ULB UNIVERSITÉ LIBRE DE BRUXELLES

Faculté des Sciences ULB

DISPERSION DES GRAINES ET DU POLLEN CHEZ UN ARBRE FORESTIER D'AFRIQUE CENTRALE *ENTANDROPHRAGMA CANDOLLEI* ET CONSÉQUENCES POUR UNE EXPLOITATION DURABLE

MÉMOIRE PRÉSENTÉ EN VUE DE L'OBTENTION DU GRADE DE MASTER EN BIOLOGIE DES ORGANISMES ET ÉCOLOGIE PAR DAMIEN VANDERHOEVEN

Service d'Évolution Biologique et Écologie

Promoteur : Olivier Hardy

SEPTEMBRE 2017

Remerciements

Cette aventure n'aurait pas été possible sans Olivier Hardy, mon cher promoteur. L'initiation du projet il y a plus d'un an m'a permis de pouvoir faire ce travail sur un sujet passionnant et dans une équipe formidable. En plus de ses engagements en tant que promoteur, qu'il a d'après moi remplis avec brio, j'aimerais remercier Olivier de m'avoir permis de découvrir un coin d'Afrique centrale. Ce fut une expérience follement enrichissante que je ne suis pas prêt d'oublier.

La personne qui a sans doute été la plus belle surprise pour moi au cours de ladite aventure est Franck Monthé, dont j'arrive enfin à correctement écrire le prénom. J'étais loin de me douter que la relation que j'allais avoir pendant plus d'un an avec un doctorant qui m'était inconnu pourrait se montrer si riche. Franck, s'il a été un support sans lequel je n'aurais sans doute jamais fini ce travail, est avant tout devenu un vrai ami. Merci pour tous ces moments de discussion scientifique, politique etc., pour ces fous rires qui allègent le quotidien et pour ta bienveillance.

J'aimerais remercier ici le reste de l'équipe EBE, qui s'est montrée accueillante et toujours prête à prêter main forte en cas de besoin. Merci Esra de m'avoir appris à me servir d'une multi-pipette. Merci Laurent de m'avoir empêché à plus d'une occasion de casser tout le laboratoire. Merci Christopher d'avoir œuvré à ce que je puisse venir travailler sept jours sur sept. À tous les doctorants et chercheurs du service, merci pour l'ambiance, pour vous être souciés de moi et pour les évènements super chouettes que vous avez mis en place.

En ce qui concerne le voyage au Cameroun, qui ne fait pas l'objet de ce mémoire, j'aimerais remercier très sincèrement la Pallisco-CIFM, son service Recherche et Reboisement et particulièrement Paul Lagoute de m'avoir accueilli au sein de leur entreprise et du milieu de la foresterie tropicale. J'aimerais aussi remercier l'association Nature + d'avoir rendu ce voyage possible et en particulier Jean-Louis Doucet et Kasso Daïnou. À l'équipe de prospecteurs, botanistes et aides de camp qui m'a suivie dans le Dja : Théophile, Cyril, Engelbert, Guiso, Alexandre, Dieudonné, Alexis, Oboré, Jean-Calvin, Joshep, Koumalo et Shoga ; merci pour votre enthousiasme, votre bonne humeur et votre acceuil.

En ce qui concerne les données qui font l'objet de ce mémoire, je me dois de remercier le projet DynAfFor et la CIB-Olam, particulièrement, Eric Forni et Sylvie Gourlet-Fleuri d'avoir rendue possible la collecte des données.

Pour poursuivre avec les organisations qui rendent des choses possibles, j'aimerais remercier le fonds Ravache d'avoir participé de manière importante à mes frais de voyage au Cameroun, la CFWB et le SSE de l'ULB pour avoir rendu ces cinq années d'études possibles et finalement Hélène Zègres, assistante sociale au SSE, de s'être occupée de tout cela pendant cinq ans.

À ma famille française, merci du fond du cœur de m'avoir communiqué l'amour de la nature et de ce qui est beau.
À mes grands-parents belges, merci de vous être investis dans mon futur et de m'avoir poussé vers l'avant.
À ma mère… merci d'avoir œuvré à faire l'homme que je suis aujourd'hui. Merci pour une éducation de laquelle je peux être fier, de m'avoir toujours soutenu et montré de quoi je suis capable ainsi que pour tout le reste, qui rentrerait difficilement dans ces quelques lignes.

À Gaëlle, merci pour toute cette douceur, ton soutient et les rires avec lesquels tu illumines mon quotidien, particulièrement en ces périodes sombres de rédaction de mémoire. J'espère sincèrement te voir un jour en train d'écrire le tien !

J'aimerais dédier ce mémoire à ma famille ; à ceux qui sont et ceux qui ont été.

Résumé

Afin de proposer des stratégies d'exploitation durable des arbres des forêts tropicales humides il faut comprendre comment ces espèces régénèrent leurs peuplements. *Entandrophragma candollei* est une essence commerciale exploitée en Afrique centrale dont la reproduction est mal connue. L'objectif de ce travail est de caractériser cette reproduction et les flux de gènes attenants pour une population du nord de la République du Congo. Des méthodes d'estimation directes et indirectes de ces flux de gènes ont été effectuées à partir de marqueurs microsatellites et des données spatiales pour tous les individus de la population. Les résultats indiquent des dispersions moyennes du pollen et des graines à une distance importante (340 et 800 m) et un sigma de 1260 m. Le taux d'autofécondation est peu élevé bien que présent et les taux de consanguinité au sein de différentes classes diamétriques présentent des motifs particuliers. Les courbes de probabilité de dispersion des gènes sont leptokurtiques et montrent une quantité importante d'évènement de dispersion à longue distance, bien que la majorité des évènements soient concentrés près des individus adultes. Les résultats montrent aussi une participation à la reproduction plus importante de la part des adultes plus âgés, à partir de 90 cm de diamètre et jusqu'à 150 cm. Il est conclu la durabilité des normes existantes dépend de leur application. Si le diamètre minimum d'exploitation, fixé à l'échelle du pays en tant que base commune à toutes les exploitations, ne semble pas être viable, le diamètre minimum d'aménagement peut l'être, à condition notamment que les taux de prélèvement ne dépassent pas une valeur seuil non estimée dans cette étude.

Summary

In order to be able to propose sustainable exploitation strategies of humid tropical forests timber species, there needs to be an understanding of the population regeneration processes of the exploited species. *Entandrophragma candollei* is a timber species from central Africa whose reproduction has not yet been studied. The goal of this study is to characterize the reproduction and gene flow of this species for a population from the North of the Republic of Congo. Direct and indirect estimates of gene flow have been obtained from SSR markers and spatial data for all individuals in the population. The results show mean dispersal of seeds and pollen of 340 and 800 m respectively and a value of sigma of 1260 m. Selfing rates are quite low, although present, and inbreeding rates among diameter classes show peculiar patterns. Probability of dispersal curves are shown to be leptokurtic. An important part of genes are thus dispersed at long distances, even though most of the dispersal events happen at a short distance from the disperser. Results also show that larger adults tend to participate more to reproduction. Although the species can disperse pollen at a diameter as low as 20 cm, the peak in reproductive success happens between DBHs of 90 and 150 cm. We conclude that the viability of the existing exploitation norms depends on their real life usage. The minimum diameter for exploitation, used throughout the country as a base for all logging companies, does not seem to be viable in regard to sustainability. The minimum diameter for management, on the other hand, can be, if a minimum rate of individuals -not estimated in this study- is left untouched after logging.

TABLE DES MATIÈRES

1 Introduction 7

1.1 Les tenants et aboutissants de l'exploitation forestière en Afrique Centrale 7

1.2 Les flux de gènes 10

1.3 Flux de gènes et durabilité de l'exploitation 14

2 Objectifs de l'étude 17

3 Matériel et méthodes 18

3.1 Modèle étudié : *Entandrophragma candollei* 18

3.2 Collecte des échantillons 19

3.3 Les marqueurs microsatellites 0

3.4 Travail de laboratoire 0

3.5 Traitement et analyse des données 2

4 Résultats 10

4.1 Consanguinité et allèles nuls 10

4.2 Notes à propos des résultats des analyses de parenté 12

4.3 Autofécondation 13

4.4 Dispersion des gènes 13

4.5 Succès reproducteur 17

4.6 Structure génétique spatiale 22

5 Discussion et conclusion 25

5.1 Le système de reproduction 25

5.2 Distance de dispersion des gènes 25

5.3 Succès reproducteur 28

5.4 Structure génétique et structure génétique spatiale 28

5.5 Conclusion 30

6 Références 31

7 Annexes 33

1 INTRODUCTION

1.1 LES TENANTS ET ABOUTISSANTS DE L'EXPLOITATION FORESTIÈRE EN AFRIQUE CENTRALE

L'Homme, comme tant d'autres animaux, a toujours dépendu des forêts pour se procurer les ressources nécessaires à sa survie. Encore aujourd'hui, un grand nombre de personnes dépend des forêts et de leurs ressources pour leur subsistance. Que ce soit pour de la nourriture, des produits médicinaux ou d'autres produits forestiers non-ligneux (PFNL). Par ailleurs, les forêts présentent des opportunités de croissance économique et de bien-être matériel pour les populations, telles que le défrichage de celle-ci pour faire place à des champs ou des pâturages, en extraire le bois pour le vendre ou se chauffer ou encore pour vendre le produit de sa chasse. L'équilibre entre exploitation et conservation est néanmoins fragile : si aujourd'hui 20% de la population mondiale dépend directement des forêts pour sa survie (IUFRO, 2015), l'impact de l'homme cause 76% des forêts tropicales d'être fragmentées ou dégradées (ISU, 2015).

Il est important de différencier déforestation, qui consiste en la coupe à blanc de la surface forestière, de dégradation des forêts, qui reprend les activités dégradant l'écosystème forestier sans pour autant changer son statut. Les deux composantes majeures de la dégradation des écosystèmes forestiers sont, d'une part, la réclamation des terres pour des activités humaine telles que l'agriculture - responsable de 80% de la déforestation dans le monde (Kissinger *et al.*, 2012) - ou l'exploitation des minerais et d'autre part, la dégradation des forêts due à la chasse, à la coupe sélective, au feu et aux effets de fragmentation (Lewis *et al.*, 2015). Parmi ces processus de dégradation, la coupe sélective de pieds pour en extraire le bois serait responsable de 70% de la dégradation des forêts dans le monde (Kissinger *et al.*, 2012).

Les grandes étendues de forêts tropicales humides sont, par ordre décroissant de superficie : le bassin amazonien, le bassin du Congo et les étendues de forêt d'Asie du sud-est. L'appellation de Bassin du Congo est ici utilisée au sens large pour reprendre l'ensemble des forêts denses

humides d'Afrique centrale. Le bassin du Congo constitue donc la deuxième plus grande superficie de forêt tropicale humide au monde, recouvrant 200 millions d'hectares de superficie et pas moins de six pays. Il est aussi un hotspot de biodiversité en Afrique (CIRAD 2011; CBD 2009). Le bassin du Congo se distingue des forêts tropicales d'autres continents par une haute concentration de grande faune (Hall *et al.* 2003), une richesse spécifique en arbres relativement plus basse que celle d'autres continents (Martin 1991) et des forêts majoritairement semi-décidues.

Des 75 millions de personnes habitant la région, 90% dépend directement des ressources forestières animales ou végétales, particulièrement les paysans pauvres (Clark et Sunderland, 2004 cité par COMIFAC 2010). De plus en plus, et par manque d'alternatives viables, les populations se tournent vers l'exploitation de la forêt pour trouver une source de travail et de revenus. La volonté d'allier développement économique et conservation du milieu a amené certains projets du type ICDP (*Integrated Conservation and Development Project*) à juxtaposer ou superposer des régions de conservation et des régions d'exploitation du bois. Ceci a des conséquences significativement néfastes pour les zones de conservation dont les effets sont observables plusieurs décennies après l'exploitation (Hall *et al.* 2003). Toutes ces observations mettent en exergue la nécessité de trouver des solutions pouvant allier l'exploitation forestière avec des programmes de gestion durable ; le développement économique avec la préservation de la biodiversité.

La République du Congo fait partie des six pays couverts par les forêts humides du bassin du Congo. Sont présents sur son territoire 20 millions d'hectares de forêt, soit 10% de l'étendue totale du bassin. Le secteur de l'exploitation forestière y représente 5% du PIB et une grande part des exportations. Le pays s'est engagé dans la lutte contre la coupe illégale forestière en 2009 en prenant part à un Accord de Partenariat Volontaire FLEGT (Application des Réglementations Forestières, Gouvernance et Echanges Commerciaux) avec la FAO et l'Union Européenne.

Bien que ses taux de déforestation soient très peu élevés - moins de 0,1% par an - la République du Congo a aussi décidé de devenir partenaire dans le programme REDD de l'ONU. Programme

qui a pour but la réduction des émissions carbone liées à la déforestation et à la dégradation des forêts dans les pays en développement.

L'aménagement du domaine forestier en République du Congo est similaire aux autres pays du bassin. Le domaine forestier permanent y est divisé en Unités Forestières d'Aménagement ou UFA, qui est l'unité de base pour l'aménagement, la conservation, la production etc. Quand celles-ci sont confiées à des exploitants, elles sont sous-divisées en assiettes de coupe annuelles où sera prélevé le volume maximum annuel de bois autorisé, par un système d'exploitation par rotation. Les plans d'aménagement des UFA sont établis soit par les services publics soit par des sociétés privées spécialisées. Dans ce dernier cas, le plan d'aménagement doit être approuvé par l'administration des eaux et forêts, suite à quoi une autorisation annuelle de coupe est délivrée. Cette autorisation détermine la liste des essences commerciales rentrant dans le volume annuel maximum et se base sur l'inventaire d'aménagement obligatoirement approuvé par l'administration. Ce dernier reprend tous les pieds dont le diamètre est supérieur à 20 cm. Les diamètres minimums exploitables (DME) sont fixés par décret pour chaque essence. En pratique, le diamètre minimum d'exploitabilité sous aménagement (DMA) est utilisé et doit être égal ou supérieur au DME. Ces diamètres minimums ont pour but d'assurer, avec la période de rotation, la reconstitution des populations exploitées. La période de rotation est calculée sur base de l'inventaire d'exploitation et est de l'ordre de 15-25 ans en fonction des assiettes de coupe. Le taux de reconstitution visé pour les essences aménagées est fixé à 50% (AFD, 2010)

Seules deux sociétés forestières parmi une dizaine sont dotées du label FSC (*Forest Stewardship Council*) en République du Congo : l'Industrielle Forestière d'Ouesso et la Congolaise Industrielle des Bois (CIB). Le FSC est une organisation non-gouvernementale à but non lucratif qui promeut la gestion durable des forêts de par le monde. En plus de volets purement écologiques, comme la régénération des espèces ou la préservation de la biodiversité, ses principes contiennent des volets socio-économiques qui visent à garantir le respect des populations locales ainsi que leur développement économique au travers de l'exploitation durable des forêts. Le FSC a développé un système de certification du bois attestant du respect de ses principes, du respect des droits des travailleurs, du suivi et de la traçabilité du bois et de bien d'autres points attenant à ses

idéaux. Les sociétés forestières qui visent à obtenir le label ont donc bien plus de contraintes tout au long de leur chaîne de production que celles établies par la législation du pays.

C'est dans ce cadre que vient se placer l'étude qui fait l'objet de ce mémoire. En effet, si des efforts importants sont faits pour comprendre au mieux les processus de régénération des espèces exploitées et l'impact de l'exploitation forestière sur ceux-ci, de nombreuses zones d'ombre demeurent. Se trouve dans cette situation le Kossipo (*Entandrophragma candollei* Harms 1896), exploité dans tout le bassin du Congo, alors que peu de choses sont connues au sujet de sa phénologie et de son écologie. Cette étude a pour but de répondre à certaines de ces interrogations par le biais d'outils génétiques permettant notamment de caractériser les flux de gènes de cette espèce. Cette information est en effet primordiale si l'on veut comprendre les conséquences de l'exploitation forestière sur la régénération de l'espèce.

1.2 LES FLUX DE GÈNES

1.2.1 Les flux de gènes : définitions

Une définition des plus simples de ce qu'est un flux de gènes pourrait se lire comme ceci : un mouvement de gènes d'une population à une autre. Ce mouvement peut se faire sous forme d'individus adultes ou juvéniles, de zygotes (cas des graines chez les plantes) ou de gamètes. Les caractéristiques de flux de gènes déterminent donc si la connectivité génétique entre populations existe et dans quelle mesure ces flux vont à l'encontre de l'isolement génétique (Sork *et al.,* 2002). L'isolement génétique permet à différentes population d'évoluer séparément et fait donc partie intégrante de du phénomène de spéciation. Chez les plantes, le mouvement d'adultes n'étant pas possible (à quelques rares cas près), les paramètres des flux de gènes sont principalement les taux de migrants et la distance parcourue par ceux-ci. Ces paramètres sont dépendants de plusieurs facteurs, notamment les moyens de dispersion du pollen et des graines, qui sont non seulement différents pour chaque espèce, mais aussi spécifiques à chaque écosystème. Les deux vecteurs de dispersion des gènes les plus couramment utilisés par les plantes sont le vent et les animaux. On parle dans le premier cas d'anémogamie si cela concerne

le pollen et d'anémochorie pour les graines. Si le vecteur de dispersion est animal on utilise alors les termes généraux de zoogamie et zoochorie. Il existe aussi des termes plus spécifiques pour chaque classe d'animaux ainsi que pour d'autres vecteurs, comme par exemple la dispersion par les cours d'eau, etc.

Les flux de gènes peuvent être estimés via des études écologiques et/ou des études génétiques. Les deux types d'études permettent de construire des courbes reprenant la quantité de pollen et de graines dispersée en fonction de la distance ainsi que les distances maximales que ceux-ci peuvent parcourir, ce qui permet d'identifier les éventuelles barrières aux flux de gènes. Ces mêmes barrières peuvent causer un isolement reproducteur partiel ou total de certaines populations, leur permettant d'évoluer indépendamment. Ce type de mécanisme est à la source même de la spéciation et fait partie intégrante du domaine d'étude de la phylogéographie.

Une **étude écologique** sur les flux de gènes complète se doit d'avoir des données regardant les paramètres définissant la dispersion des gènes d'une espèce tels que : le type de pollinisation et de dispersion des graines, les caractères écologiques des animaux participant à cette dispersion dans le cas de zoo-gamie/chorie ou les distributions des vents dans la région dans le cas d'anémo-chorie/gamie, etc. Certaines études (p. e. Jansen *et al.,* 2014) ont même été jusqu'à marquer physiquement les grains de pollen ou les graines pour les capturer après dispersion grâce à des pièges. Une fois ce type de données réunies, il est possible d'estimer, sur bases de connaissances préalables, quelles peuvent être les distances potentielles de dispersion du pollen et des graines et la distribution des probabilités de dispersion par rapport à la distance.

L'idée derrière les **études génétiques** est de reconstituer le flux de gènes d'une population grâce aux informations génétiques qu'elle contient. En effet, en connaissant la composition génétique de la population sur plusieurs locus, il devient possible, entre-autres, d'établir des coefficients de parenté entre ces individus, voir même d'établir des liens de parenté. En sachant que ces données sont la conséquence d'évènements de reproduction et en disposant des coordonnées spatiales des individus, il est possible de caractériser le flux de gènes (Ouborg *et al.,* 1999)

Afin de collecter les données nécessaires aux études écologiques, il est nécessaire de placer plusieurs sortes de pièges (pièges photographiques, pièges à pollinisateurs) autour des arbres d'intérêt afin de pouvoir identifier les disperseurs, de faire des relevés complets des données phénologiques de l'espèce étudiée - idéalement sur plusieurs saisons de reproduction - et de caractériser l'écosystème étudié à travers des paramètres tels que le type de sol ou la pluviométrie, par exemple. Ceci représente un grand investissement de temps, de moyens et de personnel. Bien que les méthodes écologiques et génétiques soient complémentaires et donnent chacune accès à des informations différentes, il est souvent plus aisé de faire recours à ces dernières, à condition d'avoir accès aux équipements et aux méthodes de manière préalable.

Plusieurs marqueurs génétiques existent permettant de faire ce type d'étude : des désuets marqueurs allozymes en passant par les microsatellites pour arriver aux modernes SNPs (*Single Nucleotide Polymorphisms*), le choix est relativement vaste et dépend notamment des moyens scientifiques et financiers disponibles, des questions scientifiques qui sont posées et du type d'organisme étudié. Cette étude a été faite en utilisant les marqueurs microsatellites car ils alliaient une grande diversité génétique avec un coût raisonnable, tout en permettant de répondre à l'ensemble des questions de départ sans compromis (Mariette *et al.,* 2002; Ouborg *et al.,* 1999b).

1.2.2 Les marqueurs microsatellites

Les marqueurs microsatellites, aussi appelés STRs (*Short Tandem Repeats*) ou SSRs (*Simple Sequence Repeats*) sont constitués d'une courte séquence d'ADN (moins de 10 nucléotides) répétée de nombreuses fois. Les séquences longues d'entre 10 et 60 nucléotides sont appelées minisatellites et forment avec les microsatellites les répétitions en tandem de l'ADN (Vieira *et al.,* 2016).

Celles-ci présentent des taux de mutation relativement élevés allant de 10^{-3} à 10^{-6} par génération cellulaire (Gemayel *et al.,* 2012). Une mutation de marqueur microsatellite est de nature différente de celle d'un gène. Les microsatellites mutent généralement en gagnant ou perdant des répétitions lors de leur réplication plutôt qu'en subissant des substitutions, additions ou

délétions de nucléotides (Gemayel *et al.,* 2012). Ceci implique donc que deux allèles microsatellites diffèrent principalement l'un de l'autre non pas par leur séquence nucléotidique mais par leur nombre de répétitions et donc leur longueur. De plus, ne présentant pas de valeur adaptative propre (les microsatellites sont dans une grande majorité des cas non-codants), ils ne subissent pas les effets de la sélection naturelle. Quelques exceptions existent néanmoins, notamment dans le cas de sélection par liaison (ou *linkage selection*), où un locus microsatellite est spatialement proche d'un gène sous sélection et partage donc son devenir. Ces caractéristiques font qu'ils sont hautement polymorphes au sein d'une même population, les rendant idéaux pour ce type d'étude où il est nécessaire de pouvoir discriminer deux individus sur base de leur génotype.

De plus, les marqueurs microsatellites sont dits « codominants », par opposition aux marqueurs dominants où un hétérozygote apparaîtra comme étant homozygote pour l'allèle dominant. Ils donnent de ce fait accès au maximum d'information possible quand utilisés sur des individus hétérozygotes (Mariette *et al.,* 2002).

En conclusion, de tous les marqueurs disponibles, les marqueurs microsatellites sont ceux présentant le plus d'avantages tout en ayant aucun réel inconvénient pour répondre aux questions qui font l'objet de cette étude. Ils sont de plus en plus utilisés afin de caractériser les flux de gènes et ceci selon deux groupes de méthodes analytiques : directes et indirectes.

1.2.3 Méthodes directes et indirectes de l'estimation des flux de gènes

Les **méthodes directes** consistent en une reconstitution des liens de parenté entre les plantules ou juvéniles et les arbres adultes. Chaque plantule présente un génotype issu de deux arbres adultes (ou d'un seul, en cas d'autofécondation). De ce fait, en disposant d'informations sur un nombre convenable de locus suffisamment polymorphes pour l'ensemble de la population, il devient possible de retrouver les parents de chaque juvénile, à la condition que ceux-ci aient été échantillonnés. Si on dispose des coordonnées spatiales des individus il devient possible d'estimer la distance entre les parents, qui équivaut à la distance de dispersion du pollen. La distance de dispersion des graines équivaut à la distance entre la mère et le juvénile. La difficulté

réside en l'identification de la mère parmi les deux parents, dans le cas d'espèces hermaphrodites. Un certain nombre d'hypothèses préalables peut permettre de faire la différence, notamment l'hypothèse que l'arbre le plus proche a une plus grande probabilité d'être la mère.

Les **méthodes indirectes** reconstituent des flux de gènes à partir des fréquences alléliques d'une population (Whitlock et Mccauley, 1999). Des modèles de génétique des populations permettent de faire le lien entre la distribution des fréquences alléliques et les effets des changements démographiques sur la structure génétique de la population (Neigel, 1997). Les caractéristiques des flux de gènes en sont ensuite déduites de cette dernière.
Un tel groupe de modèles est celui des IDB (*Isolation By Distance*). Il serait vain de tenter de détailler ici toutes les adaptations existantes des principes du modèle mais une explication du principe s'avère nécessaire. Une description simpliste des modèles IDB en génétique des populations, est qu'ils prévoient l'augmentation de la différentiation génétique entre individus quand la distance spatiale augmente. Ceci étant dû aux limites physiques et écologiques de la dispersion des gènes. Pour une introduction au sujet et des références additionnelles, se référer à Ishida (2009).

Il est important de noter que ces deux approches sont complémentaires et qu'elles ne fournissent pas les mêmes informations. L'approche directe permet d'estimer la dispersion des gènes contemporaine à travers des assignations de parenté entre adultes et juvéniles. L'approche indirecte, quant à elle, donne des informations sur les flux de gènes historiques ayant donné naissance à la population contemporaine (Oddou-Muratorio et Klein, 2008). Comme les deux approches se font sur base des mêmes données, elles ont été conjointement utilisées dans cette étude.
Les flux de gènes estimés par ces différentes méthodes visent à d'estimer quelles sont les capacités de dispersion de *E. candollei* et de juger si les normes d'exploitation en vigueur permettent à l'essence de se régénérer sur le long-terme.

1.3 FLUX DE GÈNES ET DURABILITÉ DE L'EXPLOITATION

Comme expliqué dans la sous-partie précédente, les vecteurs principaux de dispersion des gènes chez les plantes à fleur sont le pollen et les graines. La distribution de probabilités de dispersion du pollen définit si un individu est en mesure de se reproduire avec ses conspécifiques en fonction de la distance qui les sépare. De même, la distribution de probabilités de dispersion de graines permet de comprendre quelles sont les capacités de dispersion de nouveaux individus et, par conséquent, quelle peut être la structure future de la population étudiée. Ceci comprend la structure génétique tout comme les capacités de colonisation de l'espèce.

Une hypothèse de travail qui a fait l'objet de plusieurs études (Obayashi *et al.,* 2002), est que les arbres hermaphrodites qui ne disposent pas de systèmes d'autoincompatibilité efficaces auront une tendance à faire recours à l'autofécondation proportionnelle à leur isolement reproductif (et donc spatial). Ceci implique que si la dispersion du pollen ne se fait pas à assez longue distance par rapport à la densité effective d'arbres adultes, ces derniers pourraient avoir recours à l'autofécondation. Ce point, combiné au fait que l'apparemment est plus grand entre individus spatialement proches (Ewédjè *et al.,* 2017), implique qu'un flux de gènes limité en distance donnera naissance à une population de juvéniles fortement consanguine (Obayashi *et al.,* 2002). Bien que cela ne pose pas un problème pour la régénération de l'espèce en soi, les fardeaux génétiques importants de ces espèces (Alvarez-Buylla *et al.,* 1996 cité par Dick, 2001) causent des risques élevés de dépression de consanguinité et donc de perte de viabilité chez les nouvelles générations.

Comme l'exploitation sélective a pour conséquence majeure la diminution en densité de la population par l'abattage d'arbres adultes, le lien potentiel entre exploitation forestière et mauvaise régénération de l'espèce est facilement mis en évidence. On comprend donc toute l'importance que prend l'estimation des flux de gènes des espèces exploitées, ceci afin de pouvoir établir des réglementations qui permettent une exploitation durable de celles-ci. Une estimation fiable des flux permet potentiellement d'établir des densités minimales à respecter pour éviter l'isolement reproductif entre adultes ainsi que de comprendre quelles sont les capacités de l'espèce à recoloniser des territoires desquels elle a été éliminée.

Il a été précédemment mentionné que les études écologiques et génétiques étaient complémentaires et permettaient de répondre à des questions scientifiques différentes. Les réponses à ces questions constituent les objectifs que cette étude vise à atteindre. Ils sont présentés dans la partie abordée ci-après.

2 OBJECTIFS DE L'ÉTUDE

Afin de répondre à la problématique introduite relative à l'exploitation durable d'*E. candollei*, il faut pouvoir atteindre les objectifs suivants :

- Caractériser le système de reproduction de l'arbre.

  L'arbre est-il capable de s'autoféconder ?

  Le cas échéant, lequel de l'allogamie ou de l'autogamie est le système dominant ?

  La consanguinité ou la parenté au sein de la population montrent-t-elles des motifs particuliers découlant de processus biologiques ?

- Estimer la distance de dispersion des gènes.

  Quelle est la distribution des probabilités de dispersion de graines et du pollen en fonction de la distance ?

  L'estimation conjointe de la dispersion des graines et du pollen apporte-t-elle des informations supplémentaires ?

- Tester s'il existe une corrélation entre diamètre de l'arbre et efficacité de la dispersion.

  À partir de quel diamètre un arbre intervient-il de façon efficace à la reproduction ?

  Aux alentours de quels diamètres se situe l'optimum reproducteur ?

3 MATÉRIEL ET MÉTHODES

3.1 MODÈLE ÉTUDIÉ : *ENTANDROPHRAGMA CANDOLLEI*

3.1.1 Description morphologique

Entandrophragma candollei est un arbre de la famille des *Meliaceae* communément appelé Kossipo. C'est un arbre hermaphrodite de très grande taille qui peut atteindre 60 mètres de haut et 3 mètres de diamètre. Le tronc peut être muni de contreforts arrondis ou d'empattements et présente de nombreuses lenticelles. Les feuilles, composées de 5 à 9 paires de folioles, sont regroupées à l'extrémité des branches. Ses fleurs sont vert-jaune/ et ses fruits sont des capsules parfois lenticéllées contenant des akènes ailés dispersés par le vent (Meunier *et al.,* 2015).

3.1.2 Écologie

Le Kossipo est une espèce caducifoliée, semi-héliophile des forêts denses humides sempervirentes et semi-décidues. Ses graines sont dispersées en mars-avril ou septembre-octobre (Meunier *et al.,* 2015). Les graines du genre *Entandrophragma* peuvent germer dans des conditions peu ou très lumineuses. Elles se développent néanmoins au mieux quand elles sont à exposées à 25% de la lumière maximale arrivante (Hall, 2008). Il est classé comme arbre non-pionnier (Hawthorne, 1995 cité par J S Hall *et al.,* 2002).

La reproduction du Kossipo n'a pas été étudiée jusqu'à ce jour. Au niveau de sa pollinisation, on peut penser qu'il s'agisse d'une pollinisation de nature entomophile, comme la plupart des arbres de ces régions (Bawa, 1991). Il est aussi probable qu'il préfère l'échange de gènes avec d'autres individus (outcrossing ou allofécondation) à l'autofécondation comme c'est le cas de beaucoup d'autres espèces de ces régions (Ward *et al.,* 2005).Cependant, aucune donnée ne vient appuyer ces suppositions.

3.2 COLLECTE DES ÉCHANTILLONS

3.2.1 Zone d'échantillonnage

La population étudiée est située sur le site de Loudoungou, en République du Congo. La parcelle forestière qu'elle occupe se situe dans le domaine exploité par la société forestière la CIB, plus précisément dans l'assiette de coupe UFA Loundougou-Toukoulaka. Le DMA d'*E. candollei* est ici de 110 cm de diamètre, alors que le DME au niveau du Congo est de 80 cm. Un dispositif expérimental de 8 km² a été mis en place via le projet DynAfFor (www.dynaffor.org) et Franck Monthé a éffectué des collectes de matériel sur 2×3 km² au printemps 2016. Dans cette zone, tous les individus de Kossipo de plus de 20 cm de diamètre ont été exhaustivement échantillonnés. Au centre de ce dispositif, dans une zone de 200 ha, ont aussi été inventoriées les plantes de moins de 20 cm de diamètre. Malheureusement, les graines n'ont pu être échantillonnées faute de fructification à cette période.

3.2.2 Protocole d'échantillonnage

Afin d'assurer un échantillonnage exhaustif, des layons ont été tracés tous les 100 m à l'intérieur de la parcelle. Pour chaque individu, le diamètre (DBH_{130} si assez haut, sinon diamètre au collet) est mesuré et ses coordonnées GPS relevées. Un échantillon d'ADN sous la forme de feuilles est prélevé et gardé dans une enveloppe contenant les données prélevées et le nom de l'espèce. Ces enveloppes sont ensuite conservées dans des sachets hermétiques contenant du gel de silice, ceci afin de les sécher et d'éviter la dégradation des échantillons.

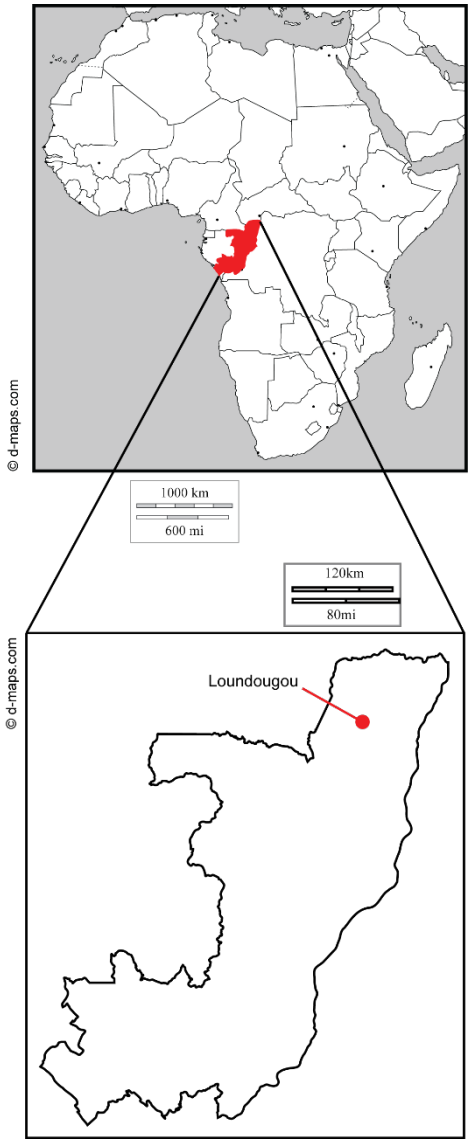
3.2.3 Résultats de l'échantillonnage

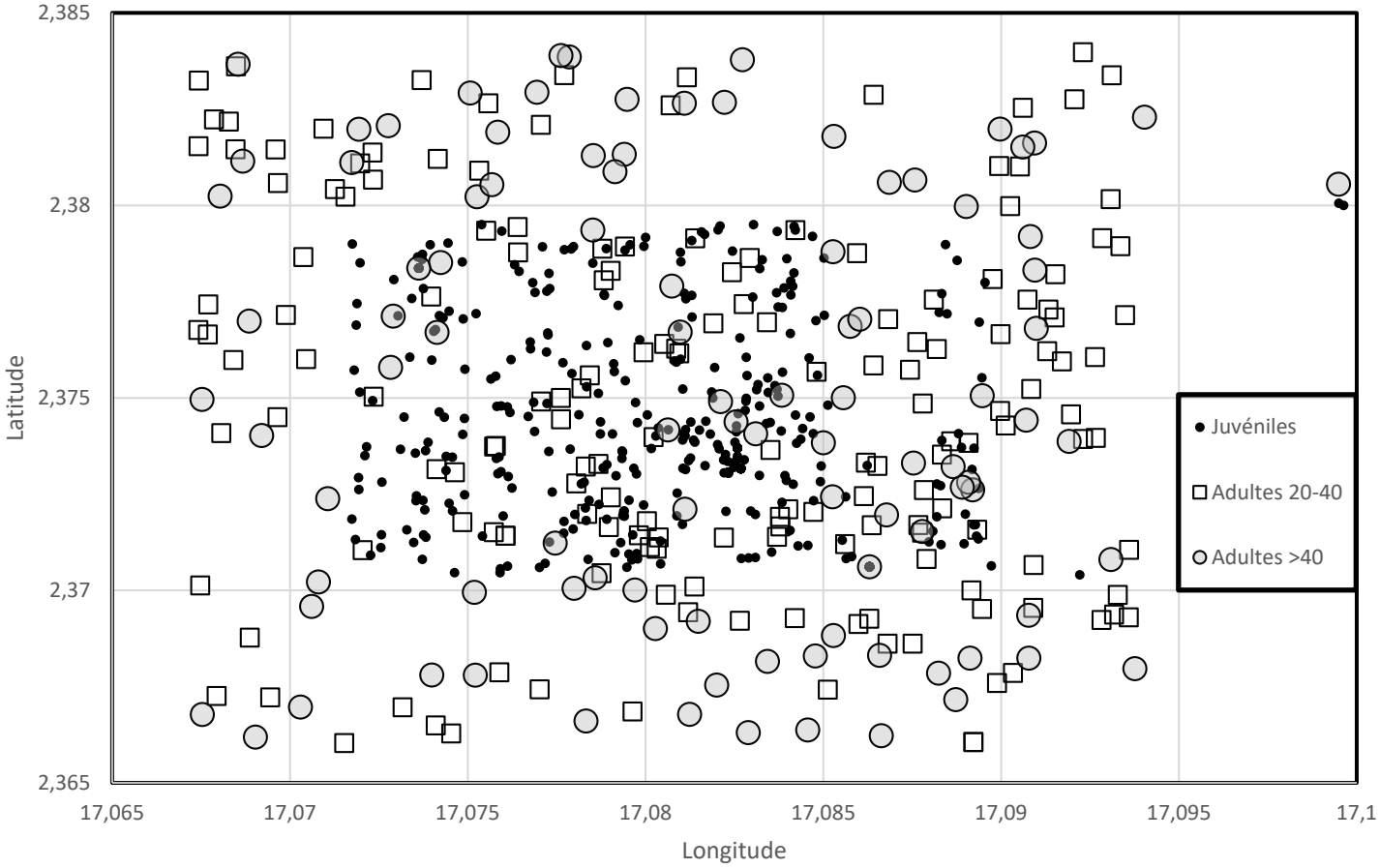
Un total de 1079 Kossipos a été échantillonné. De ce total, trois classes de diamètre ont été créées :

- Classe « juvéniles » : 10 cm de diamètre et moins, 556 individus.
- Classe « adultes » : 20 cm de diamètre et plus, 240 individus.
- Classe « intermédiaires » : plus de 10 cm et moins de 20 cm de diamètre, 282 individus.

La classe intermédiaire n'a pas été incluse dans l'étude génétique qui fait l'objet de ce mémoire. En effet, des résultats obtenus au sein du service pour une espèce du même genre montrent que cette classe contribue pas ou très peu à caractériser le flux de gènes (Monthé *et al.,* 2017) puisque les arbres sont trop jeunes pour participer à la reproduction mais trop vieux que pour être suffisamment sûr de pouvoir retrouver leurs parents. La présente étude ne comprend donc que la classe « juvéniles » et la classe « adultes » (fig. 1). Malgré l'absence de graines, des juvéniles d'un an ont été collectés sous certains individus adultes, pour un total de 8 familles. Ces familles totalisent 109 individus : 103 juvéniles et 6 mères. Une famille est donc constituée d'un arbre-mère et des plantules trouvées à son pied et pour lesquelles on a pu confirmer que l'arbre-mère est véritablement la mère (voir 3.5.3, partie 3). Si plusieurs adultes sont à proximité d'un groupe de juvéniles, ceux-ci constituent une famille mais ne présentent pas de mère assignée.

Figure 1 – Carte montrant la position de Loundougou en Afrique et en République du Congo et carte présentant la position de tous les individus d'*Entandrophragma candollei* présents sur la parcelle. Les unités de latitude et longitude sont des degrés décimaux.





3.3 LES MARQUEURS MICROSATELLITES

Les marqueurs génétiques, « marqueurs microsatellites », ont été développés par Monthe *et al.* (2017) dans le cadre de sa thèse. Ils ont mis au point 16 marqueurs polymorphes dont seulement 13 ont été choisis dans cette étude en raison de leur polymorphisme très élevé.

Les marqueurs microsatellites ont été utilisés en mélanges, ou « mix ». Les marqueurs constituants les mix sont définis en fonction de leurs tailles et d'un fluorochrome qui leur est attaché afin de différentier dans les mix les fragments de tailles similaires. Il est donc possible d'amplifier plusieurs régions de l'ADN au cours de la même PCR en utilisant plusieurs marqueurs microsatellites en même temps, ce qui permet d'optimiser les coûts et le temps de travail.

Ce processus de développement a permis de créer deux mix. Le premier contient les amorces suivantes ; ssr4, ssr13, ssr24, ssr26, ssr32, ssr42, ssr48. Le deuxième ; ssr9, ssr12, ssr16, ssr29, ssr33, ssr38.
L'annexe 1 reprend les informations génétiques sur les marqueurs utilisés.

3.4 TRAVAIL DE LABORATOIRE

3.4.1 Extraction et amplification de l'ADN

Afin d'obtenir l'ADN nécessaire à cette études, 15 à 30mg de feuille ou de cambium ont été extraits à l'aide du kit NucleoSpin® 96 plant, commercialisé par Macherey-Nagel. Chaque extraction comprend un blanc ou contrôle négatif pour vérifier la qualité des extractions réalisées.
L'amplification des séquences microsatellites a été effectuée par PCR (*Polymerase Chain Reaction*), suivant le programme mis au point par (Monthé *et al.,* 2017).

3.4.2 Génotypage

La mesure de la taille des séquences microsatellites a été effectuée par un séquenceur fonctionnant par électrophorèse sur capillaires (3730 DNA Analyzer, Applied Biosystems

Hitachi). Le marqueur de taille utilisé est le Gene Scan-500 LIZ® Size Standard. Il est à noter que le fluorochrome utilisé sur ce marqueur est d'une couleur (orange) non-utilisée dans les mix de marqueurs.

3.4.3 Lecture des résultats

L'output du séquenceur utilisé est un ensemble de fichiers d'extension « .fsa » à raison d'un fichier par puits génotypé (que celui-ci contienne un échantillon ou un contrôle). La lecture de ces fichiers a été faite à l'aide du logiciel Peak Scanner Software v1.0, qui permet de visualiser l'éléctrophérogramme de chaque échantillon sous forme d'un graphique intensité du signal/taille de la séquence (en paires de bases) (fig. 2).

Il revient alors à l'utilisateur d'effectuer la lecture des pics qui lui semblent représenter les ou les allèles présents à un locus donné. En effet, des pics parasites et autres artéfacts peuvent aussi se montrer sur l'éléctrophérogramme. Il est aussi possible que des pics représentant des allèles soient trop grands ou trop petits (en rouge sur la figure 2), traduisant alors une trop haute ou trop faible concentration de ces séquences ADN lors de la lecture par le séquenceur. Tout ceci fait de la lecture des résultats plus une affaire d'interprétation parfois ardue qu'un processus facilement automatisable. Si un individu présente un trop grand nombre de données impossibles ou difficiles à interpréter, il a soit été éliminé du jeu de données final, soit amplifié et génotypé à nouveau.

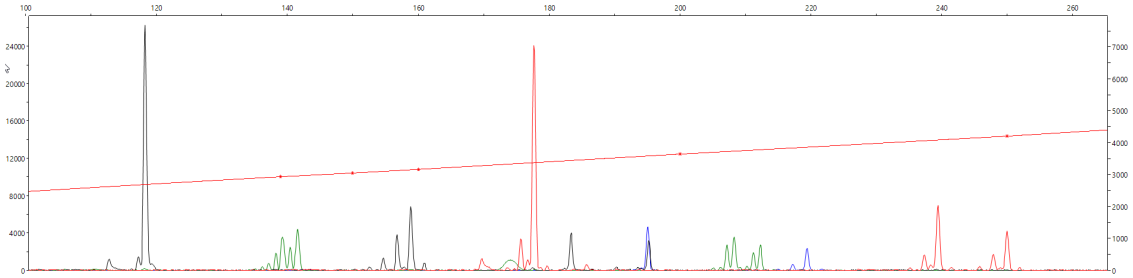


Figure 2 : exemple d'éléctrophérogramme obtenu par l'utilisation du mix 2, tel que représenté dans le logiciel Peak Scanner v1.0

3.5 TRAITEMENT ET ANALYSE DES DONNÉES

3.5.1 Traitements préliminaires

Les résultats des lectures sont dits « bruts » puisqu'ils présentent des estimations de la taille des allèles sous forme de chiffres décimaux. Une étape importante du traitement de données est donc de définir des classes de tailles brutes correspondant à un même allèle représenté par une valeur de taille entière (sans décimale), afin d'obtenir les valeurs finales utilisées dans les analyses.

De plus, tout individu présentant plus de trois allèles sur treize (un quart) pour lesquels aucune information n'a pu être récupérée n'a pas été pris en compte lors des analyses. Dix-sept individus se trouvent dans cette situation, dont un adulte. L'échantillon final utilisé lors des analyses a donc une taille de 764 individus.

3.5.2 Fréquences alléliques, consanguinité et allèles nuls

Le logiciel INEst (Chybicki et Burczyk 2009) permet de calculer les fréquences alléliques, l'hétérozygotie observée (H_o) et attendue (H_e) ainsi que le coefficient de consanguinité et ce pour chaque locus. Le F_{IS} est calculé selon la formule suivante :

$$F_{IS} = 1 - H_o/H_e$$

En plus de cela, le logiciel permet d'estimer la fréquence d'allèles nuls par locus et de fournir une valeur corrigée du F_{IS} en fonction des nouvelles fréquences alléliques. Les estimations sont faites en utilisant le modèle PIM (pour *Population Inbreeding Model*) qui fonctionne par maximum de vraisemblance. Un algorithme de rééchantillonnage jackknife est utilisé afin d'estimer l'erreur standard sur les valeurs obtenues.

Le rééchantillonnage jakknife consiste en le calcul du paramètre (ici le F_{IS}) en retirant du jeu de données une observation à la fois. Pour un échantillon de taille N, on calculera le paramètre N-1 fois. Une fois toutes les estimations calculées, l'estimation finale est faite en faisant la moyenne de celles-ci. L'erreur standard sur cette moyenne est facilement calculée. La variation de F au sein des différentes cohortes et par rapport à la population donne différentes informations. Si, par exemple, le taux de consanguinité est important au sein des juvéniles mais évolue vers une valeur plus basse avec l'âge des individus, on peut en déduire

que la population montre un phénomène de dépression de consanguinité, les individus les plus consanguins étant éliminés par les effets de la sélection naturelle.

Les travaux de Hartl et Clark (2007) montrent que le taux d'autofécondation « s » est :

$$< s > = \frac{2F}{1 - F}$$

À l'équilibre, si toute la consanguinité est due à l'autofécondation, et en l'absence de dépression de consanguinité. Ceci permet donc de calculer les taux d'autofécondation à partir de F corrigés, à condition que les conditions soient remplies.

3.5.3 Méthodes directes et analyses de parenté

Le logiciel CERVUS (Marshall *et al.,* 1998) a été utilisé pour effectuer les analyses de parenté entre arbres adultes et juvéniles. Les assignations de parenté se font par un système de maximum de vraisemblance. On attribue à chaque adulte présent un ratio de vraisemblance (ou score) qui traduit la vraisemblance de parenté de cet individu (hypothèse 1) par rapport à la vraisemblance de parenté d'un individu pris au hasard dans la population (hypothèse 2). Ces vraisemblances, étant donnés les génotypes observés, sont calculées à partir de la probabilité d'obtenir les génotypes observés sous chaque hypothèse. Deux ratios sont utilisables ; LOD (pour *Lokelihood Of Odds*) ou ratio dérivé : Delta. Le premier a été retenu comme ratio de choix lors de ces analyses. Le logiciel ne prend en compte que les génotypes des individus, faisant fi des coordonnées spatiales de ceux-ci. Quatre types d'analyses de parenté sont disponibles : maternité, paternité, et paires de parents (avec sexe des parents connus ou non). CERVUS peut aussi prendre en compte des erreurs de génotypage qui comprennent allèles nuls et valeurs fausses en un seul paramètre. Il est néanmoins incapable de considérer un missmatch potentiellement causé par un allèle nul comme autre chose qu'un missmatch, ce qui peut causer des erreurs lors des attributions de parenté en utilisant des locus en présentant. Il fonctionne en trois étapes.

1. Le calcul des fréquences alléliques

À partir des génotypes des adultes et des plantules que le chercheur fournit, le logiciel calcule les fréquences alléliques de la population, nécessaires pour ce type d'assignation par

maximum de vraisemblance (Marshall *et al.,* 1998). Sont aussi calculées lors de cette étape, les probabilités de non-exclusion (PNE) moyennes et le respect des conditions Hardy-Weinberg pour chaque locus. Les PNE sont la probabilité qu'un locus présente les mêmes allèles chez deux individus choisis aléatoirement et reflète le polymorphisme des allèles utilisés. Le calcul des PNE est fait en supposant que les locus sont en équilibre Hardy-Weinberg.

.

2. Les simulations d'analyses de parenté

Les simulations sont utiles pour deux raisons. Elles permettent tout d'abord d'évaluer la faisabilité d'une analyse de parenté sur base des fréquences alléliques données. Ensuite, elles permettent de calculer les valeurs critiques des ratios de vraisemblance afin de pouvoir donner un indice de confiance quand les analyses de parenté ont lieu. Les valeurs critiques sont calculées comme suit :

Les fréquences alléliques obtenues permettent de générer une paire de génotypes parentaux (les vrais parents), plus une série de génotypes formés de manière aléatoire et représentant des parents non-apparentés. À partir de génotypes des vrais parents et par ségrégation mendélienne, le génotype d'un juvénile est créé. L'idée est de calculer la vraisemblance relative à la parenté des vrais parents par rapport à la vraisemblance de chacun des parents non-apparentés et ceci pour un grand nombre de juvéniles simulés. Pour chaque juvénile, CERVUS identifie le parent le plus probable et enregistre son score LOD (ou Delta). En comparant la distribution des scores pour des juvéniles où le parent le plus vraisemblable était le vrai parent avec la distribution des scores où le parent était un individu non-apparentés, une valeur critique des ratios suffisamment large peut être trouvée, permettant de différencier les vrais parents des candidats non-apparentés avec un degré de confiance prédéterminé par l'utilisateur. Ici 80 et 95 pourcent ont été fixés comme degrés de confiance relaxés et stricts respectivement.

Les paramètres utilisés en entrée pour les simulations sont : Nombre de juvéniles à simuler = 100 000 ; Parents candidats = 200 % de la population réelle concernée, c.à.d. 200 pour les analyses de maternité et 460 pour les analyses de paternité et par paires ; Proportion de la population collectée = 50 % ; Proportion de loci présents = 90 % ; Proportion de loci sur laquelle des erreurs sont présentes = 4 % ; Nombre minimal de loci présents sur tous les

individus = 10 ; Tester l'autofécondation = vrai. Plus de détails concernant ces paramètres peuvent être trouvés dans le fichier d'aide de CERVUS.

3. Analyses de parenté

CERVUS calcule les ratios de vraisemblance pour chaque parent potentiel de chaque juvénile. Le logiciel génère donc un fichier tableur reprenant, pour chaque juvénile, une liste des parents analysés accompagnés de leurs LOD scores. Si ce dernier est positif, cela indique que cet individu a plus de chances d'être le parent véritable qu'un individu pris au hasard, comme expliqué précédemment. Le LOD score permet aussi de marquer les individus adultes ayant atteint le degré de confiance relaxé ou strict.

Le processus d'assignation de parenté a eu lieu en plusieurs étapes. Tout d'abord, il a fallu faire une analyse de maternité. Des observations phénologiques (K. Daïnou, comm. pers.) ont permis d'estimer que *E. candollei* est capable de fructifier à partir d'un diamètre supérieur ou égal à 40 cm. Une liste de 96 mères potentielles a donc été traitée par CERVUS. De toutes les paires juvéniles-mère qui dépassaient le niveau de confiance relaxé (80%), seules celles qui étaient distantes de moins de 300 m ont été retenues. Ceci afin d'éliminer des assignations fort peu probables qui reposeraient sur de la dispersion de graines à longue distance. La valeur de 300 m a été arbitrairement fixée sur base des résultats obtenus par le logiciel NM+ (introduit à la prochaine sous-partie). Les résultats pour les juvéniles assignés à une des 8 familles présentaient parfois une mère assignée par le logiciel différente de la mère assignée sur le terrain. Dans ce cas, l'assignation a été vérifiée manuellement à partir des génotypes mêmes et corrigée quand les missmatchs entre un juvénile et une mère assignée sur le terrain étaient potentiellement dus à des allèles nuls. Des couples juvéniles-mère ont donc été formés dans le respect des conditions expliqués ci-dessus. Quand un tel couple est formé, la mère est considérée comme un vrai parent par CERVUS, qui va alors prendre en compte son génotype lors de la recherche d'un père. Ceci force le logiciel à calculer, en plus de la vraisemblance de la paire juvénile-père, celle du trio juvénile-mère assignée-père à l'étape suivante.

Il s'ensuit donc une analyse de paternité effectuée sur tous les juvéniles. Cette fois-ci, tous les 230 adultes ont été inclus dans la liste des parents potentiels. Les résultats retenus ont été ceux présentant des trios juvénile-mère-père dont l'indice de confiance était d'au moins 80%.

Les juvéniles n'ayant pas trouvé de mère et n'ayant donc pas pu trouver de trio, ils n'ont pas pu être repris dans les données finales. Comme les résultats de ces analyses sont issus d'une analyse de maternité et paternité consécutives et afin d'éviter de longues répétitions inutiles, ils seront désignées comme issus des « analyses MatPat » dans le reste de ce travail. Il est important de noter que rien ne garantit qu'un parent identifié comme mère ne soit pas en fait le père et *vice-versa* si les deux parents ont un diamètre de plus de 40 cm et sont tous deux distants de moins de 300 m.

Finalement, une deuxième méthode d'assignation de parents a été utilisée : une analyse de paires de parents (dont le sexe n'est pas connu). Cette analyse assigne simultanément plusieurs couples de parents à un juvénile en ne prenant pas en compte les distances les séparant ou les diamètres des candidats parents. Le but de cette analyse est de comparer les différents trios obtenus selon les deux méthodes.

Une fois les trios obtenus et connaissant la position spatiale de chaque individu, il est aisé de calculer les distances entre les trois individus les composant. On sait ici que la distance séparant les deux individus parents est celle parcourue par le pollen et on peut supposer que la distance séparant l'adulte le plus proche du juvénile est celle parcourue par les graines.

3.5.4 Méthodes indirectes et modélisation du flux de gènes

Le logiciel NM+ (Chybicki et Burczyk, 2010) prend en entrée les génotypes et coordonnées spatiales des adultes et juvéniles. Il utilise le *Neighborhood Model* décrit par Adam et Birkes (1991, cité par Chybicki et Burczyk, 2010). Le *Neighborhood Model* modélise des évènements de reproduction spatialement explicites et a été conçu pour décrire les patterns de parenté de juvéniles spatialement distribués. Cette approche n'estime pas les flux de gènes d'après des assignations de parenté individuelles -bien que celles-ci soient utilisées- mais plutôt en estimant une série de paramètres propre au modèle, tels que présentés au tableau 1. Le modèle utilisé ne nécessite pas d'assignation de parenté complète et fournit à l'utilisateur des estimations des paramètres par maximum de vraisemblance. Ces estimations sont dérivées indirectement des assignations de parenté (Chybicki et Burczyk, 2010). Le modèle met aussi en relation des paramètres décrivant le succès reproducteur et des données

phénotypiques des adultes qui pourraient s'y corréler. Une telle relation est ici appelée un gradient de sélection. Dans le cette étude, seule le gradient du DBH a été estimé.

Le logiciel offre le choix entre quatre types de fonctions de dispersion de gènes décrites dans Austerlitz *et al.* (2004). Ces fonctions décrivent la distribution des probabilités des évènements de dispersion par rapport à la distance. N'a été utilisée dans cette étude que la fonction de puissance exponentielle. La probabilité qu'un grain de pollen émis au point (0, 0) participe à la dispersion au point (x, y) est donnée par :

$$p_{ep}(a, b; x, y) = \frac{b}{2\pi a^2 \Gamma(2b)} exp\left(-\left(\frac{r}{a}\right)^b\right)$$

Où : Γ est la fonction gamma (décrite dans Abramowitz et Stegun, 1964, cités par Austerlitz *et al.*, 2004) ; r est la distance de pollinisation ; a est un paramètre d'échelle et ; b est le paramètre de forme de la fonction.

Succinctement, le paramètre b donne à la fonction sa forme et définit quels sont les probabilités d'évènements de dispersion à grande distance. En effet, quand b = 1, la fonction de puissance exponentielle dégénère en la fonction exponentielle et la distribution des probabilités est normale. Quand b > 1 la fonction de dispersion décroît plus rapidement en fonction de la distance que la distribution normale, ne laissant que très peu de chances aux évènements de dispersion à longue distance. On parlera alors de queue fine, *thin tail* ou de courbe platykurtique. Au contraire, si b < 1, la chute de la fonction de dispersion est plus lente que dans le cas de la distribution normale. Il s'agit dans ce cas d'une queue épaisse, *fat tail* ou de courbe leptokurtique (Austerlitz *et al.,* 2004). L'estimation du paramètre b est très importante pour ce type d'étude.

NM+ permet aussi de calculer des gradients de sélection. Les gradients de sélection (sg) estiment des éventuelles corrélations existantes entre des variables phénotypiques et le succès reproducteur des individus. Ne disposant que du DBH en tant que variable, c'est le gradient de sélection celle-ci qui a été estimé par le logiciel.

Les allèles nuls et erreurs de génotypage sont repris sous une seule valeur (*typing errors*) qui peut être différente pour chaque locus. Cette valeur peut aussi être estimée par le logiciel.

Le logiciel peut aussi calculer l'effet de la directionnalité sur la probabilité d'être parent. Il introduit pour ça au modèle des paramètres propres à la distribution de von Mises tels que : kappa, un paramètre contrôlant l'intensité de l'effet d'anisotropie et ; l'angle, la direction prévalente. Comme la distribution de von Mises est un analogue circulaire de la distribution normale (Chybicki et Burczyk, 2010), l'angle correspond à la moyenne et kappa à la variance d'une telle distribution. Notons que la valeur de l'angle est mesurée à partir du nord, en supposant que l'axe Y suit la direction nord-sud (Chybicki et Burczyk, 2012).

Tableau 1 – Liste des paramètres estimés par le *Neighborhood Model* et leur description.

Paramètre	Description
ms / mp	Taux d'immigration (1) de graines / de pollen
ds / dp	Distance moyenne de dispersion des graines / du pollen (en mètres)
bs / bp	Paramètre définissant la forme de la fonction de dispersion des graines / du pollen
ks / kp	Paramètre définissant l'anisotropie de la fonction de dispersion des graines / du pollen
as / ap	Direction prévalente de la dispersion des graines / du pollen (angle par rapport au nord)
s	Taux d'autofécondation
sg1, sg2, …	Valeur du gradient de sélection pour la variable phénotypique 1, 2, …
Typing Error	Erreurs de génotypage pour chacun des locus. Comprend les allèles nuls

3.5.5 Structure génétique

Le logiciel SPAGeDi permet d'analyser la structure génétique d'une population (Hardy et Vekemans, 2002). Il travaille sur base de données génétiques provenant de marqueurs codominants et des coordonnées spatiales des individus. SPAGeDi peut calculer des statistiques décrivant la différentiation ou similarité entre individus ou entre populations grâce à des comparaisons par paires.

Le mode de fonctionnement général est le suivant. Afin de d'analyser comment les valeurs desdites comparaison sont liées à la distance spatiale, SPAGeDi calcule :

1. Les valeurs moyennes pour un ensemble d'intervalles de distance prédéfinis.

2. Les régressions linéaires sur la distance spatiale de statistiques calculées par paires.

Les pentes de ces régressions peuvent potentiellement être utilisées pour obtenir des estimations indirectes de paramètres définissant la distance de dispersion des gènes. Elles peuvent aussi être utilisées afin de donner une mesure synthétique de l'intensité de la structure spatiale.

Différents types de permutations permettent de tester la significativité de différents paramètres.

Ont été créées 14 intervalles de distance de manière à ce que chacun reprenne une part suffisamment importante de la population, de 150 m à 5 km. Les coefficients de parenté (Loiselle *et al.,* 2008) ont été calculés pour toute la population adulte pour chaque intervalle de distance. Une enveloppe de distribution à 95 % pour l'hypothèse nulle été calculée sur 10 000 permutations. Par ailleurs, et afin de pouvoir comparer les différentes parties de la population entre-elles, trois catégories ont été créées : juvéniles ; adultes de moins de 40 cm de diamètre (ou petits adultes) ; adultes de plus de 40 cm de diamètre (ou grands adultes). Le même coefficient de parenté a été calculé au sein des différentes catégories et entre elles. Finalement, le paramètre sigma, correspondant à la moyenne quadratique des distances parents – juvéniles, a été estimé pour différentes densités en arbres adultes, comme recommandé par Hardy *et al.* (2006). En diminuant la densité donnée en entrée au programme, on se rapproche de la densité efficace réelle pour la population.

4 RÉSULTATS

4.1 CONSANGUINITÉ ET ALLÈLES NULS

Les statistiques obtenues par le logiciel INEst sont présentées dans les tableaux 2 et 3. Elles reprennent une partie de la structure génétique de la population entière ; adultes et juvéniles confondus. Le coefficient de consanguinité corrigé moyen après la prise en compte de la présence d'allèles nuls est de $\mathbf{F_{IS} = 0,0795}$ ± 0,0234, ou plus de moins de la moitié qu'avant correction (0,184). C'est une valeur assez basse pour une population naturelle. Six des treize locus ont une fréquence d'allèles nuls estimée supérieure à 5 %. Seuls trois dépassent 10 %. Les loci sont assez fortement polymorphes, puisque seulement deux d'entre eux présentent moins de 10 allèles.

La variables relatives à la structure génétique ont aussi été calculées par INEst pour trois sous-ensembles de la population : les adultes et les juvéniles de plus et de moins de 5 cm de diamètre. Les valeurs sont reprises dans le tableau 3. Pour le F corrigé on peut y voir que les juvéniles de plus de 5 cm de diamètre sont plus consanguins (0,100) que ceux de moins de 5 cm de diamètre (0,001). La moyenne de la population adulte se situe entre ces deux groupes (0,560), bien que l'erreur standard soit assez élevée (SE = 0,800). Les valeurs de H_o et F brutes montrent aussi une valeur intermédiaire chez les adultes (0,62 et 0,17 respectivement), plus basse chez les juvéniles >5 cm DBH (0,59 et 0,21) et plus élevée chez les juvéniles <5 cm DBH (0,63 et 0,16).

Un calcul supplémentaire des variables pour la classe de moins de 5 cm de DBH, en excluant les familles de juvéniles, a été effectué afin de confirmer la valeur très basse pour lesdites valeurs. Les résultats ne montrent pas de changement entre les deux estimations pour cette cohorte.

4.1.1 Probabilités de non-exclusion et conditions Hardy-Weinberg

Les PNE pour l'ensemble de la population sont de $1,6 \times 10^{-4}$ pour un premier parent et de $1,2 \times 10^{-6}$ pour un deuxième parent. Lors du calcul de fréquences alléliques CERVUS indique un écart aux conditions de Hard-Weinberg pour 12 des 13 loci, ce qui fausse la base théorique d'obtention des PNE.

Tableau 2 - Statistiques calculées par le logiciel INEst, par locus et sur l'ensemble de la population. A = nombre d'allèles présents ; Ho = hétérozygotie observée ; He = hétérozygotie attendue ; F = coefficient de consanguinité de Wright ; A_0 = fréquence estimée d'allèles nuls ; SE = erreur standard sur l'estimation de A_0.

Locus	A	Ho	He	F	A_0	Se
ssr4	10	0,238	0,3449	0,3098	0,0701	0,0168
ssr13	19	0,4881	0,8103	0,3976	0,1476	0,0145
ssr24	13	0,4184	0,7256	0,4233	0,1429	0,0159
ssr26	18	0,7873	0,8653	0,0901	0,0031	0,0109
ssr32	8	0,4654	0,5366	0,1326	0,014	0,0153
ssr42	18	0,8541	0,8498	-0,0051	0	0
ssr48	5	0,3555	0,5594	0,3645	0,1026	0,0159
ssr9	19	0,7354	0,9097	0,1917	0,0515	0,0138
ssr12	16	0,7497	0,8211	0,087	0	0,0004
ssr16	23	0,8309	0,9173	0,0941	0,0062	0,0118
ssr29	28	0,6846	0,9226	0,2579	0,0856	0,0141
ssr33	13	0,6313	0,7079	0,1082	0,001	0,0055
ssr38	13	0,7767	0,8522	0,0885	0,0028	0,0103
Moyenne	**15,615**	**0,6166**	**0,7556**	**0,184**		

Tableau 3 – Structure génétique de la population et de différentes cohortes. H_o = hétérozygotie observée ; F = coefficient de consanguinité de Wright ; $F_{(null)}$ = F corrigé après prise en compte de la fréquence en allèles nuls par INEst ; SE = erreur standard sur $F_{(null)}$ calculée par jackknife.

Cohorte	H_o	H_e	F	$F_{(null)}$ (SE)
All population (N=764)	0,6166	0,7556	0,184	0,079 (0,023)
Adultes (N=230)	0,6184	0,7499	0,1754	0,056 (0,080)
Juvéniles >5 (N=258)	0,5952	0,7546	0,2112	0,1002 (0,022)
Juvéniles <5 (N=276)	0,6347	0,7518	0,1557	0,0014 (0,004)
Juvéniles <5 (pas de familles ; N=173)	0,6154	0,7572	0,1872	0,0013 (0,004)

4.2 NOTES À PROPOS DES RÉSULTATS DES ANALYSES DE PARENTÉ

Différents résultats ont été obtenus en fonction du système d'analyse utilisé par CERVUS (Tableau 4). Les analyses MatPat ont abouti à l'obtention d'un trio d'individus avec un niveau de confiance supérieur à 80 % pour 66 des 230 juvéniles présents, c'est-à-dire 29 % du total. Les analyses par paires ont quand-à-elles abouti sur un trio avec un niveau de confiance supérieur à 80% pour 87 juvéniles, ou 38 % du total. En comparant les trios obtenus par MatPat et obtenus par paires on remarque que :

- Quand un couple de parents a été trouvé pour un juvénile donné, et ce grâce aux deux analyses (52 cas), les deux parents sont identiques dans 98 % des cas (un seul cas n'est pas identique).
- 14 (21 %) juvéniles présents dans les résultats MatPat sont absents des résultats par paires.

  35 (40 %) juvéniles présents dans les analyses par paires sont absents des résultats MatPat.
- Des 35 juvéniles présents uniquement dans les résultats par paires, 13 présentent des trios où les deux adultes sont distants de plus de 300 m, qui est la limite arbitraire fixée pour l'assignement d'une mère dans les analyses MatPat (*cf.* 3.5.3).
- Des 87 trios obtenus par l'analyse par paires, seuls les parents de 3 trios présentent tous deux un diamètre inférieur à 40 cm.
- 16 des trios obtenus par analyse MatPat présentent deux parents dont le DBH est supérieur à 40 cm et dont la distance au juvénile est inférieure à 300 m. Il n'y a dans ce cas aucune garantie que le sexe ait correctement été attribué aux parents.

Tableau 4 – Nombre de juvéniles ayant trouvé un couple de parents en fonction de l'analyse et du degré de confiance.

Analyse	Seuil de vraisemblance	Nombre de juvéniles trouvant parents
MatPat	> 95 %	52
	> 80 %	66
Pairs	> 95 %	59
	> 80 %	87

4.3 AUTOFÉCONDATION

Trois estimations de la consanguinité ont été faites à partir de différents logiciels :

1. Tout d'abord, en utilisant la valeur de F_{IS} corrigée donnée par INEst, on sait que :

$$< s > = \frac{2F}{1 - F}$$

La fréquence d'autofécondation est de **s = 0,1727** d'après cette méthode, ce qui est une valeur assez élevée.

2. Deuxièmement, si l'on utilise le nombre d'évènements d'autofécondation obtenus d'après les analyses de CERVUS et qu'on le divise par le nombre d'évènements de parenté totaux, on obtient :

$$< s > = \frac{N_{self}}{N_{total}}$$

Qui a pour valeur $s = 7/66 = \mathbf{0,106}$ dans le cas des analyses MatPat et $s = 8/87 = \mathbf{0,092}$ dans le cas des analyses par paires. Ces deux valeurs sont fort proches.

3. Enfin, la valeur du taux d'autofécondation estimée par NM+ est de **s = 0,033** ± 0,11.

On observe que les estimations des taux d'autofécondation sont assez différentes en fonction de l'analyse effectuée.

4.4 DISPERSION DES GÈNES

4.4.1 Approche par modélisation

Les valeurs des paramètres de dispersion estimés par NM+ sont reprises au tableau 5. On y trouve des taux d'immigration importants de plus de 50 % pour les graines et le pollen 50 et 67 % respectivement). La distance de dispersion du pollen est largement plus élevée que celle des graines, avec 340 m et 800 m respectivement, bien que l'erreur standard soit assez élevée pour les deux paramètres (± 74 et 200 m). La fonction de dispersion des graines présente une constante b (*cf* 3.5.4) largement inférieure à 1 (0,23), ce qui indique que le nombre d'évènements de dispersion à grande distance est plus important que dans le cas d'une distribution normale avec une même distance moyenne de dispersion. La même observation est valable pour le pollen à un degré moindre, la valeur de bp étant de 0,71.

L'anisotropie de la dispersion des graines a été estimée par NM+ puisque celles-ci sont dispersées par le vent. Ces paramètres n'ont pas été calculés pour le pollen.

De plus, des courbes présentant la valeur cumulée de gènes dispersés par rapport à la distance à l'arbre adulte ont été obtenues à partir des valeurs des paramètres calculés par NM+ (fig. 3A et 3B). On observe que la courbe du pollen atteint la valeur cumulée de 90 % à une distance bien plus grande que celle des graines. Ceci est notamment expliqué par une valeur de dispersion moyenne bien plus élevée. Par ailleurs, on remarque que la queue de la dispersion est bien plus épaisse dans le cas des graines, comme le reflète la valeur du paramètre b.

4.4.2 Méthodes directes

Pour chacune des analyses, les distances entre individus d'un même trio ont été calculées, pour les degrés de confiance relaxés et stricts. Elles sont reprises au tableau 6. Bien que les trois types de distances évoluent de manière similaire selon les deux analyses, les distances estimées d'après l'analyse par paires sont toujours plus élevées, particulièrement dans le cas de la distance juvénile - mère/parent le plus proche où celle-ci vient presque à doubler, passant de 97 à 182 m ou de 91 à 233 m.

La figure 4 reprend la carte des évènement de dispersion des analyses de maternité, avant même les analyses de maternité. On reprend ici les juvéniles ayant trouvé une mère (de plus de 40 cm DBH donc) pour laquelle le degré de confiance dépasse 95 % et sans limitation de distance. Le centre de la carte représente la position de toutes les mères et chaque point représente les évènements de dispersion pour toutes celles-ci. On observe, une dispersion de 43,8 % des graines dans le quadrant sud-est.

Tableau 5 – Paramètres relatifs à la dispersion des graines et du pollen obtenus par le logiciel NM+ et leur erreur standard. La description des paramètres peut être trouvée au tableau 1. Le taux d'autofécondation « s » est donné ici par souci d'exhaustivité, étant déjà repris au point 4.3.

	Graines					Pollen					
Paramètre	ms	ds (m)	bs	ks	as	mp	s	dp (m)	bp	kp	ap
Valeur	0,50	338,05	0,23	-1,03	25,68	0,67	0,03	793,04	0,71	-	-
SE	0,02	73,91	0,04	0,15	7,49	0,03	0,01	199,70	0,21	-	-

Figure 3 – Valeurs cumulées de la proportion de graines (A) et de pollen (B) dispersés en fonction de la distance, d'après les paramètres du *Neighborhood Model*, ajustés par le logiciel NM+.

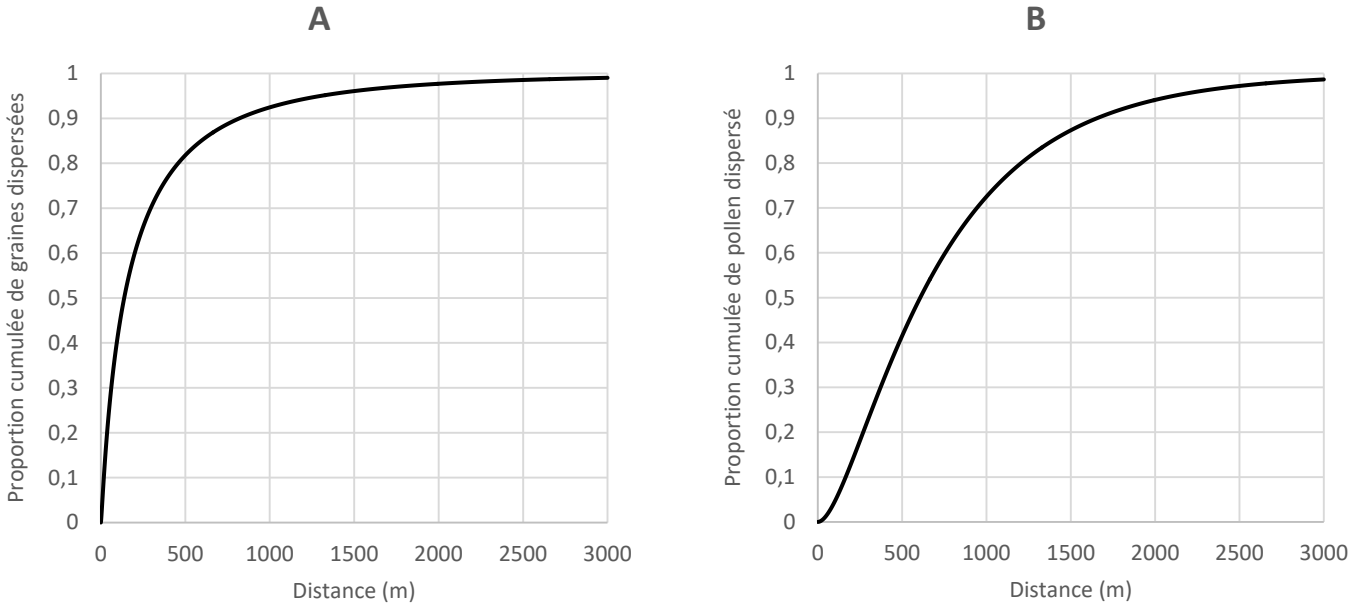
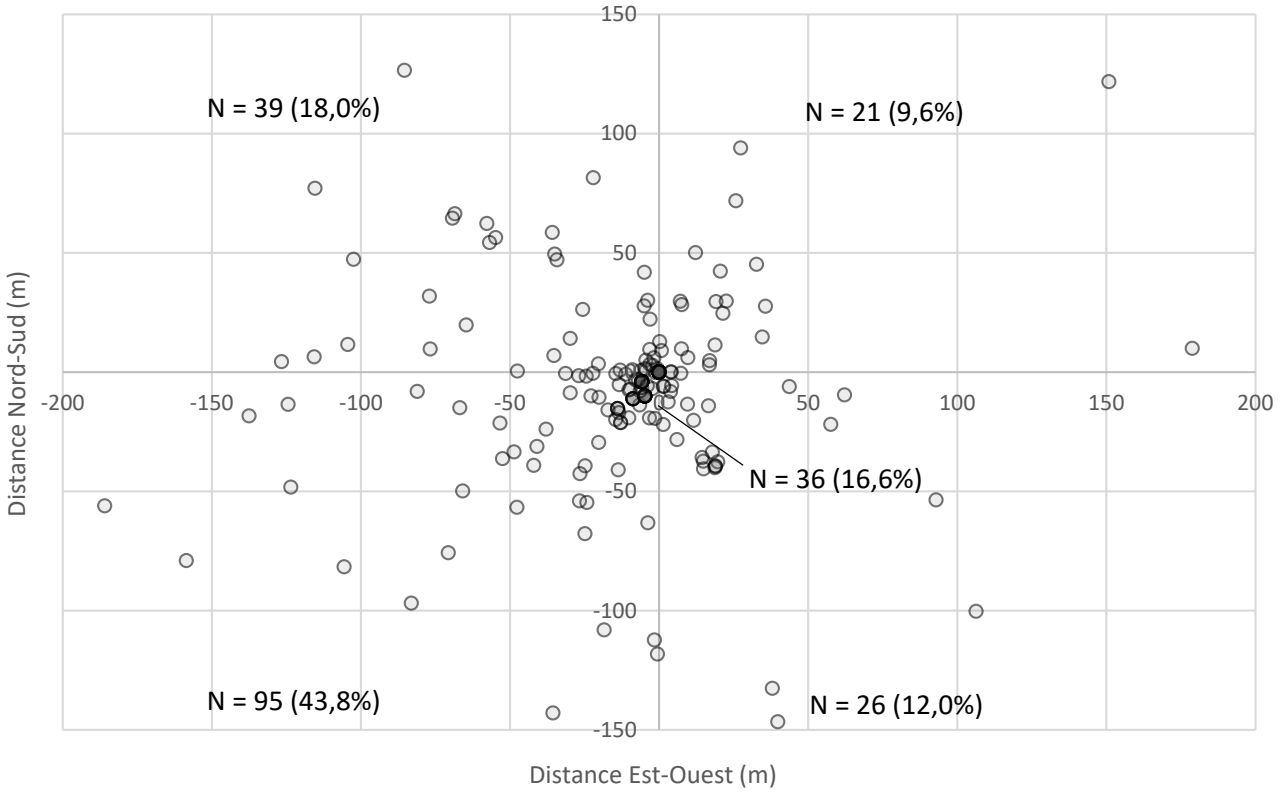


Tableau 6 – Distances moyennes -en mètres- entre individus appartenant à un même trio issu des analyses CERVUS. Les deux types d'analyses sont présentés et les valeurs moyennes de la distance sont accompagnées de leur erreur standard. o = juvénile ; p1 = mère dans le cas des analyses MatPat, ou le parent le plus proche dans le cas des analyses par paires ; p2 = père (MatPat) ou parent le plus lointain (par paires).

Analyse	D(p1 - p2)	D(o - p1)	D(o - p2)
MatPat (>80% ; N = 66)	504 (± 55)	97 (± 11)	520 (± 53)
par paires (>80% ; N = 87)	560 (± 50)	182 (± 32)	640 (± 52)
MatPat (>95% ; N = 52)	441 (± 54)	91 (± 12)	428 (± 55)
par paires (>95% ; N = 59)	402 (± 60)	233 (± 31)	469 (± 55)

Figure 4 – Répartition des évènements de dispersion autour de l'arbre mère d'après les analyses de maternité à degré de confiance > 95 % (N = 217). Chaque point représente un évènement de dispersion efficace. La coordonnée de chaque mère est à (0 ; 0). La somme des évènements de dispersion par quadrant est reprise sur la figure. Les distances sont indiquées en mètres.



4.5 SUCCÈS REPRODUCTEUR

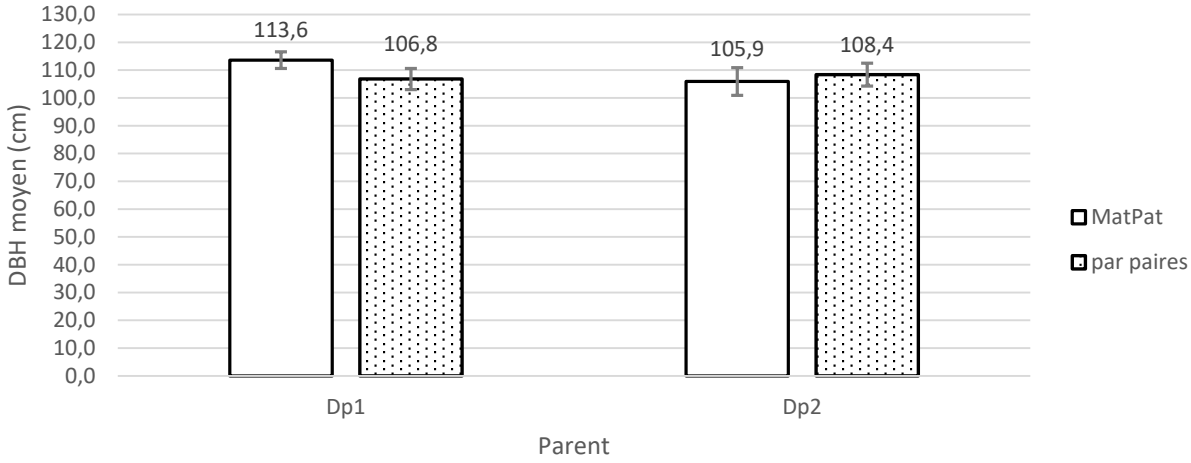
Les analyses de parenté faites par CERVUS permettent de calculer les DBH moyens des parents (fig. 5). On remarque que les différences entre analyses MatPat et par paires sont minimales. Une analyse ANOVA à deux facteurs montre qu'il n'existe aucune différence significative entre les différents groupes ou au sein de ceux-ci.

Le succès reproducteur par classe de DBH a aussi été calculé. A été calculé en premier le succès reproducteur femelle à partir des résultats des analyses de maternité (fig. 6A) et ensuite le succès reproducteur mâle d'après les résultats des analyses de paternité qui s'ensuivirent (fig. 6B). Notons que le nombre d'individus appartenant à chaque classe est le même pour les deux analyses, à la différence que l'analyse de maternité ne prend en compte les DBH inférieurs à 40 cm. Les résultats reprennent les juvéniles assignés à 19 et 20 individus adultes respectivement. Les valeurs de juvéniles assignés sont largement plus grandes dans le cas du succès reproducteur femelle et cela peut être expliqué par les 103 juvéniles classés dans des familles (*cf.* 3.2.3). On remarque que le succès reproducteur mâle et femelle est le plus important pour les classes de 90 à 140 cm de diamètre. La classe 120-130 cm DBH ne présente que 2 individus, ce qui cause probablement l'absence totale d'évènements de reproduction attribués à celle-ci, quel que soit le succès reproducteur étudié.
Parmi les 230 adultes présents sur la parcelle, on remarque que les analyses de parenté permettent d'en trouver moins de 10 % en tant que parents.

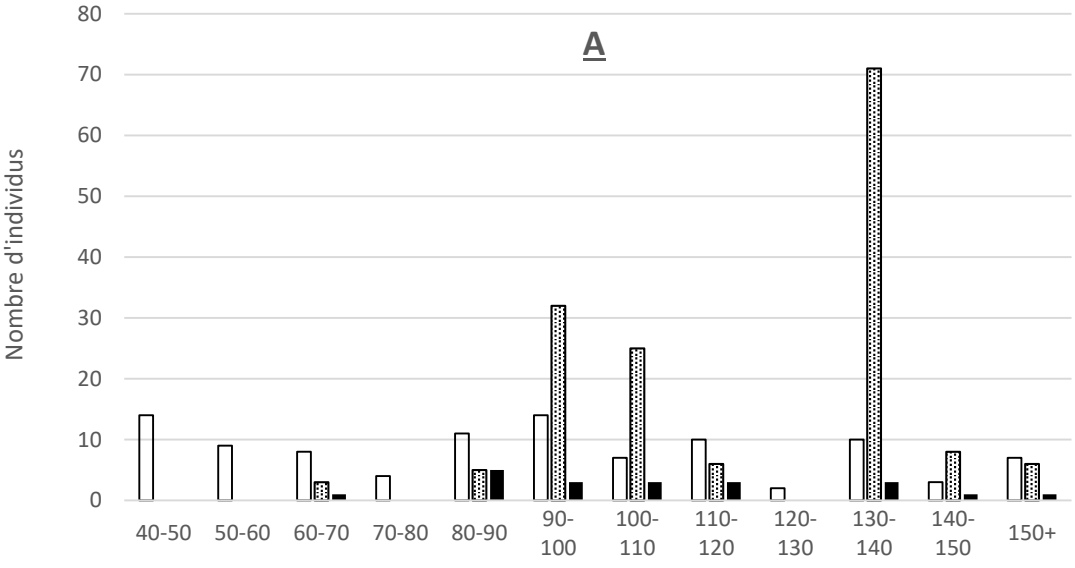
La figure 7 reprend le succès reproducteur indépendamment du sexe considéré étant donné qu'elle se base sur les résultats des analyses par paires. Elle comprend également plus de juvéniles que chacune des figures précédentes, puisque les analyses ont permis d'attribuer un couple parental à un plus grand nombre de juvéniles d'une part, et qu'il n'y a pas de division du succès reproducteur entre mâle et femelle d'autre part. La figure reprend aussi la courbe de la contribution moyenne par classe de diamètre obtenue grâce au gradient de sélection estimé par le logiciel NM+. La contribution est proportionnelle au succès reproducteur attendu pour une classe ou un individu. Cette courbe s'ajuste assez bien aux données de CERVUS.

Dans le cas des analyses de maternité, trois individus sont responsables pour 98 contributions (63%) : 60 assignations pour la classe 130-140, 12 pour la classe 100-110 et 28 pour la classe 90-110. Dans le cas des analyses de paternité, un individu est responsable pour 16 (30% du total) contributions pour la classe 130-140. Pour ce qui est des analyses par paires, quatre individus sont resposables pour 86 (49%) assignations. 49 à la classe 130-140 ; 14 à la classe 120-130 ; 10 à classe 100-110 et 13 à la classe 90-100.

Figure 5 – Diamètres moyens et erreur standard des parents obtenus par CERVUS en fonction de l'analyse. Dp1 = La mère dans le cas de l'analyse MatPat, le parent le plus proche dans le cas de l'analyse par paires ; Dp2 = le père ou le parent le plus distant.

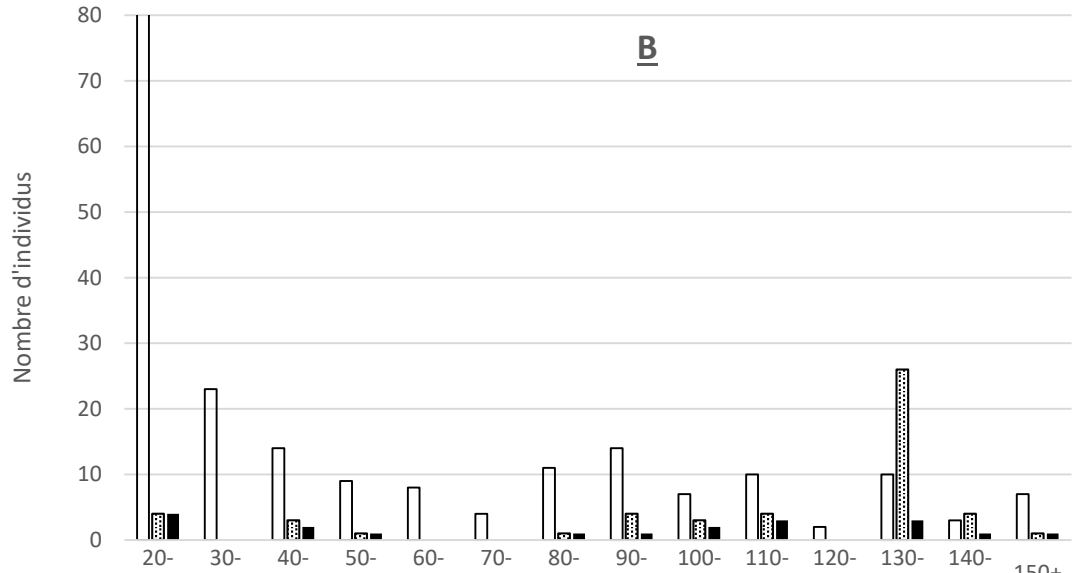


Figures 6 A et B [page 36] – Succès reproducteur femelle par classe de DBH. Sont repris dans ces graphiques : le nombre d'individus adultes appartenant à une classe donnée ; le nombre de juvéniles dont la maternité (A) ou paternité (B) a été assignée à un individu appartenant à cette classe et ; le nombre d'adultes dans cette classe à qui a été attribué au moins un évènement de maternité (A) ou de paternité (B). L'axe des ordonnées de la figure B est limité à 80 afin que cette figure soit comparable à la figure A. La valeur du nombre d'individus adultes pour la classe 20-30 cm est de 108.



	40-50	50-60	60-70	70-80	80-90	90-100	100-110	110-120	120-130	130-140	140-150	150+
# Individus	14	9	8	4	11	14	7	10	2	10	3	7
# Juvéniles assignés	0	0	3	0	5	32	25	6	0	71	8	6
# Mères	0	0	1	0	5	3	3	3	0	3	1	1

Classes de DBH



	20-30	30-40	40-50	50-60	60-70	70-80	80-90	90-100	100-110	110-120	120-130	130-140	140-150	150+
# Individus	108	23	14	9	8	4	11	14	7	10	2	10	3	7
# Juvéniles assignés	4	0	3	1	0	0	1	4	3	4	0	26	4	1
# Pères	4	0	2	1	0	0	1	1	2	3	0	3	1	1

Classe de DBH

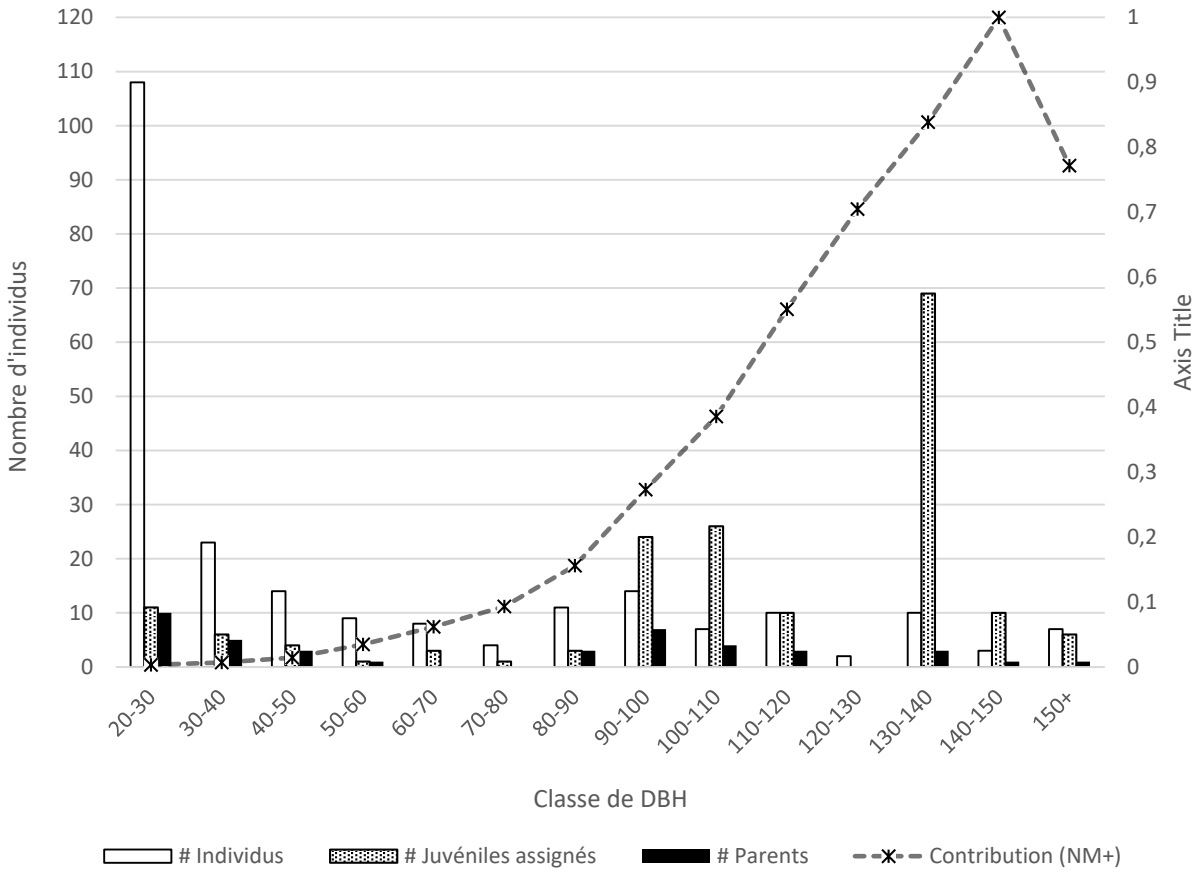


Figure 7 – Succès reproducteur par classe de DBH avec courbe du poids de ces classes superposée. Les valeurs de succès reproducteur sont obtenues grâce à CERVUS d'après une analyse de parenté par paires. Sont repris : le nombre d'individus adultes appartenant à une classe donnée ; le nombre de juvéniles dont la parenté a été assignée à un individu appartenant à cette classe et ; le nombre d'adultes dans cette classe à qui a été attribué au moins un évènement de parenté. La courbe de la contribution a été calculée d'après les paramètres relatifs aux gradients de sélection estimés par NM+. La contribution est proportionnelle au succès reproducteur attendu d'après les estimations du logiciel NM+.

4.6 STRUCTURE GÉNÉTIQUE SPATIALE

La structure génétique spatiale de la population adulte calculée par le logiciel SPAGeDi est présentée à la figure 8. La parenté entre individus adultes y est calculée par intervalle de distance et mise en graphe. L'enveloppe de distribution à 95 % pour l'hypothèse nulle y est présente. On peut observer une structure génétique spatiale forte jusqu'à 650 m de distance, la parenté entre individus chutant assez drastiquement. On peut néanmoins observer que la parenté continue de diminuer d'une pente douce jusqu'à 2750 m. Les données situées entre les pointillés ne sont pas significativement différentes de zéro.

Le tableau 7 présente les différentes valeurs de sigma calculées par SPAGeDi en fonction de la densité en adultes. La densité originale de 38.3 adultes/km² est divisée par 2, 4 et enfin 10 afin d'aboutir à ces valeurs. Le programme n'a pu faire converger la valeur de sigma pour des densités plus faibles que N/4. La valeur de sigma et de la *Neighborhood size* (Nbh) est de 1261 m et 190 individus respectivement.

La structure génétique au sein de la classe des juvéniles (fig. 9A) est très forte et diminue très fortement du premier au deuxième intervalle de distance. Ceci est probablement dû en partie à la présence de familles de juvéniles dans le jeu de données. Bien que la structure soit forte la valeur des coefficients de parenté ne dépasse pas 0,06. On observe une courbe similaire pour la parenté entre juvéniles et grands adultes, avec une parenté élevée à courte distance mais chutant rapidement avec la distance. La courbe de parenté entre juvéniles et petits adultes ne présente presque pas de structure spatiale et est toujours fort basse, n'atteignant pas 0,01.

Pour ce qui est de la parenté entre adultes (fig. 9B), on observe que la structure et la parenté est plus importante entre petits adultes qu'elle ne l'est entre grands adultes. En effet, le coefficient de parenté pour le premier intervalle de distance est plus de quatre fois plus élevé entre petits adultes qu'entre grands. La parenté entre grands et petits adultes vient se placer entre ces deux premières courbes mais dépasse à peine le seuil de parenté de 0,02.

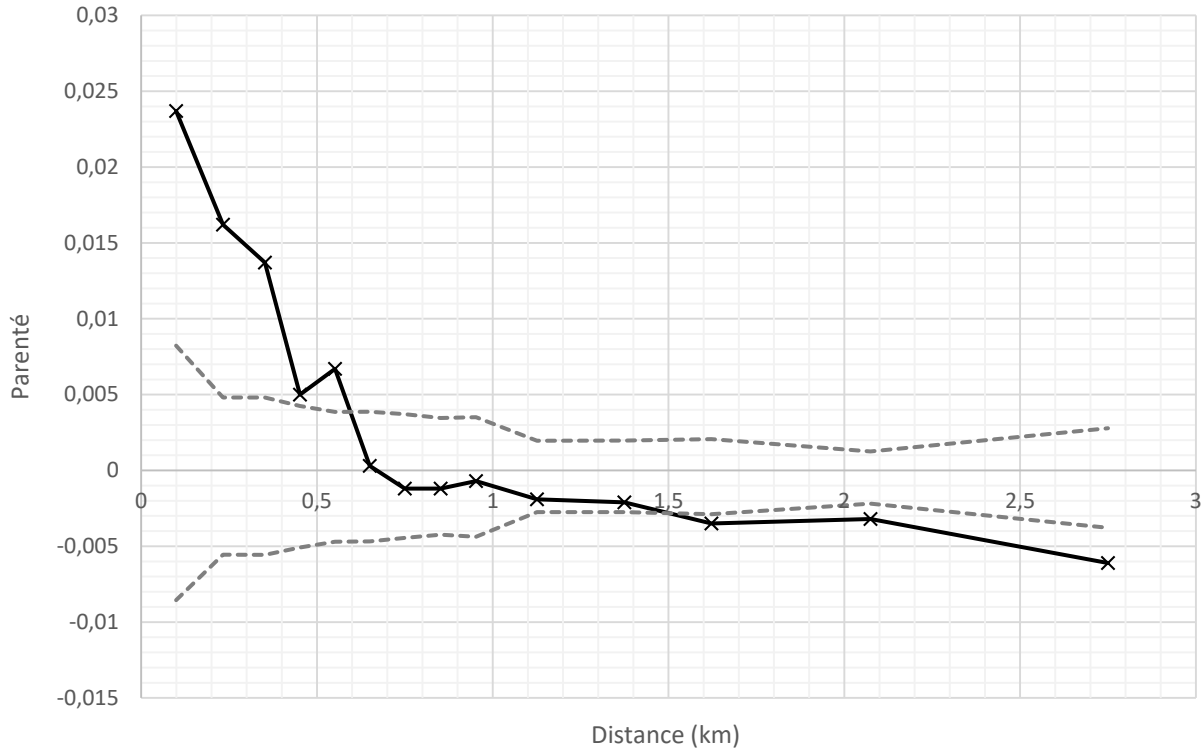


Figure 8 – Coefficient de parenté entre individus adultes et par catégorie de distance. L'enveloppe de distribution à 95 % pour l'hypothèse nulle est représentée en pointillés. L'axe des abscisses représente la distance moyenne au sein de chaque intervalle de distance.

Tableau 7 – Paramètre sigma calculé au sein de la population adulte et pour différentes densités spatiales de celle-ci. Densité efficace = différentes densités calculées à partir de la densité non-efficace réelle de 38.3 adultes/km² ; σ = moyenne quadratique des distances de dispersion des gènes ; Nbh = *Neighborhood size*, en nombre d'individus.

Densité efficace	σ (km)	Nbh (m)
D = 38,3	0,77451	288,702
D/2 = 19	0,989423	233,73
D/4 = 9,5	1,26102	189,83
D/10 = 3,8	Non-convergeant	Non-convergeant

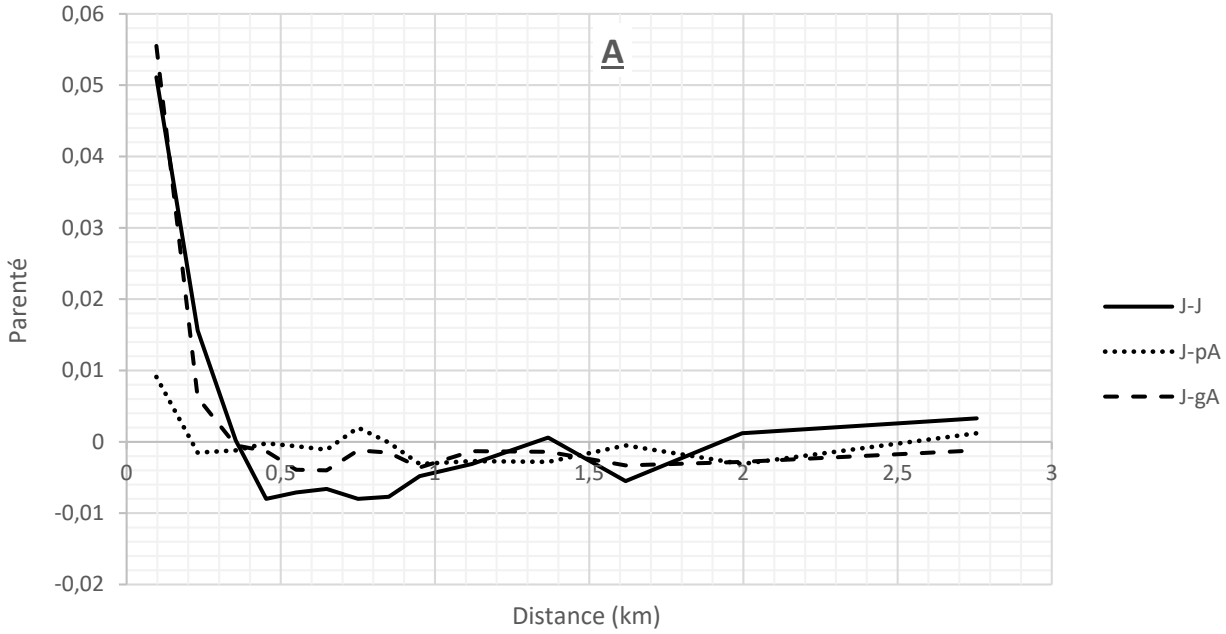


Figure 9A – Coefficients de parenté calculés entre catégories (juvéniles) et par catégorie de distance. La parenté a été calculée au sein de la catégorie juvéniles et entre celle-ci et les deux catégories adultes. J = Juvéniles ; pA = Adultes de moins de 40 cm de diamètre ; gA = Adultes de plus de 40 cm de diamètre.

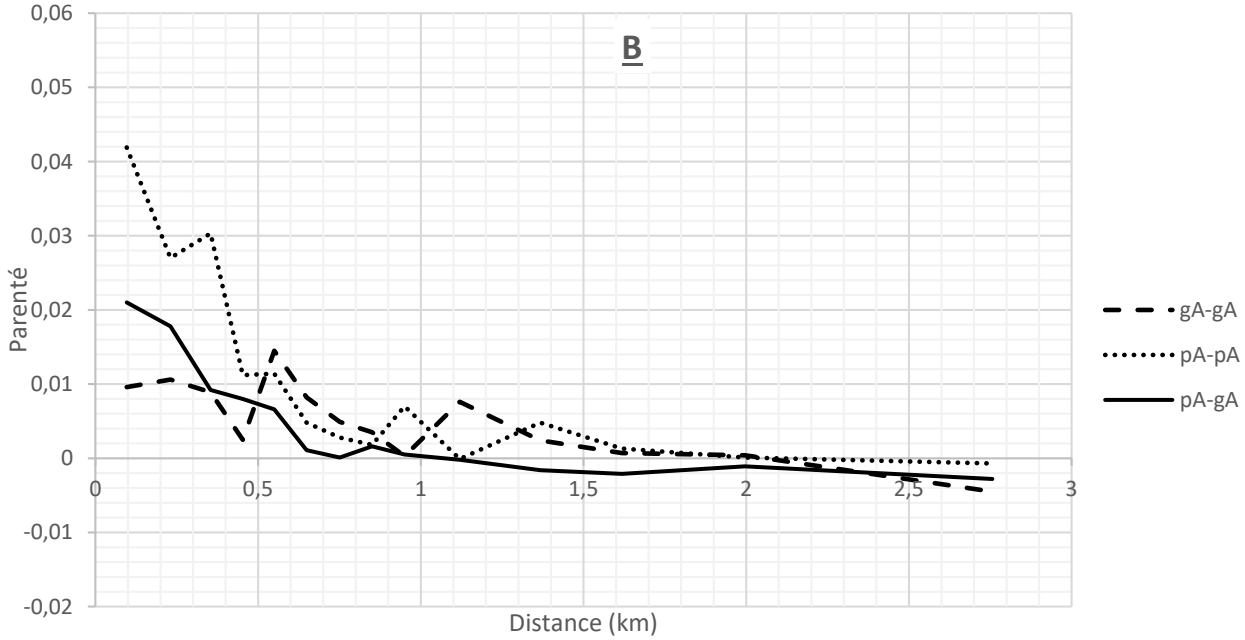


Figure 9B – Coefficients de parenté calculés entre catégories (adultes) et par catégorie de distance. La parenté a été calculée au sein des catégories d'adultes et entre celle-ci. pA = Adultes de moins de 40 cm de diamètre ; gA = Adultes de plus de 40 cm de diamètre.

5 DISCUSSION ET CONCLUSION

5.1 LE SYSTÈME DE REPRODUCTION

Entandrophragma candollei est un arbre qui se reproduit principalement par allofécondation. Son taux d'autofécondation estimé varie entre 0,033 et 0,173 en fonction du type d'analyse et/ou de modèle utilisé. Bien que cet écart soit élevé, il peut être expliqué par la manière dont les données sont utilisées par les différents logiciels et par les modèles théoriques qu'ils émulent. Les résultats issus du taux de consanguinité corrigé estimé par INEst ne sont pas utilisables puisque les conditions d'application de la formule présentée par Hartl et Clark (2007) ne sont pas respectées. On sait, par ailleurs, que les analyses de parenté CERVUS auront toujours tendance à surestimer le taux d'autofécondation. Ceci est dû au fait qu'il est plus probable de retrouver le couple parental en cas d'autofécondation qu'en cas d'allofécondation, puisqu'il suffit d'échantillonner un seul individu pour trouver un couple remplissant les conditions de confiance et de distance. Il en est déduit qu'*E. candollei* présente des mécanismes d'autoincompatibilité lui permettant de limiter les évènements d'autofécondation à une valeur inférieure à 10 %.

Un tel système de reproduction peut s'avérer un grand avantage pour l'espèce. En effet, l'allofécondation avec des individus peu apparentés est généralement garante d'une progéniture viable et d'un fitness élevé, ceci parce qu'elle évite une descendance trop consanguine et les effets d'une éventuelle dépression de consanguinité. La possibilité d'autofécondation est, néanmoins, elle aussi un avantage. La possibilité de se reproduire, même quand aucun partenaire n'est disponible, est une adaptation utile qui peut pourvoir un avantage reproducteur au risque d'une descendance plus consanguine (Murawski et Hamrick, 1991)

5.2 DISTANCE DE DISPERSION DES GÈNES

Les différentes méthodes à la source des estimations des distances de dispersion présentent des résultats différents mais compatibles. Les résultats basés sur les analyses CERVUS s'appuient uniquement sur des évènements de reproduction réalisés et uniquement à l'échelle de la parcelle inventoriée. Ceci constitue un biais, puisqu'il est virtuellement impossible de retrouver des évènements de dispersion à longue distance au sein d'une parcelle de 2×3 km². Ce biais est d'autant plus important que la proportion de gènes immigrants à la parcelle est élevée (tableau 5 ; respectivement de 50 et 67 % d'immigration pour les graines et le pollen). Ceci cause une sous-estimation des distances de dispersion. Le logiciel NM+, quant à lui, ajuste un modèle de dispersion théorique allant au-delà de la parcelle inventoriée et qui est capable de mieux prendre en compte les évènements de dispersion à longue distance, notamment par l'estimation du taux d'immigration (repris ci-dessus). Ce taux est élevé dans les résultats, ce qui justifie la sous-estimation des distances de dispersion par CERVUS. Les distances de dispersion moyennes sont élevées d'après les estimations NM+ (tableau 5 ; 338 et 793 m pour les graines et le pollen respectivement) et ce fait est renforcé par un coefficient b toujours plus bas que 1 (bs = 0,23 ; bp = 0,71), ce qui indique que des évènements de dispersion à longue distance sont d'une plus grande probabilité que dans le cas d'une distribution normale.

Le logiciel SPAGeDi montre une distance de dispersion des gènes (c'est-à-dire, une estimation conjointe pour le pollen et les graines) dont la moyenne quadratique est de 1260 m (tableau 7), ce qui semble être en accord avec le graphique de la figure 3B qui montre une part importante du pollen dispersé au-delà de un kilomètre. Les courbes de distribution des probabilités de dispersion (fig. 3A et 3B) sont aussi en accord avec la structure génétique de la population (fig. 8). Toutes deux sont compatibles avec une dispersion à longue distance significative. Les premières de par leur forme leptokurtique et la dernière de par une parenté qui diminue lentement avec la distance jusqu'à une distance de 2750 m. Il faut ici noter que la contribution à la dispersion des gènes de la part des graines est deux fois plus importante dans ce type d'analyse conjointe. En effet, celles-ci transportent une quantité de gènes deux fois plus importante (*2n*) que le pollen (*n*).

La distance de dispersion des gènes chez *E. candollei* se fait donc à grande distance. Le cas particulier des graines peut être expliqué par le type de fruit les contenant. Les akènes ailés

sont contenus dans une capsule avant leur dispersion. Les ailes de ces akènes ne sont donc pas directement exposées aux vents. Les graines peuvent se détacher soit sans action mécanique de la part du vent, se dispersant alors à courte distance à cause du manque de vent, soit lors de bourrasques d'intensité importante, forçant alors les graines hors de la capsule et les emportant à grande distance.

D'après les études de Medjibe et Hall (2002) en RCA, les graines du genre *Entandrophragma* peuvent parcourir des distances de plusieurs dizaines de mètres après avoir été dispersées. Ils montrent néanmoins que la plupart de graines qui ont fait l'objet de leur étude ont fini leur course au pied de l'arbre mère, ce qui semble être le cas de cette étude. De plus, celles-ci n'ont pas été dispersées dans le sens des vents dominants. Cette étude est en contradiction avec les travaux de Grogan (2001) et de Petrucci et Tandeau de Marsac (1994) sur d'autres arbres allochorés au Brésil (*Swietenia macrophylla*) et en RCA (*Entandrophragma cylindricum*) respectivement. Ces études avaient montré une dispersion dans le sens des vents dominants. Les facteurs influençant la dispersion des graines sont nombreux ; coïncidence avec la chute des feuilles de plusieurs autres espèces, force des vents dominants au moment de la dispersion, fréquence des vents de tempête etc. Une dispersion dans la direction des vents de tempête, aux directions différentes des vents dominants, a été mise en évidence pour *Picea glauc*a au Canada (Stewart *et al.* 1998, cité par Hall *et al.* 2003) et pourrait jouer un rôle important dans les forêts tropicales où le vent en dessous du niveau de la canopée prend des patterns complexes tout en perdant en force.

Des données concernant la distribution des vents sur une période de temps suffisante n'étant malheureusement pas disponibles, les résultats concernant l'anisotropie et la direction moyenne de la dispersion des graines (tableau 5 et figure 4) ne sont pas testables. Néanmoins, la direction de dispersion majoritaire dans le cadrant sud-est (de 90 à 180° à partir du nord, dans le sens anti-horlogique utilisé par NM+) d'après les résultats CERVUS ne comprend pas la valeur donnée par NM+ de 205°.

Il aurait été intéressant de tester si les graines sont dispersées dans le sens des vents dominants ou dans le sens des vents de tempête, afin de pouvoir comparer ce cas-ci avec les études susmentionnées. Des relevés des vents sur plusieurs mois, couvrant la période de fructification d'E*. candollei*, seraient alors nécessaires.

Pour ce qui est des distances de dispersion estimées sur d'autres arbres des forêts tropicales humides d'Afrique centrale, l'étude sur *Entandrophragma cylindricum* de Monthé *et al.* (2017), rapporte des distances de dispersion moyennes estimées par NM+ de 446 et 506 m pour les graines et le pollen respectivement. On remarque ici des distances moyennes de dispersion du pollen plus basses que pour *E. candollei*, bien que les graines d'E*. cylindricum* dispersent à plus longue distance. Les résultats issus des analyses CERVUS estiment des distances de dispersion plus élevées, de 584 et 715 m respectivement, contrairement aux résultats obtenus dans le présent travail. On note néanmoins que, de manière similaire aux résultats obtenus ici, les distances de dispersion des graines sont moindres que celles du pollen. Le sigma estimé pour *E. cylindricum* est de 1500 m. Cette valeur a été obtenue à une densité d'un quart de la densité observée (88 individus/km² ; d/4 = 22 ind/km²), la valeur la plus basse pour laquelle les algorithmes de SPAGeDi ont pu converger. De manière générale, ces résultats concordent assez bien avec cette étude sur une espèce du même genre.

Les travaux de Duminil *et al.* (2016) sur *Erythrophleum suaveolens* au Cameroun, Gabon et RDC montrent dispersion des gènes à plus courte distance, le sigma étant de 483 m à une densité efficace de 41 ind/km². Une autre étude (Duminil *et al.,* 2016), sur *Baillonella toxisperma* au Gabon, a montré une distance de dispersion du pollen estimée à partir de la structure génétique de 777 m. La distance moyenne de dispersion du pollen estimée d'après le *Neighborhood model* est ici de 311 m.

Ces résultats, obtenus pour des espèces commerciales largement exploitées dans les forêts du bassin du Congo viennent appuyer les résultats ici obtenus. Bien que les modes de dispersion de toutes ces espèces ne soient pas comparables (à l'exception notable des deux espèces d'*Entandrophragma*), les résultats obtenus semblent l'être.

5.3 VARIATION DU SUCCÈS REPRODUCTEUR

Avec un diamètre moyen des parents de plus de 100 cm (CERVUS, NM+), *E. candollei* est un arbre qui semble atteindre son optimum de succès reproducteur à un âge tardif (fig. 5 et 7). Les graphes reprenant le succès reproducteur par classe de DBH donnent accès à plus de détails (fig. 6A et 6B). S'il peut déjà disperser du pollen à un diamètre de 20-30 cm, il semble atteindre le pic de production vers 120-140 cm d'après CERVUS (fig. 6B) et 140-150 cm d'après

NM+ (fig. 7). Au niveau de la production de graines, celle-ci atteint un maximum entre 90 et 140 cm de diamètre (fig. 6A), toutes les classes recouvertes par cet intervalle participant à la reproduction de manière significative. Il faut néanmoins prendre en compte le fait que ces observations sont basées sur des résultats obtenus à partir de 19 et 20 individus adultes respectivement. L'influence que peut avoir un seul individu sur le succès reproducteur est donc majeure et les résultats pâtissent de ce type d'effet. En effet, et comme expliqué dans les résultats, une grande part des assignations de parenté sont assignées à un ou quelques individus.

On en retient donc que, si *E. candollei* peut participer à l'effort de reproduction de l'espère à partir d'un diamètre de 20 cm, ce n'est que bien plus tard, à des diamètres aux alentours d'un mètre, qu'ils sont à leur optimum de succès reproducteur.

5.4 STRUCTURE GÉNÉTIQUE ET STRUCTURE GÉNÉTIQUE SPATIALE

Les coefficients de consanguinité F, corrigés par INEst, sont assez bas pour l'ensemble de la population (tableau 2). Néanmoins, de grandes différences entre les cohortes existent (tableau 3). Si l'erreur standard sur le F_{IS} corrigé de la population adulte (0,056 ± 0,080) ne permet pas d'avoir d'estimation précise, on observe néanmoins une grande différence entre les populations de juvéniles de plus et de moins de 5 cm de diamètre. En effet, les plus jeunes juvéniles montrent un taux extrêmement bas de F_{IS} corrigé (0,100 ± 0,22) par rapport à ceux plus âgés (0,0014 ± 0,004). Quand on se penche sur les valeurs de H_o et F, on remarque le même pattern.

Il est attendu que les juvéniles présentent une consanguinité plus élevée que leur parents. Ce n'est qu'avec le temps et les effets combinés de la dépression de consanguinité et de la sélection naturelle que les individus les plus consanguins seront éliminés du pool génique, et que leur valeur de F_{IS} se rapprochera de ce qui est aujourd'hui observé chez les adultes. La valeur extrêmement basse du F corrigé ne semble s'expliquer autrement que par une estimation erronée de la fréquence d'allèles nuls ou de leur impact sur cette variable. Cependant, les valeurs de H_o et F brutes ne permettent pas non plus de conclure dans le sens de l'hypothèse décrite ci-dessus, les différences de valeurs entre cohortes étant trop faibles. Il se pourrait donc qu'*E. candollei* ne présente pas de fardeau génétique, ou en tout cas pas

dans le cas de cette population. Ce semble être aussi le cas de cas pour *Erythrophleum suaveolens* (Duminil *et al.* 2016), qui lui non plus ne présente pas de structure génétique forte entre cohortes. *E. cylindricum* ne présente pas non plus de différences entre cohortes, si ce n'est une valeur significativement positive du F_{IS} après correction pour la présence d'allèles nuls, qui pourrait montrer une dépression de consanguinité, les autres cohortes présentant des valeurs non significativement différentes de zéro (Monthé *et al.* 2017). L'étude effectuée sur *Baillonella toxisperma* montre cependant une dépression de consanguinité importante, peut-être corrélée au taux de consanguinité et d'autofécondation élevés de la population.

L'importante parenté à courte distance entre juvéniles et entre juvéniles et adultes de plus de 40 cm de diamètre (figure 9A) corrèle tout à fait avec les courbes de dispersion estimées pour l'espèce par NM+ (fig. 3A et 3B). Le contraste avec la parenté plus basse existante entre juvéniles et adultes de moins de 40 cm semble montrer que ces derniers contribuent en moindre mesure à la reproduction d'*E. candollei*, ce qui est en accord avec les résultats du succès reproducteur estimés par CERVUS et NM+.

La structure spatiale et le taux de parenté entre adultes de moins de 40 cm et les deux catégories d'adultes (fig. 9B) sont relativement plus importants qu'au sein des adultes de plus de 40 cm. Cependant, toutes les courbes montrent une parenté à longue distance plus élevée que dans le cas des juvéniles. Il en est déduit que les jeunes adultes présentent des caractéristiques relatives à la structure génétique spatiale similaires à ceux des juvéniles mais de manière moins marquée. En effet, leur taux de parenté intra-catégorie est élevé, mais moins que chez les juvéniles et leur apparentement avec les adultes de plus de 40 cm est relativement important. Aussi, on observe que la structure génétique spatiale au sein des adultes plus âgés semble s'être effacée, fait probablement dû aux effets de l'évolution sur une période de temps suffisante, tels la dérive génique, la sélection naturelle ou encore la migration.

La structure des adultes (fig. 8) reprend les caractéristiques de chaque catégorie d'adultes décrites ci-dessus. Elle est cependant plus structurée : la parenté diminue d'abord fortement jusqu'à une distance de 750 m pour ensuite diminuer de manière bien moins prononcée jusqu'à 2 km, ce qui continue d'aller dans le sens des résultats NM+, notamment par rapport aux paramètres de distance de dispersion (ds ; dp), de forme de courbe (bs ; bp) et d'immigration (ms ; mp : tableau 5). Tous ces résultats sont donc en accord avec ceux

précédemment discutés et viennent appuyer une courbe de dispersion des gènes à queue large.

5.5 CONSÉQUENCES POUR UNE EXPLOITATION DURABLE

Le DME d'*Entandrophragma candollei* en République du Congo est de 80 cm, bien que les caractéritiques de la parcelle et les résultats de l'inventaire d'aménagement donnent lieu à un DMA de 110 cm pour l'UFA sur laquelle se trouve la population étudiée.

Par manque d'informations sur le taux de prélèvement des arbres dépassant le DMA ou le DME, deux scénarios sont envisagés. Le premier voit tous les arbres dépassant le DME être coupés. Le deuxième conserve 50% des arbres sur pied après exploitation.

Pour le premier scénario, la densité en arbres adultes diminuerait de 38,3 ind/km² à 27,7 ind/km² après l'exploitation et la densité en arbres capables de fructification (de plus de 40 cm) serait de 5,9 ind/km². D'après Frankham (1995, cité par Hardy *et al.*, 2006), la densité efficace pour des populations végétales se situe généralement entre la moitié et un dixième de la population totale. Ceci nous montre donc que la densité efficace en arbres adultes capables de fructification après exploitation de tous les arbres dépassant le DME se situerait entre 2,92 et 0,59 ind/km².

Dans le cas du deuxième scénario, la densité après exploitation est de 33,6 ind/km² et la densité en arbres capables de fructification est de 10,6 ind/km². La densité finale efficace se situe donc entre 5,3 et 1,1 ind/km².

Premièrement, et en s'intéressant uniquement aux diamètres, on observe que le DME est situé à une étape du développement de l'arbre où celui-ci est encore loin d'avoir atteint son maximum reproducteur. Le DMA est plus proche de celui-ci et permet à chaque individu de connaitre plusieurs saisons de reproduction efficaces avant la coupe.

Deuxièmement, pour ce qui concerne les densités efficaces après exploitation : le premier scénario est une voie sans issue pour l'espèce. Les densités post-exploitation sont fort basses et, bien qu'elles permettent quelques évènements reproductifs, les grandes distances laissées entre des individus n'ayant pas encore atteint leur optimum reproductif sont trop grandes que pour assurer une quantité d'évènements suffisants à la régénération de l'espèce.

Le deuxième scénario laisse sur pied entre 2,35 et 0,47 individus de plus de 80 cm au kilomètre carré. Les valeurs de densité efficace, associées à une part importante de reproducteurs dans les classes présentant le plus grand succès reproducteur pourrait suffire à maintenir un taux de reproduction suffisant.

Il ne fait pas de doute que la dispersion des graines chez *E. candollei* permet à l'espèce de recoloniser les espaces desquels elle aurait été éradiquée par l'exploitation. L'enjeu de la durabilité de l'exploitation de cette espèce semble être plutôt porté sur des densités efficaces post-exploitation devant permettre un nombre d'évènements de reproduction suffisants, nombre qui ne peut être déterminé par cette étude. Il semble cependant que le DME ne soit pas assez élevée pour permettre d'atteindre ces densités. Il est donc conseillé de le revoir à la hausse, notamment grâce aux données relatives au succès reproducteur discutées ci-avant. Une utilisation conjointe d'un DMA supérieur à 100 cm et d'un taux de prélèvement inférieur à 50 % devraient permettre une régénération de l'espèce sur le court terme. Des études s'intéressant aux effets sur le long terme de l'exploitation ainsi qu'à l'écologie de cette espèce sont nécessaires, ceci afin de compléter les connaissances relatives à la régénération naturelle de l'espèce et de pouvoir statuer sur la durabilité de l'exploitation.

6 RÉFÉRENCES

Agence Française de Développement (AFD). 2010. Secteur forestier dans les pays du bassin du Congo : 20 ans d'interventions de l'AFD - Rapport final - ANNEXES.

Austerlitz, Frederic et al. 2004. Using genetic markers to estimate the pollen dispersal curve. Molecular Ecology 13(4):937-54.

Bawa, K. S. 1991. Plant-pollinator interactions in tropical rain forests. Ann. Rev. Ecol. Syst. 21(1990):399-422.

Central African Forests Commission (COMIFAC). 2010. Les forêts du bassin du Congo - Etat des Forêts 2010. Douala.

Chybicki, Igor J. et Jaroslaw Burczyk. 2009. Simultaneous estimation of null alleles and inbreeding coefficients. Journal of Heredity 100(1):106-13.

Chybicki, Igor J. et Jaroslaw Burczyk. 2010. NM+: Software implementing parentage-based models for estimating gene dispersal and mating patterns in plants. Molecular Ecology Resources 10(6):1071-75.

Chybicki, Igor J. et Jaroslaw Burczyk. 2012. NM + : the software implementing parentage-based models for estimation of gene dispersal and mating patterns based on half-sib seeds or seedlings data. The user Manual. Bydgoszcz, Poland: Kazimierz Wielki University.

CIRAD. 2011. Dynamique des forêts du bassin du Congo.

Dick, C. W. 2001. Genetic rescue of remnant tropical trees by an alien pollinator. Proceedings. Biological sciences / The Royal Society 268(1483):2391-96. Consulté (http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1088891&tool=pmcentrez&rendertype=abstract).

Duminil, J. et al. 2016. High selfing rate, limited pollen dispersal and inbreeding depression in the emblematic African rain forest tree Baillonella toxisperma – Management implications. Forest Ecology and Management 379:20-29. Consulté (http://dx.doi.org/10.1016/j.foreco.2016.08.003).

Duminil, J. et al. 2016. Relationships between population density, fine-scale genetic structure, mating system and pollen dispersal in a timber tree from African rainforests. Heredity 116(3):295-303. Consulté (http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4806568&tool=pmcentrez&rendertype=abstract).

Ewédjè, E. B. K., A. Ahanchédé, et O. J. Hardy. 2017. Breeding system, gene dispersal and small-scale spatial genetic structure of a threatened food tree species, Pentadesma butyracea (Clusiaceae) in Benin. Conservation Genetics 18(4):1-13.

Gemayel, Rita, Janice Cho, Steven Boeynaems, et Kevin J. Verstrepen. 2012. Beyond junk-variable tandem repeats as facilitators of rapid evolution of regulatory and coding sequences. Genes 3(3):461-80.

Grogan, J. E. 2001. Bigleaf mahogany (Swietenia macrophylla) in southeast Para, Brazil: a life history study with management guidelines for sustained production from natural forests. PhD Dissertation. Yale university.

Hall, J. S., V. Medjibe, G. P. Berlyn, et P. M. S. Ashton. 2002. Seedling growth of three co-occuring Entandrophragma species (Meliaceae) under simulated light environments: implications for forest management in central Africa. Forest Ecology and Management 6117:1-10.

Hall, Jefferson S. 2008. Seed and seedling survival of African mahogany (Entandrophragma spp.) in the Central African Republic: Implications for forest management. Forest Ecology and Management 255(2):292-99.

Hall, Jefferson S., David J. Harris, Vincent Medjibe, et P.Mark S. Ashton. 2003. The effects of selective logging on forest structure and tree species composition in a Central African forest: Implications for management of conservation areas. Forest Ecology and Management 183(1-3):249-64.

Hardy, O. J. et X. Vekemans. 2002. SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. Moleclar Ecology Notes 2:618-20.

Hardy, Olivier J. et al. 2006. Fine-scale genetic structure and gene dispersal inferences in 10 Neotropical tree species. Molecular Ecology 15(2):559-71.

Hartl, Daniel L. et Andrew G. Clark. 2007. Principles of Population Genetics. Sinauer. Sunderland, Massachusetts.

Ishida, Yoichi. 2009. Sewall Wright and Gustave Male ´ cot on Isolation by Distance. Philosophy of Science 76(5):784-96.

IUFRO. 2015. Forests, Trees and Landscapes for Food Security and Nutrition. A Global Assessment Report. World Seri. édité par Vira B., Wildburger C., et M. S. Vienna: UIFRO.

Jansen, Patrick A., Marco D. Visser, S. Joseph Wright, Gemma Rutten, et Helene C. Muller-Landau. 2014. Negative density dependence of seed dispersal and seedling recruitment in a Neotropical palm. Ecology Letters 17(9):1111-20.

Kissinger, Gabrielle, Martin Herold, et Veronique De Sy. 2002. Drivers of deforestation and forest degradation: A Synthesis Report for REDD+ Policymake. Kissinger, G Herold, M De Sy, Veronique 1-46.

Lewis, Simon L., David P. Edwards, et David Galbraith. 2015. Increased human dominance of Tropical Forests. Science 349(6250):827-32.

Loiselle, Bette A., Victoria L. Sork, John Nason, et Catherine Graham. 2008. Spatial Genetic Structure of a Tropical Understory Shrub , Psychotria officinalis (Rubiaceae) Author (s): Bette A . Loiselle , Victoria L . Sork , John Nason , Catherine Graham Published by : Botanical Society of America Stable URL : http://www.jstor. America 82(11):1420-25.

Mariette, S., V. Le Corre, F. Austerlitz, et A. Kremer. 2002. Sampling within the genome for measuring within-population diversity: Trade-offs between markers. Molecular Ecology 11(7):1145-56.

Marshall, T. C., J. Slate, L. E. B. Kruuk, et J. M. Pemberton. 1998. Statistical confidence for likelihood-based paternity inference in natural populations. Molecular Ecology 7:639-55.

Martin, Claude. 1991. The Rainforests of West Africa: Ecology — Threats — Conservation. Springer B.

Medjibe, Vincent et Jefferson S. Hall. 2002. Seed dispersal and its implications for silviculture of African mahogany (Entandrophragma spp.) in undisturbed forest in the Central African Republic. Forest Ecology and Management 170(1-3):249-57.

Meunier, Quentin, Carl Moumbogou, et Jean-Louis Doucet. 2015. Les arbres utiles du Gabon. édité par Q. Meunier, C. Moumbogou, et J.-L. Doucet. Gembloux: Presses Agronomiques de Gembloux.

Monthé, F. K., O. J. Hardy, J. L. Doucet, J. Loo, et J. Duminil. 2017. Extensive see and pollen dispersal and assortative mating in the rain forest tree Entandrophragma cylindricum (Meliaceae) inferred from indirect and direct analyses. Molecular Ecology 38(1):42-49.

Monthe, Franck S., Jérôme Duminil, Félicien Tosso, Jérémy Migliore, et Olivier J. Hardy. 2017. Characterization of Microsatellite Markers in Two Exploited African Trees, Entandrophragma candollei and E. utile (Meliaceae). Applications in Plant Sciences 5(2):1600130. Consulté (http://www.bioone.org/doi/10.3732/apps.1600130).

Murawski, D. a et J. L. Hamrick. 1991. The effect of the density of flowering individuals on the mating systems of nine tropical tree species. Heredity 67(2):167-74.

Neigel, Joseph. 1997. A Comparison of Alternative Strategies for Estimating Gene Flow from Genetic Markers. Annual Review of Ecology and Systematics 28:105-28.

Obayashi, Kyoko et al. 2002. Genetic Diversity and Outcrossing Rate between Undisturbed and Selectively Logged Forests of Shorea curtisii (Dipterocarpaceae) Using Microsatellite DNA Analysis. International Journal of Plant Sciences 163(1):151-58.

Oddou-Muratorio, Sylvie et Etienne K. Klein. 2008. Comparing direct vs. indirect estimates of gene flow within a population of a scattered tree species. Molecular Ecology 17(11):2743-54.

Ouborg, N. J., Y. Piquot, et J. M. van Groenendael. 1999a. Population Genetics , Molecular Markers and the Study of Dispersal in Plants. Journal of Ecology 87(4):551-68. Consulté (http://www.jstor.org/stable/2648499).

Ouborg, N. J., Y. Piquot, et J. M. van Groenendael. 1999b. Population Genetics , Molecular Markers and the Study of Dispersal in Plants. Journal of Ecology 87(4):551-68.

Petrucci, Y. et G. Tandeau de Marsac. 1994. Evolution du peuplement adulte et de la régénération acquise après interventions sylvicoles.

Secretariat of the Convention on Biological Diversity (SCBD) et Central African Forests Commission (COMIFAC). 2009. Biodiversity and Forest Management in the Congo Basin. Consulté (https://www.cbd.int/doc/books/2009/B-03188.pdf).

Sork, V. L. et al. 2002. Pollen movement in declining populations of California Valley oak, Quercus lobata: Where have all the fathers gone? Molecular Ecology 11(9):1657-68.

The Prince's Charity - International Sustainability Unit. 2015. Tropical Forests - A Review. Consulté (http://www.pcfisu.org/wp-content/uploads/2015/04/Princes-Charities-International-Sustainability-Unit-Tropical-Forests-A-Review.pdf).

Vieira, Maria Lucia Carneiro, Luciane Santini, Augusto Lima Diniz, et Carla de Freitas Munhoz. 2016. Microsatellite markers: What they mean and why they are so useful. Genetics and Molecular Biology 39(3):312-28.

Ward, M., C. W. Dick, R. Gribel, et A. J. Lowe. 2005. To self, or not to self… A review of outcrossing and pollen-mediated gene flow in neotropical trees. Heredity 95(4):246-54. Consulté (http://www.nature.com/doifinder/10.1038/sj.hdy.6800712).

Whitlock, Michael C. et David E. Mccauley. 1999. Indirect measures of gene flow and migration : F ST 1 /(4Nmɲ1). Heredity 82(November 1998):117-25. Consulté (http://www.ncbi.nlm.nih.gov/pubmed/10098262).

7 ANNEXES

Annexe 1 : informations relatives aux marqueurs génétiques utilisés.

	Mix 1							Mix 2					
Nom de l'amorce	ssr4	ssr13	ssr24	ssr26	ssr32	ssr42	ssr48	ssr9	ssr12	ssr16	ssr29	ssr33	ssr38
Couleur fluorochrome	Bleu	Jaune	Jaune	Vert	Vert	Rouge	Rouge	Bleu	Bleu	Jaune	Vert	Vert	Rouge
Nombre d'allèles	**10**	**20**	**13**	**17**	**8**	**17**	**5**	**20**	**16**	**24**	**29**	**13**	**13**
Allèles (en paires de bases)	193	118	184	129	205	218	170	209	147	153	135	233	190
	195	124	187	131	206	219	178	210	149	155	141	237	192
	206	128	189	135	208	220	180	213	154	160	143	240	194
	208	130	190	137	210	224	186	216	156	163	146	242	196
	211	133	193	139	212	227	188	217	158	165	150	244	198
	215	135	195	142	214	231		219	160	166	152	246	200
	217	136	197	144	218	232		221	162	168	154	248	202
	219	137	199	146	220	235		222	164	170	156	250	204
	221	139	201	148		237		223	166	171	160	252	207
	224	141	203	150		239		225	168	172	162	254	209
		144	204	153		241		227	170	174	164	256	211
		146	206	155		244		229	174	176	166	258	213
		148	209	159		246		231	176	178	168	260	215
		150		161		248		233	181	180	170		
		152		163		250		235	183	182	172		
		155		165		252		237	185	184	174		
		157		167		254		239		186	176		
		159						241		188	178		
		161						243		190	180		
		163						253		192	182		
										194	184		
										196	186		
										198	188		
										200	190		
											191		

											193		
											195		
											201		
											203		

)