaire de caractériser génétiquement la diversité phoenicole afin de recenser les ressources génétiques et d'organiser leur conservation et leur sauvegarde.

De nombreux travaux visant à caractériser la diversité des palmeraies algériennes ont été réalisés à l'aide de marqueurs morphologiques, phénotypiques, biochimiques ou moléculaires.

Les travaux de Bennaceur *et al.* (1991) basés sur des systèmes enzymatiques ont été d'un grand apport dans la discrimination des variétés de palmier dattier.

Des études de marqueurs RAPD (Random Amplified Polymorphic DNA) ont été réalisées par Bouchireb & Clark (1997) en utilisant 6 amorces arbitraires de 39 individus de 8 variétés, et ont montré la variabilité intra-variétale pour certaines variétés.

Avant le développement des microsatellites pour le palmier dattier, 24 amorces microsatellites de Cocotier (*Cocos nucifera* L.) et 18 amorces microsatellites de palmier à huile (*Elaeis guineensis* Jac.) ont été testées sur 50 individus de 10 variétés algériennes de dattiers. Les résultats ont montré le polymorphisme et la transférabilité de ces marqueurs au palmier dattier (Bennaceur *et al.* 2000). Ces auteurs ont pu analyser la diversité génétique de quelques variétés de palmier dattier. L'utilisation de 12 paires d'amorces chloroplastiques universelles de tabac (*Nicotiana tabacum*) sur 60 variétés de dattiers algériens par la méthode SSCP et CAPS a généré un faible polymorphisme chez ces cultivars (Bennaceur *et al.* 2002).

L'utilisation de ces marqueurs biochimiques, enzymatiques et moléculaires a été insuffisante pour discriminer la diversité du dattier en Algérie. Par la suite, il a été primordial de standardiser les procédures d'identification des génotypes à l'aide de marqueurs moléculaires plus puissants qui permettront de caractériser l'agrobiodiversité et de mettre en place des stratégies de conservation.

Billote *et al.* (2004) ont développé les premiers microsatellites de palmier dattier. Leur utilisation par de nombreux chercheurs a permis une percée dans l'étude de la diversité génétique du palmier dattier et de l'identification des variétés clés dans plusieurs pays (Zehdi *et al.* 2004). Les marqueurs moléculaires de type microsatellite ont montré par la suite un intérêt primordial dans l'analyse phylogénétique du palmier dattier dans différents pays phoenicicoles (Al-Ruqaishi *et al.* 2008 ; Akkak *et al.* 2009 ; Elshibli & Korpelainen 2011 ; Elmeer *et al.* 2011 ; Elmeer & Mattat 2012 ; Bodian *et al.* 2012 ; Hamza *et al.* 2013 ; Al-Faifi *et al.* 2016).

Le séquençage partiel du génome du palmier dattier récemment publié (Al-Mssallem *et al.* 2013 ; Hazzouri *et al.* 2015), a offert la possibilité d'accéder à un grand nombre de ces marqueurs et d'évaluer leur pertinence dans la perspective de la caractérisation de l'agrobiodiversité de l'espèce.

En Algérie, une étude récente basée sur les marqueurs moléculaires ISSR et RAPD a été réalisée pour quantifier la diversité génétique et établir une relation phylogénétique entre 20 variétés algériennes collectées dans l'oasis de Bou-Saada (Guettouchi *et al.* 2017). La culture du palmier dattier dans l'oasis de Bou-Saada représente l'une des cultures traditionnelles de l'Atlas saharien qui soutient son système agroécologique et fournit des sources de nourriture et de revenus à ses habitants.

Aucune étude rapportant la diversité génétique des variétés de dattier en Algérie, en se basant sur les marqueurs microsatellites, n'a été jusque là effectuée.

L'objectif de cette étude est de mettre au point une clé d'identification des variétés de palmier dattier en utilisant des marqueurs microsatellites nucléaire et chloroplastique.

Résultats

Dans le but d'identifier les variétés de *Phoenix dactylifera* rencontrées en Algérie, une collection de 135 variétés algériennes a été échantillonnée à partir de différents oasis du sud algérien et comparée avec des variétés du Maroc (2 variétés), Tunisie (18 variétés), Irak (6 variétés), Pakistan (5 variétés), Soudan (2 variétés), Arabie Saoudite (2 variétés), Italie (2 variétés), Espagne (1 variété), Niger (1 variété), Mauritanie (1 variété), Lybie (1 variété), Egypte (1 variété), Californie (1 variété) et Oman (1 variété).

Six locus nucléaires et un locus chloroplastique ont été utilisés pour établir la clé d'identification variétale. Le principe de la classification hiérarchisée des locus selon leur nombre d'allèles a été utilisé.

Cette clé a permis la séparation des génotypes, et a donné de l'information sur les affinités des individus. Elle présente un maximum de fiabilité par la séparation des variétés selon le chlorotype. La structure de cette clé a été résumée dans un arbre de distance génétique établi par sept locus et identifié les clones, avec une lecture complète de l'agrobiodiversité (Figure 20).

Selon la valeur de puissance discriminante pour chaque locus, nous avons testé différentes combinaisons commençant par les plus discriminants et en ajoutant un locus à chaque étape. Le locus chloroplastique CpfM12 regroupe les génotypes en deux grands groupes : groupe à haplotype 3 et groupe à haplotype 4. L'analyse des résultats montre que les locus CpfM12, mpdIRD031, mpdIRD033 permettent d'identifier 7% des génotypes étudiés. Le locus mpdIRD031a généré 2 génotypes, tandis que le locus mpdIRD033 a donné huit autres génotypes. La combinaison des locus CpfM12, mpdIRD031, mpdIRD033, mpdIRD040, PdCUC3-ssr2, PdAP3-ssr-F4 permet de différencier 92% de génotypes. Avec un dernier locus mPdCIR78, il était possible de différencier génétiquement toutes les variétés étudiées sur la base de leur génotype multilocus (Figure 20).

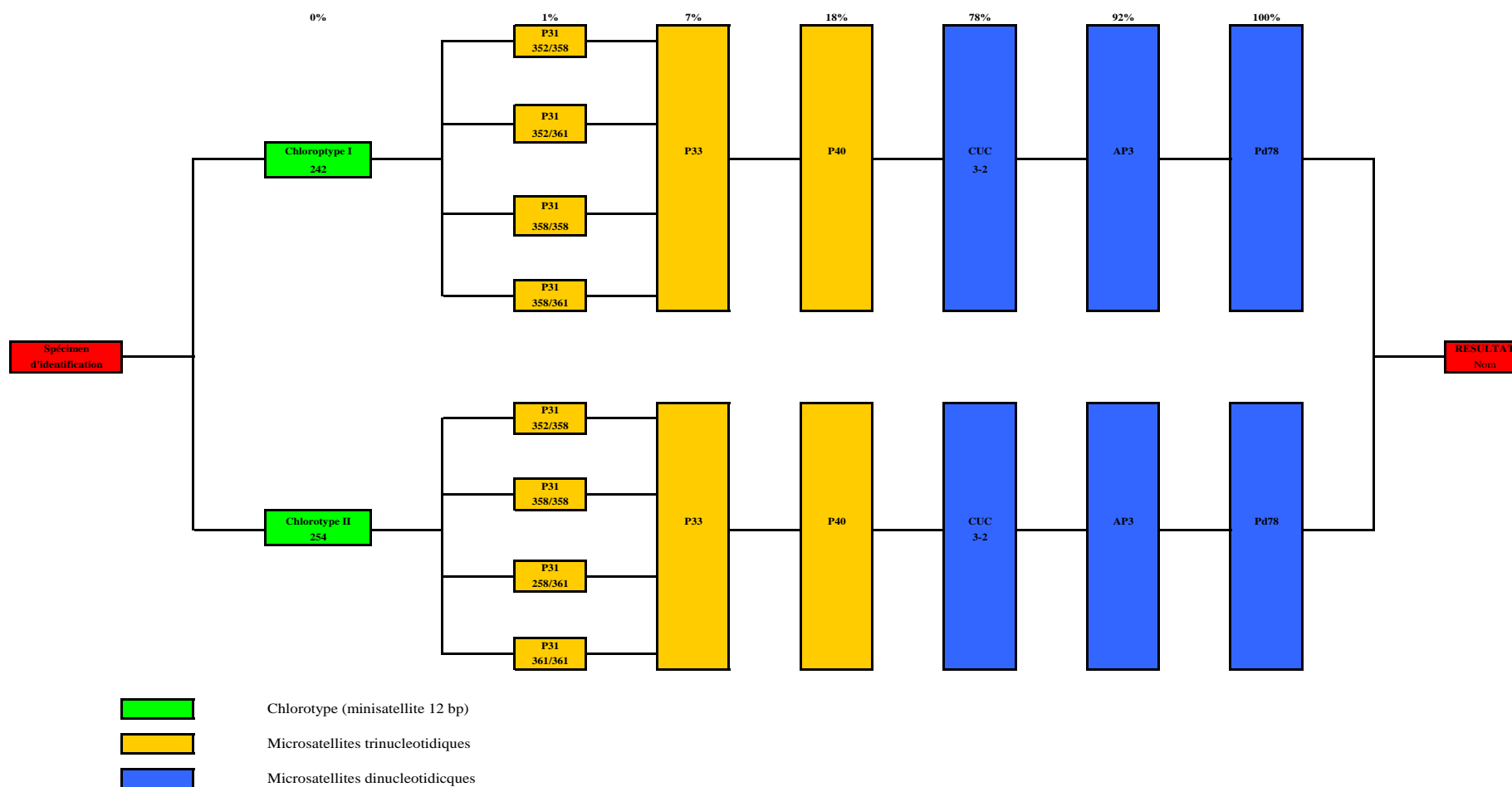


Figure 20 : Fonctionnement du système d'identification du dattier à 7 marqueurs.

La clé d'identification est présentée dans la figure 21. Ce diagramme a montré un pouvoir de résolution de 100% puisqu'il a permis de différencier sans ambiguïté les 177 génotypes étudiés. Ce dendrogramme a donné une image globale de la classification des variétés de dattier, y compris les homonymes et les synonymes (Figure 21).

Les synonymes ont été observés pour seize génotypes, avec deux à trois dénominations par génotype. Nous avons trouvé ce phénomène de synonymie entre les variétés des différents pays (Tableau 16).

Tableau 16 : Synonymes des dénominations variétales des pays.

Appellation algérienne	Appellation étrangère
Arechti	Arichti (Tunisie)
Fagous	Bou feggous (Maroc)
Amari	Amari (Tunisie)
Horra	Horra (Tunisie)
Halimi, Masri	Zaidi (Irak), Zaidi (Pakistan)
Deglet Nour	Deglet Nour (Tunisie, Californie)
Zohdi	Halawi (Irak)
Ghars	Angou (Tunisie)
	Medjool (Maroc), La Confitera (Espagne)

Ainsi des synonymes ont été mis en évidence dans une oasis donnée et d'une oasis à une autre (Tableau 17).

Tableau 17 : Synonymes des dénominations variétales des oasis algériennes.

Biskra	Oued Souf	Touggourt	Ouargla	Ghardaïa	Timimoun	Adrar	Béni Abbès
					Chikh	Chikh Mhammed	Hamuri
	Alig / Bu'Rus			Our'rous			
						Ben Cherk	Cherka
	Tati / Tacherwint						
	Kesba				Sokrya		
					Azizaou / Adam Zrak		
						Aharthan	Harthan / Harthan Oumazer
Tanteboucht	Tanteboucht	El Kayed	Tanteboucht				

La clé d'identification a été utilisée également pour détecter les variétés portant le même nom mais en réalité elles sont différentes. Les variétés qui ont été identifiées comme homonymes sont :

- La variété Ghars (Algérie) est homonyme à Ghars (Tunisie)
- Bou Feggous (Maroc) est homonyme à Bou Feggous (Tunisie)
- Alig (Algérie) est homonyme à Alig (Tunisie)
- Kenta (Algérie) est homonyme à Kenta (Tunisie)
- Bayd Hmam (Algérie) est homonyme à Bidh Hmam (Tunisie)
- Tanteboucht (Algérie) est homonyme à Tantabecht (Tunisie).

Les homonymes élucidés à travers cette étude illustrent la confusion dans la nomenclature qui peut être observée entre les pays.

CpfM12 mPdIRD031 mPdIRD033 mPdIRD040 PdCUC3-ssr2 PdAP3-ssr-F4 mPdCIR078

					146:152	Halwa
		184:193	333:333	257:259		
					146:154	Halwaya
352:358	198:198		301:365	234:260	133:136	Ashrasi (Irak)
		187:193	317:261	257:259	146:146	Kentichi (Tunisie)
			317:331	257:259	146:146	Mech Degla
352:361	198:198	184:193	287:331	243:247	136:146	Dguel Mghas
	189:198	187:193	331:331	257:259	117:146	Dokhar
		193:196	261:333	243:247	117:152	Arécthi (Algérie)
	195:198					Arichti (Tunisie)
		193:193	287:319	243:247	117:117	Deglet Jdir
			331:351	249:259	117:146	Barakawi (Soudan)
			333:333	243:257	122:146	Outaghsait
			287:329	247:247	117:117	Buféa
			309:331	247:247	133:146	Adam Goas (Niger)
		184:193				
			319:331	247:257	117:146	Deglet Wlad Mahmoud
			261:287	257:257	117:127	Ghars (Tunisie)
				243:257	117:136	Akerbouche
			261:351			
				247:257	117:117	Baydir
			261:319	243:257	117:122	Tinjdel
			303:331	247:257	117:117	Abbad
			307:333	247:257	123:146	Timakur
		187:193				
			309:333	247:259	123:146	Fagous (Algérie)

					Bou Feggous (Maroc)	
			317:333	247:257	117:146	Alig (Tunisie)
			317:329	251:259	123:146	Thorry
			331:351	243:247	117:146	Kenta (Algérie)
			333:333	245:247	117:123	Medjool (Maroc)
						La Confitera (Espagne)
				241:247	117:133	Bidh-Hmam (Tunisie)
			261:351			
				247:257	117:117	Kenta (Tunisie)
	358:358		187:196			
			329:351	247:257	117:117	Chikh Hamuri Chikh Mhammed
		198:198				
					133:146	Alig (Algérie)
			261:333	247:247		Bu'Rus Our'rous
					117:146	Tamazouart
			261:331	257:257	117:136	Deglet Gurara
			261:287	247:257	117:117	Lagou (Tunisie)
			303:317	243:257	117:133	Bayd Hmam (Algérie)
			307:331	234:259	120:146	Bitamoda (Soudan)
			193:193			
			309:329	247:249	142:152	Bent Cherk Cherka
			317:329	241:257	117:123	Bukhannus
			317:331	241:257	117:123	Degla Baida
			319:333	247:257	117:122	Tinasser
			287:351	243:257	117:152	Timedjwel
			303:329	247:257	122:123	Takhoudrayt
242						

		193:196	307:333	247:249	146:146	Bou Feggous (Tunisie)
			333:351	249:259	138:152	Bessir (Tunisie)
			261:351	247:247	117:146	Deglet Bey (Tunisie)
			261:331	245:259	117:152	Sekani (Mauritanie)
		184:187	331:331	245:249	117:133	Talees (Libye)
		193:193	329:329	243:257	123:152	Tinazidhane
	198:201	187:187	331:331	249:257	127:146	Tinhamour II
	201:201	187:193	287:329	257:257	127:133	Tinhamour I
	189:198	193:193	287:331	257:257	117:138	Hmara (Tunisie)
		187:196	261:331	243:257	117:117	Gasbi (Tunisie)
	195:198	193:196	319:332	243:257	117:117	Mes'udiya
			303:329	247:257	114:117	Tadmamt
		187:193	309:317	245:259	123:152	Tindukan
	358:361		351:351	247:257	117:117	Amari (Algérie) Amari (Tunisie)
			309:319	247:259	146:152	Takerboucht
		193:193	319:329	247:257	117:152	Tazizawt
			309:329	247:257	117:133	Ouarglia
		193:196	319:319	243:259	123:146	Aghares
	198:198	184:193	287:329	247:257	117:117	Goundi (Tunisie)
		187:196	331:351	247:247	123:152	Horra (Algérie) Horra (Tunisie)
			303:303	234:251	133:133	Rachna (Pakistan)
	189:198					

352:358		193:193				Halimi (Algérie)	
			319:351	234:247	133:136	Masri (Algérie)	
						Zaidi (Irak)	
						Zaidi (Pakistan)	
		198:198	196:196	331:331	247:257	117:152	Romana (Italie)
		198:201	187:187	261:331	247:257	123:146	Dokhar
		189:189	187:193	319:351	234:247	133:133	Aseel (Pakistan)
			187:193	319:319	234:251	137:142	Barhi (Irak)
		189:198					
			193:193	319:351	241:243	122:135	Tadala
			187:193	287:319	243:259	133:133	Mcharret
		195:198		333:333	243:259	117:123	Taghera
			193:193				
				303:351	243:247	117:136	Tati
							Tacherwint
			184:184	307:332	247:247	117:117	Timrisa
				261:331	247:257	123:152	Deglet Nour (Algérie)
							Deglet Nour (Tunisie)
			184:187				Deglet Nour (Californie)
				309:319	259:259	117:117	Adam Hror
			261:309	247:257	120:152	D'guel El Hadj	
		184:193	319:329	247:247	122:133	Banekhluf	
			319:331	243:247	122:127	Bukezzine	
			303:319	234:247	137:152	Zohdi (Algérie)	
		187:187				Halawi (Irak)	
			319:331	249:259	123:133	Adam Timu	
				243:247	117:133	Chataya	
			261:261				
				247:247	120:133	Tawadan	

			261:303	243:247	117:136	Yellow Lulu
			261:319	247:247	117:122	Bouldjib
			261:331	243:247	123:146	Aghaliane
			261:331	243:247	122:152	Ain Zbib
				243:257	136:146	Tazerzist
			261:351	243:247	117:152	Deglet Jito
			287:319	247:259	117:133	Adam Bchir
			287:331	249:259	120:133	Ma'tug
			301:301	234:234	136:142	Ajwa (Arabie Saoudite)
			303:303	251:253	129:129	Basra (Irak)
			303:3019	245:251	131:133	Fardh (Egypte)
			303:331	234:259	117:136	Kesba
						Sokrya
			307:319	249:251	133:136	Khadrawi (Irak)
		187:193	309:351	243:247	120:152	Deglet Mech'a
			319:319	251:251	133:136	Khalass (Arabie Saoudite)
	198:198		319:329	243:247	122:152	Ajoullil
			319:351	243:247	117:117	Azizaou
					Adam Zrak	
					117:122	Tafeziwin
	358:358		331:331	247:247	146:152	Adam Deglet Nour
				247:259	117:122	Timliha
			331:351	243:257	117:122	Adam Esof

254		261:303	243:257	117:117	Dguel Sidi Khilil
		287:287	243:259	117:117	Ain Zbib
				122:146	Ououcht
			243:259	133:146	Tinhud
		287:319	243:243	117:146	Tawragha
			243:247	122:133	Tissibi
		287:331	247:257	117:123	Ighes N'wagada
		287:351	243:247	117:117	Tilemsu
				146:149	Hilali (Oman)
		301:317	234:251	147:147	Hussaini (Pakistan)
		309:319	243:259	117:122	Gharas
		319:319	245:257	117:128	Aghammu
			237:243	136:148	Bent Qbala
		319:331		135:152	Nasser Oussaleh
			241:247	122:122	Bawa'adhim
				122:146	Timliha
		319:332	247:257	117:120	Tifazwin
		319:333	237:243	117:136	Tanslit
		319:351	243:247	117:136	Hmira
			247:247	117:146	Taghera
		331:331		117:122	Tgaza
			247:259		
		333:333	247:257	122:146	Asemmat
		351:351	243:247	117:146	Adam Bulla

193:193

			261:333	247:257	117:146	Hamraya
	193:196		319:333	243:259	136:148	Tinrigh
			331:331	243:247	117:136	Khadri
	193:202		319:331	243:247	122:146	Tazarzayet
	184:193		333:333	257:259	146:146	Ebrea (Italie)
	187:193	261:331		243:259	117:133	Andekly
				257:257	117:123	Litima
198:201			331:331	243:243	122:122	Harthan
	193:193					Harthan Oumazer
			319:351	243:247	133:136	Timjouhart
189:198	193:196	261:309	247:257	120:152		Red Lulu
195:198	193:196	287:331	247:259	117:122		Ouzamig
	184:193	303:333	243:257	136:146		Tachlilat
		261:261	247:257	117:146		Dimolo
	187:193	261:331	243:247	117:120		Tawraghet
		309:331	243:247	117:120		Tazougart
		261:303	241:243	136:146		Tantabecht (Tunisie)
			243:257	117:136		Tinicine
		261:309	243:247	120:120		Hamraya
			247:247	117:120		Ghars (Algérie)
						Angou (Tunisie)
198:198		303:329	241:243	136:146		Tanteboucht (Algérie)
						El Kayed
258:361	193:193	303:331	243:243	136:136		Abdel Azzaz

			303:332	241:243	117:136	Bacheir
			309:309	247:247	120:123	Zog El Moggar
			309:317	241:247	120:123	Taramount
			309:329	247:257	117:117	Bahdid
			309:351	247:247	117:120	El Gachouche
			333:333	257:257	117:120	D'guel M'rass
		193:196	261:303	243:257	136:152	Takermoust
		187:193	261:319	243:247	117:136	Tizouyadj
	198:201	193:193	303:329	241:257	122:123	Zaghraya
361:361	198:198	184:193	261:309	243:257	120:136	Dfor Lgot

Figure 21 : Clé d'identification de 155 génotypes de palmier dattier (*Phoenix dactylifera*), basée sur six marqueurs microsatellites nucléaires et un marqueur minisatellite chloroplastique.

Nos résultats fournissent la première clé d'identification pour les génotypes algériens, qui offre un outil intéressant pour les phoeniculteurs afin de gérer et certifier le matériel végétal. Cette clé d'identification a permis la discrimination de toutes les accessions étudiées, a confirmé la fiabilité des marqueurs microsatellites pour l'empreinte digitale des génotypes, et la conservation du matériel génétique du dattier algérien. En outre, étant donné que de nombreux écotypes sont cultivés dans le monde entier et que les échanges de matériel entre les phoeniculteurs sont courants, la disponibilité d'un système d'identification basé sur les SSR est intéressante.

Discussion

a. Concepts de variété, de cultivar

La variété en botanique (du latin *varietas* ‘qui diverge’) est un rang taxonomique de niveau inférieur au rang de l’espèce. Elle rassemble un groupe de plantes que l’on peut distinguer des autres variétés de la même espèce grâce à leurs caractéristiques. Le terme cultivar est synonyme de ‘variété cultivée’, et plus communément ‘variété’. C’est un ensemble de clones d’un même individu qui a été propagé végétativement la plupart de temps et sélectionné, parfois depuis plusieurs millénaires, pour certaines de ses caractéristiques que l’on a voulu transmettre d’une génération à l’autre. En matière de gestion des ressources génétiques, la variété est souvent considérée comme l’unité minimale de perception et de conservation de l’agro-biodiversité (McKey *et al.* 2001).

Les caractéristiques importantes lors de l’identification d’une variété sont l’origine historique, l’adaptation locale et l’identité (Camacho Villa *et al.* 2006), ce qui peut être confirmé en utilisant un système de marqueur moléculaire d’une puissante capacité de discrimination. Généralement, la plupart des efforts de conservation se sont concentrés sur des cultures agricoles importantes.

Chez le palmier dattier l’identification, l’inventaire et la classification des variétés était une étape primordiale pour caractériser son agro-biodiversité (Elshibli 2010). Les études portant sur l’agro-biodiversité de palmier dattier à l’échelle des oasis, des pays ou de sa zone de distribution dans le monde constituent en fait la majeure partie de la littérature du dattier.

Cependant, très peu d’études fournissent un concept explicite de variété également appelée cultivar et, en fait, derrière cette terminologie commune existe une grande complexité et parfois une confusion. El Houmaizi *et al.* (2006), a également introduit le terme de landrace (races de pays), partiellement clonal ou entièrement multiplié par graine, mais conservant par consanguinité un ensemble de traits morphologiques stables.

Du point de vue des phoeniculteurs, la variété se distingue par des critères morphologiques. Elle correspond en ce sens à un ‘morphotype’ (Barnaud *et al.* 2007). Cette conception de la variété contraste avec la vision du généticien qui se fonde sur le génotype pour identifier des variétés.

Les travaux effectués à ce jour montrent que la description morphologique et la représentation basée sur les noms des variétés constitue un point d'entrée important pour l'évaluation de l'agro-biodiversité. Actuellement, le concept de la variété est perçu à un niveau génétique.

Les études portant sur plusieurs individus de même variété du dattier (Zehdi *et al.* 2004 ; El Houmaizi *et al.* 2006 ; Bodian *et al.* 2012) permettent de classer les variétés dans différentes catégories génétiques. Les variétés sont précisément définies lorsqu'elles sont strictement clonales. Dans ce cas, tous les individus appartenant à la même variété sont génétiquement identiques. De telles variétés sont dominantes dans l'agro-biodiversité des régions où la pratique de la multiplication végétative par des rejets est la règle, comme l'Algérie, la Tunisie, l'Oman et l'Irak. La majorité des variétés d'élites propagées massivement appartiennent à cette catégorie, y compris 'Deglet Nour', 'Zaidi', 'Halawi', 'Medjoul' et 'Aseel'. Ainsi, ces variétés élites faisant l'objet de culture intensive peuvent ne représenter chacune qu'un seul clone, abondamment multiplié.

Les seuls problèmes trompeurs pour les variétés sont nomenclaturaux, c'est-à-dire l'utilisation de différents noms pour le même clone. Ces variétés qui forment le même clone ne sont pas nécessairement reconnaissables morphologiquement en raison de modalités de sélection différentielles. Par exemple, selon notre étude la variété 'Medjool' au Maroc est la même que 'La Confitera' d'Elche, ainsi que 'Ghars' en Algérie et 'Angou' en Tunisie. Ces variétés, bien que génétiquement très affines, sont morphologiquement bien distinctes. Ainsi, la variété 'Bitamoda' au Soudan (Elshibli & Korpelainen 2008) est dénommé 'Bertamoda' en Egypte (Saker *et al.* 2006) ; ces variétés identiques portent des noms légèrement modifiés.

Aussi, il existe des problèmes d'homonymie, c'est-à-dire que des variétés différentes peuvent porter un même nom. Cependant, le nombre exact des variétés d'un pays reste difficile à évaluer.

Encadré 6 : Définitions des termes ‘cultivar et ‘landrace’

Landrace : Depuis Von Rünker en 1908, le premier qui a utilisé le terme ‘Landrace’ couramment traduit par ‘variété local’ ou ‘variété du pays’ (Zeven 1998), divers termes qui varient par leur précision et leur applicabilité ont été associés à la notion de ‘Landrace’ ; cultivar primitif, variété primitive, variété paysanne, variété traditionnelle, variété locale et variété populaire. Tous ces termes ont été utilisés comme des synonymes pour désigner ‘Landrace’, mais à chaque terme, il y a une incohérence dans l’application (Camacho Villa *et al.* 2006).

Basés sur une large revue, Camacho Villa *et al.* (2006) ont proposé une définition consensuelle, intégrant les concepts de l’agro-biodiversité : « Une variété locale ‘Landrace’ est une population dynamique d’une plante cultivée qui a une origine historique, une identité distincte et n’est pas issue de programme de sélection génétique moderne, souvent génétiquement diversifiée, adaptée localement et associée à des systèmes agricoles traditionnels ».

Cultivar : est une contraction du terme anglais cultivated variety (variété cultivée) qui présentent des caractères souhaitables et qui a reçu un nom unique. Le terme cultivar n’appartient pas au système de classification taxonomique, il est surtout employé par les sélectionneurs qui est défini par les trois critères de l’UPOV : Distinct, Homogène et Stable (Critères DHS ; UPOV 2002).

Le concept de cultivar répond aux besoins des pratiques de l’horticulture, l’agriculture, la foresterie, etc.

L’utilisation de termes tels que ‘variété’ et ‘cultivar’ pour le terme ‘Landrace’ est la source de confusion car ils se réfèrent plus précisément à sa forme améliorée.

b. Etablissement de la clé d'identification variétale

La zone de production de la datte en Algérie a augmenté de façon spectaculaire au cours des dernières décennies et devrait continuer à augmenter. Le palmier dattier fait partie des rares plantes qui peuvent prospérer dans les milieux arides et peut fournir d'énormes ressources aux tribus nomades et aux populations locales. De plus, les dattes ont une grande importance culturelle et traditionnelle dans les régions du sud algérien. L'installation de palmeraies monovariétales encouragée par des circonstances commerciales favorables pour la variété élite « Deglet Nour » est actuellement très avancée dans la région du sud-est, ce qui a accentué le phénomène d'érosion génétique et le déséquilibre de l'environnement oasien de cette région. L'analyse de la diversité des palmeraies algériennes est d'un apport pour la conservation de notre patrimoine phoénicien.

Les ressources génétiques des palmiers dattiers algériens sont principalement représentées par des variétés traditionnelles qui sont des femelles sélectionnées par les agriculteurs en fonction de leur productivité, la valeur marchande, la valeur nutritive, les qualités gustatives et leur résistance à la sécheresse et aux maladies, notamment le bayoud. Les variétés reçoivent des noms vernaculaires locaux, très souvent, un nom berbère pour le fruit peut indiquer un emplacement géographique, le nom du village ou même un propriétaire de ferme individuelle. La sélection aléatoire par les phoeniculteurs algériens au cours des millénaires a conduit à l'évolution du nombre de variétés. Les palmeraies algériennes conservent une diversité importante; plus d'un millier de variétés traditionnelles sont recensées à partir des critères morphologiques (Benkhalifa *et al.* 2004), chacune adaptée à des types légèrement différents de sol, de température et d'humidité.

A partir de notre collecte dans différentes oasis du Sahara, nous avons révélé que les variétés montrent une répartition très marquée dans les parties Est, Centre et Ouest du Sahara du pays. La plupart des variétés sont endémiques à leur région d'origine ; chaque palmeraie possède une composition variétale qui lui est propre, résultant d'une sélection locale selon les caractères du fruit, au sein d'une oasis déterminée. Brac de la Perrière & Benkhalifa (1989) ont trouvé un taux très élevé d'endémisme de 70% pour les dattes du sud-ouest et plus de 60% en moyenne dans celles des parties sud-est de l'Algérie.

La nomenclature des variétés de dattier est confuse en raison de la longue histoire de la culture, des échanges importants de matériel génétique, du dioïsme et de la propagation par

semis. Pour de nombreuses variétés, il existe un grand nombre de synonymes et d'homonymes entre et à l'intérieur des pays phoenicicoles.

Comme exemple ; les variétés Ghars, Tanteboucht, Alig et Kenta (Algérie) sont des homonymes des variétés Ghars, Tanteboucht, Alig et Kenta (Tunisie). Ainsi, les variétés Arechti, Amari et Horra (Algérie) sont des synonymes des variétés Arechti, Amari et Horra (Tunisie). La variété Bou-Feggous (Maroc) est synonyme de Fagous (Algérie), ces deux sont homonymes de Bou-Feggous (Tunisie). Aussi, les variétés Halimi et Masri sont synonymes de Zaidi (Irak, Pakistan). Ce phénomène de synonymie se reflèterait par la redondance des appellations des variétés de dattier d'une oasis à l'autre, voire d'un pays à l'autre. Un problème pour de telles variétés est également une variation somaclonale (variabilité intra-variétale) (Devanand et Chao 2003 ; El Hadrami *et al.* 2011), c'est-à-dire l'accumulation de mutations somatiques dans les lignées clonales, dont on sait très peu à l'heure actuelle.

Bien qu'il existe de nombreuses variétés de dattes, certains sont devenus prééminents sur le marché mondial (Krueger 2001). 'Deglet noir', qui signifie 'date de la lumière' en arabe, vient de notre Sahara algérien et est l'une des principales variétés cultivées en Afrique du Nord et en Californie. La variété Deglet noir est identique en dépit de son origine géographique, c'est le même clone en Algérie, Tunisie et en Californie. Ce qui est vérifié par des écrits concernant la Californie qui rapporte le transfert de cette variété de El Arfiane (Biskra).

'Halawi', signifiant doux en arabe, est une date douce et de haute qualité avec une saveur riche de l'Irak. Cette variété Halawi d'Irak est un synonyme de la variété Zohdi en Algérie.

'Khadrawy', signifiant vert en arabe, est un arbre court et moyennement productif avec des fruits mous du sud de l'Irak. 'Medjool', datte originaire du Maroc, produit de gros fruits largement acceptés par les marchés du monde entier. Ces deux variétés, Khadrawi (Irak) et Medjool (Maroc), les plus importantes produites dans leurs pays d'origine, n'ont pas de synonyme en Algérie.

Les synonymes et les homonymes parmi les arbres fruitiers ont été largement rapportés (Saddoud *et al.* 2007 ; Ahtak *et al.* 2009 ; Gouta *et al.* 2012 ; Jbir *et al.* 2012).

En conséquence, tous les individus d'une variété donnée, sont identiques entre eux et ils dérivent les uns des autres à partir d'une même plante mère pour constituer un clone, vu que le palmier dattier est essentiellement multiplié végétativement par rejets. Toutefois, les facteurs environnementaux ainsi que certaines mutations pourraient se produire tout en étant à l'origine des variations observées concernant les caractères végétatifs de la plante et ou du

fruit. En effet, il a été constaté que certaines variétés délocalisées de leur lieu d'origine (Deglet Nour) produisaient des fruits différents sous l'effet de conditions environnementales différentes

Les microsatellites sont les marqueurs les plus puissants en génétique des végétaux. Pour cela, nous avons adopté ces derniers afin de caractériser moléculairement les palmeraies algériennes.

Pour identifier efficacement toutes les accessions de dattier, nous avons proposé un ensemble constitué d'un locus minisatellite chloroplastique et six loci SSR nucléaires. Cet ensemble a été choisi principalement en fonction de la puissance discriminante des marqueurs SSR individuels. Premièrement, nous avons choisi un locus chloroplastique pour discriminer les deux grands groupes de génotypes (génotype à haplotype 3 et à haplotype 4). Puis, nous avons vérifié la liaison d'équilibre entre les marqueurs SSR nucléaires et sélectionné les six loci les plus discriminants qu'on a ajouté pour discriminer tous les génotypes étudiés.

L'utilisation d'un locus chloroplastique et six locus microsatellites nucléaires étaient suffisants pour établir les empreintes génétiques de nos variétés étudiées.

Comme les allèles d'un locus ne sont pas liés aux allèles des autres loci, les six loci SSR avec le locus CpfM12 ont pu montrer à travers le diagramme un pouvoir de résolution de 100%, puisque ce dernier a permis de différencier sans ambiguïté les écotypes étudiés, et indiquant ainsi la netteté du modèle de clé d'identification obtenu.

Il convient de souligner que ce pourcentage obtenu est semblable à celui rapporté dans l'étude de Zehdi *et al.* (2004, 2012) pour la Tunisie. Cependant, parmi les six loci SSR proposés et utilisés pour caractériser les variétés algériennes, un a été utilisé dans la clé d'identification des variétés tunisiennes du dattier, il s'agit du locus mPdCIR78 (Zehdi *et al.* 2004, 2012). Ces auteurs ont pu différencier 49 écotypes tunisiens sur la base de trois locus microsatellites nucléaires (Zehdi *et al.* 2004), puis ils ont établi le typage moléculaire de 102 accessions différentes avec cinq locus avec une efficacité absolue de 100% (Zehdi *et al.* 2012).

Ce pourcentage d'efficacité de 100% obtenu est supérieur à celui trouvé par l'utilisation des marqueurs isoenzymatiques avec un pourcentage de 93% (Ould Mohamed Salem *et al.* 2001). Des marqueurs moléculaires de type CAPS utilisés pour l'identification de 38 variétés ont abouti à un pourcentage d'efficacité de 71% (Sakka 2004). Des pourcentages inférieurs à 100% mais presque similaires ont été trouvés dans d'autres espèces de culture fruitière tel que l'abricotier et le figuier (Krichen *et al.* 2006 ; Saddoud *et al.* 2007 ; Baraket *et al.* 2010).

L'élargissement du nombre de microsatellites et des écotypes étudiés a été d'un grand intérêt pour confirmer l'efficacité de notre clé d'identification.

c. L'efficacité de la clé d'identification

L'efficacité des clés d'identification à l'aide de microsatellites a été rapportée chez plusieurs espèces de plantes comme l'olivier (Pasqualone *et al.* 2004), l'abricotier (Krichen *et al.* 2006), le figuier (Saddoud *et al.* 2007 ; Ahtak *et al.* 2009), l'amandier (Gouta *et al.* 2012), le grenadier (Jbir *et al.* 2012), et l'oranger (Mahjbi *et al.* 2015), mais aussi chez le dattier en Tunisie (Zehdi *et al.* 2004). Cette dernière étude était basée uniquement sur le dépistage de dix loci SSR, alors que dans la présente étude, nous avons testé 18 loci SSR et sélectionné les six marqueurs les plus efficaces.

Quelles nouvelles connaissances la présente étude nous donne-t-elle sur la caractérisation des variétés de dattier? Auparavant, Zehdi *et al.* (2004) ont caractérisé une collection de dattier tunisiennes à l'aide de trois loci SSR et ont proposé la première base de données moléculaires pour la gestion du germoplasme du dattier. Cependant, cette étude a été limitée à un petit nombre de loci SSR, qui ont été sélectionnés parmi un ensemble de 10 microsatellites développés dans le dattier (Billotte *et al.* 2004). Dans la présente étude, nous avons proposé une clé d'identification moléculaire après des étapes de sélection et de validation des marqueurs. D'abord, à l'aide de critères moléculaires et génétiques, nous avons sélectionné 18 loci SSR développés par plusieurs équipes de recherche (Billotte *et al.* 2004 ; Adam *et al.* 2010 ; Ludeña *et al.* 2011 ; Aberlinc *et al.* 2014 ; Zehdi *et al.* 2015). Après, en fonction du pouvoir discriminant des marqueurs SSR individuels et de la liaison d'équilibre entre les marqueurs, nous avons sélectionné la meilleure combinaison de loci SSR pour l'identification des variétés du dattier. Enfin, nous avons montré l'efficacité de la caractérisation moléculaire de la combinaison d'un locus chloroplastique et des loci SSR nucléaires sélectionnés en identifiant 155 génotypes sur une collection de variétés algériennes et étrangères du palmier dattier.

Nos résultats témoignent la pertinence du choix de jeu de sept marqueurs sur la collection de variétés algériennes, par l'établissement des profils de ces variétés.

Grace à nos marqueurs puissants, nous possédons actuellement une base de données de génotypage de référence qui a permis l'identification de 135 variétés algériennes. Ces données générées sont utilisées pour l'analyse génétique du patrimoine phoenicicole algérien. La clé d'identification des variétés de dattier, construite à partir des données de génotypage, est proposée à la fois comme un outil de conservation et de caractérisation de l'agro-biodiversité de dattier en Algérie. Cependant, ce travail sur l'établissement de la clé d'identification est

d'un grand intérêt pour le typage moléculaire et l'identification des variétés, la certification des rejets échangés et la conformité des vitroplants issus par la technique de culture *in vitro*. Il permettrait une meilleure gestion de nos ressources en envisageant une caractérisation précoce des rejets sans attendre leur fructification et généralement destinés à une extension de la palmeraie algérienne.

Chapitre 5

Diversité génétique du palmier dattier dans l'ancien monde : Origine et processus de domestication

Introduction

Un processus essentiel dans l'histoire des sociétés a été le passage d'une économie de prédation (chasse, cueillette, pêche) à une économie fondée sur la domestication consciente de certaines plantes et certains animaux à des fins d'augmentation des productions (Zeder *et al.* 2006). Ce passage dans les modes d'existence des sociétés a conduit à une transformation progressive des espèces et de leur environnement. La domestication des plantes est une sélection de traits qui correspondent aux besoins de l'homme pouvant aboutir à des modifications génétiques. Ce processus a conduit à une série de changements phénotypiques et biologiques. Au cours du processus de domestication des plantes, les actions combinées d'élevage, de sélection, de migration et de mélange ont donné naissance à des populations cultivées génétiquement et phénotypiquement distinctes des pools génétiques ancestraux (Doebley *et al.* 2006).

Des études antérieures concernant la domestication des plantes indiquaient que les cultures proviennent d'un seul événement de domestication (Vavilov 1992) et, en effet, de nombreuses plantes combinent un modèle comme le maïs et le soja (Matsuoka *et al.* 2002 ; Guo *et al.* 2010). Cependant, selon des études récentes, plusieurs plantes cultivées peuvent avoir de nombreuses origines telles que la vigne, l'olivier ou la noix de coco (Arroyo-Garcia *et al.* 2006 ; Haouane *et al.* 2011 ; Gunn *et al.* 2011).

Les premières civilisations agraires reposent principalement sur la culture des céréales (le blé et l'orge), apparue d'abord au Moyen-Orient, et qui ont rapidement diffusé en Asie occidentale, en Afrique et en Europe (Wadley & Martin 1993). Quant à la culture des arbres fruitiers, elle semble avoir commencé trois mille ans après, encore au Moyen-Orient, et les légumes et d'autres cultures ont suivi (Zohary & Hopf 2000).

Le palmier dattier est considéré comme l'un des premiers arbres fruitiers domestiqué dans l'Ancien Monde (Zohary *et al.* 2012). Il est d'une importance majeure pour les populations passées et présentes vivant dans les zones arides car il s'agit d'une espèce clé de l'agrosystème oasien. En tant que l'un des plus anciens arbres fruitiers cultivés, le palmier dattier est lié à l'histoire de la migration humaine. Sur la base du dossier archéologique, on croit souvent que sa domestication, ou la culture la plus ancienne, s'est produite dans le croissant fertile (Zohary & Spiegel-Roy 1975). Le palmier dattier est néanmoins issu d'un seul événement de domestication qui a eu lieu dans le golfe Persique, qui est censé représenter la zone de distribution naturelle selon les restes archéologiques (Zohary *et al.* 2012). Les premières

preuves archéologiques indiquent sa culture dans cette région au cours du 3^e millénaire av. J.-C. (Tengberg 2012). En Afrique du Nord, les restes archéologiques sont rares (sauf en Égypte) dus à un manque de recherche archéologique intensive (Tengberg 2012). La tradition orale stipule que la culture du palmier dattier s'est répandue en Afrique du Nord en association avec la propagation de l'islam, mais les restes archéologiques démontrent qu'il est beaucoup plus vieux au moins en Égypte et en Libye (Tengberg 2012 ; Tengberg *et al.* 2014).

En outre, les études génétiques du palmier dattier ont été principalement limitées à un pays ou à une petite région, y compris un nombre relativement restreint de variétés, et ne sont donc pas représentatives de l'agro- biodiversité globale. Des résultats plus récents semblent montrer l'existence d'une forte différenciation génétique entre les cultivars de l'Est et de l'Ouest sur la base des marqueurs nucléaires (Zehdi *et al.* 2015) et cytoplasmiques (Pintaud *et al.* 2013). Ainsi, de nombreuses hypothèses pour l'origine du palmier dattier et sa domestication ont été présentées (Gros-Balthazard *et al.* 2013 ; Zehdi *et al.* 2015 ; Mathew *et al.* 2015).

Des études génétiques utilisant des marqueurs SSR ont d'abord été effectuées pour l'analyse de la diversité génétique du palmier dattier en Tunisie (Zehdi *et al.* 2004, 2012), puis au Soudan (Elshibli & Korpelainen 2008), Oman (Al-Ruqaishi *et al.* 2008), Qatar (Ahmed & Al-Qaradawi 2009 ; Elmeer *et al.* 2011), Irak (Khierallah *et al.* 2011), Iran (Arabnezhad *et al.* 2012), Émirats Arabes Unis (Chaluvadi *et al.* 2014), et récemment en Algérie (Moussouni *et al.* 2017). Cependant, toutes ces études étaient basées sur un nombre relativement restreint d'accessions centrées sur les pays et, par conséquent, ne sont pas représentatives de la diversité génétique globale du palmier dattier.

Rhouma-Chatti *et al.* (2014) ont utilisé des variétés du dattier tunisiennes et ont inclus certaines issues du Golfe Persique. Ces auteurs n'ont pas détecté un signal fort pour la séparation génétique de leurs échantillons en fonction de leur origine géographique. Une enquête génétique a été menée sur un vaste ensemble de cultivars de palmier dattier (Chaluvadi *et al.* 2014); cependant, la plupart des variétés proviennent de la région du Golfe Persique. En revanche, les enquêtes récentes qui ont inclus des palmiers dattiers de plusieurs régions différentes ont montré des données soutenant deux antécédents génétiques distincts de l'Afrique du Nord (Ouest) et du Golfe Persique (Est) (Zehdi *et al.* 2015 ; Mathew *et al.* 2015). En outre, les études du chromosome XY ont identifié une séparation similaire avec deux allèles distincts de Y séparant les cultivars de l'Est et de l'Ouest (Cherif *et al.* 2013). Certaines études utilisant des méthodes telles que la métabolomique ou la morphométrie n'ont pas

montré de séparation claire entre les deux groupes géographiques (Khierallah *et al.* 2011 ; Terral *et al.* 2012 ; Farag *et al.* 2014).

Ainsi, un réexamen de l'origine du palmier dattier semble être nécessaire en fonction d'un échantillon provenant d'une plus vaste zone de distribution. Cette analyse plus large de la diversité est nécessaire pour démêler les relations génétiques entre les groupes géographiques du dattier. En outre, la localisation des centres de domestication des palmiers reste encore incertaine.

Afin d'étudier la structure et la diversité génétique du palmier dattier, nous avons rassemblé un plus grand échantillonnage sur une vaste zone de distribution géographique (une plus grande distribution non étudiée). Dans le but de déchiffrer l'origine géographique et la domestication du dattier, nous avons combiné les marqueurs nucléaires et chloroplastiques afin de tester l'hypothèse d'une structure géographique possible dans la diversité génétique. Les données nucléaires, en particulier les microsatellites, ont prouvé leur efficacité pour tracer les origines des plantes (Matsuoka *et al.* 2002). Nous avons analysé un minisatellite chloroplastique (SSRCp), qui est un marqueur haploïde avec un taux de mutation plus faible que les SSR nucléaires (Provan *et al.* 1999) et évolue ainsi différemment. En effet, étant hérité et transmis par la mère, le génome des chloroplastes nous fournit des informations sur l'origine maternelle de la plante. La combinaison des données chloroplastes et nucléaires est donc essentielle pour avoir une image complète d'une possible différenciation géographique de l'agro-biodiversité du palmier dattier. Ainsi, la diversité génétique des palmeraies algériennes a également été comparée à la diversité génétique mondiale afin d'identifier le pool génétique auquel elles appartiennent et en apprendre davantage sur leur origine.

L'objectif de ce chapitre est donc d'analyser la diversité génétique du dattier présente au niveau de 13 pays, en utilisant, d'une part, des marqueurs nucléaires et, d'autre part, un marqueur chloroplastique.

Résultats

1. Polymorphisme des marqueurs microsatellites

Au total, 240 allèles ont été identifiés. Le nombre d'allèles par locus varie de 2 (PdCUC3-ssr1) à 35 (PdAG1-ssr), avec une moyenne de 13.33. Le nombre de génotypes par locus variait de 2 (PdCUC3-ssr1) à 106 (PdAG1-ssr), avec une moyenne de 40.28 génotypes par locus. Les valeurs de PIC varient de 0.01 (PdCUC3-ssr1) à 0.89 (PdCUC3-ssr2), avec une moyenne de 0.60. Le locus le plus informatif était PdCUC3-ssr2 (0.89). La fréquence de l'allèle le plus commun à chaque locus varie de 18% (PdAG1-ssr) à 99% (PdCUC3-ssr1). En moyenne, 48% des 262 accessions partagent un allèle majeur commun à chaque locus considéré (Tableau 18). L'hétérozygotie attendue varie de 0.01 (PdCUC3-ssr1) à 0.82 (mPdCIR010), avec une moyenne de 0.58 par locus, tandis que l'hétérozygotie observée varie de 0.01 (PdCUC3-ssr1) à 0.90 pour (PdCUC3-ssr2), avec une moyenne de 0.63 par locus, ce qui suggère que le germoplasme de palmier dattier examiné a des niveaux élevés de polymorphisme. Le coefficient de consanguinité (F_{IS}) varie de -0.086 (mPdCIR025) à 0.221 (mPdCIR035), avec une moyenne de 0.015 par locus. Un des 18 loci SSR (mPdCIR035) a affiché un déficit significatif d'hétérozygotie ($P < 0.01$: valeur F_{IS} significative). Un des 18 loci SSR (mPdCIR035) a affiché un déficit significatif d'hétérozygotie ($P < 0.01$: valeur F_{IS} significative).

Tableau 18 : Diversité génétique des 481 accessions de dattier au niveau de 18 loci SSR.

Marqueur	Nombre de génotype total	Nombre de génotype identifié	Fréquence de l'allèle le plus commun	Nombre de génotype identifié par locus	Nombre d'allèles par locus	Hétérozygotie observée (Ho)	Hétérozygotie attendue (He)	PIC	F _{IS}
mPdIRD031	481.00	481.00	0.86	7.00	4.00	0.25	0.26	0.24	-0.041
mPdIRD033	481.00	481.00	0.88	8.00	4.00	0.22	0.20	0.21	0.017
mPdIRD040	481.00	481.00	0.65	13.00	5.00	0.51	0.51	0.45	-0.020
PdCUC3-ssr2	481.00	481.00	0.17	92.00	29.00	0.90	0.79	0.89	0.124
PdAP3-ssr-F4	481.00	481.00	0.28	49.00	13.00	0.84	0.80	0.82	-0.055
mPdCIR078	481.00	481.00	0.21	100.00	27.00	0.89	0.81	0.88	0.038
mPdCIR015	481.00	481.00	0.37	36.00	12.00	0.79	0.73	0.76	0.020
mPdCIR016	481.00	481.00	0.45	11.00	5.00	0.67	0.58	0.61	0.061
mPdCIR032	481.00	481.00	0.36	42.00	14.00	0.79	0.75	0.77	0.048
mPdCIR035	481.00	471.00	0.59	28.00	11.00	0.60	0.44	0.56	0.221*
mPdCIR057	481.00	481.00	0.57	24.00	10.00	0.63	0.59	0.60	-0.051
mPdCIR085	481.00	481.00	0.25	69.00	20.00	0.85	0.80	0.84	-0.006
PdAG1-ssr	481.00	481.00	0.18	106.00	35.00	0.89	0.81	0.88	0.026
mPdCIR010	481.00	481.00	0.29	52.00	15.00	0.83	0.82	0.81	0.009
mPdCIR025	481.00	481.00	0.29	50.00	16.00	0.80	0.80	0.77	-0.086
mPdCIR063	481.00	481.00	0.32	33.00	15.00	0.76	0.69	0.73	0.001
PdCUC3-ssr1	481.00	481.00	0.99	2.00	2.00	0.01	0.01	0.01	-0.019
mPdIRD013	481.00	481.00	0.98	3.00	3.00	0.03	0.03	0.03	-0.012
Moyenne	481.00	480.44	0.48	40.28	13.33	0.63	0.58	0.60	0.015

PIC : Polymorphism Information Content. test significatif à * $p < 0,01$.

2. Analyse Bayésienne de regroupement

L'analyse effectuée par l'approche bayésienne de regroupement, implémentée dans le programme STRUCTURE (Pritchard *et al.* 2000), a été utilisée pour étudier la structure génétique des génotypes de palmier dattier pour un nombre K de groupes entre $K = 2$ à $K = 6$, en effectuant 10 simulations par valeur de K .

La procédure d'Evanno *et al.* (2005), qui permet d'évaluer le nombre probable K de clusters en se basant sur le ΔK , montre que la structure génétique la plus représentative est obtenue pour $K = 2$ ($\Delta K = 1635.28$) (Figure 22).

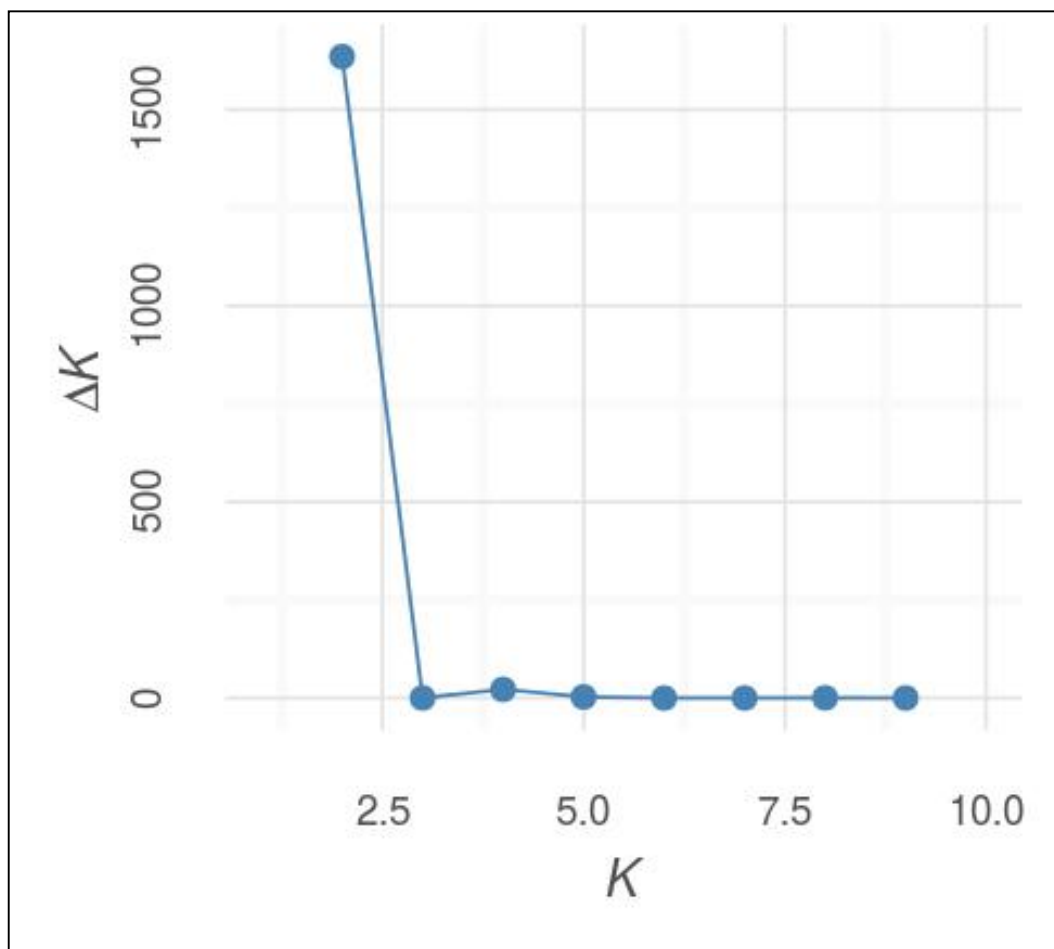


Figure 22 : Résultat de l'approche bayésienne de regroupement avec l'identification de la valeur de K la plus vraisemblable suivant la méthode du Delta K en fonction du nombre possible de sous-populations.

$K = 2$, les accessions de *Phoenix dactylifera* ont été différenciées en deux groupes géographiques. Le premier groupe nommé 'groupe Occidental' est composé d'adhésions en provenance d'Afrique du Nord à l'exception de ceux de Djibouti. Ce groupe est constitué de 323 palmiers dattiers de l'Algérie, du Maroc, de la Tunisie, de la Mauritanie, du Niger, de la Libye, de l'Égypte et du Soudan. Le deuxième groupe appelé 'Oriental' est composé d'accessions originaires du bassin oriental d'Asie et de celles-ci de Djibouti. Il est constitué de 158 palmiers dattiers du Golfe Persique (Irak, Oman, Émirats Arabes Unis), du Pakistan et de Djibouti (Figure 23).

$K = 3$, les accessions de l'Algérie et de la Mauritanie ont été séparées de ceux du pool occidental. $K = 4$, un quatrième groupe a été identifié, comprenant les adhésions du Niger, l'Égypte et la Libye. À $K = 5$, les accessions de Djibouti ont été séparées de ceux recueillis en Asie et un mélange a été observé dans le groupe pakistanais. À $K = 6$, la structure génétique du palmier dattier dans les deux pools de gènes principaux n'a pas été modifiée et aucun groupe supplémentaire significatif n'a été observé (Figure 23).

Il est important de noter que les groupes Omanais et Émirats Arabes Unis ont montré un signal de mélange minimal et ont été clairement maintenus pour toutes les valeurs de K . Cela nous a permis de considérer que les groupes omanais et émirats arabes unis sont les meilleurs représentants du pool oriental. Ainsi, le groupe tunisien montre un signal de mélange minimal dans le pool occidental. Ce groupe tunisien est le plus caractéristique de la diversité des groupes occidentaux.

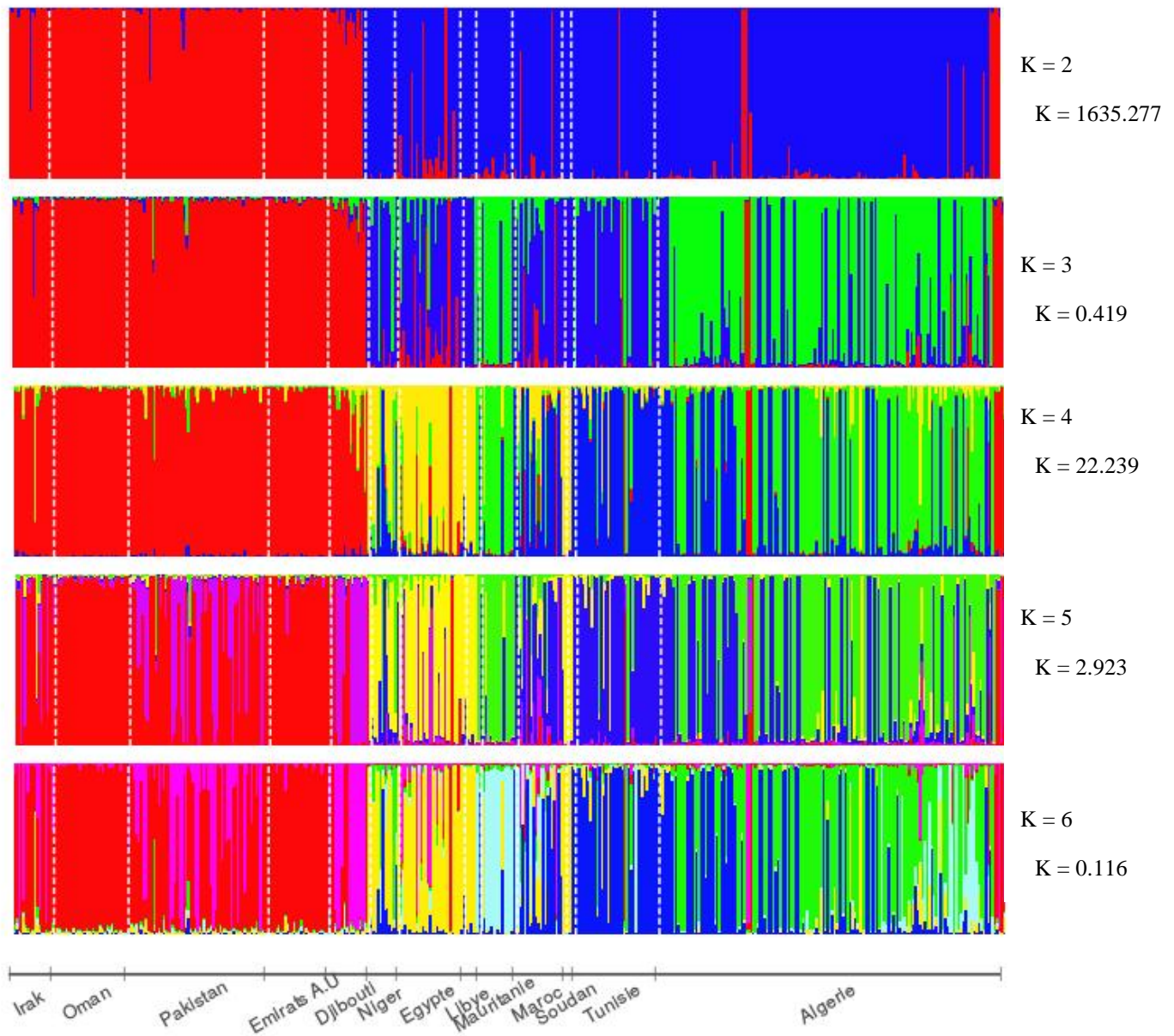


Figure 23 : Résultat de l'approche bayésienne de regroupement avec l'assignement de 481 individus suivant leur localisation géographique pour un nombre de sous-population de $K = 2$ à $K = 6$. Chaque couleur représente une sous-population génétique, les lignes verticales représentent les individus. Un individu défini par plusieurs couleurs, est considéré comme ayant des ancêtres issus de plusieurs sous-populations génétiques. Les groupes de palmier dattier sont séparés par une ligne pointillée blanche. K mesure l'indice de stabilité d'Evanno *et al.* (2005).

3. Diversité et structure génétique entre les groupes géographiques du palmier dattier

Les indices de diversité génétique ont été calculés au sein des treize groupes géographiques, et pour les deux groupes mesurés par STRUCTURE et déduits pour $K = 2$ (groupe Oriental et Occidental).

Dans le groupe occidental, l'Égypte a affiché les valeurs d'hétérozygotie attendues les plus élevées avec 0.611, tandis que le Soudan avait la valeur la plus basse avec 0.457 (Tableau 19). L'hétérozygotie observée était la plus élevée pour l'Algérie (0.606) et la plus faible pour la Mauritanie (0.460). Comme le nombre d'allèles observés dans un groupe dépend fortement de la taille de l'échantillon, la richesse allélique a été calculée pour chaque pays et groupe. La richesse allélique la plus élevée a été détectée pour le groupe égyptien, avec 3.620, tandis que la valeur la plus basse a été notée en Mauritanie (2.797). Les valeurs de la richesse allélique étaient équivalentes entre les deux groupes oriental et occidental (Tableau 19). La même chose est observée avec l'hétérozygotie attendue et observée ; nous n'avons trouvé aucune différence entre les deux groupes de palmier dattier.

Tableau 19 : La diversité génétique du palmier dattier dans 13 pays pour les 18 SSR microsatellites.

Population	Groupe occidental									Groupe oriental					
	Algérie	Maroc	Tunisie	Niger	Égypte	Libye	Soudan	Mauritanie	Total	Djibouti	Pakistan	Irak	Oman	Émirats Arabes Unis	Total
N°	188	22	37	20	29	7	4	16	323	18	62	18	33	27	158
Na	7.889	6.000	5.944	5.444	6.944	3.667	3.056	4.167	5.389	5.833	8.222	5.167	5.500	5.778	6.100
Ho	0.606	0.597	0.577	0.525	0.569	0.595	0.528	0.465	0.558	0.559	0.582	0.590	0.559	0.589	0.576
He	0.592	0.595	0.567	0.574	0.611	0.556	0.457	0.475	0.553	0.564	0.595	0.577	0.536	0.539	0.562
Ra	3.452	3.517	3.233	3.289	3.620	3.144	3.056	2.797	3.264	3.325	3.483	3.261	3.116	3.236	3.284
F_{IS}	-0.006	0.004	-0.003	0.110	0.085	0.006	-0.013	0.052	0.029	0.038	0.030	0.01	-0.03	-0.073	-0.005
P value F_{IS}	0.26	0.583	0.469	1	1	0.608	0.537	0.916	0.672	0.882	0.959	0.607	0.162	0.001	0.522

N° : Nombre de génotype ; Na : Nombre d'allèles par locus ; Ho : Hétérozygotie observée ; He : Hétérozygotie attendue ; Ra : Richesse allélique ; F_{IS} : Indice de fixation.

Les relations génétiques entre les groupes géographiques de palmier dattier prédéfinis (13 pays) ont également été évaluées en fonction des distances génétiques 'D_{AS}' (distance des allèles partagées ou shared allele distance) et de l'algorithme de Neighbor Joining (Figure 24). Chaque diagramme circulaire représenté avec un pays sur l'arbre phylogénétique représente la probabilité que sa population appartienne à un des deux groupes trouvés avec le logiciel STRUCTURE à K = 2, avec le rouge indiquant le groupe Oriental et le bleu le groupe Occidental.

Cette analyse génétique a classé les treize groupes géographiques de palmier dattier en deux groupes: le groupe I (dont la Mauritanie, la Libye, l'Égypte, le Soudan, le Niger, le Maroc, l'Algérie et la Tunisie) correspondant au cluster Occidental de STRUCTURE, qui a été clairement distingué du groupe II (le Djibouti, l'Oman, l'Irak, les Émirats Arabes Unis et le Pakistan) correspondant au cluster Oriental de STRUCTURE. Cet arbre phylogénétique a différencié deux groupes de palmier dattier grâce à un support de bootstrap élevé de 80%.

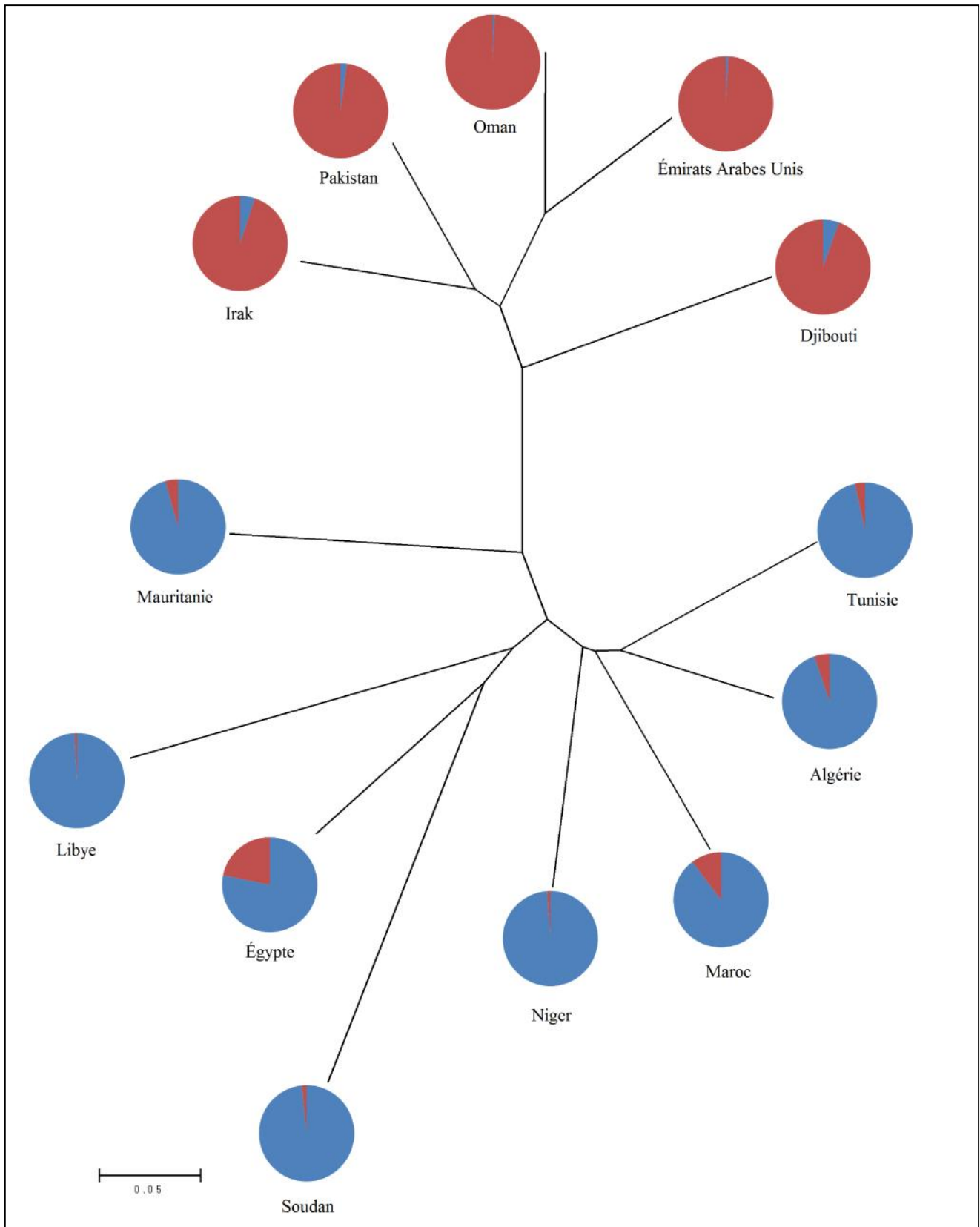


Figure 24 : Arbre Neighbor Joining des groupes géographiques basé sur les valeurs de distance génétique D_{AS} , ainsi que la répartition des deux clusters génétiques définis par l'analyse STRUCTURE à $K = 2$ au sein de chaque pays. Le cluster occidental en bleu et le cluster oriental en rouge.

La DAPC a été effectuée sur des fréquences alléliques, et les accessions de *Phoenix dactylifera* ont été projetées selon deux axes avec une proportion de variance des données.

La DAPC sur l'ensemble des données génétiques des 13 pays, montre que le premier axe a différencié les groupes géographiques Orientaux (Irak, Pakistan, Émirats Arabes Unis et Oman) avec Djibouti, des Occidentaux (Mauritanie, Maroc, Tunisie, Niger, Libye, Soudan, Algérie et Égypte) (Figure 25) et le deuxième axe a différencié le Soudan et la Libye des autres groupes Occidentaux. La distribution de groupe observée reflète leur situation géographique. Sur la région de l'Est, Oman et Émirats Arabes Unis apparaissent génétiquement plus proches.

Une deuxième DAPC est effectuée uniquement sur les données génétiques des pays constituant le groupe Occidental avec plus de détail sur l'Algérie (Figure 26). Cette DAPC concerne les accessions des oasis algériennes (Biskra, Touggourt, Oued Souf, Ouargla, Ghardaïa, El Méria, Timimoun, Adrar, Béni Abbès et Tamanrasset) et celles en provenance des pays de la zone géographique d'ouest (Mauritanie, Maroc, Tunisie, Niger, Libye, Soudan et Égypte). On observe sur cette DAPC que la diversité génétique et la distribution des oasis algériennes s'exprime par un continuum qui suit un axe géographique Est-Ouest. Les oasis de l'est de l'Algérie (Biskra, Touggourt, Oued Souf, Ouargla) s'accrochent avec la Tunisie, les oasis de l'ouest de l'Algérie (El Méria, Timimoun, Adrar et Béni Abbès) s'accrochent avec le Maroc et la Mauritanie, et Tamanrasset s'accroche avec le Niger. Les deux axes de l'DAPC ont différencié la Libye, l'Égypte et le Soudan des autres régions du groupe Occidental (Figure 26).

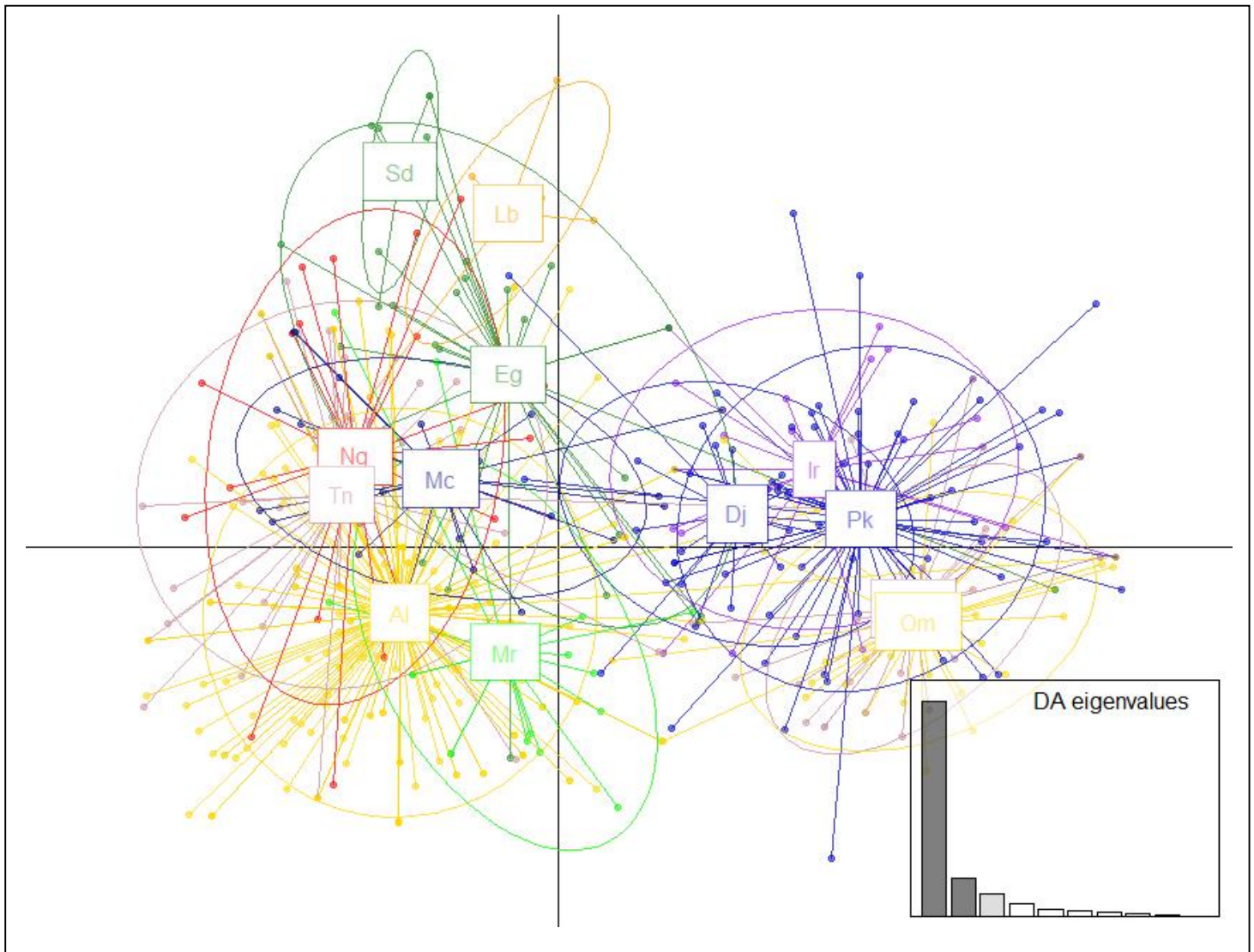


Figure 25 : Analyse discriminante des composantes principales (DAPC) sur les données SSR des génotypes mondiaux.

Al : Algérie, Mc : Maroc, Tn : Tunisie, Ng : Niger, Eg : Égypte, Dj : Djibouti, Lb : Libye, Sd : Soudan, Pk : Pakistan, Ir : Irak, Om : Oman, Em : Émirats Arabes Unis, Mr : Mauritanie.

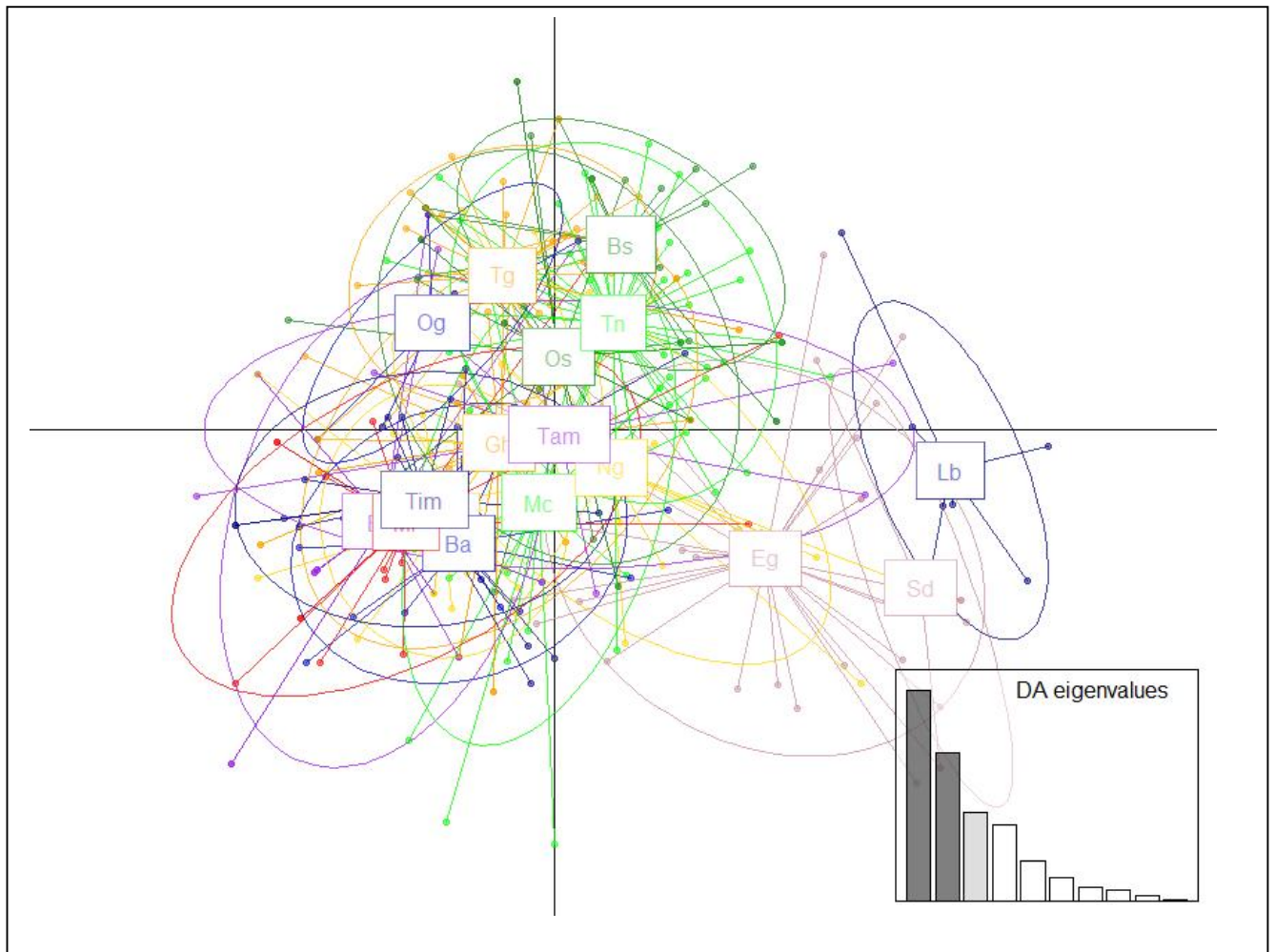


Figure 26 : Analyse discriminante des composantes principales (DAPC) sur les données SSR des géotypes occidentaux.

Bs : Biskra, Tg : Touggourt, Os : Oued Souf, Og : Ouargla, Gh : Ghardaïa, Tam : Tamanrasset, El : El Ménia, Tim : Timimoun , Ad : Adrar, Ba : Béni Abbès, Mc : Maroc, Tn : Tunisie, Ng : Niger, Eg : Égypte, Dj : Djibouti, Lb : Libye, Sd : Soudan, Pk : Pakistan, Ir : Irak, Om : Oman, Em : Émirats Arabes Unis, Mr : Mauritanie.

4. Diversité génétique du minisatellite chloroplastique

Seuls deux allèles (3 et 4 répétitions du minisatellite de 12 pb) ont été observés dans l'ensemble du palmier dattier, définissant deux chlorotypes. Leur répartition a montré une structure géographique forte (Figure 27).

Le chlorotype dit Oriental (254 pb) était présent à une fréquence d'allèle très élevée de 95.6% dans le palmier dattier de l'Est (Figure 27). Au total, 151 individus sur 158 ont montré ce chlorotype et seulement 7 accessions avaient le chlorotype Occidental.

Les palmiers dattiers africains présentaient les deux chlorotypes; cependant, l'haplotype de 242 pb (chlorotype Occidental) est légèrement plus abondant, avec une fréquence d'allèle de 51.4% (166 échantillons sur 323). Il est important de noter que l'Algérie présente un chlorotype Oriental majoritaire (62.77%) (Figure 27).

Les fréquences de chlorotypes étaient significativement différentes entre les deux régions géographiques de palmier dattier (test G, p values < 0.05). Entre autres, nous avons observé un gradient de leurs fréquences d'Ouest en Est. La fréquence du chlorotype Occidental diminue lorsqu'on se dirige vers l'Est, tandis que la fréquence du chlorotype Oriental augmente jusqu'à 100% en Oman.

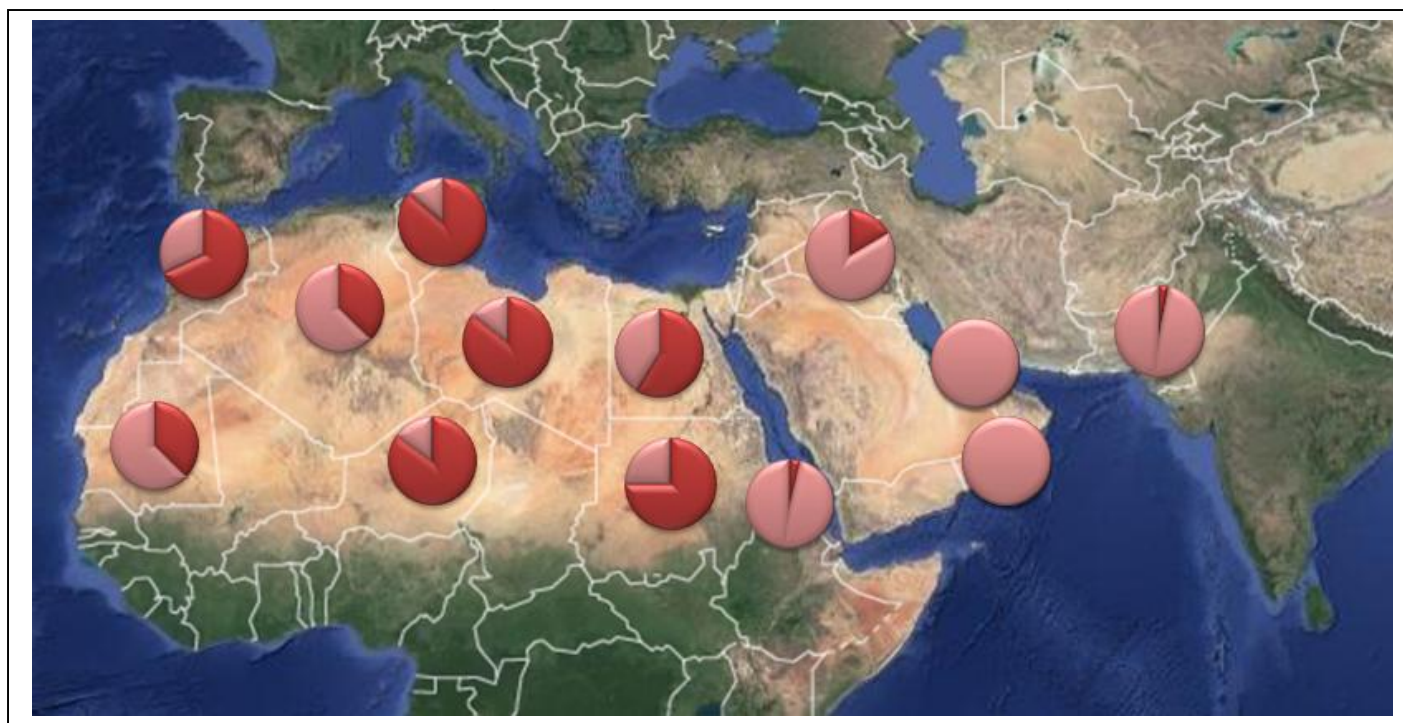


Figure 27 : Distribution des deux chlorotypes du palmier dattier au sein des 481 accessions selon leur origine géographique.

Rouge : 3 répétitions du minisatellite de 12 pb

Rose : 4 répétitions du minisatellite de 12 pb.

Discussion

a. Structure et diversité génétique de l'ADN nucléaire et chloroplastique du palmier dattier

Cette étude nous a permis de valider le scénario de la structure génétique et l'origine du palmier dattier, mis en place avec des ressources génétiques provenant de treize pays, en utilisant à la fois des marqueurs SSR nucléaires et chloroplastiques.

Les microsatellites sont des marqueurs codominants et très polymorphes (Litt & Luty 1989). En raison de leur niveau élevé de polymorphisme, ils constituent un outil très puissant dans la détermination de la parenté (Blouin *et al.* 1996).

Au total, 481 accessions ont été analysées à l'aide de 18 loci SSR nucléaires qui ont révélé 240 allèles avec un nombre moyen de 13 allèles par locus. Ce résultat met en évidence un degré élevé de diversité génétique dans l'échantillonnage de palmier dattier étudié et montre un allèle moyen par locus significativement plus élevé que celui obtenu dans plusieurs études centrées sur un seul pays, comme l'Irak (8.54) (Khierallah *et al.* 2011), le Qatar (7.7) (Elmeer *et al.* 2011), la Tunisie (6.94) (Zehdi *et al.* 2016) et l'Algérie (7.9) (Moussouni *et al.* 2017).

Les analyses de la structure génétique de la population de palmiers dattiers, en utilisant l'approche de regroupement bayésien (Figure 23), et les analyses phylogénétiques basées sur le NJ et la DACP (Figure 24 et 25), sont clairement compatibles et ont révélé une forte différenciation géographique entre les palmiers dattiers du Moyen-Orient (Est) et d'Afrique (Ouest). Les palmiers dattiers en provenance du Moyen-Orient, du Pakistan et du Djibouti affichent principalement le génome nucléaire dit oriental, tandis que les palmiers dattiers d'Afrique affichent principalement le génome nucléaire appelé Occidental.

Nos résultats démontrent une structuration géographique en deux groupes. Le premier, appelé le groupe de l'Est, se compose des adhésions de Djibouti, d'Oman, d'Irak, des Émirats Arabes Unis et du Pakistan ; et le second, appelé le groupe de l'Ouest, se compose des adhésions restantes d'Afrique, dont la Mauritanie, le Soudan, le Maroc, l'Algérie, la Tunisie, la Libye, le Niger et l'Égypte.

Plusieurs analyses génétiques effectuées sur le palmier dattier ont précédemment suggéré l'existence de pools de gènes orientaux et occidentaux (Pintaud *et al.* 2013). Elshibli & Korpelainen (2008) ont comparé la diversité génétique du palmier dattier du Soudan et du Maroc et ont montré une différenciation significative entre eux. Arabnezhad *et al.* (2012) ont obtenu le même résultat en utilisant une nouvelle série de marqueurs SSR dans une autre collection de palmier dattier d'Iran, d'Irak et d'Afrique du Nord. Al-Ruqaishi *et al.* (2008) ont

également utilisé des marqueurs microsatellites pour analyser la diversité génétique parmi les génotypes de palmier dattier d'Oman, de Bahreïn et d'Irak, et un génotype marocain 'Medjool' qui est apparu très distinct des accessions du Moyen-Orient.

Nos résultats sont cohérents avec ceux de Cherif *et al.* (2013), qui ont identifié deux lignées paternelles ancestrales masculines distinctes de type Y, et avec ceux de Zehdi *et al.* (2015) sur une collection du palmier dattier couvrant dix pays. Le travail de Mathew *et al.* (2015) sur les données génotypiques des SNP nucléaire du palmier dattier, a soutenu fermement le modèle de deux régions indépendantes distinctes de la culture ancienne du palmier dattier centrées en Afrique du Nord et dans le Golfe persique.

Des résultats similaires ont été rapportés dans les analyses de structuration génétique d'autres cultures méditerranéennes. En effet, de nombreuses études génétiques sur les oliviers ont révélé une différenciation génétique entre les régions méditerranéennes Occidentale et Orientale (Breton *et al.* 2009 ; Haouane *et al.* 2011 ; Besnard *et al.* 2013) et ont montré la structuration du germoplasme d'olivier méditerranéen en trois principaux pools de gènes, correspondant fortement à trois zones géographiques distinctes, à savoir les régions de la Méditerranée de l'Ouest, du Centre et de l'Est.

Le génome chloroplastique se caractérise par un taux de mutation très faible permettant de détecter des événements anciens, par l'absence de recombinaison et par la transmission uniparentale maternelle. Ces propriétés rendent ce polymorphisme pertinent pour mieux retracer les voies de diffusion maternelle, et pourrait apporter également un éclairage sur les phénomènes d'introgession. L'ensemble de ces caractéristiques est complémentaire au polymorphisme nucléaire pour une étude de structuration génétique pertinente.

Dans notre étude, le polymorphisme du marqueur chloroplastique était aussi important que celui trouvé dans l'ADN nucléaire car deux allèles ont été détectés dans le minisatellite, comme indiqué précédemment (Pintaud *et al.* 2013 ; Zehdi *et al.* 2015 ; Moussouni *et al.* 2017). Nous avons observé un gradient géographique des fréquences des deux chlorotypes. Étant donné que la distribution naturelle du palmier dattier couvrait l'Afrique et le Moyen-Orient, et que la distribution des chlorotypes avant la domestication est inconnue, deux hypothèses sont possibles: l'haplotype Occidental n'a été trouvé qu'en Afrique et l'Oriental était restreint au Moyen-Orient, ou les deux chlorotypes étaient présents dans l'ensemble de la zone de distribution et la domestication a conduit au profil géographique actuel.

b. Origines géographiques et Histoire génétique du palmier dattier cultivé indiquent deux centres de domestication indépendants

L'hypothèse acceptée à propos de l'origine du palmier dattier est basée sur des restes archéologiques et suggère qu'elle provient d'un seul événement de domestication autour du golfe Persique suivi de la diffusion du palmier dattier cultivé vers l'Égypte et l'Afrique du Nord à l'ouest, et à l'est jusqu'à l'Inde (Munier 1973 ; Zohary *et al.* 2012). Cela implique que les variétés africaines ne contiennent qu'un sous-ensemble de la diversité des variétés du Moyen-Orient, lors de la diffusion à travers l'Afrique. Ce n'est pas ce que nous avons observé: la diversité génétique (richesse allélique et hétérozygotie) n'est pas faible dans les variétés africaines par rapport aux variétés du Moyen-Orient (Tableau 19).

Cependant, dans la présente étude, les marqueurs SSR nucléaires ont montré que les accès de l'Ouest (Occidentaux) différaient fortement de ceux de l'Est (Orientaux). En supposant une seule origine orientale, ce modèle de diversité ne serait observé que dans le cas d'un fort goulet d'étranglement lors de l'expansion vers l'Ouest. Cela aurait entraîné une dérive génétique importante et donc une perte ou une fixation d'allèles dans le groupe occidental qui aurait donc montré un modèle allélique très différent de celui du groupe de l'Est. Cependant, les niveaux de diversité observés dans les groupes de l'Est et de l'Ouest apparaissent semblables, mettant en cause la précision de l'hypothèse d'origine unique, conformément aux données morphométriques précédentes (Terral *et al.* 2012).

Malgré les preuves archéologiques d'un centre de recherche autour du Golfe Persique, certaines preuves récentes ont montré deux sous-populations distinctes du palmier dattier dans le Golfe persique et l'Afrique du nord (Arabnezhari *et al.* 2012 ; Zehdi *et al.* 2015 ; Mathew *et al.* 2015). Nos analyses appuient deux événements de domestication indépendants en Afrique et au Moyen-Orient.

La forte structure géographique entre les groupes Oriental et Occidental trouvée dans notre étude suggère une autre hypothèse, soit l'existence de deux origines de culture, l'une à l'Est et l'autre à l'Ouest. Le golfe Persique n'aurait peut-être pas été le seul centre de domestication pour le palmier dattier. L'Afrique du Nord pourrait également être un centre de domestication, soit indépendant (domestication primaire), soit après la diffusion des variétés d'Est à l'Ouest, suivie d'une introgression avec des variétés locales (domestication secondaire), tels que rapporté dans l'olivier (Besnard *et al.* 2002).

La localisation précise des deux événements de domestication reste à faire l'objet d'une enquête. Au Moyen-Orient, le golfe Persique est l'hypothèse privilégiée basée sur les données

archéologiques (Tengberg 2012). En Afrique, les données archéologiques concernant la culture du palmier dattier sont absentes sauf en Égypte et en Libye (Tengberg 2012, Tengberg *et al.* 2014). L'Égypte a cependant été éliminée comme un centre de domestication pour le palmier dattier puisque les données archéologiques indiquent que ce dernier a probablement été introduit pour la culture (Tengberg *et al.* 2014). On devrait remarquer que de multiples événements de domestication peuvent avoir eu lieu en Afrique et en Asie. En effet, ce sont de vastes zones et les signaux génétiques de multiples événements de domestication peuvent être perdus dans un modèle de domestication prolongé en raison des flux de gènes entre les populations cultivées (Allaby *et al.* 2008).

L'origine de cette structuration en deux centres de domestication indépendants est néanmoins imparfaitement connue en raison de l'information encore limitée sur la répartition et la structure génétique des populations sauvages du palmier dattier dans ces deux centres de diversification (Gros-Balthazard *et al.* 2013).

Bien que le génome nucléaire semble parfaitement différencié entre les palmiers dattiers de l'Est et de l'Ouest, les données sur le génome chloroplastique fournissent des informations supplémentaires sur cette structure. Les deux chlorotypes précédemment identifiés, Oriental et Occidental (Henderson *et al.* 2006 ; Pintaud *et al.* 2013) ont également été récupérés dans la présente étude. Étant donné qu'il est hérité par la mère, les données chloroplastiques donnent une indication sur l'origine maternelle de la plante et donc sur la diffusion des graines et des rejets.

Le groupe de l'Est a montré une prédominance du chlorotype Oriental avec une moyenne de 96,2% et atteignait 100% au Émirats Arabes Unis et Oman. Le groupe Occidental, cependant, a révélé les deux haplotypes mais avec une prédominance du chlorotype Occidental et avec une valeur maximale en Tunisie et au Niger. L'hypothèse la plus parcimonieuse est que chaque groupe a d'abord eu son haplotype avant que les mouvements de populations et les actions anthropiques n'aboutissent à la dissémination du chlorotype oriental dans le groupe Occidental, et vice versa dans une moindre mesure. Le degré élevé de diversité génétique et la coexistence des deux chlorotypes dans le groupe Occidental, d'autre part, indiquent que la diffusion des accessions de l'Est à l'Ouest est plus importante, comme en témoignent les autres cultures fruitières comme le raisin (Arroya-Gracia *et al.* 2006) et l'olivier (Besnard *et al.* 2002).

Alors que l'étude du génome nucléaire utilisant des marqueurs SSR a révélé que cette composante est fortement structurée entre les palmiers dattiers de l'Ouest et de l'Est, les

chlorotypes identifiés précédemment (Occidental et Oriental) n'étaient pas toujours liés à la région, en particulier dans l'Ouest. Par conséquent, certains groupes présentent un génome nucléaire Occidental associé au chlorotype Oriental, ce qui est particulièrement le cas en Algérie et en Mauritanie. Ce modèle, de différentes origines des génomes nucléaires et chloroplastiques dans une seule plante, pourrait être observé si les génotypes femelles ont été initialement importés des régions de l'Est vers l'Ouest et ont ensuite été pollinisés par des mâles locaux (présentant le chlorotype Occidental) au fil des générations, ce qui a entraîné de nouvelles populations avec un génome nucléaire à prédominance Occidental et un chlorotype Oriental. Cela serait conforme au fait que la culture du palmier dattier a toujours été basée sur la sélection des accessions femelles.

CONCLUSION GENERALE

Le palmier dattier présente une importance économique, écologique et socioculturelle dans les pays d'Afrique du Nord et du Moyen-Orient, et il a beaucoup attiré l'attention des chercheurs de ces pays dans différents domaines de la connaissance tel que son origine, la diffusion et l'évolution du phénotype du dattier qui demeurent fragmentaires, sa diversité génétique à l'échelle de chaque pays.

L'objectif de notre travail était donc d'améliorer les connaissances sur la diversité génétique, l'origine et l'histoire évolutive de l'espèce *Phoenix dactylifera* L. Pour ce faire, nous avons étudié les 14 espèces du genre *Phoenix* par le biais d'analyses génétiques. Aussi, nous avons caractérisé les variétés du dattier (clé d'identification) et étudié la diversité génétique de *Phoenix dactylifera* en Algérie et dans le vieux monde. Les analyses génétiques ont mis en jeu des séquences chloroplastiques ainsi que des microsatellites nucléaires et minisatellites chloroplastiques. Elles ont consisté en des analyses phylogénétiques, de polymorphisme, de génétique des populations et d'approche Bayésienne.

Les relations phylogénétiques du genre *Phoenix* ont été inférées par les microsatellites nucléaires et les minisatellites chloroplastiques. La place du dattier au sein de genre *Phoenix* a été précisée par cette analyse, et ses parents proches ont été identifiés (*P. sylvestris* et *P. atlantica*). En 1973, Munier signale que le palmier dattier est le résultat de l'hybridation de plusieurs *Phoenix* existant encore dans le voisinage de son aire de dispersion, en particulier le *Phoenix sylvestris* et *P. atlantica*. Nous avons pu confirmer que ces espèces ne sont pas des ancêtres sauvages du palmier dattier. Selon nos résultats, le *Phoenix dactylifera* cultivé possède un nombre important d'allèles privés, suggérant qu'il a été domestiqué à partir de populations sauvages de la même espèce.

Le génome chloroplastique présente de nombreux avantages. Son seul défaut est d'être peu variable chez les palmiers. Le premier marqueur chloroplastique pour lequel des amorces universelles ont été définies, *trnL-F* (Taberlet *et al.* 1991), l'un des marqueurs les plus utilisés chez les plantes s'est avéré très décevant chez les palmiers où son pouvoir de résolution phylogénétique est très faible (Baker *et al.* 2000).

Un minisatellite avec un motif de 12 paires de bases répété de 2 à 6 fois dans le genre *Phoenix* a été identifié dans l'espaceur *psbZ-trnfM*. La variation de ce minisatellite permet d'identifier le chlorotype chez *Phoenix dactylifera* de façon très simple, avec de grandes conséquences sur la compréhension de l'histoire de la domestication de cette espèce. L'identification des

deux chlorotypes chez *Phoenix dactylifera* a été consolidée par le biais de l'analyse de séquences chloroplastiques (4 locus chloroplastiques).

Les informations sur la diversité mondiale du palmier dattier ont révélé une forte différenciation géographique. Les dattiers d'Afrique sont parfaitement différenciés des dattiers du Moyen-Orient. La diversité génétique (richesse allélique et hétérozygotie) chez les dattiers africains, en particulier en Algérie, est similaire à la diversité présente dans le Moyen Orient. Ces résultats indiquent qu'il y a eu au moins deux évènements de domestication naturelle du dattier, et suggèrent deux origines, l'une à l'Est et l'autre à l'Ouest de la distribution du dattier. La signature génétique de deux zones de domestication indépendantes (Méditerranéo-Saharienne et Arabo-Persique) est également visible au niveau du segment chromosomique Y non recombinant (Cherif *et al.* 2013), de même que le génome chloroplastique permet de retracer le parcours des lignées femelles. Cependant au Sahara, le pool génétique nucléaire Oriental, très minoritaire (provenant sans doute de quelques graines de dattes ou rejets importés de la péninsule arabique ou de la Mésopotamie), systématiquement pollinisé avec les génotypes Nord-Africains, est essentiellement détectable par son génome chloroplastique caractéristique.

La structure génétique du palmier dattier dans le vieux monde révèle une forte différenciation entre les populations Orientales et Occidentales. Cette étude basée sur les microsatellites nucléaires a montré que le génome nucléaire apparaît parfaitement différencié entre les palmiers dattiers Orientaux et Occidentaux, et l'Algérie n'est pas particulièrement différente des autres pays Occidentaux. En conséquence, certains groupes de palmiers en Algérie présentent un génome nucléaire Occidental associé à un chlorotype Oriental. Une hypothèse pour expliquer cette tendance particulière, est que les semences et les rejets ont été importés du Moyen-Orient à l'Afrique du Nord, et après des générations de croisement avec les palmiers locaux (génome nucléaire Occidental), les variétés ont acquis un génome nucléaire Occidental tandis que le chlorotype Oriental a été maintenu. Ces variétés portent donc l'haplotype Oriental. La combinaison de données nucléaires et chloroplastiques démontre une forte diffusion de l'Est vers l'Ouest. Une autre explication possible est que les deux haplotypes Oriental et Occidental étaient répandues dans le Sahara Algérien, et que les agriculteurs ont sélectionné les individus hétérozygotes. Cela suggère que la sélection humaine a joué un rôle important dans le maintien de la diversité variétale de palmier dattier en Algérie.

Le palmier dattier est cultivé dans les oasis algériennes dans la plupart des régions au sud des montagnes de l'Atlas saharien. Près de 1 000 cultivars issus de rejets ont été inventoriés et leur distribution montre une répartition très marquée dans les parties Sud-Est, centrale et Sud-Ouest du pays. L'exploration de cette diversité est une condition préalable à l'identification et à la caractérisation du cultivar algérien pour contrôler l'érosion génétique afin d'aider à sa conservation et à sa protection. Six marqueurs microsatellites (SSR) et un minisatellite chloroplastique ont été utilisés pour étudier le polymorphisme et discriminer 135 géotypes algériens du palmier dattier. Sur la base des géotypes multilocus, une clé d'identification variétale a été établie et a permis de différencier sans ambiguïté les variétés avec un pouvoir discriminant de 100%. Nos résultats obtenus ont été discutés en termes d'établissement et de gestion d'une collection algérienne de variétés du palmier dattier, de contrôles de conformité, d'identification d'homonymes et de synonymes, et de criblage des ressources locales.

L'ensemble de nos travaux, incluant un large échantillonnage de *Phoenix dactylifera*, et combinant les données génétiques chloroplastiques et nucléaires, a permis de faire des avancées majeures sur l'histoire de la diversité du dattier et sa diffusion par l'Homme.

En perspective, nos travaux de thèse ouvrent des voies pour des études plus approfondies sur la distribution et la domestication du palmier dattier en Algérie et dans le vieux monde. Des études génétiques supplémentaires sont nécessaires à l'identification de populations sauvages de palmier dattier au Hoggar, permettant la comparaison entre les compartiments sauvages et cultivés.

REFERENCES BIBLIOGRAPHIQUES

- Aberlenc-Bertossi F., Castillo K., Tranchant-Dubreuil C., Chérif E., Ballardini M., Abdoukader S., Gros-Balthazard M., Chabrillange N., Santoni S., Mercuri A., Pintaud J.C. (2014). In silico mining of microsatellites in coding sequences of the date palm (*Arecaceae*) genome, characterization, and transferability. *Applications in Plant Sciences* 2: 1300058.
- Achtak H., Oukabli A., Ater M., Santoni S., Khadari B. (2009). Microsatellite Markers as Reliable Tools for Fig Cultivar Identification. *J. Amer. Soc. Hort. Sci.* 134(6): 624–631.
- Adam H., Marguerettaz M., Qadri R., Adroher B., Richaud F., Collin M., Thuillet A.C., Vigouroux Y., Laufs P., Tregear J.W., Jouannic S. (2010). Divergent expression patterns of miR164 and CUP-SHAPED COTYLEDON genes in palms and other Monocots: Implication for the evolution of meristem function in Angiosperms. *Molecular Biology and Evolution* 28:1439-1454.
- Ahmed T.A. & Al-Qaradawi A.Y. (2009). Molecular phylogeny of Qatari date palm genotypes using Simple Sequence Repeats markers. *Biotechnology* 8: 126-131.
- Akkak A., Scariot V., Torello Marinoni D., Boccacci P., Beltramo C., Botta R. (2009). Development and evaluation of microsatellite markers in *Phoenix dactylifera* L. and their transferability to other *Phoenix* species. *Biologia Plantarum* 53 (1): 164-166.
- Al-Faifi S.A., Migdadi H.M., Algamdi S.S., Altaf Khan M., Ammar M.H., Al-Obeed R.S., Al-Thamra M.I., EL-Harty E.H., Jakse J. (2016). Development, characterization and use of genomic SSR markers for assessment of genetic diversity in some Saudi date palm (*Phoenix dactylifera* L.) cultivars. *Electronic Journal of Biotechnology* 21: 18–25.
- Allaby R.D., Fuller D.Q., Brown T.A. (2008). The genetics expectations of a protracted model for the origins of domesticated crops. *Proceedings of the National Academy of Sciences, USA.* 105(37): 13982-13986.
- Al-Mssallem I.S., Hu S., Zhang X., Lin Q., Liu W., Tan J., Yu X., Liu J., Pan L., Zhang T., Yin Y., Xin C., Wu H., Zhang G., Ba Abdullah M.M., Huang D., Fang Y., Alnakhli Y.O., Jia S., Yin A., Alhuzimi E.M., Alsaihati B.A., Al-Owayyed S.A., Zhao D., Zhang S., Al-Otaibi N.A., Sun G., Majrashi M.A., Li F., Tala., Wang J., Yun Q., Alnassar N.A., Wang L., Yang M., Al-Jelaify R.F., Liu K., Gao S., Chen K., Alkhalidi S.R., Liu G., Zhang M., Guo H., Yu J. (2013). Genome sequence of the date palm *Phoenix dactylifera* L. *Nature Communications* 4: 2274.
- Al-Qurainy F., Khan S., Al-Hemaid F.M., Ajmal Ali M., Tarroum M., Ashraf M. (2011). Assessing Molecular Signature for Some Potential Date (*Phoenix dactylifera* L.) Cultivars from Saudi Arabia, Based on Chloroplast DNA Sequences *rpoB* and *psbA-trnH*. *International Journal of Molecular Sciences* ISSN 1422-0067.
- Al-Ruqaishi I.A., Davey M., Alderson P., Mayes S. (2008). Genetic relationships and genotype tracing in date palms (*Phoenix dactylifera* L.) in Oman, based on microsatellite markers. *Plant Genetic Resources* 6: 70-72.
- Angaji A. (2009). Single nucleotide polymorphism genotyping and its application on mapping and marker-assisted plant breeding. *African Journal of Biotechnology* 8(6): 908-914.

- Arabnezhad H., Bahar M., Mohammadi H., Latifia M. (2012). Development, characterization and use of microsatellite markers for germplasm analysis in date palm (*Phoenix dactylifera* L.). *Scientia Horticulturae* 134: 150-156.
- Arroyo-García R., Ruiz-García L., Bolling L., Ocete R., López M.A., Arnold C., Ergul A., Söylemezo lu G., Uzun H.I., Cabello F., Ibáñez J., Aradhya M.K., Atanassov A., Atanassov I., Balint S., Cenis J.L., Costantini L., Goris-Lavets S., Grando M.S., Klein B.Y., McGovern P.E., Merdinoglu D., Pejic I., Pelsy F., Primikirios N., Risovannaya V., Roubelakis-Angelakis K.A., Snoussi H., Sotiri P., Tamhankar S., This P., Troshin L., Malpica J.M., Lefort F., Martinez-Zapater J.M. (2006). Multiple origins of cultivated grapevine (*Vitis vinifera* L. ssp *sativa*) based on chloroplast DNA polymorphisms. *Molecular Ecology* 15(12): 3707-3714.
- Asmussen C. B. & Chase M.W. (2001). Coding and noncoding plastid DNA in palm systematics. *American Journal of Botany* 88: 1103-1117.
- Asmussen C.B., Dransfield J., Deichmann V., Barfod A., Pintaud J.C., Baker W.J. (2006). A new subfamily classification of the palm family (Arecaceae) : evidence from plastid DNA phylogeny. *Botanical Journal of Linnean Society* 151: 15-38.
- Baker W.J., Asmussen C.B., Barrow S., Dransfield J., Hedderson T.A. (2000). A phylogenetic study of the palm family (Palmae) based on chloroplast DNA sequences from the *trnL-trnF* region. *Plant Systematics and Evolution* 219: 111-126.
- Baker W.J., Savolainen V., Asmussen-Lange C.B., Chase M., Dransfield J., Forest F., Harley M., Uhl N., Wilkinson M. (2009). Complete Generic-level Phylogenetic Analyses of Palms (Arecaceae) with Comparisons of Supertree and Supermatrix Approaches. *Systematic Biology* 58: 240–256.
- Ballardini M., Mercuri A., Littardi C., Abbas S., Couderc M., Ludeña B., Pintaud J.C. (2013). The chloroplast DNA locus *psbZ-trnfM* as a potential barcode marker in *Phoenix* L. (Arecaceae). *ZooKeys* 365: 71–82.
- Baraket G., Chatti K., Saddoud O., Ben Abdelkarim A., Mars M., Trifi M., Salhi Hannachi A. (2010). Comparative assessment of SSR and AFLP markers for evaluation of genetic diversity and conservation of Fig, *Ficus carica* L., genetic resources in Tunisia. *Plant Mol Biol Rep.* doi: 10.1007/s11105-010-0217-x.
- Barnaud A., Monique D., Garine E. (2007). Local genetic diversity of sorghum in a village in northern Cameroon: structure and dynamics of landraces. *Theor Appl Genet* 114: 237-248.
- Barrow S. (1998). A monograph of *Phoenix* L. (Palmae: Coryphoideae). *Kew Bulletin* 53: 513-575.
- Baur (1909). Das Wesen und die Erblichkeitsverhältnisse der « Varietates albomarginatae hort von *Pelargonium zonale* ». *Ztschr f ind Abstamm u Vererbungel* 1: 330-351.
- Beccari O. (1890). Revista monografica delle species del genera *Phoenix* L. *Malesia* 3: 345-416.
- Belkhir K., Borsa P., Chikhi L., Raufaste N., Bonhomme F. (2004). GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations. Montpellier, France: Laboratoire Génome, Populations, Interactions, CNRS UMR 5171, Université de Montpellier II.

- Benkhalifa A. (1999). Gestion de la diversité génétique du palmier dattier en Algérie. Atelier 'Construction et organisation d'équipes de recherche scientifique dans les domaines de foresterie et des arbres fruitiers', Marrakech.
- Benkhalifa A., Bennaceur M., Chevallier M.H., Bouguedoura N., Riveill N. (2004). Impacts des pratiques humaines et des biotechnologies sur la gestion et la conservation des ressources génétiques du palmier dattier (*Phoenix dactylifera* L.) en Algérie. 8^{ème} journée scientifique de l'AUF LOME.
- Bennaceur M., Lanaud C., Chevalier M.H., Bounaga N. (1991). Genetic diversity of the date palm (*Phoenix dactylifera* L.) from Algeria revealed by enzyme markers. *Plant Breeding* 107: 56-57.
- Bennaceur M., Lebrun P., Billote N. (2000). Polymorphisme du palmier dattier (*Phoenix dactylifera* L.) à l'aide de marqueurs microsatellites du cocotier (*Cocos nucifera* L.) ou du palmier à huile (*Elaeis guineensis* Jacq.). VII^{es} journées scientifiques: Des modèles biologiques à l'amélioration des plantes. *Ed Hamon S*, pp 599–600.
- Bennaceur M., Chevallier M.H., Bouguedoura N., Benkhalifa A. (2006). Transférabilité des marqueurs microsatellites du cocotier (*Cocos nucifera* L.) et du palmier à huile (*Elaeis guineensis* Jacq.) chez le palmier dattier (*Phoenix dactylifera* L.). Conférence régionale : Mutagenèse induite et biotechnologie d'appui pour la production du palmier dattier contre le bayoud.
- Besnard G., Khadari B., Baradat P., Bervillé A. (2002). *Olea europaea* (Oleaceae) phylogeography based on chloroplast DNA polymorphism. *Theoretical and Applied Genetics* 104: 1353-1361.
- Besnard G., Khadari B., Navascués M., Fernández-Mazuecos M., El Bakkali A., Arrigo N., Baali-Cherif D., Brunini-Bronzini de Caraffa V., Santoni S., Vargas P., Savolainen V. (2013). The complex history of the olive tree: from Late Quaternary diversification of Mediterranean lineages to primary domestication in the northern Levant. *Proceedings of the Royal Society B: Biological Sciences* 280: 20122833.
- Bilote N., Marseillac N., Brottier P., Noyer J.L., Jacquemoud-Collet J.P., Moreau C., Couvreur T., Chevalier M.H., Pintaud J.C., Risterucci A.M. (2004). Nuclear microsatellite markers for the date palm (*Phoenix dactylifera* L.): characterization and utility across the genus *Phoenix* and in other palm genera: *Molecular Ecology Notes* 4: 256-258.
- Birky C.W.J. (1995). Uniparental inheritance of mitochondrial and chloroplast genes: mechanisms and evolution. *Proceedings of the National Academy of Science* 92: 11331-11338.
- Blouin M.S., Parsons M., Lacaille V., Lotz S. (1996). Use of microsatellite loci to classify individuals by relatedness. *Molecular Ecology* 5: 393-401.
- Bodian A., Ndoye Ndir K., Diedhiou P.M., Borgel A., Sane D., Sagna M. (2012). Analysis of genetic diversity of date palm (*Phoenix dactylifera* L.) cultivars from Mauritania using microsatellite markers. *International Journal of Science and Advanced Technology* 2: 36-43.
- Bonavia E. (1885). The date-palm. *The Gardeners' Chronicle* XXIV: 178-211.

- Botstein D., White R.L., Skolnick M., Davis R. (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am J Hum Genet* 32: 314-331.
- Bouchireb N. & Clark M.S. (1997). The application of biotechnology to date palm culture. In: Watanabe KN, Pehu E (eds) Plant biotechnology and plant: genetic resources for sustainability and productivity. *Academic*, San Diego, pp 183–195.
- Bouguedoura N., Bennaceur M., Benkhalifa A. (2008). Le palmier dattier en Algérie Situation, contraintes et apports de la recherche. Actes du 3^e Séminaire du réseau AUF-BIOVEG «Biotechnologie du palmier du palmier dattier » Montpellier, France.
- Bouguedoura N. (2012). Le palmier dattier: développement morphogénétique des appareils végétatif et reproducteur. Editions universitaires européennes. *Book Editions Universitaires et Internationales*.
- Bouguedoura N., Bennaceur M., Babahani S., Benziouche S.E. (2015). Date Palm Status and Perspective in Algeria. 4: 125-168. In Al-Khayri JM, Jain SM, Johnson DV: Date Palm Genetic Resources and Utilization. Volume 1: *Africa and the Americas*.
- Brac de la Perrière R.A. & Benkhalifa A. (1989). L'Identification de cultivars de dattiers (*Phoenix dactylifera* L.) du sud-ouest algérien. *Plant Gen Res Newsl* FAO/IBPGR 78/79: 13-20.
- Breton C., Terral J.F., Pinatel C., Médail F., Bonhomme F., Bervillé A. (2009). The origins of the domestication of the olive tree. *Comptes Rendus Biologies* 332: 1059–1064.
- Burrows P.A., Sauuiov L.A., Svab Z., Maliga P., Nixon P.J. (1998). Identification of a functional respiratory complex in chloroplasts through analysis of tobacco mutants containing disrupted plastid *ndh* genes. *EMBO J.* 17: 868-876.
- Butcher P.A., Decroocq S., Gray Y., Moran G.F. (2000). Development, inheritance and cross-species amplification of microsatellite markers from *Acacia mangium*. *Theoretical and Applied Genetics* 101: 1282-1290.
- Camacho Villa T.C., Maxted N., Scholten M., Ford-Lloyd B. (2006). Defining and identifying crop landraces. *Plant Genetic Resources* 3(3): 373–384.
- Campbell N.A., Mathieu R. (1995). La photosynthèse, dans Biologie. éditeurs : Éditions du Renouveau Pédagogique inc., pp 199-217 (The Benjamin/Cummings *Publishing Company Inc.*, Canada.
- Camps G. (1989). Les chars sahariens. *Antiquités africaines* 25: 11-40.
- Chaluvadi S.R., Khanam S., Aly M.A.M., Bennetzen J.L. (2014). Genetic diversity and population structure of native and introduced Date palm (*Phoenix dactylifera* L.) germplasm in the United Arab Emirates. *Tropical Plant Biology* 7: 30–41.
- Chang C.C., Lin H.C., Lin I.P., Chow T.Y., Chen H.H., Chen W.H., Cheng C.H., Lin C.Y., Liu S.M., Chang C.C., Chaw S.M. (2006). The chloroplast genome of *Phalaenopsis aphrodite* (Orchidaceae): comparative analysis of evolutionary rate with that of grasses and its phylogenetic implications. *Mol Biol Evol* 23: 279–291.
- Chao C.C.T. & Krueger R.R. (2008). *Phoenix dactylifera*, date palm. pp: 138-151 in Janick J, Paull RE (eds.): The encyclopedia of fruit and nuts. CAB International, Cambridge University Press, Cambridge, U.K.

- Cherif E., Zehdi S., Castillo K., Chabrillange N., Abdoukader S., Pintaud J.C., Santoni S., Salhi-Hannachi A., Glémin S., Aberlenc-Bertossi F. (2013). Male-specific DNA markers provide genetic evidence of an XY chromosome system and recombination arrest, and allow to trace paternal lineages in date palm. *New Phytologist* 197: 409-415.
- Chumley T.W., Palmer J.D., Mower J.P. Fourcade H.M., Calie P.J., Boore J.L., Jansen R.K. (2006). The complete chloroplast genome sequence of *Pelargonium x hortorum*: organization and evolution of the largest and most highly rearranged chloroplast genome of land plants. *Mol Biol Evol* 23: 2175–2190.
- Clegg M.T., Learn G.H., Golenberg E.M. (1991). Molecular evolution of chloroplast DNA. In “Evolution at the molecular level” edited by Selander R.K., Clark A.G., Whittman T.S. *Sinauer associates INC*. Publishers, chapter 7: 135-149.
- Clegg M.T. & Zurawski G. (1992). Chloroplast DNA and the study of plant phylogeny: present status and future prospects. *Molecular Systematics of Plants*, 1: 1-13.
- Cleuzio S. & Constantini L. (1982). A l’origine des oasis. *La recherche* 13: 1180-1182.
- Cornicquel B. & Mercier L. (1996). Identification of date palm (*Phoenix dactylifera* L.) cultivars by RFLP: partial characterization of a cDNA probe that contains a sequence encoding for a zinc finger motif. *International Journal of Plant Science* 158: 152-156.
- Correns C. (1909). Vererbungsversuche mit blass (gelb) grünen und buntblättrigen Sippen bei *Mirabilis*. *Urtica und Lunaria* 1: 291–329.
- Corriveau J.L. & Coleman A.W. (1988). Rapid screening method to detect potential biparental inheritance of plastid DNA and results for over 200 angiosperms. *American Journal of Botany* 75: 1443-1458.
- Costantini L. (1985). Considerazioni su alcuni reperti di palma da dattero e sul centro di origine e l’area di coltivazione della *Phoenix dactylifera* L. In Gnoli G. & Lanciotti L. (Ed.), *Orientalia Josephi Tucci Memoriae Dicata*. Serie Orientale 56 (1), Rome, Istituto Italiano per il Medio ed Estremo Oriente: 209-218.
- Costantini L., Audisio L. (2000). Plant and insect remains from the Bronze age site of Ra’s al-Jinz (RJ-2), Sultanate of Oman. *Paleorient* 26(1): 143-156.
- Couvreur T.L.P., Hahn W.J., Granville J.J., Pham J.L., Ludeña B., Pintaud J.C. (2007). Phylogenetic relationships of the cultivated Neotropical palm *Bactris gasiapes* (Arecaceae) with its wild relatives inferred from chloroplast and nuclear DNA polymorphisms. *Systematic Botany* 32: 519-530.
- Couvreur T., Forest F., Baker W.J. (2011). Origin and global diversification patterns of tropical rain forests: inferences from a complete genus-level phylogeny of palms. *BMC Biology* 9: 44.
- de Vienne D. (1998). Biotechnologies végétales. Les marqueurs moléculaires et leurs applications. CNED. Paris. 118pp.
- Delwiche C.F. (1999). Tracing the thread of plastid diversity through the tapestry of life. *Am Nat* 154: 164-177.
- Demesure B., Sodzi N., Petit R.J. (1995). A set of universal primers for amplification of polymorphic non-coding regions of mitochondrial and chloroplast DNA in plants. *Molecular Ecology* 4: 129-131.

- Devanand P.S. & Chao C.T. (2003). Identification of Genetic Strains of 'Medjool' and 'Deglet Noor' Date Palm Cultivars in California Using Amplified Fragment Length Polymorphism (AFLP) Markers. *Plant Genetic Resources. Acta Hort.* 623, ISHS.
- Doebley J.F., Gaut B.S., Smith B.D. (2006). The molecular genetics of crop domestication. *Cell* 127: 1309-1321.
- Downie S.R. & Palmer J.D. (1992). Use of chloroplast DNA rearrangements in reconstructing plant phylogeny. *Molecular Systematics of Plants* 2: 14-35.
- Dransfield J., Uhl N.W., Asmussen C.B., Baker W.J., Harley M.M., Lewis C.E. (2008). Genera Palmarum. The evolution and classification of palms. *Royal Botanic Gardens, Kew, Richmond, U.K.*, 732 p.
- Drummond A., Ashton B., Buxton S., Cheung M., Cooper A., Duran C., Field M., Heled J., Kearse M., Markowitz S. (2010). Geneious v4.7.6 Available: <http://www.geneious.com>.
- El Hadrami A., Daayf F., Elshibli, S., Lain S.M., El Hadrami I. (2011). Somaclonal variation in date palm. p.183-203. In: Jain S.M. *et al.* (eds.), Date palm biotechnology, *Springer B.V.*
- El Houmaizi M.A., Devanand P.S., Fang J., Chao C.C.T. (2006). Confirmation of 'Medjool' date as a landrace variety through genetic analysis of 'Medjool' accessions in Morocco. *J Amer So Hort Sci* 131: 403-407.
- Ellegren H., Moore S., Robinson N., Byrne K., Ward W., Sheldon B.C. (1997). Microsatellite evolution: a reciprocal study of repeat lengths at homologous loci in cattle and sheep. *Molecular Biology and Evolution* 14: 854-860.
- Ellegren H. (2004). Microsatellites: simple sequences with complex evolution. *Nature Reviews Genetics* 5: 435-445.
- Elmeer K., Sarwath H., Malek J., Baum M., Hamwiah A. (2011). New microsatellite markers for assessment of genetic diversity in date palm (*Phoenix dactylifera* L.). *3 Biotech* 1: 91-97.
- Elmeer K. & Mattat I. (2012). Marker-assisted sex differentiation in date palm using simple sequence repeats. *3 Biotech* 2: 241-247.
- Elshibli S. & Korpelainen H. (2008). Microsatellite markers reveal high genetic diversity in date palm (*Phoenix dactylifera* L.) from Soudan. *Genetica* 135: 251-260.
- Elshibli S. & Korpelainen H. (2009). Biodiversity of date palms (*Phoenix dactylifera* L.) in Sudan: chemical, morphological and DNA polymorphisms of selected cultivars. *Plant Genetic Resources.*
- Elshibli S. (2010). Biodiversity in date palm (*Phoenix dactylifera* L.). p. 58-71. In: MARHABA, The Yearbook of the Finnish-Arabic Society. The Finnish-Arabic Society, Helsinki.
- Elshibli S. & Korpelainen E. (2011). Biodiversity in date palm: molecular markers as indicators. p. 371-406. In: Jain S.M. *et al.* (eds.), Date palm biotechnology, *Springer B.V.*
- Estoup A., Jarne P., Cornuet J.M. (2002). Homoplasy and mutation model at microsatellite loci and their consequences for population genetics analysis. *Molecular Ecology* 11: 1591-1604.
- Evanno G., Regnaut S., Goudet J. (2005). Detecting the number of clusters of individuals using the software Structure: A simulation study. *Molecular Ecology* 14: 2611-2620.

- Farag M.A., Mohsen M., Heinke R., Wessjohann L.A. (2014). Metabolomic fingerprints of 21 date palm fruit varieties from Egypt using UPLC/PDA/ESI-qTOF-MS and GC-MS analyzed by chemometrics. *Food Research International* 64: 218–226.
- Felsenstein J. (1989). PHYLIP: Phylogeny Inference Package (Version 3.2). *Cladistics* 5: 164–166.
- Frankham R. (2003). Genetics and conservation biology. *Comptes Rendus Biologies* 326: 22–29.
- Frankham R. (2005). Genetics and extinction. *Biological Conservation* 126: 131–140.
- Gaiotto F.A., grattapaglia D., Venkovsky R. (2003). Genetic structure, mating system, and long-distance gene flow in heart of palm (*Euterpe edulis* Mart.). *Journal of Heredity* 94: 399–406.
- Gao L., Yi X., Yang Y.X., Su Y.J., Wang T. (2009). Complete chloroplast genome sequence of a tree fern *Alsophila spinulosa*: insights into evolutionary changes in fern chloroplast genomes. *BMC Evol Biol* 9: 130.
- Gaut B.S., Muse S.V., Clark W.D., Clegg M.T. (1992). Relative rates of nucleotide substitution at the *rbcL* locus of monocotyledonous plants. *Journal of Molecular Evolution* 35: 292–303.
- Goldstein D.B. & Schlötterer C. (1999). *Microsatellites: evolution and applications* Oxford University Press, Oxford.
- Goudet J. (2002). Fstat Software, v 2.9.3.2. [http://www2.unil.ch/popgen/softwares/fstat .htm](http://www2.unil.ch/popgen/softwares/fstat.htm).
- Gouta H., Ksia E., Buhner-Zaharieva T., Mliki A., Gogorcena Y. (2012). Development of an SSR-based identification key for Tunisian local almonds. *Sci Agric* 69: 108–113.
- Govaerts R. & Dransfield J. (2005). World checklist of palms. *Royal Botanic Gardens, Kew, U.K.*
- Grivet D., Heinze B., Vendramin G.G., Petit R.J. (2001). Genome walking with consensus primers: application to the large single copy region of chloroplast DNA. *Molecular Ecology Notes* 1: 345–349.
- Gros-Balthazard M., Newton C., Ivorra S., Tengberg M., Pintaud J.C., Terral J.F. 2013). Origines et domestication du palmier dattier (*Phoenix dactylifera* L.). *Revue d'ethnoécologie* 4.
- Guettouchi A., Elshibli S., Haider N., Nabulsi I., Ykhlef N. (2017). Molecular diversity in date palm (*Phoenix dactylifera* L.) cultivars from Algeria indicated by RAPD and ISSR polymorphisms. *Plant Cell Biotechnology and Molecular Biology* 18:76–89.
- Gunn B.F., Baudouine L., Olsen K.M. (2011). Independent Origins of Cultivated Coconut (*Cocos nucifera* L.) in the Old World Tropics. *PLOS ONE*. Doi.org/10.1371/journal.pone.0021143.
- Guo S.W., Thompson E.A. (1992). Performing the exact test of Hardy-Weinberg proportion for multiple alleles. *Biometrics* 48: 361–372.
- Guo J., Wang Y., Song C., Zhou J., Qiu L., Huang H., Wang Y. (2010). A single origin and moderate bottleneck during domestication of soybean (*Glycine max*): implications from microsatellites and nucleotide sequences. *Ann Bot* 106 (3): 505–514.
- Hahn W.J. (2002). A molecular phylogenetic study of the Palmae (Arecaceae) based on *atpB*, *rbcL*, and 18S nrDNA sequences. *Systematic Biology* 51: 92–112.

- Hamrick J.L., Murawski D.A., Nason J. (1993). The influence of seed dispersal mechanisms on the genetic structure of tropical tree populations. *Vegetatio* 107/108: 281-297.
- Hamza H., Ben Abederrahim M.A, Elbekkay M., Ferchichi A. (2013). Comparison of the effectiveness of ISSR and SSR markers in determination of date palm (*Phoenix dactylifera* L.) agronomic traits. *Australian Journal of Crop Science* 7 (6): 763-769.
- Hansen A.K., Escobar L.K., Gilbert L.E., Jansen R.K. (2007). Paternal, maternal and biparental inheritance of the chloroplast genome in *Passiflora* (Passifloraceae): implications for phylogenetic studies. *American Journal of Botany* 94: 42-46.
- Haouane H., El Bakkali A., Moukhli A., Tollon C., Santoni S., Oukabli A., El Modafar C., Khadari B. (2011). Genetic structure and core collection of the World Olive Germplasm Bank of Marrakech: towards the optimised management and use of Mediterranean olive genetic resources. *Genetica* 39:1083-94.
- Hardy G.H. (1908). Mendelian proportions in a mixed population. *Science* 28: 49-50.
- Hardy O.J., Vekemans X. (1999). Isolation by distance in a continuous population: reconciliation between spatial autocorrelation analysis and population genetics models. *Heredity* 83: 145-154.
- Hardy O.J., Charbonnel N., Fréville H., Heuertz M. (2003). Microsatellites allele sizes: a simple test to assess their significance on genetic differentiation. *Genetics* 163: 1467-1482.
- Harley M.M. (2006). A summary of fossil records for Arecaceae. *Bot J Linn Soc* 151: 39-67.
- Hazzouri K.M., Flowers J.M., Visser H.J., Khierallah H.S.M., Rosas U., Pham G.M., Meyer R.S., Johansen C.K., Fresquez Z.A., Masmoudi K., Haider N., El Kadri N., Idaghdour Y., Malek J.A., Thirkhill D., Markhand G.S., Krueger R.R., Zaid A., Michael D., Purugganan M.D. (2015). Whole genome re-sequencing of date palms yields insights into diversification of a fruit tree crop. *Nature Communications* 6: 8824.
- Helentjaris T., King G., Slocum M., Sidenstrang C., Wegman S. (1985). Restriction fragment length polymorphism as probes for plant diversity and their development as tools for applied plant breeding. *Plant Mol Biol* 5: 109-118.
- Henderson S.A., Billotte N., Pintaud J.C. (2006). Genetic isolation of Cape Verde Island *Phoenix atlantica* (Arecaceae) revealed by microsatellite markers. *Conservation Genetics* 7: 213–223.
- Hilgeman R.H. (1972). History of date culture and research in Arizona. *Date growers instate Report* 49: 11-14.
- Hussein A.A., Nawal M.G., Nour G.M., Habib S.S. (1984). Evaluation of some date palm varieties grown in South Sinai Governorate. *Agric Res Review* 62: 289-303.
- Hutter C.M., Schug M.D., Aquadro C.F. (1998). Microsatellite variation in *Drosophila melanogaster* and *Drosophila simulans*: a reciprocal test of the ascertainment bias hypothesis. *Molecular Biology and Evolution* 15: 1620-1636.
- Jahiel M. (1996). Phénologie d'un arbre méditerranéen acclimaté en région tropicale : le dattier au sud du Niger et son appropriation par la société *Manga*. Thèse de doctorat de l'Université de Montpellier II, 268 p.
- Jain S.M., Al-Khayri J.M., Johnson D.V. (2011). Date palm biotechnology. Berlin : *Springer*.

- Jarne P. & Lagoda P.J.L. (1996). Microsatellites, from molecules to populations and back. *Trends in Ecology and Evolution* 11: 424-429.
- Jbir R., Zehdi S., Hasnaoui N., Ben Dhiab A., Mars M., Salhi Hannachi A. (2012). Microsatellite polymorphism in Tunisian pomegranates (*Punica granatum* L.): Cultivar genotyping and identification. *Biochemical Systematics and Ecology* 44: 27–35.
- Jeanson M.L., Labat J.N., Little D.P. (2011). DNA barcoding: a new tool for palm taxonomists? *Ann. Bot.* 108: 1445-1451.
- Jeffreys A.J., Wilson V., Thein S.L. (1985). Hypervariable ‘minisatellite’ regions in human DNA. *Nature* 314: 67-73.
- Jehan T. & Lakhanpaul S. (2006). Single nucléotide polymorphism (SNP) – Méthodes end applications in plant genetics : a review. *Indian journal of biotechnology* 43-459.
- Jombart T., Devillard S., Balloux F. (2010). Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC genet* 11: 94.
- Jones A.G., Small C.M., Paczolt K.A., Ratterman N.L. (2010). A practical guide to methods of parentage analysis. *Molecular Ecology Resources* 10: 6-30.
- Kelchner S.A. & Wendel J.F. (1996). Hairpins create minute inversions in non-coding regions of chloroplast DNA. *Current Genetics* 30: 259-262.
- Khan A., Khan I.A., Heinze B., Azim M.K. (2012). The Chloroplast Genome Sequence of Date Palm (*Phoenix dactylifera* L. cv. ‘Aseel’). Springer-Verlag. *Plant Mol Biol* 30:666–678
- Khierallah H.S.M., Bader S.M., Baum M., Hamwiah A. (2011). Genetic diversity of Iraqi Date palms revealed by microsatellite polymorphism. *Journal of the American Society for Horticultural Science* 136: 282–287.
- Kim K. & Lee H.L. (2005). Widespread occurrence of small inversions in the chloroplast genomes of land plants. *Molecules and Cells* 16: 104–113.
- Kimura M. (1968). Evolutionary Rate at the Molecular Level. *Nature* 217: 624-626.
- Kimura M. (1983). *The neutral theory of molecular evolution* Cambridge University Press.
- Krichen L., Mnejja M., Arús P., Marrakchi M., Trifi F.N. (2006). Use of microsatellite polymorphisms to develop an identification key for Tunisian apricots. *Genetic Resources and Crop Evolution* 53: 1699–1706.
- Krueger R.R. (2001). Date palm germplasm: Overview and utilization. *Proceedings, First International Conference on Date Palm, United Arab Emirates University (Al Ain, U.A.E).* in USA, Al Afifi M.A.R. & Al-Badawy A.A., pp 2–37.
- Kumar S., Tamura K., Nei M. (2004). MEGA3: integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. *Brief Bioinform* 5: 150–163.
- Lhote H. (1964). Gravures rupestres d’Aguennar (Ahaggar). *Journal de la Société des Africanistes* 34:35-84.
- Li Y.C., Korol A.B., Fahima T., Beiles A., Nevo E. (2002). Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review. *Molecular Ecology* 11: 2453-2465.
- Litt M. & Luty J.A. (1989). A hypervariable microsatellite revealed by in-vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. *Am J Hum Genet* 44: 397-401.

- Liu K. & Muse S.V. (2005). Power Marker: an integrated analysis environment for genetic marker analysis. *Bioinformatics* 21: 2128–2129.
- Ludeña B., Chabrillange N., Aberlenc-Bertossi F., Adam H., Tregear J.W., Pintaud J.C. (2011). Phylogenetic utility of the nuclear genes *Agamous 1* and *phytochrome B* in palms (Arecaceae): an example within Bactridinae. *Annals of Botany* 108: 1433–1444.
- Maddison W. & Maddison D. (2007). Mesquite 2. <http://mesquiteproject.org>.
- Mahjbi A., Oueslati A., Baraket G., Salhi-Hannachi A., Zehdi Azouzi S. (2015). Assessment of genetic diversity of Tunisian orange, *Citrus sinensis* (L.) Osbeck using microsatellite (SSR) markers. *Genetics and Molecular Research* 15 (2): gmr.15026564.
- Mantel N. (1967). The detection of disease clustering and a generalized regression approach. *Cancer Research* 27: 209-220.
- Mardanov A.V., Ravin N.V., Kuznetsov B.B., Samigullin T.H., Antonov A.S., Kolganova T.V., Skyabin K.G. (2008). Complete sequence of the duckweed (*Lemna minor*) chloroplast genome: structural organization and phylogenetic relationships to other angiosperms. *J Mol Evol* 66: 555-564.
- Mathew L.S., Seidel M.A., George B., Mathew S., Spannagl M., Haberer G., Torres M.F., Al-Dous E.K., Al-Azwani E.K., Diboun I., Krueger R.R., Mayer K.F.X., Ali Mohamoud Y., Suhre K., Malek J.A. (2015). Genome-Wide Survey of Date Palm Cultivars Supports Two Major Subpopulations in *Phoenix dactylifera*. *G3* 5:1429-1438.
- Matsuoka Y., Vigouroux Y., Goodman M.M., Sanchez J.G., Buckler E.S., Doebley J.F. (2002). A single domestication for maize shown by multilocus microsatellite genotyping. *Proceedings of the National Academy of Sciences, USA*. 99: 6080-6084.
- McKey, D., Emperaire L., Elias M., Pinton F., Robert T., Desmoulière S., Rival L. (2001). Gestions locales et dynamiques régionales de la diversité variétale du manioc en Amazonie. *Genet Sel Evol* 33: 465-490.
- Mogensen H.L. (1996). The hows and whys of cytoplasmic inheritance in seed plants. *American Journal of Botany* 83: 383-404.
- Moore H. E.J. (1973). The major groups of palms and their distribution. *Gentes Herb* 11(2): 27-141.
- Moran P.A.P. (1950). "Notes on continuous stochastic phenomena", *Biometrika* 37: 17-23.
- Moussouni S., Pintaud J.C., Vigouroux Y., Bouguedoura N. (2017). Diversity of Algerian oases date palm (*Phoenix dactylifera* L., Arecaceae): Heterozygote excess and cryptic structure suggest farmer management had a major impact on diversity. *PLOS ONE* <https://doi.org/10.1371/journal.pone.0175232>.
- Munier P. (1973). Le palmier dattier. Ed. Maisonneuve. *Edward Arnold* (Publishers) Ltd, London. 217p.
- Neel M.C. & Ellstrand N.C. (2001). Patterns of allozym diversity in the threatened plant *erigeron parishii* (Asteraceae). *American Journal of Botany* 88(5): 810–818.
- Nei M. (1973). Analysis of gene diversity in subdivided populations. *Proc Natl Acad Sci USA* 70: 3321-3323.
- Nei M., Chesser R. (1983). Estimation of fixation indices and gene diversities. *Annals of human genetics* 47: 253-259.

- Nei M., Tajima F., Tatenno Y. (1983). Accuracy of estimated phylogenetic trees from molecular data. *J Mol Evol* 19: 153-170.
- Newton C., Terral J.F., Ivorra S., Gros-Balthazard M., Tito de Morais C., SPicq S., Tengberg M., Pintaud J.C. (2013). Graines d'histoire : Approche morphométrique de l'agrobiodiversité du palmier dattier, actuelle et d'Égypte ancienne. *Revue d'ethnoécologie* 4: 2267-2419.
- Ohyama K., Fukuzawa H., Kohchi T., Shirai H., Sano T., Sano S., Umesono K., Shiki Y., Takeuchi M., Chang Z., Aota S.I, Inokuchi H., Ozeki H. (1986). Chloroplast gene organization deduced from complete sequence of liverwort *Marchantia polymorpha* chloroplast DNA. *Nature* 322: 572-574.
- Ouenoughi M. & Dubost D. (2005). Le voyage forcé du dattier en Nouvelle-Calédonie. Article scientifique. *Sécheresse* 16 (4): 241-6.
- Ould Mohamed Salem A., Trifi M., Rhouma A., Marrakchi M. (2001). Genetic inheritance of four enzymes in date palm (*Phoenix dactylifera* L.). *Genetic resources and crop evolution* 48: 361-368.
- Ozenda P. (1958). Flore du Sahara septentrional et central. Paris, Centre National de Recherches Scientifiques.
- Palmer J.D. & Thompson W.F. (1982). Chloroplast DANN rearrangements are more frequent when a large inverted repeat sequence is lost. *Cell* 29: 537-550.
- Palmer J.D. (1991). Plastid chromosomes: structure and evolution. In: Bogorad L, Vasil IK (eds) Molecular biology of plastids and the photosynthetic apparatus. *Academic Press, New York*, pp5-53.
- Pasqualone A., Montemurro C., Caponio F., Blanco A. (2004). Identification of virgin olive oil from different cultivars by analysis of DNA microsatellites. *J Agr Food Chem* 52: 1068–1071.
- Petit R.J., El Mousadik A., Pons O. (1998). Identifying populations for conservation on the basis of genetic markers. *Conservation Biology* 12: 844-855.
- Pintaud J.C., Couvreur T., Lara C., Billotte N., Ludena B., Pham J.L. (2006). Structure et dynamique de la diversité génétique dans un complexe sauvage-cultivé tropical : le cas du palmier sud-américain *Bactris gasipaes*. *Les Actes du BRG* 355-370.
- Pintaud J.C., Couvreur T.L.P., Lara C., Ludeña B., Pham, J.L. (2008). Reciprocal introgression between wild and cultivated peach palm (*Bactris gasipaes* Kunth, Arecaceae) in western Ecuador. p.296-308. In: Maxted N., Ford-Lloyd B.V., Kell S.P., Iriondo J.M., Dulloo M.E., Turok J. (Eds.), Crop wild relatives conservation and use. CAB International, U.K.
- Pintaud J.C., Zehdi S., Couvreur T., Barrow S., Henderson S., Aberlenc-Bertosi F., Tregear J., Billotte N. (2010). Species delimitation in the genus *Phoenix* (Arecaceae) based on SSR markers, with emphasis on the identity of the Date Palm (*Phoenix dactylifera* L.). 2010; 267-286. In: Seberg O, Petersen G, Barfod A, Davis J (eds.): Diversity, phylogeny, and evolution in the Monocotyledons, Aarhus University Press, Denmark.
- Pintaud J.C., Ludeña B., Aberlenc-Bertossi F., Zehdi S., Gros-Balthazard M., Ivorra S., Terral J.F., Newton C., Tengberg M., Abdoukader S., Daher A., Nabil M., Saro Hernández I., González-Pérez M.A., Sosa P., Santoni S., Boughedoura N., Moussouni S., Si-Dehbi F. (2013). Biogeography of the Date Palm (*Phoenix dactylifera* L., Arecaceae):

- Insights on the Origin and on the Structure of Modern Diversity. *First IS on Date Palm, Acta Horticulturae* 994, ISHS.
- Popenoe P. (1973). The date palm. Miami, *Field Research Projects* 259 p.
- Pritchard J.K., Stephens M., Donnelly P. (2000). Inference of population structure using multilocus genotype data. *Genetics* 155: 945–959.
- Provan J., Soranzo N., Wilson N.J., Goldstein D.B., Powell W. (1999). A low mutation rate for chloroplast microsatellites. *Genetics* 153: 943–947.
- Provan J., Powell W., Hollingsworth P.M. (2001). Chloroplast microsatellites: new tools for studies in plant ecology and evolution. *Trends in Ecology and Evolution* 16: 142–147.
- Provan J., Biss P., McMeel D., Mathews S. (2004). Universal primers for the amplification of chloroplast microsatellites in grasses (Poaceae). *Molecular Ecology Notes* 4: 262–264.
- Pujol B., Patrice David P., McKey D. (2005). Microevolution in agricultural environments: how a traditional Amerindian farming practice favours heterozygosity in cassava (*Manihot esculenta* Crantz, Euphorbiaceae). *Ecology Letters* 8: 138–147.
- R Development Core Team. (2008). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>.
- Raubeson L.A., Peery R., Chumley T.W., Dziubek C., Fourcade H.M., Boore J.L, Jansen R.K. (2007). Comparative chloroplast genomics: analyses including new sequences from the angiosperms *Nuphar advena* and *Ranunculus macranthus*. *BMC Genomics* 8: 174.
- Reynolds J., Weir B.S., Cockerham C.C. (1983). Estimation of the coancestry coefficient: Basis for a short-term genetic distance. *Genetics* 105: 767–779.
- Rhouma-Chatti S., Choulak S., Zehdi-Azzouzi S., Chatti K., Said K. (2014). Molecular polymorphism and phylogenetic relationships within Tunisian date palm (*Phoenix dactylifera* L.): Evidence of non-coding trnL-trnF regions of chloroplast DNAs. *Scientia Horticulturae* 170: 32–38.
- Rieseberg L.H. & Soltis D.E. (1991). Phylogenetic consequences of cytoplasmic gene flow in plants. *Evolutionary Trends in Plants* 5: 65–84.
- Rivera D., de Castro C.O., Carreno E., Inocencio C., Alcaraz F., Rios S., Palazon J.A., Vazquez L., Laguna E. (2007). Morphological systematics of date-palm diversity (*Phoenix*, Arecaceae) in Western Europe and some preliminary molecular results. In 5th International Symposium on the Taxonomy of Cultivated Plants, Wageningen, NETHERLANDS. *ISHS Acta Horticulturae* 799 : 97–104.
- Roset J.P. (1982). Tagalagal : un site à céramique au Xe millénaire avant notre ère dans l’Aïr (Niger). *Comptes rendus des séances de l’Académie des Inscriptions et Belles lettres* 126: 565–570.
- Rossignol-Strick M. (2003). Climat et végétation sur le plateau iranien. *Dossiers d’Archéologie* 287: 5–17.
- Saddoud O., Chatti K., Salhi-Hannachi A., Mars M., Rhouma A., Marrakchi M., Trifi M. (2007). Genetic diversity of Tunisian figs (*Ficus carica* L.) as revealed by nuclear microsatellites. *Hereditas* 144: 149–157.
- Saitou N. & Nei M. (1987). The neighbour-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* 4: 406–425.

- Saker M., Adwy S., Mohamed A., El-Itriby H. (2006). Monitoring of cultivar identity in tissue culture-derived date palms using RAPD and AFLP analysis. *Biol Plant* 50: 198-200.
- Sakka H., Zehdi S., Ould Mohamed Salem A., Rhouma A., Marrakchi M., Trifi M. (2004). Genetic polymorphism of plastid DNA in Tunisian date palm germplasm (*Phoenix dactylifera* L.) detected with PCR-RFLP. *Genetic Resources and Crop Evolution* 51: 479-487.
- Sambrook J., Fritsch E.F., Maniatis T. (1989). Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY.
- Santoni V., Kieffer S., Desclaux D., Masson F., Rabilloud T. (2000). Membrane proteomics: use of additive main effects with multiplicative interaction model to classify plasma membrane proteins according to their solubility and electrophoretic properties. *Electrophoresis* 21: 3329–3344.
- Sato S., Nakamura Y., Kaneko T., Asamizu E., Tabata S. (1999). Complete structure of the chloroplast genome of *Arabidopsis thaliana*. *DNA Res* 6 (5): 283-90.
- Scarcelli N., Barnaud A., Eiserhardt W., Treier U., Seveno M., d’Anfray A., Vigouroux Y., Pintaud J.C. (2011). A set of 100 DNA primer pairs to study population genetics and phylogeny in Monocotyledons. *PLOS ONE* 6(5): e19954. doi:10.1371/journal.pone.0019954.
- Sedra M.H., Lashermes P., Trouslot P., Combes M.C. (1998). Identification and genetic diversity analysis of date palm (*Phoenix dactylifera* L.) varieties from Morocco using RAPD markers. *Euphytica* 103: 75-82.
- Selkoe K.A. & Toonen R.J. (2006). Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecology Letters* 9: 615-629.
- Shaw J., Lickey E.B., Beck J.T., Farmer S.B., Liu W., Miller J., Siripun K.C., Winder C.T., Schilling E.E., Small R.L. (2005). The tortoise and the hare II: relative utility of 21 non-coding chloroplast DNA sequences for phylogenetic analyses. *American Journal of Botany* 92: 142-166.
- Shikanai T., Endo T., Hashimoto T., Yamada Y., Asada K., Yokota A. (1998). Directed dimorphism of the tobacco *ndhB* gene impairs cyclic electron flow around photosystem I. *Proc Natl Acad Sci USA* 95: 9705-9709.
- Shinozaki K., Ohme M., Tanaka M., Wakasugi T., Hayashida N., Matsubayashi T., Zaita N., Chunwongse J., Obokata J., Yamaguchi-Shinozaki K., Ohto C., Torazawa K., Meng B.Y., Sugita M., Deno H., Kamogashira T., Yamada K., Kusuda J., Takaiwa F., Kato A., Tohdoh N., Shimada H., Sugiura M. (1986). The complete nucleotide sequence of the tobacco chloroplast genome: its gene organization and expression. *EMBO J* 5: 2043–2049.
- Slatkin M. (1985). Gene flow in natural populations. *Annual Review of Ecology and Systematics* 16: 393-430.
- Smouse P., Long J., Sokal R. (1986). Multiple regression and correlation extensions of the Mantel test of matrix correspondence. *Systematic Zoology* 35: 627-632.
- Smouse P. & Peakall R. (1999). Spatial autocorrelation analysis of individual multiallele and multilocus genetic structure. *Heredity* 82: 561-573.
- Sugiura M. (1992). The chloroplast genome. *Plant Mol. Biol* 19: 149-157.

- Taberlet P., Gielly L., Pautou G., Bouvet J. (1991). Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Molecular Biology* 17: 1105-1109.
- Tengberg M. (2012). Beginnings and early history of date palm garden cultivation in the Middle East. *Journal of Arid Environments* 86: 139-147.
- Tengberg M., Newton C., Battesti V. (2014). « L'arbre sans rival ». Palmiers dattiers et palmeraies au Moyen-Orient et en Égypte de la préhistoire à nos jours. *Revue d'ethnoécologie* 4: 1575.
- Terral J.F., Newton C., Ivorra S., Gros-Balthazard M., Tito de Morais C., Picq S., Tengberg M., Pintaud J.C. (2012). Insights into the historical biogeography of the date palm (*Phoenix dactylifera* L.) using geometric morphometry of modern and ancient seeds. *Journal of Biogeography* 39: 929-941.
- Tirichine I., Belguedj M., Benkhalifa A., Guerradi M. (2011). Application des indicateurs ethnobotaniques de la diversité au palmier dattier. *Algerian journal of arid environment* ISSN 2170-1318.
- Turmel M., Otis C., Lemieux C. (1999). The complete chloroplast DNA sequence of the green algae *Nephroselmis olivacea* : insights into the architecture of ancestral chloroplast genomes. *Proc. Nd. Acad Sci USA* 96: 10248-1 0253.
- Uhl N.W. & Dransfiel, J. (1987). Genera palmarum. *Allen Press*, Lawrence, Kansas, U.S.A.
- Umesono K. & Ozeki H. (1987). Chloroplast gene organization in plants. *TIG* 3(10): 281-287.
- Vavilov N.I. (1992). Origin and geography of cultivated plants (English translation), Cambridge University Press, New York, New York, USA.
- Vigouroux Y., Glaubitz J.C., Matsuoka Y., Goodman M.M., Sánchez J.G., Doebley J. (2008). Population structure and genetic diversity of New World maize races assessed by DNA microsatellites. *American Journal of Botany* 95(10): 1240–1253.
- Vos P., Hogers R., Bleeker M., Reijmans M., Van de Lee T., Homes M., Frijters A., Pot J., Peleman J., Kupier M. and Zabeau M. (1995). AFLP: a new technique for DNA fingerprinting. *Nucl Acids Res* 23: 4407-4414.
- Wadley G. & Martin A. (1993). The origins of agriculture – a biological perspective and a new hypothesis. *Australian Biologist* 6: 96–105.
- Walsh J.B. (1993). The effects of different levels of genetic exchange on organelle evolution. *Journal of heredity* 84: 415-418.
- Wang R.L., Stec A., Hey J., Lukens L., Doebley J. (1999). The limits of selection during maize domestication. *Nature* 398: 236-239.
- Weinberg W. (1908). Über den Nachweis der Vererbung beim Menschen. *Jahreshefte des Vereins Varterländische Naturkdunde in Württemberg* 64: 369-382.
- Weir B., Cockerham C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution* 38: 1358-1370.
- Welsh J. & McClelland M. (1990). Fingerprinting genomes using PCR with arbitrary primers. *Nucl Acids Res* 18: 7213-7218.
- Weng M.L., Ruhlman T.A., Jansen R.K. (2017). Expansion of inverted repeat does not decrease substitution rates in *Pelargonium* plastid genomes. *New Phytologist*, 214(2):842-851.

- Werth E. (1933). Zur kultur der dattel Palme und die Frageihrer. *Herkunft Ber Deutsch Bot Ger* 51: 10.
- Williams J.G.K., Kubelik A.R., Livak K.J., Rafalski J.A. and Tingey S.V. (1990). DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucl. Acids Res.* 22: 6531-6535.
- Wilson M.A., Gaut B., Clegg M.T. (1990). Chloroplast DNA evolves slowly in the palm family (Arecaceae). *Molecular Biology and Evolution* 7: 303-314.
- Wolfe K.H., Li W.H., Sharp P.M. (1987). Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. *Proc Natl Acad Sci USA* 84: 9054-9058.
- Wright S. (1931). Evolution in mendelian populations. *Genetics* 16: 97-159.
- Wright S. (1951). The genetical structure of populations. *Ann. Eugen (Lond)* 15: 323-354.
- Wright S. (1969). Evolution and the genetics of populations: a treatise in four volumes. Chicago Press, USA.
- Yang M., Zhang X., Liu G., Yin Y., Chen K., Yun Q., Zhao D., Al Mssallem I. (2010). The complete chloroplast genome sequence of date palm (*Phoenix dactylifera* L.). *PLOS ONE* 5: e12762.
- Zeder M.A., Emshwiller E., Smith B.D., Bradley D.G. (2006). Documenting domestication: the intersection of genetics and archaeology. *Trends in Genetics* 22(3): 139-155.
- Zeder M.A. (2011). The origins of agriculture in the Near East. *Current Anthropology* 52 S4 DOI: 10.1086/659307.
- Zehdi S., Trifi M., Billotte N., Marrakchi M., Pintaud J.C. (2004). Genetic diversity of Tunisian date palms (*Phoenix dactylifera* L.) revealed by nuclear microsatellite polymorphism. *Hereditas* 141: 278-287.
- Zehdi S., Pintaud J.C., Billote N., Ould Mohamed A., Sakka H., Rhouma A., Marrakchi M., Trifi M. (2006). Etablissement d'une clé d'identification variétale chez le palmier dattier (*Phoenix dactylifera* L.) par les marqueurs microsatellites. *Plant Genetic Resources Newsletter* 145: 11-18.
- Zehdi S., Cherif E., Rhouma S., Santoni S., Salhi Hannachi A., Pintaud J.C. (2012). Molecular polymorphism and genetic relationships in date palm (*Phoenix dactylifera* L.): The utility of nuclear microsatellite markers. *Scientia Horticulturae* 148: 255–263.
- Zehdi Azouzi S., Cherif E., Moussouni S., Gros-Balthazard M., Abbas-Naqvi S., Ludeña B., Castillo K., Chabrillange N., Bouguedoura N., Bennaceur M., Si-Dehbi F., Abdoukader S., Daher A., Terral J.F., Santoni S., Ballardini M., Mercuri A., Bensalah M., Kadri K., Othmani A., Littardi C., Salhi Hannachi A., Pintaud J.C., Aberlenc-Bertossi F. (2015). Genetic structure of the date palm (*Phoenix dactylifera* L.) in the Old World reveals a strong differentiation between eastern and western populations. *Annals of Botany* 116: 101–112.
- Zehdi-Azouzi S., Cherif E., Guenni K., Ben Abdelkrim A., Bermil A., Rhouma S., Ben Salah M., Santoni S., Pintaud J.C., Aberlenc-Bertossi F., Salhi Hannachi A. (2016). Endemic insular and coastal Tunisian date palm genetic diversity. *Genetica* 144: 181–190.

- Zeven A.C. (1998). Landraces: A review of definitions and classifications. *Euphitica* 104(2):127-139.
- Zhang D.X. & Hewitt G.M. (2003). Nuclear DNA analyses in genetic studies of populations: practice, problems and prospects. *Molecular Ecology* 12: 563-584.
- Zohary D. & Spiegel-Roy P. (1975). Beginning of fruit growing in the old world. *Sciences* 187: 319-32.
- Zohary D. & Hopf N. (1988). Fruits trees and nuts in domestication of plants in the old word. Oxford Soc. *Clarendon Press Publications*, pp 146-149.
- Zohary D. & Hopf M. (2000). Domestication of plants in the Old World. The origin and spread of cultivated plants in the West Asia, Europe and the Nile Valley. 3rd Ed. Oxford University Press, New York.
- Zohary D., Hopf M., Weiss E. (2012). Domestication of plants in the Old World: the origin and spread of domesticated plants in Southwest Asia, Europe and the Mediterranean bassin (4rd ed.). Oxford University Press, Oxford, U.K.

ANNEXES

Annexe 1 : Liste des échantillons utilisés dans le géotypage SSR nucléaire.

Accession	Nom de la variété	Synonyme	Oasis	Latitude	Longitude	Origine /distribution
MS001	Deglet Nour		Biskra	34.85	5.73	Algérie, Tunisie
MS002	Deglet Nour		Biskra	34.85	5.73	Algérie, Tunisie
MS011	Ghars (Algérien)		Biskra	34.85	5.73	Algérie
MS021	Degla Baida		Biskra	34.85	5.73	Algérie
MS022	Degla Baida		Biskra	34.85	5.73	Algérie
MS031	Arichti		Biskra	34.85	5.73	Algérie, Tunisie
MS032	Arichti		Biskra	34.85	5.73	Algérie, Tunisie
MS041	Halwa		Biskra	34.85	5.73	Algérie
MS042	Litima		Biskra	34.85	5.73	Algérie
MS047	Dokhar (male)		Biskra	34.85	5.73	Algérie
MS048	Dokhar (male)		Biskra	34.85	5.73	Algérie
MS057	Mech Degla		Biskra	34.85	5.73	Algérie, Tunisie
MS058	Mech Degla		Biskra	34.85	5.73	Algérie, Tunisie
MS059	unknown		Biskra	34.85	5.73	Algérie
MS062	unknown		Biskra	34.85	5.73	Algérie
MS064	Halwaya		Biskra	34.85	5.73	Algérie
MS068	Zog El Moggar		Biskra	34.85	5.73	Algérie
MS070	Tanteboucht		Biskra	34.85	5.73	Algérie
MS071	Hamraya I		Biskra	34.85	5.73	Algérie
MS072	Tinicine		Biskra	34.85	5.73	Algérie
MS074	Tinicine		Touggourt	33.10	6.07	Algérie
MS077	Tanteboucht		Touggourt	33.10	6.07	Algérie
MS081	Tanteboucht	El Kayed	Touggourt	33.10	6.07	Algérie
MS082	Tanteboucht	El Kayed	Touggourt	33.10	6.07	Algérie
MS084	Tissibi		Touggourt	33.10	6.07	Algérie
MS085	Tissibi		Touggourt	33.10	6.07	Algérie
MS088	unknown		Touggourt	33.10	6.07	Algérie
MS089	Hamraya II		Touggourt	33.10	6.07	Algérie
MS090	Deglet Mech'a		Touggourt	33.10	6.07	Algérie
MS091	Tanslit		Touggourt	33.10	6.07	Algérie
MS092	Tifazwin		Touggourt	33.10	6.07	Algérie
MS094	Bouldjib		Touggourt	33.10	6.07	Algérie
MS095	Tazougart		Touggourt	33.10	6.07	Algérie
MS097	Taramount		Touggourt	33.10	6.07	Algérie
MS098	Takhoudrayt		Touggourt	33.10	6.07	Algérie
MS100	Tazougart		Touggourt	33.10	6.07	Algérie
MS101	Dimolo		Touggourt	33.10	6.07	Algérie
MS102	D'guel M'rass		Touggourt	33.10	6.07	Algérie
MS104	Takermoust		Touggourt	33.10	6.07	Algérie
MS105	Deglet Jito		Touggourt	33.10	6.07	Algérie
MS107	Tachlilat		Touggourt	33.10	6.07	Algérie

MS110	Chataya		Touggourt	33.10	6.07	Algérie
MS112	Deglet Nour (graine)		Touggourt	33.10	6.07	Algérie
MS113	Bayd Hmam		Touggourt	33.10	6.07	Algérie, Tunisie
MS115	D'guel El Hadj		Touggourt	33.10	6.07	Algérie
MS116	Tawadant		Touggourt	33.10	6.07	Algérie
MS118	Tawraghet		Touggourt	33.10	6.07	Algérie
MS119	Bahdid		Touggourt	33.10	6.07	Algérie
MS120	Kesba	Sokrya	Touggourt	33.10	6.07	Algérie
MS121	Bacheir		Touggourt	33.10	6.07	Algérie
MS122	Aghaliane		Touggourt	33.10	6.07	Algérie
MS124	Deglet Gurara		Touggourt	33.10	6.07	Algérie
MS127	Zaydi	Halimi	Oued Souf	33.37	6.87	Irak
MS128	Zaydi	Masri	Oued Souf	33.37	6.87	Irak
MS130	Halawi	Zohdi	Oued Souf	33.37	6.87	Irak
MS132	Tanteboucht		Oued Souf	33.37	6.87	Algérie
MS135	Abdel Azzaz		Oued Souf	33.37	6.87	Algérie
MS140	Dguel Sidi Khilil		Oued Souf	33.37	6.87	Algérie
MS141	Dguel Sidi Khilil		Oued Souf	33.37	6.87	Algérie
MS142	Amari		Oued Souf	33.37	6.87	Algérie, Tunisie
MS143	Amari		Oued Souf	33.37	6.87	Algérie, Tunisie
MS144	Alig (Algérien)		Oued Souf	33.37	6.87	Algérie
MS145	Alig (Algérien)		Oued Souf	33.37	6.87	Algérie
MS146	Hamraya		Oued Souf	33.37	6.87	Algérie
MS150	Tati		Oued Souf	33.37	6.87	Algérie
MS151	Tati		Oued Souf	33.37	6.87	Algérie
MS154	Halwa		Oued Souf	33.37	6.87	Algérie
MS156	Bukhannus		Oued Souf	33.37	6.87	Algérie
MS157	Bukhannus		Oued Souf	33.37	6.87	Algérie
MS158	Khadri		Oued Souf	33.37	6.87	Algérie
MS160	Dguel Mghas		Oued Souf	33.37	6.87	Algérie
MS162	Dfor Lgot		Oued Souf	33.37	6.87	Algérie
MS163	Dfor Lgot		Oued Souf	33.37	6.87	Algérie
MS164	Zaghraya		Oued Souf	33.37	6.87	Algérie
MS166	Horra		Oued Souf	33.37	6.87	Algérie, Tunisie
MS167	Horra		Oued Souf	33.37	6.87	Algérie, Tunisie
MS168	Alig (Algérien)	Bu'Rus	Oued Souf	33.37	6.87	Algérie
MS169	Alig (Algérien)	Bu'Rus	Oued Souf	33.37	6.87	Algérie
MS170	Tati		Oued Souf	33.37	6.87	Algérie
MS172	Lulu red		Oued Souf	33.37	6.87	Algérie
MS175	Lulu yellow		Oued Souf	33.37	6.87	Algérie
MS178	Kesba		Oued Souf	33.37	6.87	Algérie
MS179	Deglet Nour		Ouargla	31.96	5.35	Algérie, Tunisie
MS189	Ghars (Algérien)		Ouargla	31.96	5.35	Algérie
MS199	Tanslit		Ouargla	31.96	5.35	Algérie

MS205	Takermoust		Ouargla	31.96	5.35	Algérie
MS206	Tissibi		Ouargla	31.96	5.35	Algérie
MS208	Bent Qbala		Ghardaia	32.49	3.68	Algérie
MS209	Bent Qbala		Ghardaia	32.49	3.68	Algérie
MS218	Timjouhart		Ghardaia	32.49	3.68	Algérie
MS219	Timjouhart		Ghardaia	32.49	3.68	Algérie
MS224	Tafeziwin		Ghardaia	32.49	3.68	Algérie
MS225	Tafeziwin		Ghardaia	32.49	3.68	Algérie
MS229	Akerbouche		Ghardaia	32.49	3.68	Algérie
MS230	Akerbouche		Ghardaia	32.49	3.68	Algérie
MS231	Akerbouche		Ghardaia	32.49	3.68	Algérie
MS232	Akerbouche		Ghardaia	32.49	3.68	Algérie
MS233	Akerbouche		Ghardaia	32.49	3.68	Algérie
MS234	Akerbouche		Ghardaia	32.49	3.68	Algérie
MS238	Tissibi		Ghardaia	32.49	3.68	Algérie
MS242	Tazarzayet		Ghardaia	32.49	3.68	Algérie
MS243	Tazarzayet		Ghardaia	32.49	3.68	Algérie
MS245	Tamazouart		Ghardaia	32.49	3.68	Algérie
MS246	Ajoujil		Ghardaia	32.49	3.68	Algérie
MS247	Ajoujil		Ghardaia	32.49	3.68	Algérie
MS249	Tadmam (Algérien)		Ghardaia	32.49	3.68	Algérie
MS253	Tadala		Ghardaia	32.49	3.68	Algérie
MS257	Tazizawt		Ghardaia	32.49	3.68	Algérie
MS258	Tazarzayet		Ghardaia	32.49	3.68	Algérie
MS260	Ououcht		Ghardaia	32.49	3.68	Algérie
MS262	Baydir		Ghardaia	32.49	3.68	Algérie
MS263	Baydir		Ghardaia	32.49	3.68	Algérie
MS266	Alig (Algérien)	Our'rous	Ghardaia	32.49	3.68	Algérie
MS270	Nasser Oussaleh		Ghardaia	32.49	3.68	Algérie
MS272	El Gachouche		Ghardaia	32.49	3.68	Algérie
MS273	El Gachouche		Ghardaia	32.49	3.68	Algérie
MS274	unknown		Tamanrasset	22.78	5.52	Algérie
MS275	unknown		Tamanrasset	22.78	5.52	Algérie
MS276	unknown		Tamanrasset	22.78	5.52	Algérie
MS277	unknown		Tamanrasset	22.78	5.52	Algérie
MS278	unknown		Tamanrasset	22.78	5.52	Algérie
MS279	unknown		Tamanrasset	22.78	5.52	Algérie
MS280	unknown		Tamanrasset	22.78	5.52	Algérie
MS281	Timjouhart		El Menia	30.59	2.88	Algérie
MS291	Deglet Nour		El Menia	30.59	2.88	Algérie, Tunisie
MS296	Ghars (Algerian)		El Menia	30.59	2.88	Algérie
MS301	Hmira		El Menia	30.59	2.88	Algérie
MS302	Hmira		El Menia	30.59	2.88	Algérie
MS307	unknown		El Menia	30.59	2.88	Algérie

MS316	Deglet Jdir		El Menia	30.59	2.88	Algérie
MS317	Deglet Jdir		El Menia	30.59	2.88	Algérie
MS326	Ouarglia		El Menia	30.59	2.88	Algérie
MS329	Ouarglia		El Menia	30.59	2.88	Algérie
MS331	Tinasser		El Menia	30.59	2.88	Algérie
MS332	Tinasser		El Menia	30.59	2.88	Algérie
MS336	Takerboucht		El Menia	30.59	2.88	Algérie
MS337	Takerboucht		El Menia	30.59	2.88	Algérie
MS338	Takerboucht		El Menia	30.59	2.88	Algérie
MS345	Tawragha		El Menia	30.59	2.88	Algérie
MS346	Tawragha		El Menia	30.59	2.88	Algérie
MS349	Ouzamig		Timimoun	29.25	0.23	Algérie
MS351	Chikh		Timimoun	29.25	0.23	Algérie
MS352	Adam Esof		Timimoun	29.25	0.23	Algérie
MS353	Adam Bulla		Timimoun	29.25	0.23	Algérie
MS356	Tgaza		Timimoun	29.25	0.23	Algérie
MS357	Azizaou	Adam Zrak	Timimoun	29.25	0.23	Algérie
MS358	Abbad		Timimoun	29.25	0.23	Algérie
MS360	Ighes N'wagada		Timimoun	29.25	0.23	Algérie
MS361	Gharas		Timimoun	29.25	0.23	Algérie
MS362	Tinrigh		Timimoun	29.25	0.23	Algérie
MS363	Azizaou	Adam Zrak	Timimoun	29.25	0.23	Algérie
MS365	Bukezzine		Timimoun	29.25	0.23	Algérie
MS366	Tinjdel		Timimoun	29.25	0.23	Algérie
MS367	Outaghsaït		Timimoun	29.25	0.23	Algérie
MS368	Timedjwel		Timimoun	29.25	0.23	Algérie
MS369	Tinazidhane		Timimoun	29.25	0.23	Algérie
MS372	Tinhud		Timimoun	29.25	0.23	Algérie
MS373	Aghammu		Adrar	27.87	-0.28	Algérie
MS374	Aghares		Adrar	27.87	-0.28	Algérie
MS375	Tinasser		Adrar	27.87	-0.28	Algérie
MS376	Deglet Wlad Mahmoud		Adrar	27.87	-0.28	Algérie
MS377	Chikh		Adrar	27.87	-0.28	Algérie
MS378	Banekhluf		Adrar	27.87	-0.28	Algérie
MS379	Mes'udiya		Adrar	27.87	-0.28	Algérie
MS380	Takerboucht		Adrar	27.87	-0.28	Algérie
MS381	Takerboucht		Adrar	27.87	-0.28	Algérie
MS382	Takerboucht		Adrar	27.87	-0.28	Algérie
MS383	Timakur		Adrar	27.87	-0.28	Algérie
MS384	Aharthan		Adrar	27.87	-0.28	Algérie
MS385	Asemmat		Adrar	27.87	-0.28	Algérie
MS388	Tilemsu		Adrar	27.87	-0.28	Algérie
MS389	Takerboucht		Adrar	27.87	-0.28	Algérie

MS390	Bent Cherk		Adrar	27.87	-0.28	Algérie
MS391	Bawa'adhim		Adrar	27.87	-0.28	Algérie
MS393	Fagous		Adrar	27.87	-0.28	Algérie, Maroc
MS394	Timliha		Adrar	27.87	-0.28	Algérie
MS395	Andekly		Adrar	27.87	-0.28	Algérie
MS396	Tindukan		Beni Abbes	30.13	-2.17	Algérie
MS397	Adam Bchir		Beni Abbes	30.13	-2.17	Algérie
MS398	Fagous		Beni Abbes	30.13	-2.17	Algérie, Maroc
MS399	Bent Cherk	Cherka	Beni Abbes	30.13	-2.17	Algérie
MS400	Adam Hror		Beni Abbes	30.13	-2.17	Algérie
MS401	Adam Tirnu		Beni Abbes	30.13	-2.17	Algérie
MS403	Harthan		Beni Abbes	30.13	-2.17	Algérie
MS404	Mcharret		Beni Abbes	30.13	-2.17	Algérie
MS405	Buféa		Beni Abbes	30.13	-2.17	Algérie
MS406	Kenta (Algerian)		Beni Abbes	30.13	-2.17	Algérie
MS407	Aharthan		Beni Abbes	30.13	-2.17	Algérie
MS408	Timrisa		Beni Abbes	30.13	-2.17	Algérie
MS410	Ma'tug		Beni Abbes	30.13	-2.17	Algérie
MS411	Chikh	Hamuri	Beni Abbes	30.13	-2.17	Algérie
MS413	Timliha		Beni Abbes	30.13	-2.17	Algérie

Annexe 2 : Inventaire variétal dans les 13 pays phoenicicoles.

Pays	Variétés
Irak	Ashrasi, Basra, Bdmalki, Benosh, Khadrawi, Khara Basra, Khastawi, Maktoom, Saidi, Saylani, Bahram, Barhi, Halawi, Zaidi, Khadraoui, Amir_Hajj, Badrayah.
Oman	Abou_Da'an, Abu_Sabaa, Aqiba, Bahlani_mâle, Barni, Fard, Hassass, Hob, Khalass Oman, Khorî, Mabseli, Naghal, Naghal Hilali, Naghayli, Nashu el Khashba, Nashu el Wahra, Om el Khalass, Qash al Wali, Qash Hamid bin Gharib, Qash zabad, Rajab, Ramli, Seima, Siyali, Umm issi, Fardh 4, Hilali, Jabri, Khalass el Dhahira, Khasab, Khunaizi, Qash Bismin.
Pakistan	Ajwa, Akhrot, Aseel, Begum J, Begun, Be Rehmi, Chapshoq, Choti Sa, Desi, Dhaidi wali, Dhakki, Fasli, Gulistan, Gurbago, Halawi, Hardasht, Hawwa wali, Jaman, Jhahjri, Karblain, Kharoch, Khurmo, Klair, Ko-harba, Kokna, Laloo, Makran, Mobai, Mozati, Naz Taba, Nisri, Noor Muh, Otakanr, Pashpag, Pasht Kh, Peli Sundar, Roghni, Rukkan, Sajho Wali, Seib, Shado, Shakri, Shamran, Surkh, Surkh Be, Wahn wali, Zard Kha, Zardo, Zaydi, Zerî, Tabarzal, Chohara, Goknah, Hussaini, Kasho wali, Khupra, Lasora, Lolo, Nisri, Qantar, Qasab, Rachna.
Émirats Arabes Unis	Bagal abdullah, Bagal Abyad, Bobdiar, Bouman, Chichi, Fardh Ahmar, Hilali, Jesh Naghal, Jesh Ramli, Jibri, Loulou, Marjaban, Moussalli, Nameishi, Qash Abash, Qash Arabi, Rizaz, Sheikhani, Barhi, Khalass, 0372_DAC51, Aksaf, Aleyni, Bouma, Dabbas, Ibri, Nafs-es-Seif.
Djibouti	Femelle, mâle, AF4, AF5, AF7, AGM1, AM1, AMM4, AMM5, CF1, CF2, CF4, CF5, CF6, CM1, CM7, CM8, PB144.
Niger	Babrw46, Maga, Jale7, Galbi41, Zinzime, Namiji1, Ja1, Ja5, Jahaske2, Massara1, Kila72, Ja maga, Koji74, Mad5, Fari2, Mayduhu3, Tiloum, Jahaske1, Baki, Washa.
Égypte	Aghram Aghzâl, Amhat, Arebi, Beda Rufaya, Deqqna, Frehi, Ghazali, Hamra Eg, Hayani, Malakabi, Om el FraH, Oushek, Qaraqoda, Rotob, Samany, Tagtagt, Wardi, Zaghloul, Taj, 2031 DAC545, 2038 DAC547, Barthamoda, Bint Aïsha, El Shabal, Fardh 4 male, Qandila, Sakkoty, Sayidi, Siwi.
Libye	Amreer, Aqdool, Awreeq, L Deglet, Khmag, Taghiyat, Talees.
Mauritanie	Ahmar, Edaghed, Kaylanya, Sekani, Ahmar se, Ahmar1, Ahmar2, Ahmar3, Ahmar4, Ahmar5, Ahmar10.
Maroc	Bouskri, Bou Faggous marocain, Medjoul, Aguelid, Ahardane, Black Bou Sthammi, Bourer, Bouslikh, Bouzegar, Hafsa, Jihel, Mahalbit, Najda, Oum Nhel, Tadmant, Thorry, 2080 DAC559, 2100 DAC571, Bit Djej, Iklane, Oskri, Tarzaoua.
Soudan	Barakawi, Barthamo, Dogna, Gondaila.
Tunisie	Alig tunisien, Amari, Angou, Arehti, Bent Ess, Besser Helou, Bidh Hmam, Bou Feggous tunisien, Cheddakh, Chek Bar, Deglet Bey tunisien, D Mbarka, Deglet Nour, Fazani, Gasbi, Goundi, Hamra, Horra, Kenta tunisien, Kentichi, Khalt Hm, Lagou, Malti, Rhaimya, Tantabec, Tarchegui, Tazerzist, Deg Hasan, Medjhoul Crrao, Lemsi mâle, mâle, mâle, Alig (Ftimi), Aman, Bessir, Deglet Rachid, Ghars tunisien.

Annexe 3 : Extraction d'AND (Isolation of DNA from Plant Tissue Using the DNeasy Plant Mini Kit Qiagen)

Protocole

- Placer le matériel végétal (45mg de folioles coupées en très petits morceaux) dans des tubes de 2ml.
- Ajouter 2 billes d'acier et un peu de sable.
- Broyer les échantillons à l'aide d'un broyeur à billes (TissueLyser). Après broyage, ouvrir les microtubes et y transférer 600µl de Buffer AP1 (Tampon de lyse), refermer les tubes et homogénéiser.
- Ajouter 4µl RNAase.

Extraction et purification de l'ADN

- Incuber les microtubes au bain marie à 65C° pendant 15mn.
- Ajouter 195µl de Buffer AP2, puis agiter les microtubes.
- Mettre les microtubes dans la barquette de glace pendant 10mn.
- Centrifuger les microtubes pendant 5mn à 13000rpm, cette centrifugation permet de précipiter les débris cellulaires.
- Les surnageant sont ainsi prélevés et transférés dans des microtubes stériles à filtres. Les microtubes contenant les culots cellulaires et les billes d'acier sont jetés.
- Centrifuger les microtubes à filtres contenant le surnageant pendant 2mn à 13000rpm.
- Transférer 450µl du surnageant filtré dans un nouveau microtube, et ajouter 675µl de Buffer AP3.
- Prélever à l'aide d'une micropipette 650µl de la solution et la transférer dans des colonnes, puis centrifuger pendant 10 secondes à 13000rpm, et jeter la solution filtrée.
- Rajouter la deuxième partie du volume de la solution, et centrifuger pendant 2mn à 13000rpm.
- Jeter la solution filtrée et ajouter sur le même microtube à filtre 500µl de Buffer AW.
- Centrifuger les microtubes à filtre pendant 1mn.
- Jeter la solution obtenue et centrifuger les microtubes à colonnes vide pendant 2mn à 13000rpm pour bien sécher la membrane du filtre.
- Mettre les colonnes sur des nouveaux microtubes et ajouter 50µl de Buffer AE, laisser 5mn puis centrifuger 1mn, à 8000rpm.

- Rajouter sur les mêmes microtubes à colonnes 50µl de Buffer AE, laisser 5mn puis centrifuger 1mn à 8000rpm.
- L'ADN total extrait est conservé sous forme de solution (100µl) à 20C° avant analyse.

Qualification de l'ADN

La qualité d ADN est estimée visuellement sous rayons ultraviolets après migration sur gel d agarose 1 % en présence de bromure d'éthidium (1µg/ml).

Quantification de l'ADN par spectrophotométrie

L'ADN peut être mesuré directement dans des solutions aqueuses sous forme diluée ou non diluée en mesurant l'absorption A (également définie comme état la densité optique, DO) en lumière ultraviolette selon la méthode spectrophotométrique par le NanoDrop et NanoQuant. La concentration d'acides nucléiques est généralement déterminée par une mesure effectuée à 260 nm contre un échantillon appelé « blanc ». L'interférence par des contaminants se reconnaît par calcul d'un « ratio ». Les protéines absorbant à 280 nm, le ratio A_{260}/A_{280} est utilisé pour estimer la pureté de l'acide nucléique. L'ADN pur devrait avoir un ratio d'environ 1,8.

Annex 4 : Genetic assignment of the sample cultivars according to chlorotype and STRUCTURE grouping for K = 3.

CV name	N° of samples*	Chlorotype	Ancestry G1	Ancestry G2	Ancestry G3
Abbad	1/1	occidental	06.7	93.3	00.0
Abdel Azzaz	1/5	oriental	100.0	00.0	00.0
Adam Bchir	1/1	oriental	01.7	05.0	93.3
Adam Bulla	1/1	oriental	03.3	06.7	90.0
Adam Esof	1/1	oriental	03.3	06.7	90.0
Adam Hror	1/1	oriental	06.7	06.7	86.6
Adam Tirnu	1/1	oriental	01.7	03.3	95.0
Aghaliane	1/2	oriental	51.7	48.3	00.0
Aghammu	1/1	oriental	03.3	16.7	80.0
Aghares	1/1	occidental	01.7	03.3	95.0
Aharthan	2/2	oriental	01.0	00.0	99.0
Ajoujil	2/3	oriental	0.66	03.0	96.4
Akerbouche	6/8	oriental	100	00.0	00.0
Alig (Algerian)	5/8	occidental	01.0	99.0	00.0
Amari	2/2	occidental	08.3	91.7	00.0
Andekly	1/1	oriental	01.6	06.7	91.7
Arichti	2/10	occidental	56.7	43.3	00.0
Asemmat	1/1	oriental	01.6	41.7	56.7
Azizaou	2/2	oriental	03.3	00.0	96.7
Bacheir	1/1	oriental	96.7	01.6	01.7
Bahdid	1/1	oriental	96.7	03.3	00.0
Banekhluf	1/1	oriental	01.7	08.3	90.0
Bawa'adhim	1/1	oriental	01.7	00.0	98.3
Bayd Hmam	1/2	occidental	53.3	46.7	00.0
Baydir	2/4	occidental	01.7	98.3	00.0
Bent Cherk	2/2	occidental	03.3	01.7	95.0
Bent Qbala	2/11	oriental	01.7	01.6	96.7
Bouldjib	1/1	oriental	30.0	01.7	68.3

Buféa	1/1	occidental	01.7	15.0	83.3
Bukezzine	1/1	oriental	01.7	03.3	95.0
Bukhannus	2/2	occidental	05.0	95.0	00.0
Chataya	1/2	oriental	80.0	15.0	05.0
Chikh	3/3	occidental	03.3	70.0	26.7
Degla Baida	3/19	occidental	01.0	99.0	00.0
Deglet Gurara	1/3	occidental	23.3	68.4	08.3
Deglet Jdir	2/10	occidental	01.7	01.6	96.7
Deglet Jito	1/2	oriental	36.7	50.0	13.3
Deglet Mech'a	1/1	oriental	94.0	05.0	01.0
Deglet Nour	4/25	oriental	01.0	99.0	00.0
Deglet Wlad Mahmoud	1/1	occidental	01.7	08.3	90.0
Dfor Lgot	2/2	oriental	99.3	00.7	00.0
Dguel Mghas	1/2	occidental	70.0	15.0	15.0
Dguel Sidi Khilil	2/2	oriental	93.3	01.6	01.7
Dimolo	1/1	oriental	93.3	06.7	00.0
D'guel El Hadj	1/1	oriental	63.3	36.7	00.0
D'guel M'rass	1/1	oriental	76.7	23.3	00.0
El Gachouche	2/2	oriental	95.0	05.0	00.0
Fagous	1/2	occidental	01.7	76.7	21.6
Gharas	1/2	oriental	03.3	00.0	96.7
Ghars (Algerian)	3/25	oriental	99.0	00.0	01.0
Halawi	1/2	oriental	16.7	03.3	80.0
Halwa	2/3	occidental	00.0	100.0	00.0
Halwaya	1/4	occidental	00.0	100.0	00.0
Hamraya I	1/1	oriental	73.3	26.7	00.0
Hamraya II	1/1	oriental	100.0	00.0	00.0
Hamraya	1/4	occidental	50.0	23.3	26.7
Harthan	1/1	occidental	01.0	06.7	92.3
Hmira	2/7	oriental	01.7	00.0	98.3
Horra	2/2	oriental	00.0	100.0	00.0
Ighes N'wagada	1/1	oriental	00.7	09.3	90.0

Kenta (Algerian)	1/1	occidental	01.7	05.0	93.3
Kesba	2/2	oriental	86.7	03.3	11.7
Khadri	1/2	oriental	91.7	05.0	03.3
Litima	1/1	oriental	00.7	94.0	05.3
Lulu red	1/3	oriental	83.3	16.7	00.0
Lulu yellow	1/3	oriental	90.0	05.0	05.0
Ma'tug	1/1	oriental	25.0	01.7	73.3
Mcharret	1/1	oriental	03.3	01.0	95.7
Mech Degla	2/7	occidental	01.0	99.0	00.0
Mes'udiya	1/1	occidental	20.0	13.3	66.7
Nasser Oussaleh	1/2	oriental	06.7	43.3	50.0
Ouarglia	2/5	occidental	03.3	96.7	00.0
Ououcht	1/2	oriental	03.3	10.0	86.7
Outaghsait	1/1	occidental	06.7	10.0	83.3
Ouzamig	1/1	oriental	01.0	00.0	99.0
Tachlilat	1/3	oriental	95.0	03.3	01.7
Tadala	1/4	oriental	16.7	03.3	80.0
Tadmamt (Algerian)	1/5	occidental	13.3	86.7	00.0
Tafeziwin	2/5	oriental	01.7	01.6	96.7
Takerboucht	7/14	occidental	03.3	01.7	95.0
Takermoust	2/3	oriental	93.4	03.3	03.3
Takhoudrayt	1/2	occidental	06.7	86.7	06.6
Tamazouart	1/1	occidental	05.0	95.0	00.0
Tanslit	2/7	oriental	15.7	01.0	83.3
Tanteboucht	5/9	oriental	100.0	00.0	00.0
Taramount	1/1	oriental	63.3	36.7	00.0
Tati	3/4	oriental	100.0	00.0	00.0
Tawadant	1/2	oriental	53.3	46.7	00.0
Tawragha	2/5	oriental	01.7	01.6	96.7
Tawraghet	1/1	oriental	78.3	10.0	11.7
Tazarzayet	3/5	oriental	00.0	00.0	100.0
Tazizawt	1/3	occidental	03.3	20.0	76.7

Tazougart	2/2	oriental	93.3	06.7	00.0
Tgaza	1/	oriental	-	-	-
Tifazwin	1/2	oriental	26.7	10.0	63.3
Tilemsu	1/1	oriental	03.3	03.3	93.4
Timakur	1/1	occidental	01.7	96.7	01.6
Timedjwel	1/1	occidental	10.0	16.7	73.3
Timjouhart	3/6	oriental	01.0	00.0	99.0
Timliha	2/3	oriental	00.7	01.0	98.3
Timrisa	1/1	oriental	05.0	95.0	00.0
Tinasser	3/7	occidental	03.3	06.7	90.0
Tinazidhane	1/1	occidental	01.0	43.3	55.7
Tindukan	1/1	occidental	03.3	85.0	11.7
Tinhud	1/1	oriental	00.7	20.0	79.3
Tinicine	2/4	oriental	100.0	00.0	00.0
Tinjdel	1/1	occidental	10.0	10.0	80.0
Tinrigh	1/1	oriental	03.3	06.7	90.0
Tissibi	4/9	oriental	03.3	00.0	98.7
Zaghraya	1/2	oriental	16.7	10.0	73.3
Zaydi	2/3	oriental	03.3	05.0	91.7
Zog El Moggar	1/2	oriental	85.0	10.0	05.0

*First data ABI genotyping of nuclear and chloroplast SSR, second data Qiaxcel genotyping of chloroplast minisatellites

Annex 5 : Linkage disequilibrium results for 18 markers.

	loc1	loc2	r2
1	mPdIRD031	mPdIRD033	0.000382717269438089
2	mPdIRD031	mPdIRD040	0.000822388876452006
3	mPdIRD031	PdCUC3_ssr2	0.0271795508479885
4	mPdIRD031	PdAP3_ssr_F4	0.000649587886271771
5	mPdIRD031	mPdCIR078	0.000196879685691566
6	mPdIRD031	mPdCIR015	0.00354594048629652
7	mPdIRD031	mPdCIR016	0.00435305063161379
8	mPdIRD031	mPdCIR032	0.0686420404739059
9	mPdIRD031	mPdCIR035	0.000181738826494775
10	mPdIRD031	mPdCIR057	0.0402700890143889
11	mPdIRD031	mPdCIR085	0.0117550249019509
12	mPdIRD031	PdAG1_ssr	0.00627085409601604
13	mPdIRD031	mPdCIR010	0.00865406596572999
14	mPdIRD031	mPdCIR025	0.0100155045173372
15	mPdIRD031	mPdCIR063	2.99714210511709e-06
16	mPdIRD031	PdCUC3_ssr1	0
17	mPdIRD031	mPdIRD013	0.00990365788133385
18	mPdIRD033	mPdIRD040	0.000843644544431946
19	mPdIRD033	PdCUC3_ssr2	3.06683815971309e-06
20	mPdIRD033	PdAP3_ssr_F4	0.00584623997585936
21	mPdIRD033	mPdCIR078	0.000216711229324077
22	mPdIRD033	mPdCIR015	0.00233237065236185
23	mPdIRD033	mPdCIR016	0.0343991294727815
24	mPdIRD033	mPdCIR032	0.0102293113101658
25	mPdIRD033	mPdCIR035	0.0255571918678574
26	mPdIRD033	mPdCIR057	0.000216581146283063
27	mPdIRD033	mPdCIR085	0.00757045925036749
28	mPdIRD033	PdAG1_ssr	0.0153438440752749
29	mPdIRD033	mPdCIR010	0.00293977149699714
30	mPdIRD033	mPdCIR025	0.00894121385924665
31	mPdIRD033	mPdCIR063	0.00141873290422829
32	mPdIRD033	PdCUC3_ssr1	0
33	mPdIRD033	mPdIRD013	0.00115635704975388
34	mPdIRD040	PdCUC3_ssr2	0.0435248620910081
35	mPdIRD040	PdAP3_ssr_F4	0.000403065384481567
36	mPdIRD040	mPdCIR078	0.011015234113419
37	mPdIRD040	mPdCIR015	0.00268197237008597
38	mPdIRD040	mPdCIR016	0.011526992356459
39	mPdIRD040	mPdCIR032	0.00254219346770406
40	mPdIRD040	mPdCIR035	0.00249918523283886
41	mPdIRD040	mPdCIR057	0.00918871463574244
42	mPdIRD040	mPdCIR085	0.000516988472570942

43	mPdIRD040	PdAG1_ssr	0.0311696794492949
44	mPdIRD040	mPdCIR010	0.0467214366507295
45	mPdIRD040	mPdCIR025	0.010526814050473
46	mPdIRD040	mPdCIR063	0.0466379677354182
47	mPdIRD040	PdCUC3_ssr1	0
48	mPdIRD040	mPdIRD013	0.0192680424124957
49	PdCUC3_ssr2	PdAP3_ssr_F4	0.00030016447764237
50	PdCUC3_ssr2	mPdCIR078	0.00992944811978705
51	PdCUC3_ssr2	mPdCIR015	4.15350934417593e-07
52	PdCUC3_ssr2	mPdCIR016	0.0235915181063782
53	PdCUC3_ssr2	mPdCIR032	0.0670358531957313
54	PdCUC3_ssr2	mPdCIR035	0.00192334753781248
55	PdCUC3_ssr2	mPdCIR057	0.000393077224223495
56	PdCUC3_ssr2	mPdCIR085	0.0118622175417075
57	PdCUC3_ssr2	PdAG1_ssr	0.000241370997203613
58	PdCUC3_ssr2	mPdCIR010	0.0273619660351858
59	PdCUC3_ssr2	mPdCIR025	0.0109160087438354
60	PdCUC3_ssr2	mPdCIR063	0.0106657183395729
61	PdCUC3_ssr2	PdCUC3_ssr1	0
62	PdCUC3_ssr2	mPdIRD013	0.00197914801612925
63	PdAP3_ssr_F4	mPdCIR078	0.000353416952938039
64	PdAP3_ssr_F4	mPdCIR015	0.012868985144043
65	PdAP3_ssr_F4	mPdCIR016	0.0372238328421604
66	PdAP3_ssr_F4	mPdCIR032	0.00366220145939701
67	PdAP3_ssr_F4	mPdCIR035	0.0327862259673023
68	PdAP3_ssr_F4	mPdCIR057	0.0382637570740716
69	PdAP3_ssr_F4	mPdCIR085	0.0040597506840609
70	PdAP3_ssr_F4	PdAG1_ssr	0.0138334991148758
71	PdAP3_ssr_F4	mPdCIR010	0.0202039758803603
72	PdAP3_ssr_F4	mPdCIR025	0.000944403052736939
73	PdAP3_ssr_F4	mPdCIR063	0.00696588071637712
74	PdAP3_ssr_F4	PdCUC3_ssr1	0
75	PdAP3_ssr_F4	mPdIRD013	0.00017387752182595
76	mPdCIR078	mPdCIR015	0.157724289513712
77	mPdCIR078	mPdCIR016	0.00072929331010063
78	mPdCIR078	mPdCIR032	0.00517569841564749
79	mPdCIR078	mPdCIR035	0.0106390655888397
80	mPdCIR078	mPdCIR057	0.0362462955919366
81	mPdCIR078	mPdCIR085	0.0464112771768544
82	mPdCIR078	PdAG1_ssr	0.0102459907840628
83	mPdCIR078	mPdCIR010	0.0512471812614954
84	mPdCIR078	mPdCIR025	0.0560281233385465
85	mPdCIR078	mPdCIR063	0.0128823849747096
86	mPdCIR078	PdCUC3_ssr1	0
87	mPdCIR078	mPdIRD013	0.000744843698296168

88	mPdCIR015	mPdCIR016	0.00655701532439451
89	mPdCIR015	mPdCIR032	0.0447549248719361
90	mPdCIR015	mPdCIR035	0.000624867549370315
91	mPdCIR015	mPdCIR057	0.0404302402188928
92	mPdCIR015	mPdCIR085	0.00612864323465014
93	mPdCIR015	PdAG1_ssr	0.00636494491910748
94	mPdCIR015	mPdCIR010	0.0355834138025111
95	mPdCIR015	mPdCIR025	0.0905960203476982
96	mPdCIR015	mPdCIR063	0.0340767947492435
97	mPdCIR015	PdCUC3_ssr1	0
98	mPdCIR015	mPdIRD013	0.0211152271082638
99	mPdCIR016	mPdCIR032	0.0118907476415877
100	mPdCIR016	mPdCIR035	0.0411819886423851
101	mPdCIR016	mPdCIR057	0.00972361861470236
102	mPdCIR016	mPdCIR085	0.000698231943996028
103	mPdCIR016	PdAG1_ssr	0.00369878937633725
104	mPdCIR016	mPdCIR010	0.00596824398785828
105	mPdCIR016	mPdCIR025	0.112522498167946
106	mPdCIR016	mPdCIR063	4.29213978634287e-05
107	mPdCIR016	PdCUC3_ssr1	0
108	mPdCIR016	mPdIRD013	6.37577093696913e-05
109	mPdCIR032	mPdCIR035	5.9466931961294e-05
110	mPdCIR032	mPdCIR057	0.0167377165727428
111	mPdCIR032	mPdCIR085	0.000878357866051523
112	mPdCIR032	PdAG1_ssr	0.00442372272925404
113	mPdCIR032	mPdCIR010	0.01973657148106
114	mPdCIR032	mPdCIR025	0.0121918852818
115	mPdCIR032	mPdCIR063	0.00916681777934644
116	mPdCIR032	PdCUC3_ssr1	0
117	mPdCIR032	mPdIRD013	0.00237260182553454
118	mPdCIR035	mPdCIR057	0.0682931543791107
119	mPdCIR035	mPdCIR085	0.0146866909790032
120	mPdCIR035	PdAG1_ssr	0.00371496081308918
121	mPdCIR035	mPdCIR010	0.0159795591517684
122	mPdCIR035	mPdCIR025	0.00687452935415063
123	mPdCIR035	mPdCIR063	1.71184419650046e-05
124	mPdCIR035	PdCUC3_ssr1	0
125	mPdCIR035	mPdIRD013	0.0107932495818951
126	mPdCIR057	mPdCIR085	0.0191295102702965
127	mPdCIR057	PdAG1_ssr	0.00364340394894195
128	mPdCIR057	mPdCIR010	0.000516585409228642
129	mPdCIR057	mPdCIR025	0.000319631533350089
130	mPdCIR057	mPdCIR063	0.000764036872971146
131	mPdCIR057	PdCUC3_ssr1	0
132	mPdCIR057	mPdIRD013	0.000870814710220407

133	mPdCIR085	PdAG1_ssr	0.00012700155375995
134	mPdCIR085	mPdCIR010	0.000588382595782908
135	mPdCIR085	mPdCIR025	0.010525856088996
136	mPdCIR085	mPdCIR063	0.000141900076560041
137	mPdCIR085	PdCUC3_ssr1	0
138	mPdCIR085	mPdIRD013	0.0076380458861821
139	PdAG1_ssr	mPdCIR010	0.0239631488581693
140	PdAG1_ssr	mPdCIR025	0.054852635725666
141	PdAG1_ssr	mPdCIR063	0.00735688774245966
142	PdAG1_ssr	PdCUC3_ssr1	0
143	PdAG1_ssr	mPdIRD013	0.0225096562443351
144	mPdCIR010	mPdCIR025	0.000987657148325623
145	mPdCIR010	mPdCIR063	0.00459492875906347
146	mPdCIR010	PdCUC3_ssr1	0
147	mPdCIR010	mPdIRD013	0.00599525735464234
148	mPdCIR025	mPdCIR063	0.00461608664522858
149	mPdCIR025	PdCUC3_ssr1	0
150	mPdCIR025	mPdIRD013	0.0227130562326093
151	mPdCIR063	PdCUC3_ssr1	0
152	mPdCIR063	mPdIRD013	0.0389128572687169
153	PdCUC3_ssr1	mPdIRD013	0

Annex 6 : P-value of allelic richness between oases calculated by the Wilcoxon test.

	Biskra	Touggourt	Oued Souf	Ouargla	Ghardaia	Tamanrasset	El Menia	Timimoun	Adrar	Beni Abbes
Biskra		0.8176	0.05688	0.1901	0.6777	0.2247	0.8176	0.3289	0.6112	0.05054
Touggourt			0.1148	0.1055	0.8176	0.2891	0.816	0.306	0.5791	0.06483
Oued Souf				0.6603	0.1743	0.5872	0.1272	0.8176	0.05688	0.1594
Ouargla					0.2435	0.3935	0.1272	0.463	0.2633	1.0000
Ghardaia						0.4874	0.9632	0.4874	0.9265	0.05224
Tamanrasset							0.518	1.0000	0.4874	0.3318
El Menia								0.306	0.89	0.07968
Timimoun									0.2247	0.1202
Adrar										0.002487
Beni Abbes										

The bonferroni corrected p-value for 40 tests is 0.00125.

Annex 7 : P-value of expected heterozygosity between oases calculated by the Wilcoxon test.

	Biskra	Touggourt	Oued Souf	Ouargla	Ghardaia	Tamanrasset	El Menia	Timimoun	Adrar	Beni Abbes
Biskra		1.0000	0.2583	0.2659	0.9773	0.7332	0.7041	0.485	0.6599	0.1472
Touggourt			0.1727	0.4953	0.5698	0.8506	0.6492	0.3384	0.6359	0.2242
Oued Souf				0.1052	0.191	0.3485	0.205	0.776	0.2761	0.5505
Ouargla					0.4344	0.5756	0.9095	0.1126	0.32	0.06826
Ghardaia						0.7367	0.7536	0.9176	0.7982	0.32
Tamanrasset							0.7959	0.6191	0.5681	0.118
El Menia								0.4923	0.3277	0.1928
Timimoun									0.6699	0.2452
Adrar										0.1084
Beni Abbes										

The bonferroni corrected p-value for 40 tests is 0.00125.

Annex 10 : P-value of allelic richness between Structure populations calculated by the Wilcoxon test.

	population 1	population 2	population 3
population 1		0.007734	0.003159
population 2			0.1743
population 3			

Annex 11 : P-value of expected heterozygosity between Structure populations calculated by the Wilcoxon test.

	population 1	population 2	population 3
population 1		0.4341	0.103
population 2			0.2774
population 3			

Annex 12 : P-value of observed heterozygosity between Structure populations calculated by the Wilcoxon test.

	population 1	population 2	population 3
population 1		0.4486	1.0000
population 2			0.2442
population 3			

Annex 13 : P-value of F_{IS} between Structure populations calculated by the Wilcoxon test.

	population 1	population 2	population 3
population 1		0.03182	0.2242
population 2			0.2242
population 3			

نخيل التمر هو أساس زراعة الواحات. استعملنا عينات من 10 واحات في الجزائر لفهم بنية التنوع الوراثي لنخيل التمر. 18 مؤشرات التتابعات القصيرة المتكررة النووية و مؤشر التتابع الطويل المتكرر بلاستيدي لتوصيف 414 يمثلون 114 . وجدنا معامل القرابة سلبي، مما يدل على التحديد النشط من المزارعين للأفراد المتخالفة. هذا التحليل بإنشاء مفتاح التقرير الجزئي للأصناف الجزائرية. . ان التباين داخل وبين الأصناف يقدم قراءة كاملة للتنوع البيولوجي الزراعي للنخيل الجزائري.

تنوع والهيكل الوراثي للنخيل في العالم، قمنا بتقييم تباينات المؤشرات النووية على عينة كبيرة من النخيل على توزيع جغرافي واسع للكشف عن العلاقة بين مجموعات النخيل في مختلف البلدان. قمنا أيضا بتحليل مؤشر التتابع الطويل رر بلاستيدي الذي هو مؤشر انفرادي مع انخفاض معدل الطفرة لديه مقارنة مع المؤشرات النووية، وهكذا تطوره . لأنه موروث من الأم، جينوم بلاستيديات الخضراء يوفر لنا معلومات حول أصل النبتة الأم.

تسلسل مؤشر التتابع الطويل المتكرر الذي يحتوي على اثني عشر قاعدة ازوتية و الموجود في منطقة بين الجينات يظهر الاختلاف بين ثلاثة وأربعة تكرار في النخيل. وقد أظهرت دراسة هذا المؤشر التتابع المتكرر الطويل، عند أربعة عشر نوع فينيكس وجود ستة نسخ، مع وجود بعض النسخ في العديد من أنواع فينيكس.

نسختين جينوم بلاستيديات

ضع جيني للبلستيديات الخضراء في 24 في النخيل، وتتميز بمجموعة من الأشكال المرتبطة.

Abstract

The date palm (*Phoenix dactylifera* L.) is the mainstay of Oasis agriculture. To understand the genetic structure of date palm, we have sampled 10 Oasis in Algeria. Eighteen molecular markers and one chloroplastic marker were used to characterize 414 individuals which correspond to 114 named varieties. We found that, the negative consanguinity coefficient is significant and involved active heterozygote selection by the farmers. This analyze allow us to identify the molecular key to characterize Algerian varieties. However, the intra and inter variability study furnish the complet knowledge of agro-biodiversity of Algerian date palm.

In order to study the genetic structure and diversity of date palm in the world, we have evaluated the nuclear SSR marker polymorphism on large number of sample cover important geographic distribution to make relation among date palm population collected in different countries. We have also analyze a chloroplast minisatellite marker (CpSSR), which is known as haploid with mutation rate less than nuclear SSRs markers and so evolve differently. This chloroplast minisatellite marker, as mother heritage, gives information on the origin of the plant.

The sequencing of this minisatellite marker of 12 base pairs in the non-coding region in the level of *psbZ-trnFM* shows variability within three or four tendems in date palm. The study of this minisatellite in 14 species of *Phoenix* has shown the presence of six haplotyps and some of them are present in many species.

The sequencing of four chloroplastic loci in 24 varieties allowed identifying the two chlorotyps in date palm characterized by linked polymorphism.

Keywords: *Phoenix dactylifera* L., genetic diversity, microsatellites, chloroplast minisatellite.