

ASSOCIATION DES AMIS DE LA MASSANE

# RÉSERVE NATURELLE DE LA MASSANE

# TRAVAUX 100

**IMPACT DU DÉPÉRISSEMENT ET DE L'ISOLEMENT  
REPRODUCTEUR SUR LA QUALITÉ DE LA RÉGÉNÉRATION DU  
HÊTRE (*FAGUS SYLVATICA*)**

Une étude dans une population naturelle et marginale

---

par Sylvie ODDOU-MURATORIO, Matthieu LINGRAND,  
Joseph GARRIGUE, Jean-André MAGDALOU,  
Christophe HURSON, Elodie MAGNANOU



## REMERCIEMENTS

---

Nous remercions Norbert Turion, Olivier Gilg et Medhi Pringarbe de l'INRA d'Avignon pour leur expertise et leur aide lors des échantillonnages. Florence Lespine, Directrice de la Fédération des Réserves Naturelles Catalanes et Fabienne Savignac sont venues prêter main forte sur le terrain, nous en sommes reconnaissants.

### Directeur de publication

Gilles Boeuf,  
Président de l'Association des  
Amis de la Massane

### Auteurs

Sylvie Oddou-Muratorio,  
Matthieu Lingrand, Joseph  
Garrigue, Jean-André Magdalou,  
Christophe Hurson,  
Elodie Magnanou

### Comité de relecture

Joseph Travé, Marie-Thérèse  
Panouse, Élodie Magnanou

### ISSN

2418-134X

### Mise en page

Diane Sorel, Réserve Naturelle  
Nationale de la Forêt de la  
Massane 2015

### Illustration de couverture

*Plantule de Hêtre* - E. Magnanou

### Référence bibliographique à citer

Oddou-Muratorio, Sylvie et coll.  
« Impact du dépérissement et  
de l'isolement reproducteur sur  
la qualité de la régénération du  
hêtre (*Fagus sylvatica*) », Travaux  
de la Massane, Tome n°100  
(2015).

# IMPACT DU DÉPÉRISSEMENT ET DE L'ISOLEMENT REPRODUCTEUR SUR LA QUALITÉ DE LA RÉGÉNÉRATION DU HÊTRE (*FAGUS SYLVATICA*) Une étude dans une population naturelle et marginale

par Sylvie ODDOU-MURATORIO, Matthieu LINGRAND, Joseph GARRIGUE, Jean-André MAGDALOU, Christophe HURSON, Élodie MAGNANOU

<b>Impact du dépérissement et de l'isolement reproducteur sur la qualité de la régénération du Hêtre .....</b>	<b>5</b>
Contexte général.....	6
Problématique de cette étude .....	10
Intérêt des approches génétiques pour répondre à ces questions .....	12
Démarche adoptée.....	16
<b>Résultats .....</b>	<b>21</b>
Variabilité génétique.....	22
Structure génétique spatiale (SGS) et flux de gènes historiques.....	24
Flux de gènes efficaces et succès reproducteurs réalisés .....	26
Régime de reproduction, noyaux de dispersion primaires et fertilités intrinsèques .....	34
Effets de la phénologie et du dépérissement sur les fertilités .....	40
<b>Discussion .....</b>	<b>45</b>
Régime de reproduction, flux de gènes et taille efficace .....	46
Effet des caractères des arbres adultes leurs fertilités relatives .....	49
Conclusion.....	51
<b>Matériel et Méthodes .....</b>	<b>53</b>
Site d'étude et suivi des arbres adultes .....	54
Échantillonnage du matériel biologique et marquage des individus.....	56
Analyses de diversité et de structure génétique spatiale .....	59
Estimation indirecte des flux de gènes .....	60
Analyse de parenté/maternité .....	61
Estimation de la fonction de dispersion du pollen et des graines et des fécondités individuelles.....	72
Analyse de l'effet des variables écologiques sur les fécondités individuelles.....	76
<b>Annexes .....</b>	<b>79</b>
Liste des Figures, Tableaux, et Appendices.....	80
Références bibliographiques .....	82
Appendices.....	85
Glossaire.....	94







# Impact du dépérissement et de l'isolement reproducteur sur la qualité de la régénération du Hêtre (*FAGUS SYLVATICA*)

---

Les capacités de dispersion des gènes par pollen et par graine, et les variations de fécondité entre individus en réponse aux facteurs climatiques sont des paramètres importants de la réponse des populations aux changements environnementaux. Dans cette étude, nous avons utilisé des approches de génétique des populations pour caractériser le régime de reproduction, les capacités de dispersion du pollen et des graines et les variations de fécondité du Hêtre dans la forêt de la Massane. Pour cela, nous avons cartographié et génotypé des semis établis et des adultes reproducteurs sur deux placettes (chacune couvrant ~0.9 ha) situées dans la réserve naturelle et différant pour leur structure en âge.

Le régime de reproduction est essentiellement allogame, avec un taux d'autofécondation négligeable. La distance de dispersion des graines est de  $\delta_s = 9.6$  m, celle du pollen est de  $\delta_p = 92$  m. Les capacités de dispersion à longue distance du pollen apparaissent cependant non négligeables, en accord avec les taux de migration par pollen élevés dans chacune des placettes ( $m_p \sim 68\%$ ). Ces caractéristiques de dispersion des gènes par pollen et par graines sont cohérentes avec la structure génétique spatiale significative observée (c'est-à-dire l'agrégation spatiale d'individus plus apparentés qu'attendu au hasard jusqu'à des distances de 35-40 m). Les capacités de dispersion ne diffèrent pas entre placettes.

Les contributions des adultes à la reproduction ne sont pas équilibrées, et se traduisent par un rapport de la densité observée de reproducteurs à la densité efficace supérieure à 1 :  $d_{obs}/d_e = 4.4$  en moyenne pour les contributions par voie de graine et  $d_{obs}/d_e = 2.2$  pour les contributions par voie de pollen. Les variations interindividuelles du diamètre sont un facteur important expliquant les variations interindividuelles des fécondités mâles et femelles. Mais nous avons identifié en outre un effet négatif du dépérissement foliaire (mesuré par la présence de branches mortes) sur la fécondité femelle (et pas sur la fécondité mâle). La présence du champignon lignicole *Oudemansiella mucida* n'a en revanche pas d'effet systématique sur la fécondité, mâle ou femelle.

## CONTEXTE GÉNÉRAL

---

**Des impacts majeurs du changement climatique (CC) sont attendus sur le fonctionnement et la dynamique des forêts, et de leurs organismes clés de voûte : les arbres.**

D'importants phénomènes de dépérissement voire de mortalité sont d'ores et déjà observés en réponse à l'augmentation des températures et de l'intensité ou de la durée des sécheresses (Allen *et coll.*, 2010). De tels dépérissements (qui se traduisent par la perte de vitalité progressive et la mortalité d'organes pérennes -branches, rameaux) devraient conduire à la mortalité de l'arbre, si les réserves carbonées ou d'eau deviennent trop limitantes, si des dysfonctionnements hydrauliques deviennent trop importants ou si les mécanismes de défense contre les parasites sont dépassés (la combinaison de ces facteurs accroissant encore le risque de mortalité). Le climat affecte aussi directement les capacités de reproduction des arbres (Morin *et coll.*, 2008), avec des conséquences majeures sur la régénération des peuplements forestiers. En effet, diverses études attestent du rôle majeur de la température ou de l'humidité de l'air sur les différentes étapes de la formation de la graine, depuis le développement des gamétophytes jusqu'à la maturation du fruit (Bykova *et coll.*, 2012). Les étapes post-reproduction de la phase de régénération (de la dispersion de la graine au recrutement d'un nouveau semis) pourraient également être affectées par les conditions climatiques, avec un découplage possible des conditions climatiques optimales pour les semis et pour les adultes (Zhu *et coll.*, 2014).

**Les changements attendus des capacités de reproduction peuvent concerner la quantité totale de graines viables produites, mais aussi leur qualité.**

La qualité des graines peut notamment être appréhendée par leurs performances (germination, survie), et aussi par la variabilité génétique de l'ensemble des graines produites, avec l'hypothèse généralement admise qu'une variabilité génétique élevée garantit un potentiel adaptatif plus grand en réponse à des conditions environnementales changeantes. Ainsi définie, la qualité des graines dépend étroitement des paramètres du régime de reproduction, notamment du taux d'autofécondation\* (qui détermine le coefficient de consanguinité\*), du nombre de parents contribuant à la reproduction (qui détermine l'apparentement entre les graines produites), et de la distance de dispersion du pollen et des graines (qui détermine l'intensité du brassage génétique). Les changements dans le temps de la quantité/qualité de la régénération sont complexes à étudier, en partie parce que beaucoup d'arbres montrent une forte variation interannuelle dans la production de graines (ou masting\*, Kelly & Sork, 2002).



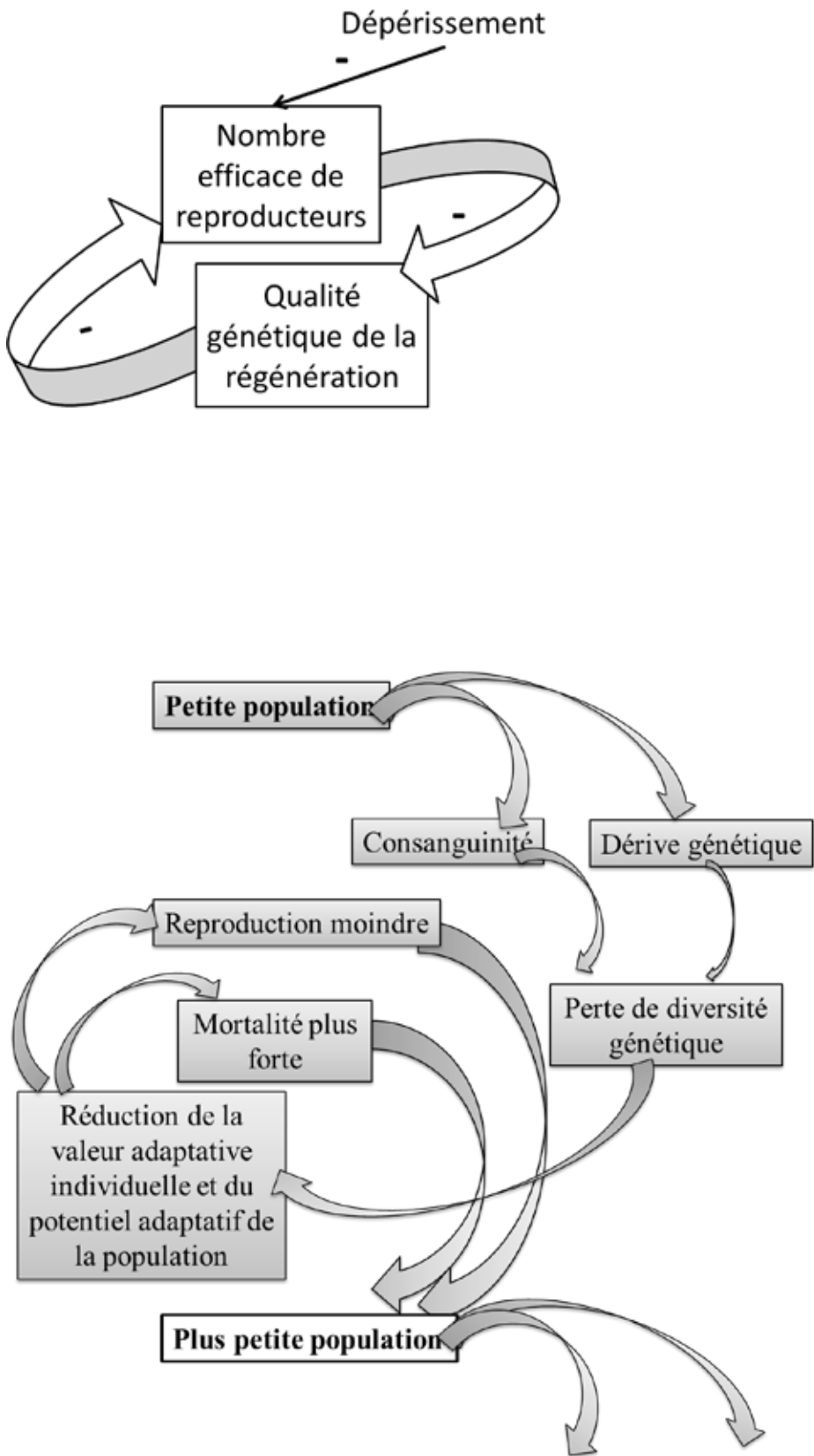
En revanche, différentes études ont porté sur la variabilité spatiale de la quantité/qualité de la régénération entre populations, en particulier pour comprendre comment ces processus limitent l'aire de distribution des espèces. Certaines études mettent en évidence une diminution substantielle des quantités de graines produites dans les populations marginales (Jump & Woodward, 2003; Rasmussen & Kollmann, 2004), en particulier aux marges Nord de l'aire de distribution des espèces tempérées de l'hémisphère nord. Ces diminutions sont généralement associées à des dates de floraison et de maturité plus tardives, une plus forte variabilité interannuelle de production et une proportion plus importante de graines non viables que dans les populations du centre de l'aire. Par ailleurs, une augmentation du taux d'autofécondation et une réduction de nombre de parents contribuant effectivement à la reproduction (appelée taille efficace\*) ont aussi pu être observées dans les populations marginales (Mimura & Aitken, 2007), en lien avec des densités ou des nombres de reproducteurs plus faibles.

Le nombre (ou la densité) d'adultes contribuant à la reproduction apparait donc comme un paramètre clé de la quantité et de la qualité de la régénération. Or, le dépérissement d'un arbre, même s'il n'aboutit pas toujours à sa mortalité, est susceptible d'être associé à une limitation de l'investissement individuel à la reproduction (via la limitation des ressources carbonées et le stress hydrique) et donc de réduire la taille efficace de la population. **L'interaction entre dépérissement et reproduction peut donc se traduire par une boucle de rétroaction négative, telle que la réduction du nombre efficace de reproducteurs induit une diminution de la qualité génétique de la régénération, qui peut elle-même provoquer une baisse de la taille de la population à la génération suivante** (Fig. 1). Mais l'hypothèse inverse est aussi formulée : certains auteurs rapportent une **augmentation possible de la production de graines en réponse au stress** (Lee, 1988). Une telle réponse repose probablement sur une réallocation des ressources consacrées respectivement à la croissance végétative et à la reproduction, deux fonctions bien connues pour être en compétition pour l'utilisation des ressources carbonées. Sous cette hypothèse alternative, l'interaction entre dépérissement et reproduction pourrait freiner les effets négatifs de la mortalité sur la taille efficace de populations. L'importance de telles interactions en réponse aux facteurs climatiques dans des populations initialement de grande taille reste largement méconnue actuellement, en partie parce que leur étude requiert des approches interdisciplinaires, et des suivis de long terme.



Figure 1 : A. Boucle de rétroaction négative attendue entre dépérissement et qualité génétique de la régénération.

La localisation de telles « spirales » ou « vortex d'extinction » est étudiée depuis longtemps en biologie de la conservation (B), typiquement dans des populations de petite taille où le facteur aggravant initial (au lieu du dépérissement lié au climat) est généralement la fragmentation de l'habitat.



## PROBLÉMATIQUE DE CETTE ÉTUDE

---

Le Hêtre commun (*Fagus sylvatica*) est une espèce d'arbre tempéré généralement dominante dans les peuplements forestiers de plaine du centre et l'ouest de l'Europe. Son aire de distribution (Fig. 2) s'étend du nord de l'Espagne au sud de la Norvège, et de la côte Atlantique à la Mer Caspienne (von Wuehlisch, 2008). Cette espèce est décrite comme tolérante à l'ombre, et sensible à la sécheresse en particulier estivale (Michelot *et coll.*, 2012).

La hêtraie du Massif de l'Albera se compose d'une mosaïque de milieux et constitue une île d'habitat ceinturée de végétation méditerranéenne. Cette configuration, à un carrefour biogéographique et climatique, explique l'extrême diversité faunistique et floristique rencontrée sur le site et le classement en Réserve Naturelle Nationale de la forêt de la Massane (336 hectares au sein de ce massif). Les données de pédo-anthracologies recueillies sur le site de la Massane ont permis de retracer la présence du hêtre jusqu'à 4580 ans avant nos jours (Talon *et coll.*, in prep). Le site n'est pas exploité depuis plus d'un siècle et le bois dépérissant ou mort participe au cycle de vie de centaines d'espèces. Le hêtre constitue l'espèce clé de voûte de cet écosystème. Pourtant la position de cette forêt en marge sud de l'aire de distribution du hêtre et à basse altitude impose de fortes contraintes climatiques, accrues par la tendance actuelle à la réduction des précipitations et à l'augmentation des températures.

Un suivi forestier basé sur la cartographie exhaustive du peuplement arboré de 50,000 individus a été initié il y a plus de dix ans. Il concerne 15% de l'habitat forestier de la réserve (soit 28,8 ha), correspondant à la ripisylve et la réserve intégrale (une zone de 10 hectares clôturée pour être protégée du pâturage bovin). Le diamètre, le dépérissement foliaire, la présence de champignons et la mortalité des arbres sont répertoriés. Le dépérissement foliaire (Fig. 3) correspond à la présence de branches mortes (ou couvertes de feuilles mortes) dans la partie haute du houppier (descente de cime), et peut être l'indicateur d'une réponse au stress hydrique, ou plus généralement d'une perte de vitalité de l'arbre. Les champignons lignicoles, généralement abondants sur les arbres morts et dépérissants, sont un élément important de la biodiversité, car ils hébergent une faune importante d'invertébrés. Certaines espèces sont des parasites pouvant compromettre la survie de l'arbre alors que d'autres constituent des pourritures exploitant le bois mort et accélérant le recyclage des éléments minéraux (Dajoz, 1998). Le suivi à la Massane met en évidence un champignon lignicole largement majoritaire, *Oudemansiella mucida* (la mucidule visqueuse, Fig. 4), dont la présence est généralement favorisée par le stress hydrique ou par la perte de vitalité. Cette réserve constitue donc un

cas d'étude particulièrement adapté pour caractériser l'interaction entre dépérissement et qualité génétique de la régénération dans une population de la marge Sud de l'aire de distribution du hêtre.

Figure 2 : Distribution européenne du Hêtre commun (*Fagus sylvatica*)

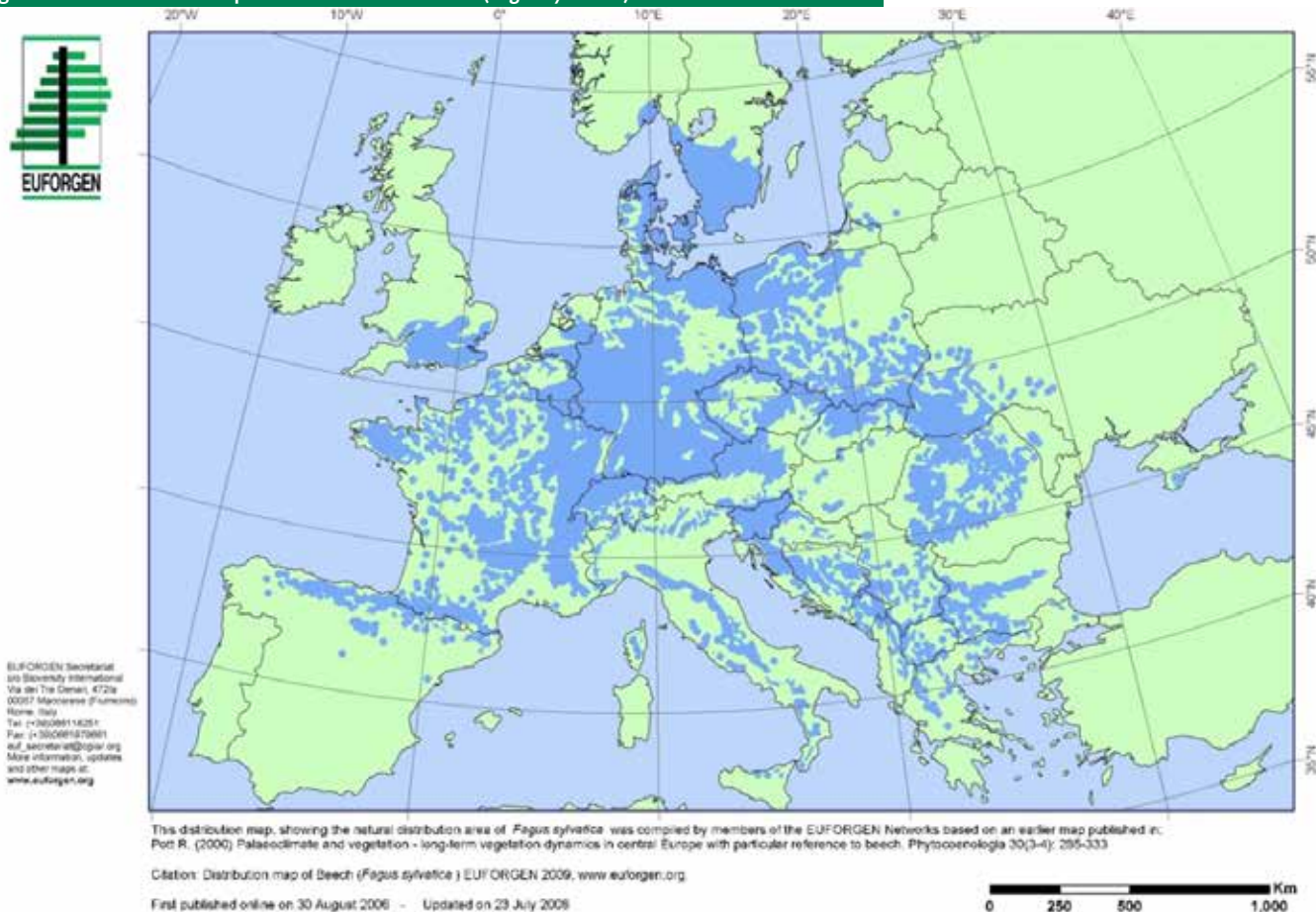


Figure 3 : Dépérissement foliaire chez le Hêtre.

La présence de branches mortes (ou couvertes de feuilles mortes) dans la partie haute du houppier (descente de cime) peut indiquer une réponse au stress hydrique, ou plus généralement d'une perte de vitalité de l'arbre.

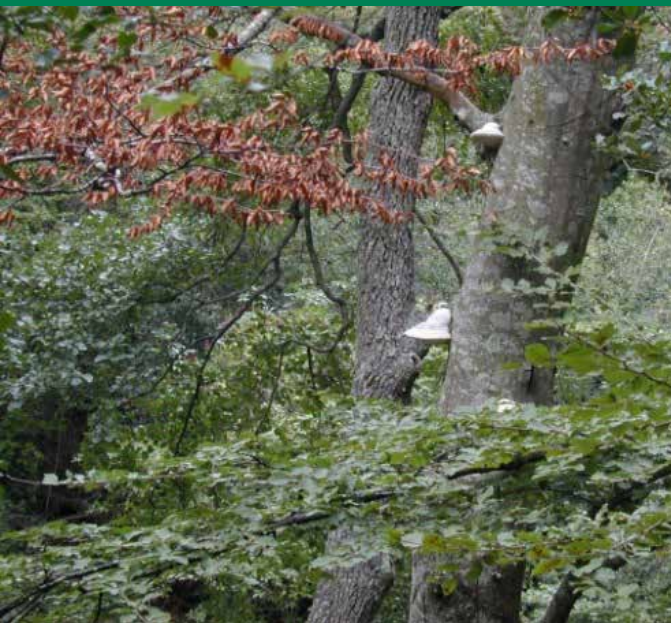


Figure 4 : Le champignon lignicole *Oudemansiella mucida* est l'espèce majoritairement recensée sur les hêtres adultes faisant partie du suivi forestier.

Fin 2012, sur 4 795 hêtres vivant sur pied (et de diamètre supérieur à 1m30), 293 présentent au moins un sporophore de champignon lignicole. Dans 70% des cas (207 individus), il s'agit d'*O. mucida*. Les autres espèces lignicoles recensées sont *Fomes fomentarius*, *Pleurotus sp.*, *Armillaria mellea*, *Pholiota sp.*, ou *Hypholoma*.

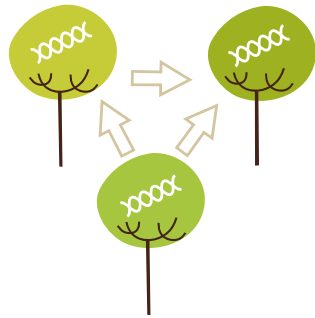




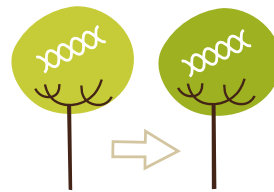


# GÉNOTYPAGE DES PARENTS POTENTIELS ET SEMIS

## MÉTHODES



Méthodes indirectes d'analyse des apparentements génétiques entre individus



Méthodes directes d'assignation de parenté



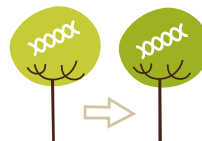
## CONCEPTS



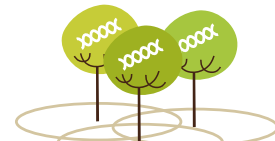
Structure génétique spatiale



Modèle d'isolement par la distance



Assignation catégorique



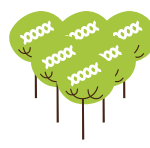
Croisement dans le voisinage spatialement explicite (SEMM)



Flux de gènes historique



Dispersion ou flux de gènes efficace



Succès reproducteur efficace



Dispersion primaire



Fertilité intrinsèque



## PARAMÈTRES ESTIMÉS

# INTÉRÊT DES APPROCHES GÉNÉTIQUES POUR RÉPONDRE À CES QUESTIONS

---

Les approches génétiques basées sur les marqueurs moléculaires\* permettent de caractériser le niveau de diversité génétique neutre, ainsi que d'estimer l'intensité de la **dérive génétique\*** et les **flux de gènes\* historiques** (c'est-à-dire cumulés sur un grand nombre de générations). C'est notamment un des objectifs des méthodes classiques de génétique des populations dites **indirectes** (Sork *et coll.*, 1999), qui permettent d'estimer le nombre efficace de migrants ( $N_e m$ ) entre populations, à partir de la structure génétique de ces populations et sous les hypothèses du modèle en îles\*.

Par ailleurs, le développement depuis les années 1980 de marqueurs moléculaires hypervariables a stimulé l'utilisation et le développement de **méthodes dites directes** qui permettent de caractériser les flux de gènes\* et la dérive génétique à une échelle de temps **contemporaine**, typiquement via la reconstitution de filiation (analyse de parenté) entre deux générations (Broquet & Petit, 2009; Jones *et coll.*, 2010). Les **analyses de parenté catégoriques** consistent à génotyper les reproducteurs (parents potentiels) et des descendants à l'intérieur d'une zone définie, afin de retrouver le parent ou le couple de parents de chaque descendant. Chez les plantes, les filiations ainsi établies permettent, sous certaines hypothèses, de reconstituer ensuite les **flux de pollen et de graines réalisés** dans la population étudiée, ainsi que de caractériser le régime de reproduction (taux d'autofécondation) et les **succès reproducteurs efficaces** ou réalisés des adultes.

Plus récemment, d'autres **méthodes directes** issues de l'analyse de parenté ont été proposées pour estimer simultanément les flux de pollen et de graines, le régime de reproduction, et la variance des fertilités mâles et femelles à partir de données génétiques et spatiales (Burczyk *et coll.*, 2006; Moran & Clark, 2011; Oddou-Muratorio & Klein, 2008). Ces approches, basées sur des modèles de croisement dans un voisinage spatialement explicite (SEMM pour Spatially Explicit Mating Model) tiennent compte explicitement des positions spatiales des adultes et des semis, et permettent ainsi de séparer les contributions aux variations de fertilité entre individus des facteurs écologiques et spatiaux, contrairement aux méthodes plus basiques d'assignations de parenté. Les variations interindividuelles de succès reproducteurs efficaces (estimées par ces dernières) sont en effet dues à la fois à des processus spatiaux (position des arbres adultes et dispersion du pollen et des graines) et à des variations de fertilité entre individus (production intrinsèque de graine ou de pollen). Par ailleurs, parce que les approches de type SEMM modélisent explicitement le processus de dispersion, elles permettent d'estimer le **noyau de dispersion primaire** du pollen et des graines, indépendamment des processus de mortalité post-dispersion (Encadré 1).

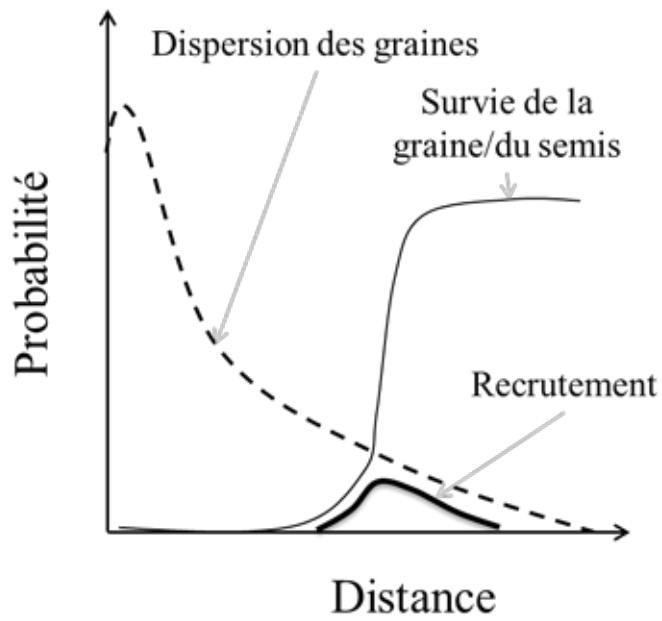
#### Encadré 1 : Notion de noyau de dispersion primaire et efficace

chez les plantes, la dispersion des graines est un processus majeur qui détermine la distribution spatiale des individus de la nouvelle génération, en interaction avec la distribution spatiale des parents. Les processus post-dispersion, comme la germination des graines, ou la mortalité des semis peuvent éventuellement transformer la distribution primaire des graines générée par la dispersion, pour produire une distribution efficace qui sera celle des adultes à la génération suivante. En particulier, les semis subissent en général des processus de mortalité densité-dépendante, pilotés par la densité de semis ou d'arbres adultes dans le voisinage via la compétition ou la prédation. Comme la dispersion des graines se fait préférentiellement à courte distance, les conditions de survie sont d'autant plus défavorables que le semis est proche de l'arbre-mère. Dans certains cas, l'intensité de la mortalité densité-dépendante peut être assez forte pour que la densité de semis survivants soit maximale à une certaine distance de l'arbre-mère (effet de Janzen-Connell, fig. A). Même une mortalité densité-dépendante moins intense aboutit en général à une augmentation de la distance efficace de dispersion (celle qui sépare les semis survivants de leur mère) au cours du recrutement.

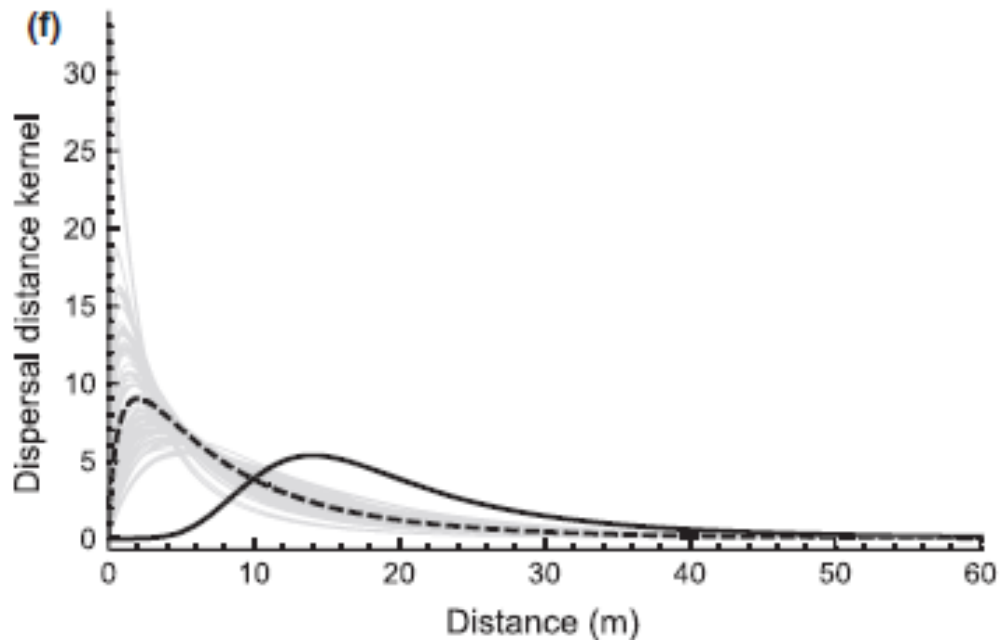
Les noyaux de dispersion\* primaire\* et efficace\* sont deux courbes utilisées pour caractériser la dispersion, et intégrer ces processus. De façon formelle, le noyau de dispersion des graines est la fonction de probabilité qu'un arbre-mère situé à la position (0,0) ait émis une graine retrouvée à la position (x,y). Le noyau de dispersion primaire est défini en ne considérant que le mouvement et la survie des graines pendant la dispersion. Le noyau de dispersion efficace inclut en plus la survie des graines post-dispersion, qui conduit au recrutement. Quand la survie des graines dépend de la distance aux adultes les plus proches, les deux noyaux primaire et efficace sont différents (fig. B). Par simulations, Klein *et coll.* (2013) ont montré que les méthodes de type SEMM estiment toujours le noyau de dispersion primaire, même quand elles sont appliquées sur des semis ayant subi de la mortalité densité-dépendante. Ainsi, elles ne sont pas affectées par la survie de la graine après sa dispersion.



A. Schéma de l'hypothèse de Janzen-Connell (Janzen, 1970).



B. Noyau de dispersion primaire (ligne noire pointillée), efficace (ligne noire pleine) et estimé (ligne grise) avec le SEMM



## DÉMARCHE ADOPTÉE

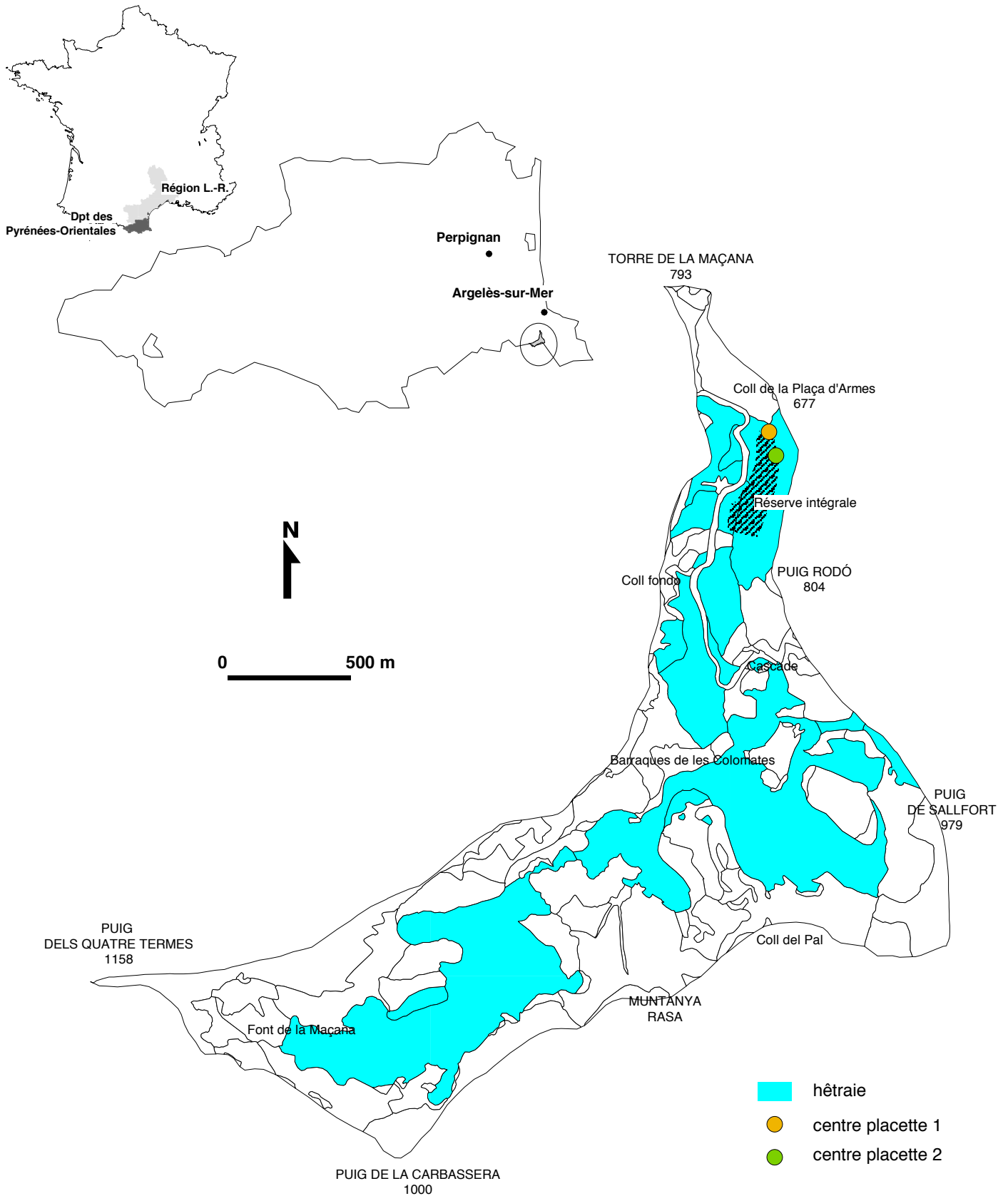
---

Dans cette étude, nous avons délimité deux placettes à l'intérieur de la réserve intégrale de la Massane (Fig. 5). Nous y avons cartographié 1 048 individus appartenant à deux cohortes (semis établis et arbres adultes, Tab 1). Nous avons génotypé ces individus avec 21 marqueurs microsatellites (Append. 1), et nous avons utilisé une combinaison d'approches génétiques pour :

1. Caractériser l'intensité de la dérive et des flux de gènes historiques\* dans la population. Pour cela nous avons utilisé une méthode indirecte basée sur l'analyse de la structure génétique spatiale (SGS) dans le cadre du modèle d'isolement par la distance\*.
2. Caractériser par des approches directes les flux de gènes contemporains\* par pollen et par graines, ainsi que les variations de fertilités individuelles qui affectent la dérive génétique contemporaine. Pour cela, nous avons utilisé d'une part des assignations de parenté catégoriques, pour estimer les flux de gènes et les succès reproducteurs efficaces ; et d'autre part une méthode de type SEMM pour estimer les noyaux de dispersion primaires du pollen et des graines, et les fertilités intrinsèques des adultes.
3. Tester si les fertilités des adultes sont affectées par le dépérissement foliaire et la présence de champignons lignicoles.

**Figure 5a : Vue d'ensemble de la réserve naturelle de la Massane, et des deux placettes étudiées.**

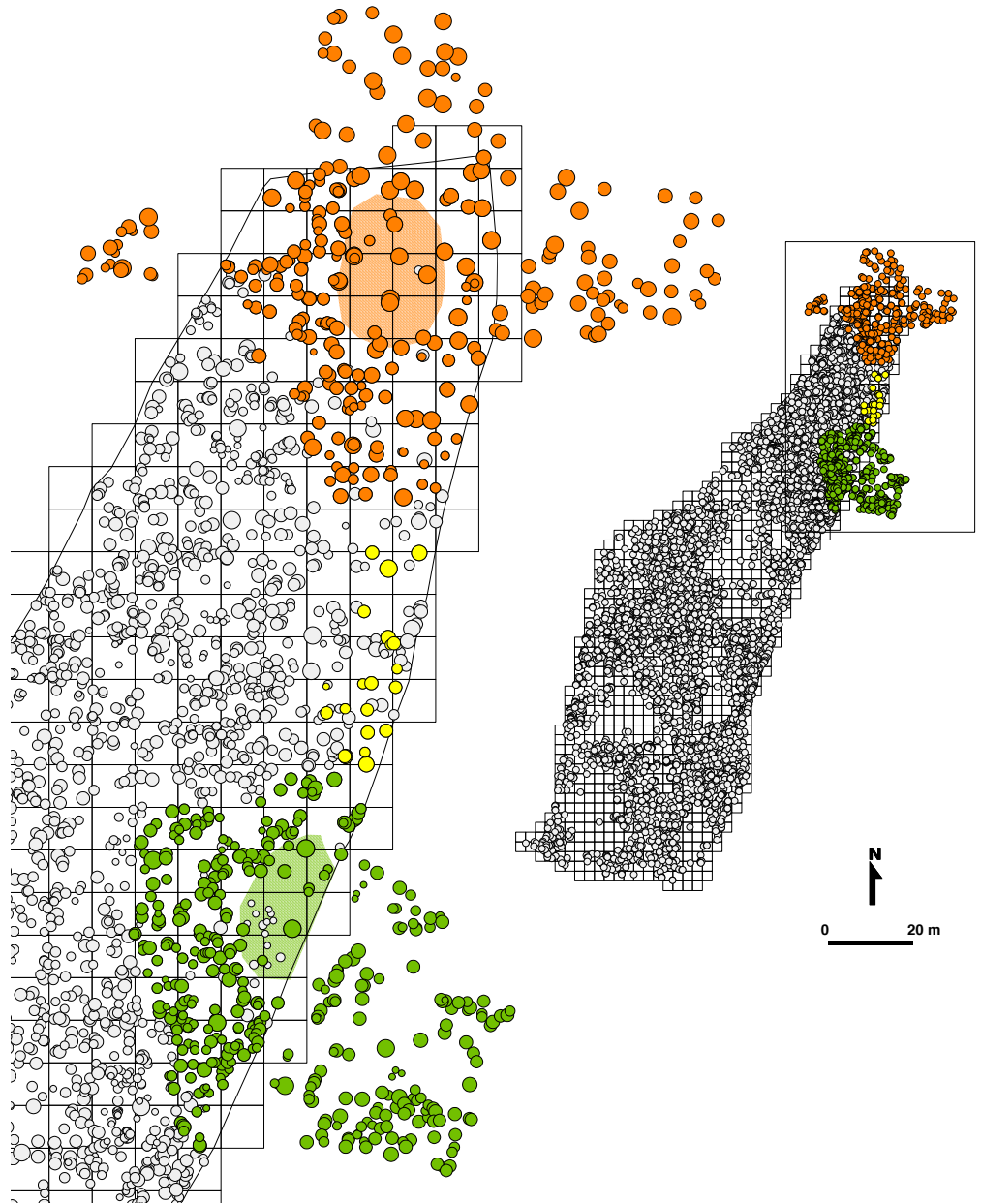
La zone quadrillée (faisant l'objet d'un traitement SIG) correspond à la réserve intégrale, qui comprend 8 578 arbres sur pied (hauteur > ou = 1,3 m) dont 4795 *Fagus sylvatica*.





**Figure 5b. Détail de la cartographie de la placette 1 et la placette 2, avec pour chacune les adultes (figuré de couleur) et la zone d'échantillonnage des semis (zone hachurée).**

On notera que certains hêtres adultes ne faisaient pas partie du suivi forestier de la réserve intégrale et ont été ajoutés à la cartographie à l'occasion de cette étude.



**Légende**

○ Individus vivants sur pied	○ 70 - 200
● Individus génotypés (placette 1)	○ 60 - 70
● Individus génotypés (placette 2)	○ 50 - 60
● Individus génotypés (corridor)	○ 40 - 50
■ Placette 1 - prélèvement semis et plantules	○ 30 - 40
■ Placette 2 - prélèvement semis et plantules	○ 20 - 30
	○ 10 - 20
	○ 1 - 10
	○ 0 - 1

**Tableau 1 : Nombre d'adultes reproducteurs et de descendants échantillonnés.**

NA est le nombre total d'adultes ; sur chacun, le diamètre à 1.30m (DBH) a été mesuré (sont donnés la moyenne=  $\mu$ ., les valeurs minimales et maximales=min max et l'écart-type= $\sigma$ ). NAS est le nombre d'adultes appartenant à la Réserve intégrale pour lesquels le dépérissement foliaire (DEF) et la présence de champignon (OUD) ont été suivis chaque année entre 2002 et 2010, par une note annuelle (0 : absence de symptôme ; 1 : symptôme). Sont donnés ici la moyenne et l'écart-type de la note cumulée.

Placette	Adultes reproducteurs								Descendants	
	$N_A$	$\mu$ DBH (min-max)	$\sigma$ DBH	$N_{A-S}$	$\mu$ DEF	$\sigma$ DEF	$\mu$ OUD	$\sigma$ OUD	Plantules	Semis
1	266	38.68 (3.5 – 119)	22.3	169	1.2	2.57	0.49	1.8	100	100
2	397	21.35 (1.9-96)	13	248	1.6	3.0	0.18	1.2	100	65
Corridor	21	34.28 (13.1-90)	18	21					-	-
Total	684	28.6 (1.9 – 119)	19.4	438	1.4	2.8	0.29	1.4	200	165





## Résultats

---



# VARIABILITÉ GÉNÉTIQUE

Nous avons comparé les indices de diversité moyens (sur les 21 marqueurs) entre les combinaisons placettes/cohortes étudiées (Tab. 2). Ni le nombre d'allèles par marqueurs ni la diversité génétique de Nei ne diffèrent entre cohortes ou entre placette. L'indice de consanguinité n'est jamais significativement différent de 0, même s'il tend à être positif dans les cohortes de reproducteurs (dénotant un excès d'homozygotes, vraisemblablement dû au mélange de différentes cohortes au sein de chaque cohorte de reproducteurs). Enfin, seule la cohorte de reproducteurs de la placette 1 est significativement génétiquement différenciée de deux des autres combinaisons placettes/cohortes étudiées (Tab. 3).

**Tableau 2 : Diversité génétique moyenne des combinaisons cohortes/placettes étudiées, et structure génétique spatiale\* (SGS).**

Na est le nombre d'allèles différents, RA la richesse allélique pondérée (le nombre d'allèles dans un échantillon aléatoire de 36 individus), He l'indice diversité génétique de Nei\*, et  $F_{IS}$  l'indice de consanguinité\* (NS=non significativement différent de 0). La SGS est caractérisée par la pente  $b_{log}$  de la régression entre coefficients d'apparentement et logarithme de la distance, l'indice moyen d'apparentement dans la première classe de distance (F1) et la Sp-Stat (avec l'erreur standard de l'estimation entre parenthèses).

Cohorte	Placette	Diversité				SGS					
		Na	RA	He	$F_{IS}$	$b_{log}$	F1	Sp-stat			
Adultes.	1	7.10	5.67	0.66	0.016NS	-0.0195	(0.0021)	0.0723	(0.007)	0.021	(0.0082)
	2	6.24	5.54	0.67	0.017NS	-0.0174	(0.0013)	0.0663	(0.0043)	0.0186	(0.0054)
Semis	1	6.33	5.44	0.64	-0.03NS	-0.0146	(0.0045)	0.0337	(0.0056)	0.0151	(0.0183)
	2	6.33	5.56	0.64	-0.011NS	-0.0198	(0.0076)	0.0272	(0.0071)	0.0203	(0.0307)

**Tableau 3 : Différentiation génétique moyenne (FST) entre paires de cohortes/placettes.**

	Adultes, PI2	Semis, PI 1	Semis, PI2
Adultes, PI1	0.006 NS	0.016 *	0.016 *
Adultes, PI2		0.019 NS	0.015 NS
Semis, PI 1			0.006 NS

Figure 6 : Structure génétique spatiale\* des cohortes de semis (A) et de reproducteurs (B). Les points représentent les valeurs moyennes des coefficients de Loiselle et coll. (1995) dans chaque classe de distance, et les barres en pointillé les intervalles de confiance à 95%.

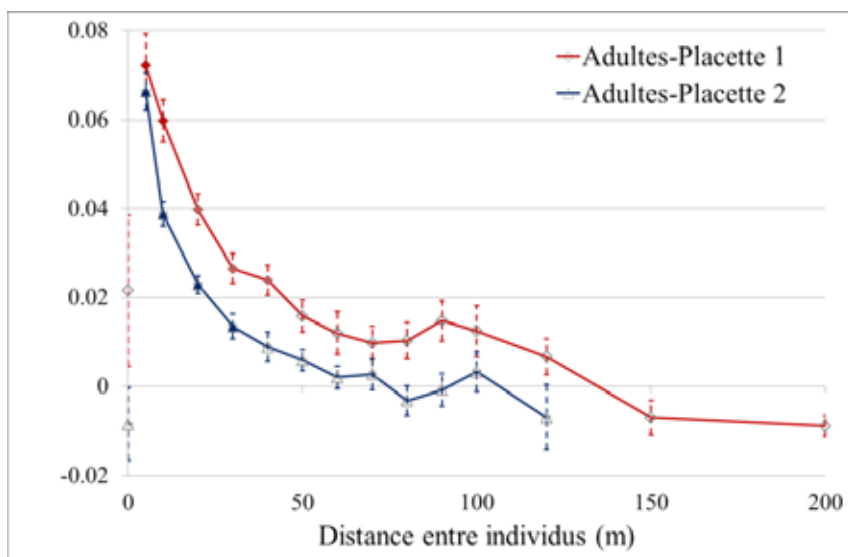
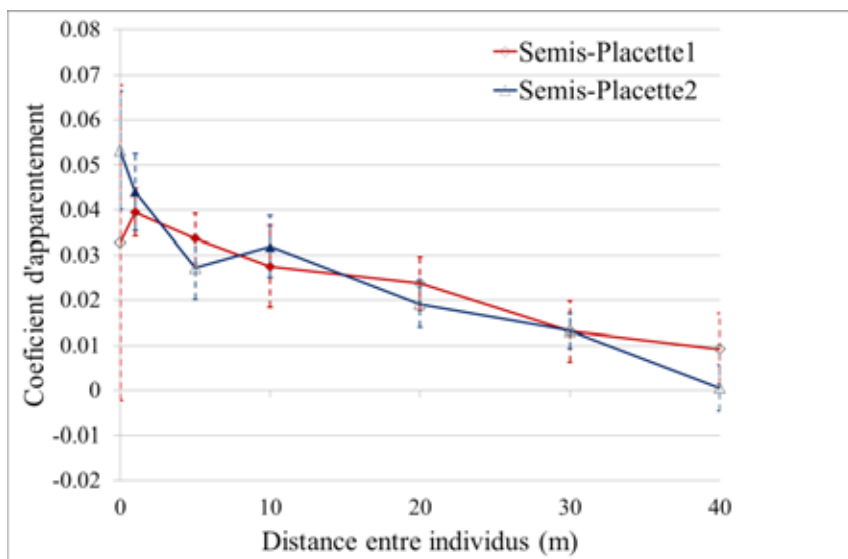


Tableau 4 : Distance moyenne de dispersion des gènes ( $\sigma$ , en m) estimée par l'approche indirecte sur la cohorte des semis ou des adultes. NC indique que la procédure itérative ne converge pas.

$d_{obs}/d_e$	$d_e$	$\sigma$ semis	$\sigma$ adulte
1	0.014	20.2	19.2
5	0.074	NC	26.8
10	0.014	NC	65.7

## STRUCTURE GÉNÉTIQUE SPATIALE (SGS) ET FLUX DE GÈNES HISTORIQUES

---

Dans toutes les combinaisons cohortes/placette, on observe une diminution significative de l'apparement génétique entre paires d'individus quand la distance les séparant augmente (Fig. 6), ce qui dénote une structure génétique spatiale\* (SGS) significative. De plus, la pente de la régression des coefficients d'apparement contre la distance ( $b_{log}$  dans Tab. 4) ne diffère pas significativement entre placettes ou entre cohortes. Ce résultat indique que l'intensité de la SGS est similaire entre placettes et cohortes. Dans les différentes classes de distance, les niveaux moyens d'apparement génétiques sont globalement très similaires entre les semis (placette 1 et 2) et les adultes de la placette 2 ; les adultes de la placette 1 montrent des niveaux moyens d'apparement génétique légèrement plus élevés que les autres combinaisons cohortes/placette (Fig. 6).

L'intensité de la SGS permet également d'estimer les flux de gènes historiques\* et la dérive; en effet  $b_{log}$  est lié au produit de la variance axiale de dispersion des gènes ( $\sigma^2$ ) et de la densité efficace de reproducteur ( $d_e$ ) (eq 1 dans le Matériel et méthode).  $\sigma^2$  reflète l'intensité des flux de gènes\* historiques dans la population, et  $d_e$  reflète la force de la dérive génétique à l'œuvre dans la population. En première approximation on peut estimer  $d_e$  par la densité observée de reproducteurs ( $d_{obs}$ ), mais une estimation plus rigoureuse requiert de tenir compte de tous les facteurs à l'origine des variations de succès reproducteur entre individus (qui vont diminuer  $d_e$ ). Pour estimer  $\sigma^2$ , nous avons utilisé une procédure itérative (voir Matériel et méthode) basée sur la SGS mesurée sur l'ensemble des arbres adultes (placettes 1 et 2), et en considérant différentes valeurs du rapport  $d_{obs}/d_e$  (respectivement 1, 2 et 10). La distance moyenne de dispersion des gènes ainsi estimée (Tab. 4) varie entre 19.2 m (en supposant,  $d_{obs}/d_e = 1$ ) et 65.7 m (en supposant,  $d_{obs}/d_e = 10$ ).



Tableau 5 : Nombre (et pourcentage) de descendants issus de dispersion de graine et de pollen inter- ou intra-placette. « Corridor » désigne les cas de dispersion à partir de parents échantillonnés entre les placettes 1 & 2, et « Extérieur » désigne les cas de dispersion à partir de parents non échantillonnés.

A- Résultats des assignations de parenté, avec PropEch=100% ; Terr=3% et Nm<sub>q</sub>=21.

Dispersion des graines	Dispersion du pollen	Placette 1	Placette 2	Total
Intra-placette	<b>Total</b>	<b>185 (92.5 % )</b>	<b>158 (95.8 % )</b>	<b>343 (94.2 % )</b>
	Intra-placette	111 (55.5 % )	106 (64.2 % )	217 (59.6 % )
Inter-placette	Inter-placette	71 (35.5 % )	49 (29.7 % )	120 (33 % )
	Corridor	3 (1.5 % )	3 (1.8 % )	6 (1.6 % )
	<b>Total</b>	<b>12 (6 % )</b>	<b>6 (3.6 % )</b>	<b>18 (4.9 % )</b>
Extérieur	Intra-placette	12 (6 % )	6 (3.6 % )	18 (4.9 % )
		<b>1 (0.5 % )</b>	<b>0 (0 % )</b>	<b>1 (0.3 % )</b>
Corridor		<b>2 (1 % )</b>	<b>1 (0.6 % )</b>	<b>3 (0.8 % )</b>
<b>Total général</b>		<b>200 (100%)</b>	<b>165 (100%)</b>	<b>365 (100%)</b>

B- Résultats des assignations de parenté/maternité, avec PropEch=80% ; Terr=0% et Nm<sub>q</sub>=21.

Dispersion des graines	Dispersion du pollen	Placette 1	Placette 2	Total
Intra-placette	<b>Total</b>	<b>163 (81.5%)</b>	<b>144 (87.3%)</b>	<b>307 (84.1%)</b>
	Intra-placette	81 (40.5%)	89 (53.9%)	170 (46.6%)
	Inter-placette	43 (21.5%)	29 (17.6%)	72 (19.7%)
	Extérieur	34 (17%)	24 (14.5%)	58 (15.9%)
	Corridor	2 (1%)	6 (3.6%)	8 (2.2%)
Inter-placette	<b>Total</b>	<b>14 (7%)</b>	<b>6 (3.6%)</b>	<b>20 (5.5%)</b>
	Intra-placette	3 (1.5%)	2 (1.2%)	5 (1.4%)
Extérieur	Extérieur	2 (1%)	3 (1.8%)	5 (1.4%)
		<b>21 (10.5%)</b>	<b>6 (3.6%)</b>	<b>27 (7.4%)</b>
Corridor		<b>200 (100%)</b>	<b>165 (100%)</b>	<b>365 (100%)</b>
<b>Total général</b>		<b>14 (7%)</b>	<b>6 (3.6%)</b>	<b>20 (5.5%)</b>

# FLUX DE GÈNES EFFICACES ET SUCCÈS REPRODUCTEURS RÉALISÉS

---

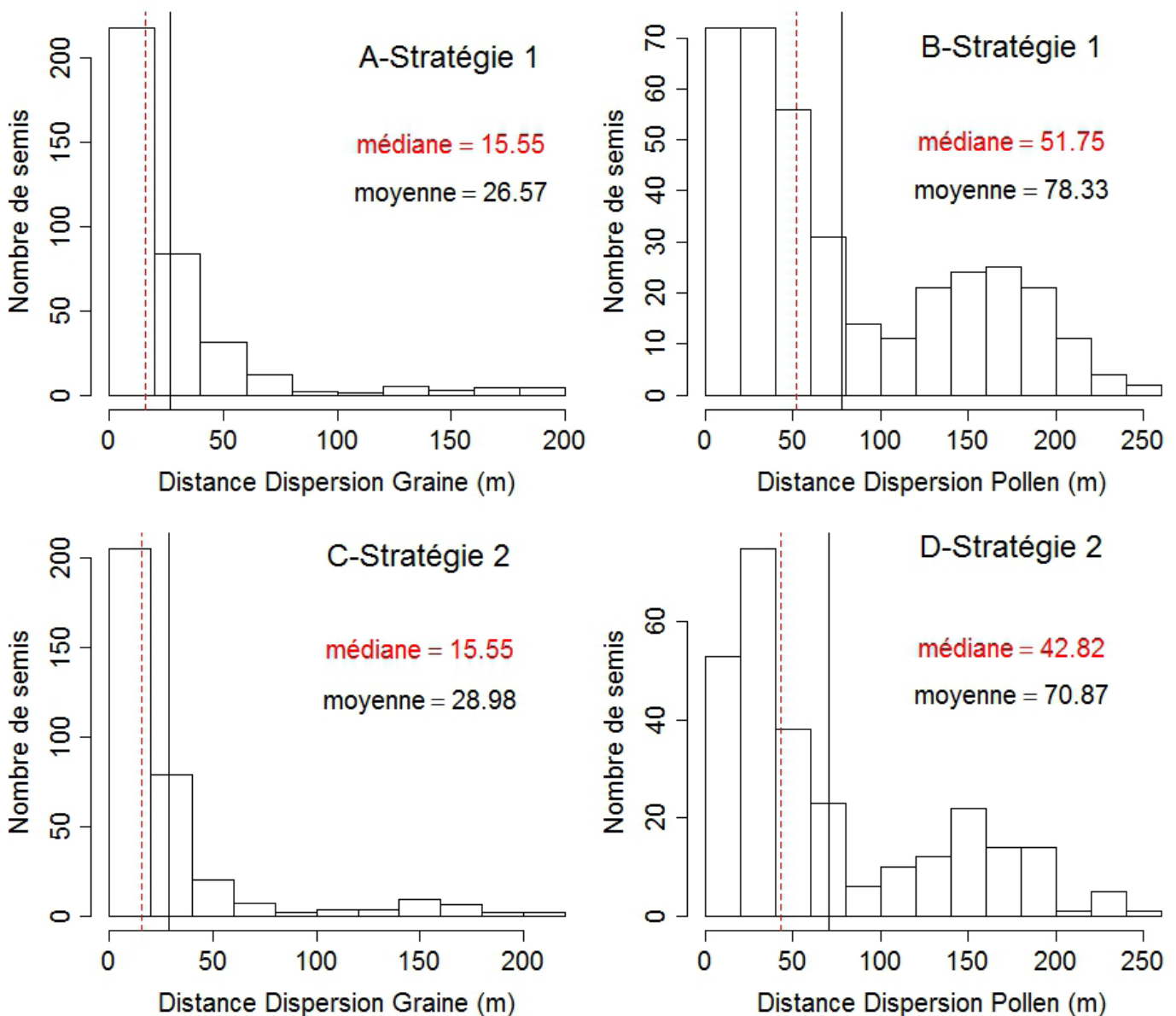
Les assignations catégoriques de parenté/maternité nous ont permis de reconstituer les flux de graines et de pollen efficaces et d'estimer les succès reproducteurs mâles et femelles réalisés. Sur la base de l'analyse de sensibilité (voir Matériel et méthode), **nous avons choisi deux stratégies alternatives d'assignation catégorique. La première stratégie tient compte des erreurs de génotypage, mais néglige le sous-échantillonnage des adultes reproducteurs.** Pour cela, nous avons artificiellement modifié les génotypes aux quatre marqueurs suspectés de présenter des allèles nuls, en transformant tous les génotypes homozygotes XX (X étant un allèle quelconque) en génotypes hétérozygotes XZ (Z étant un nouvel allèle fictif), une stratégie suggérée par Kalinowski & Taper (2006). Ensuite, nous avons fixé les paramètres de l'analyse à PropEch=100% et Terr=3% (voir Matériel et méthode). Mais nous avons éliminé du jeu de résultats les assignations incluant plus de 3 incompatibilités entre descendant et parent(s).

Avec cette première stratégie, un couple de parents peut être assigné pour 364 des 365 semis (Tab 5A), avec seulement une graine provenant d'une mère non échantillonnée. Il y a en moyenne 0.25 locus non compatibles entre le descendant et chaque parent (max=3), et 0.7 locus non compatibles entre le descendant et le couple de parents (max =4). Un seul événement d'autofécondation est identifié. Les descendants sont assignés dans 94.2% des cas à une mère de la même placette (185 descendants/ 199 dans la placette 1 ; 158 descendants/ 165 dans la placette 2, Tab. 5A). Pour les descendants issus de dispersion de graines intra-placette, 59.6 % des pères sont également trouvés à l'intérieur de la placette (111 descendants/ 190 dans la placette 1 ; 106 descendants/ 147 dans la placette 2), alors que 33% des pères sont situés sur l'autre placette, et 1.6% dans le corridor reliant les deux placettes (et aucun à l'extérieur). Pour les 4.9% de descendants issus de dispersion de graines inter-placette, les pères sont toujours retrouvés sur la même placette que la mère (c'est-à-dire que la dispersion du pollen s'est faite à l'intérieur de la placette).

**La deuxième stratégie d'assignation tient compte du sous-échantillonnage des adultes reproducteurs, mais néglige les erreurs de génotypage.** Pour cela, nous avons fixé les paramètres de l'analyse à PropEch=80% et Terr=0%. Avec cette deuxième stratégie, un couple de parents ou une mère peut être assigné pour 338 des 365 semis, les 27 autres semis (7.4% des semis) provenant donc d'un arbre-mère non échantillonné (Tab 5B). Aucun événement d'autofécondation n'est identifié. Les descendants sont assignés dans 84.4% des cas à une mère de la même placette (160 descendants/ 200 dans la placette 1 ; 148 descendants/ 165 dans la placette 2, Tab. 10B). Pour les descendants issus de dispersion de graines intra-placette, 46.6% des pères sont également trouvés à l'intérieur de la placette, 19.7% des pères sont trouvés sur l'autre placette, 2.2% dans le corridor reliant les deux placettes et 15.9% des pères n'ont pas été échantillonnés. Pour les 6.8% de descendants issus de dispersion de graines inter-placette, les pères sont encore le plus souvent retrouvés sur la même placette que la mère (dispersion du pollen intra-placette). Ces résultats (notamment les 7.4% de semis qui ne peuvent être assignés à aucun parent ou les 15.9% de semis qui ne peuvent être assignés à aucun père) reflètent un peu mieux le fait qu'une large proportion des parents reproducteurs à l'extérieur des placettes n'a pas été échantillonnée.

Les deux stratégies alternatives d'assignation ont un effet limité sur la distribution des distances de dispersion efficace du pollen et des graines (Fig. 7). Ces résultats s'accordent sur le fait que la dispersion des graines et du pollen se fait à courte distance. La distance médiane de dispersion efficace des graines est  $d_{\text{eff-graine}} = 15.6$  m avec les deux stratégies. La forme de l'histogramme de  $d_{\text{eff-graine}}$  est très semblable pour les deux stratégies, et montre que  $d_{\text{eff-graine}}$  décroît rapidement avec la distance (Fig. 7A, C). La distance médiane de dispersion efficace du pollen est estimée avec la première stratégie ( $d_{\text{eff-pollen}} = 51.8$ m) est supérieure à celle estimée avec la deuxième stratégie ( $d_{\text{eff-pollen}} = 42.8$ m), ce qui est cohérent avec le plus faible pourcentage d'assignations de pères inter-placette avec cette 2<sup>ème</sup> stratégie. La bosse à 150 m sur l'histogramme de  $d_{\text{eff-pollen}}$  correspond aux événements de dispersion inter-placette (Fig. 7B,D), et elle est un peu moins marquée avec la deuxième stratégie.

Figure 7 : Distribution des distances de dispersion efficace A,C : des graines et B,D : du pollen sur l'ensemble des deux placettes. A,B : Résultats avec PropEch=100% ; Terr=3% et Nm<sub>q</sub>=21. C,D : Résultats avec PropEch=80% ; Terr=0% et Nm<sub>q</sub>=21.





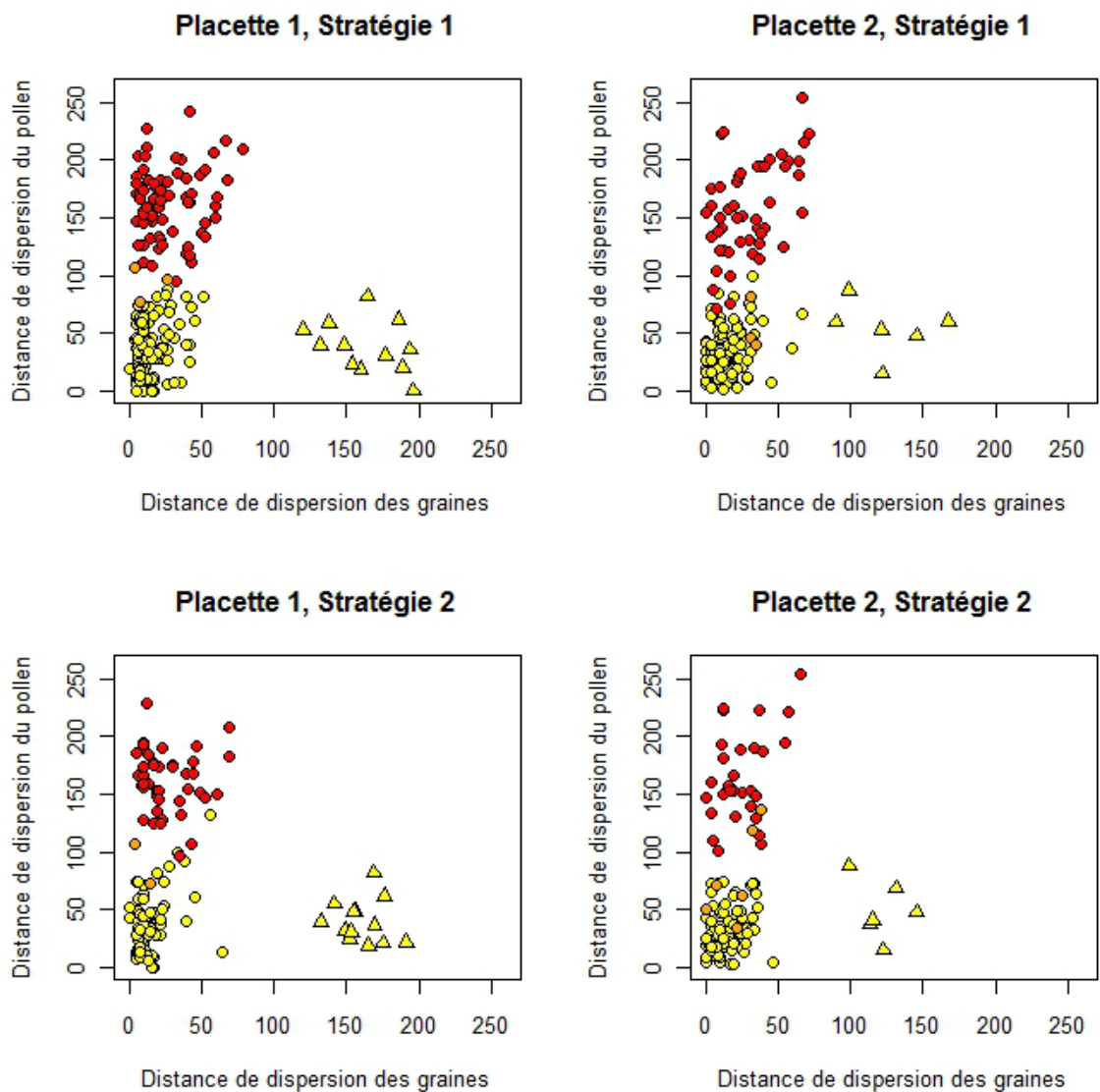
**Ainsi, quelle que soit la stratégie d'assignation, nos résultats mettent en évidence que la dispersion du pollen est moins limitée par la distance que la dispersion des graines (Fig. 8). La distribution conjointe des distances de dispersion des graines (en abscisse, fig. 8) et du pollen (en ordonnée, fig. 8) permet de visualiser par ordre d'importance les événements de dispersion :**

- des graines et du pollen à l'intérieur de la placette (points jaunes)
- des graines à l'intérieur de la placette et du pollen entre placettes (points rouges)
- des graines entre placettes et de pollen à l'intérieur de la placette (triangles jaunes)

Ces graphiques illustrent la rareté des événements de dispersion de graine inter-placettes par rapport aux événements de dispersion de pollen inter-placettes quelles que soient la placette et la stratégie.

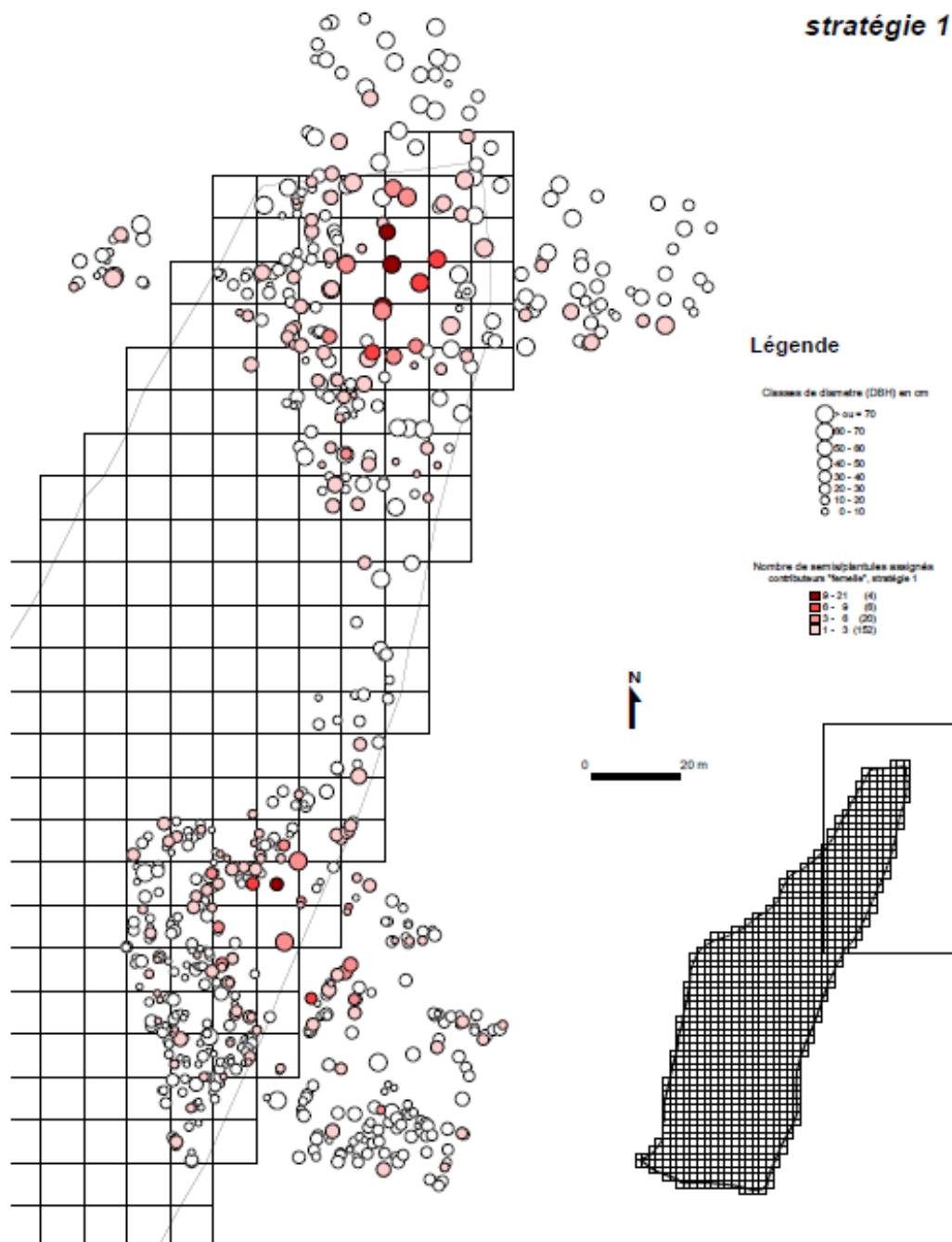
**Figure 8 : Distribution conjointe des distances de dispersion des graines et du pollen par placette et par stratégie.**

Les figurés ronds représentent les événements de dispersion de graine intra-placette, les figurés triangulaires les événements de dispersion de graine inter-placette. La couleur jaune correspond aux événements de dispersion de pollen intra-placette, la couleur orange aux événements de dispersion de pollen impliquant un parent du corridor, la couleur rouge aux événements de dispersion de pollen inter-placette.



Les deux stratégies alternatives d'assignation ont un effet limité sur la variation entre individus des succès reproducteurs femelles et mâles estimés, et sur leur distribution spatiale des (Fig. 9, Append. 2). Ces succès reproducteurs varient fortement entre individus (de 0 à 21 semis par voie femelles ; de 0 à 6 semis par voie mâle) et dans l'espace, les plus forts succès reproducteurs étant observés pour les adultes les plus proches des semis échantillonnés (i.e. ceux au centre de chaque placette).

**Figure 9 : Distribution spatiale des succès reproducteurs individuels**  
A- femelles et B- males estimés par l'analyse de parenté/maternité avec la stratégie 1.



## Stratégie 1

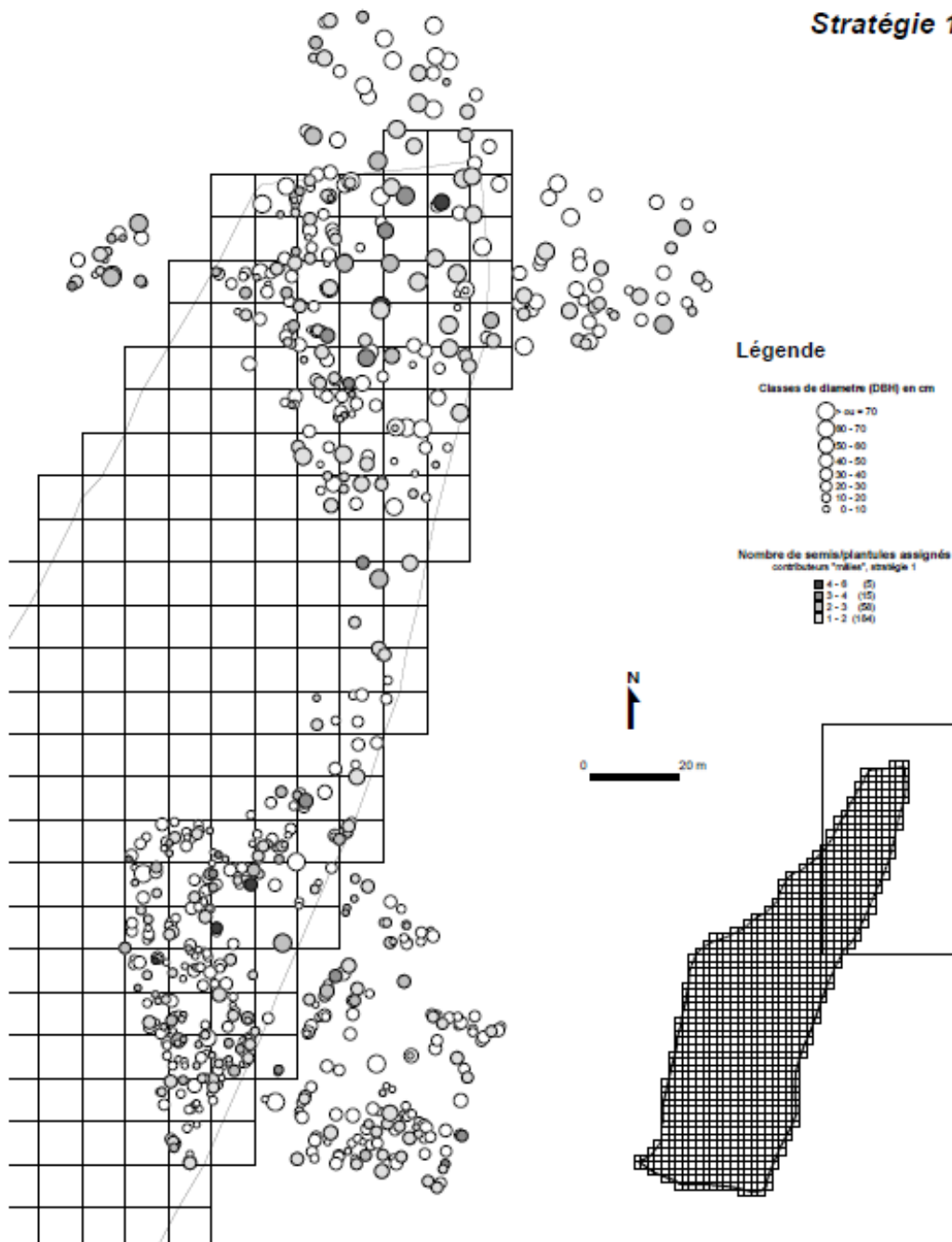


Figure 10 : Comparaison des succès reproducteurs efficaces femelle et mâle estimés avec les deux stratégies

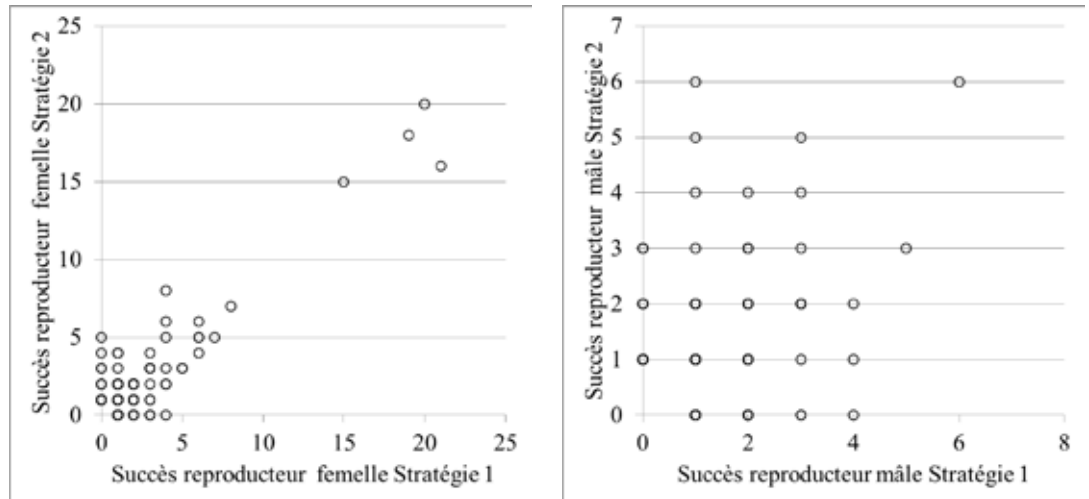
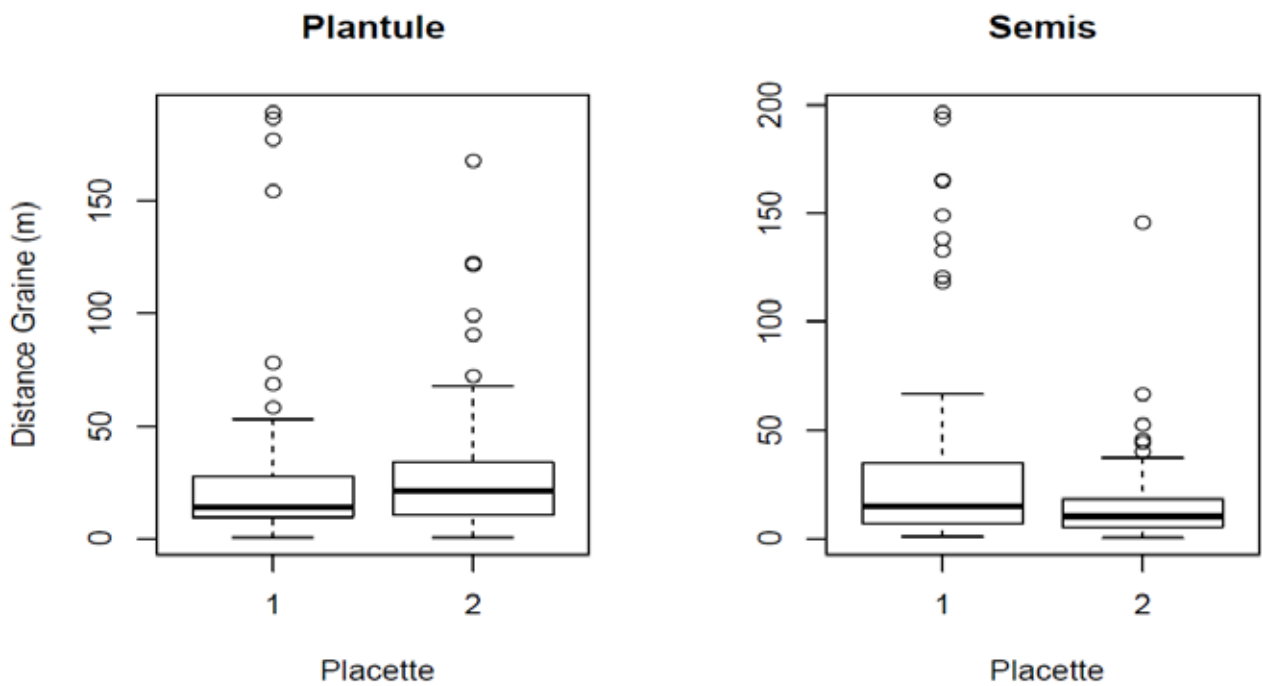


Figure 11 : Distances de dispersion efficace des graines

A. pour les plantules sur chaque placette ; B. pour les semis de l'année sur chaque placette. Résultats avec PropEch=100% ; Terr=3% et Nm<sub>q</sub>=21





Les succès reproducteurs individuels femelles et mâles estimés par chacune des stratégies peuvent également être comparés (Fig. 10). Pour les succès reproducteurs individuels femelles, les deux stratégies donnent des résultats relativement concordants (Fig 10A), avec une corrélation de  $r=0.88$  entre les deux estimations. En revanche, pour les succès reproducteurs individuels mâles, les estimations obtenues par les deux stratégies diffèrent fortement ( $r=0.15$ , Fig 10B).

**Pour les deux stratégies, les distances moyennes de dispersion efficaces des graines ou du pollen ne diffèrent pas entre placettes.**

Par exemple, pour la stratégie 1 (Fig. 11) :  $d_{\text{eff-graineP1}} = 28.58\text{m}$  (placette 1) et  $d_{\text{eff-graineP2}} = 24.15\text{m}$  (placette 2) ;  $d_{\text{eff-pollenP1}} = 82.6\text{ m}$  (placette 1) et  $d_{\text{eff-pollenP2}} = 73.17\text{m}$  (placette 2).

Les distances de dispersion efficaces des graines ne diffèrent également pas entre cohortes (c'est-à-dire entre semis de l'année et plantules), quand on considère les deux placettes ensemble. En revanche, sur la placette 2,  $d_{\text{eff-graine}}$  est significativement plus grande pour les plantules ( $d_{\text{eff-grainePlant}} = 28.7\text{ m}$ ) que pour les semis de l'année ( $d_{\text{eff-graineSemis}} = 17.1\text{ m}$ ).

Tableau 6 : Régime de reproduction, noyaux de dispersion primaires du pollen et des graines, et variance des fertilités intrinsèques mâles et femelles pour chaque placette. Les paramètres estimés sont le taux d'autofécondation (s), les taux de migration par pollen et par graine (mP et mS), les distances moyennes de dispersion par pollen et par graine ( $\delta P$  et  $\delta S$ ) et le paramètre de forme du noyau (bP et bS), ainsi que le rapport  $d_{\text{obs}}/d_{\text{e-theo}}$  empirique (calculé à partir des variances de fertilité dans la population) et théorique (calculé d'après le paramètre de la distribution des fertilités, ici une loi Gamma). Estimation sans ou avec erreur de génotypage.

	Placette 1		Placette 2	
	Médiane	IC	Médiane	IC
$\delta_s$	11	[9.7; 14]	8.2	[6.6; 10]
$b_s$	2.1	[1; 3.4]	0.97	[0.9; 1.2]
$m_s$	0.21	[0.14; 0.28]	0.17	[0.096; 0.27]
$\delta_p$	126	[20; 5405]	58	[23; 5843]
$b_p$	0.2	[0.1; 0.68]	0.62	[0.14; 2.5]
$m_p$	0.7	[0.61; 0.78]	0.66	[0.55; 0.77]
s	$1 \cdot 10^{-8}$	[ $<10^{-23}$ ; 0.004]	$2 \cdot 10^{-18}$	[ $<10^{-28}$ ; $9 \cdot 10^{-5}$ ]
$d_{\text{obs}}/d_{\text{e-theo}}$ ♀	4	[1.9; 8.8]	2.8	[1; 6.5]
$d_{\text{obs}}/d_{\text{e-emp}}$ ♀	5	[4; 9]	3.7	[2.2; 7.1]
$d_{\text{obs}}/d_{\text{e-theo}}$ ♂	1	[1; 5.7]	1	[1; 3]
$d_{\text{obs}}/d_{\text{e-emp}}$ ♂	2.1	[1.8; 6.4]	2.2	[1.8; 4.1]

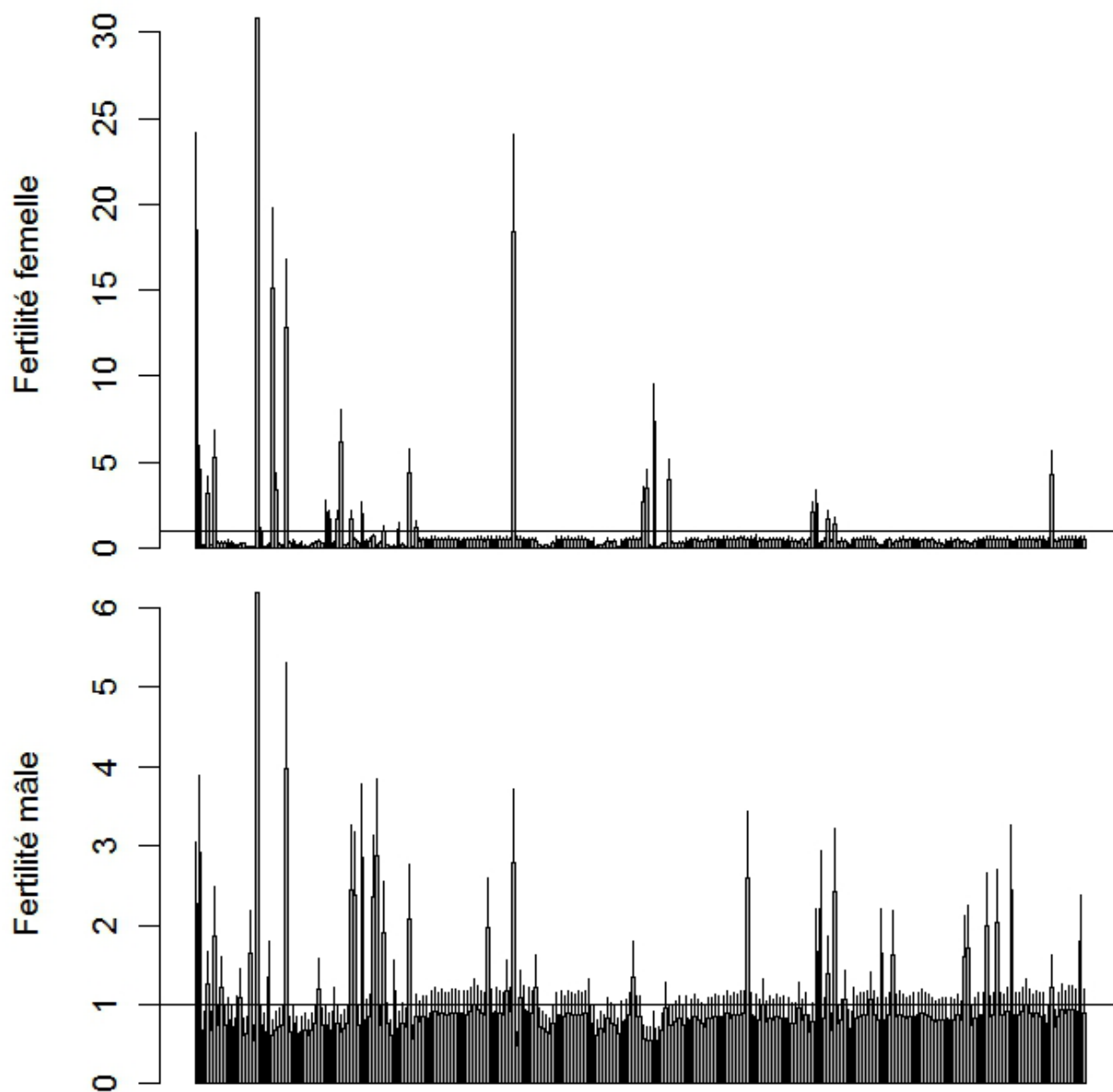
# RÉGIME DE REPRODUCTION, NOYAUX DE DISPERSION PRIMAIRES ET FERTILITÉS INTRINSÈQUES

---

Nous avons utilisé le modèle de croisement dans un voisinage spatialement explicite (SEMM) pour estimer les paramètres du régime de reproduction, des noyaux de dispersion primaires du pollen et des graines, et la distribution des fertilités intrinsèques mâles et femelles. Seuls les arbres adultes de diamètre supérieur à 10 cm (placette 1 : 247 ; placette 2 : 303) ont été considérés, du fait de la production anecdotique de fânes observée sur les arbres de diamètre inférieur à ce seuil (Append. 3).

**Les paramètres du régime de reproduction et de dispersion des graines et du pollen estimés pour les placettes 1 et 2 sont relativement similaires** (Tab. 6). Tout d'abord, le taux d'autofécondation estimé est négligeable ( $<0.001\%$ ). Les taux de migration des graines ( $m_{SPlac1}=0.21$ ,  $m_{SPlac2}=0.17$ ) sont inférieurs aux taux de migration par pollen ( $m_{PPlac1}=0.7$ ,  $m_{PPlac2}=0.66$ ). Les distances moyennes de dispersion des graines sont faibles ( $\bar{\delta}_{SPlac1} = 11$  m,  $\bar{\delta}_{SPlac2} = 8.2$  m). La forme du noyau de dispersion des graines est proche de celle d'une gaussienne sur la placette 1 ( $b_{SPlac1} = 2.1$ ) et de celle d'une exponentielle sur la placette 2 ( $b_{SPlac2} = 0.97$ ), indiquant des événements de dispersion à longue distance plutôt rares. Les distances moyennes de dispersion du pollen sont plus élevées que celles des graines ( $\bar{\delta}_{PPlac1} = 126$  m,  $\bar{\delta}_{PPlac2} = 58$  m). La forme du noyau de dispersion du pollen est celle d'une exponentielle puissance ( $b_{PPlac1} = 0.2$ ;  $b_{PPlac2} = 0.62$ ), ce qui indique des événements de dispersion à longue distance non négligeables. En outre, ces paramètres du régime de reproduction et de dispersion sont peu affectés par la prise en compte des erreurs de génotypage (résultats non présentés).

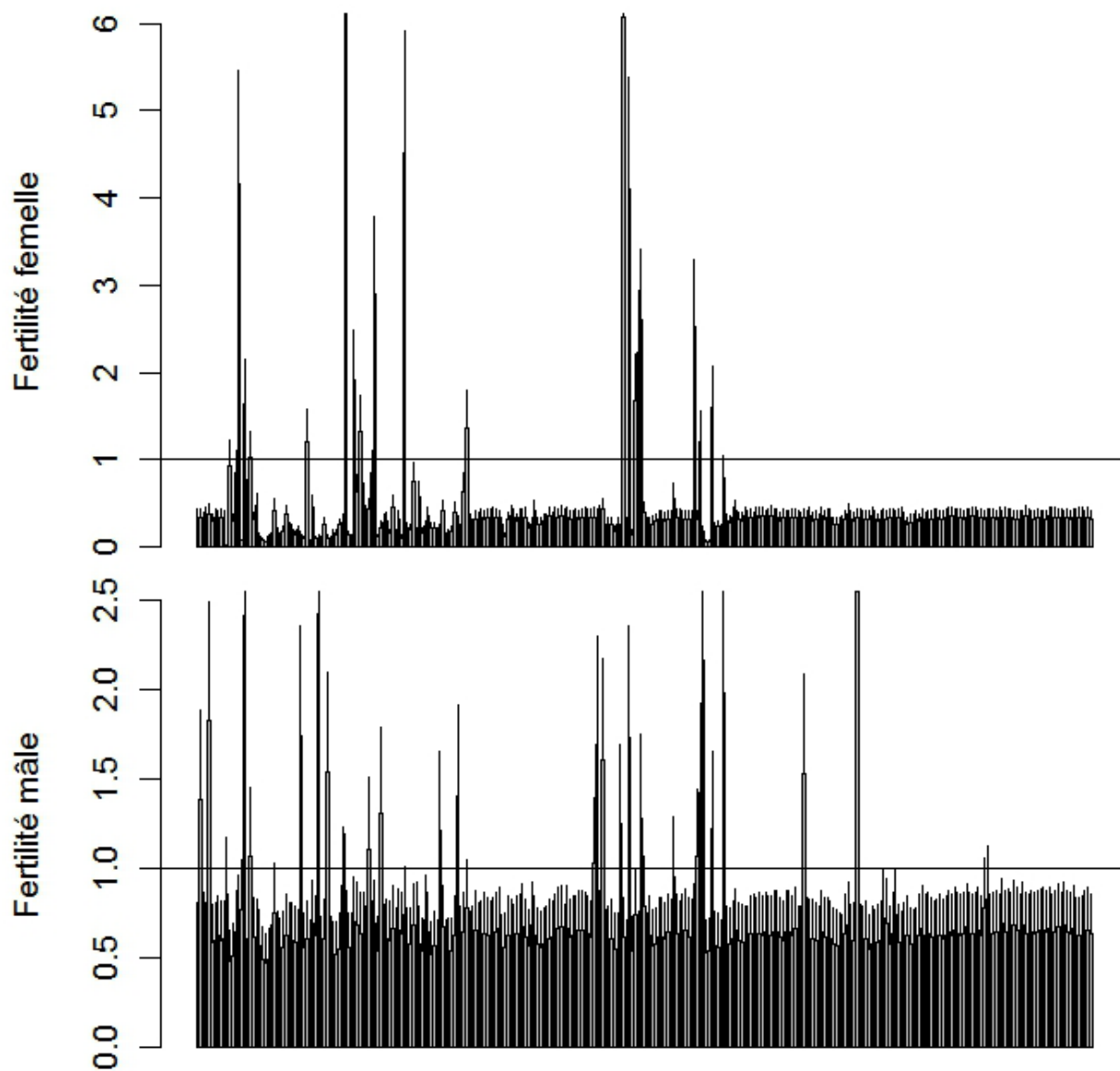
Figure 12 : Distribution empirique des fertilités femelles et mâles sur la placette 1  
Les barres représentent les fertilités femelles(en haut) et mâles (en bas) médianes des arbres adultes, estimées sans tenir compte des erreurs de génotypage. La barre horizontale (Fertilité médiane=1) représente la médiane à priori.



**Le SEMM permet aussi d'estimer les fertilités relatives intrinsèques femelles et mâles**, d'une part à travers le paramètre de la distribution théorique de ces fertilités (décrite par une loi Gamma), et d'autre part à travers la distribution empirique des fertilités individuelles estimées (Fig12, Fig13). L'inspection de la distribution empirique des fertilités femelles individuelles montre **que la distribution des fertilités est déséquilibrée, avec peu d'individus présentant une forte fertilité femelle, et beaucoup d'individus présentant une faible fertilité femelle** (Fig. 12, 13). La médiane a posteriori de la fertilité femelle estimée varie de 0.02 à 11.6 entre individus sur la placette 1 et de 0.01 à 6.1 sur la placette 2. Les distributions théorique et empirique des fertilités peuvent être comparées en calculant à partir de chacune le ratio  $d_{obs}/d_e$ , qui est d'autant plus fort que les fertilités diffèrent entre individus. Nos résultats montrent que les ratios  $d_{obs}/d_e$  estimés à partir des distributions théorique ou empirique des fertilités femelles sont similaires : sur la placette 1,  $d_{obs}/d_{e\text{théo-plac1}}=4$  et  $d_{obs}/d_{e\text{emp-plac1}}=5$  ; sur la placette 2,  $d_{obs}/d_{e\text{théo-plac2}}=2.8$  et  $d_{obs}/d_{e\text{emp-plac2}}=3.7$  (Tab 6). En outre, ces valeurs de  $d_{obs}/d_e$  suggèrent une dérive génétique\* relativement limitée sous l'effet des variations de fertilité femelles entre individus. En effet, les densités observées de reproducteurs (ie d'arbres de diamètre >10cm) sont de 130 individus/ha sur la placette 1 et de 303 individus/ha sur la placette 2, ce qui donne des densités efficaces de reproducteurs par voie femelle de 26 et 82 individus/ha respectivement pour les placettes 1 et 2 (en tenant compte des ratio  $d_{obs}/d_{e\text{emp}}$ ). Enfin, les ratios  $d_{obs}/d_e$  sont similaires sur les deux placettes, montrant que le déséquilibre des contributions femelles à la reproduction est comparable entre placettes.

L'inspection de la distribution empirique des fertilités mâles individuelles montre également que **la distribution des fertilités est déséquilibrée, avec peu d'individus présentant une forte fertilité mâle, et beaucoup d'individus présentant une faible fertilité mâle** (Fig. 12, 13). La médiane a posteriori de la fertilité mâle estimée varie de 0.3 à 4.1 entre individus sur la placette 1 et de 0.45 à 2.55 sur la placette 2. Les ratios  $d_{obs}/d_e$  estimés à partir des distributions théorique ou empirique sont similaires, et similaires entre placettes : sur la placette 1,  $d_{obs}/d_{e\text{théo-plac1}}=1$  et  $d_{obs}/d_{e\text{emp-plac1}}=2.1$  ; sur la placette 2,  $d_{obs}/d_{e\text{théo-plac2}}=1$  et  $d_{obs}/d_{e\text{emp-plac2}}=2.2$ . Ces valeurs de  $d_{obs}/d_e$  suggèrent un effet de dérive très limité sous l'effet des variations de fertilité mâles entre individus, avec des densités efficace de reproducteurs par voie mâle de 61 et 138 individus/ha respectivement pour les placettes 1 et 2.

Figure 13 : Distribution empirique des fertilités femelles et mâles sur la placette 2.  
Les barres représentent les fertilités mâles médianes des arbres adultes, estimées sans tenir compte des erreurs de génotypage (en haut) ou en en tenant compte (en bas). Les barres d'erreurs sont les intervalles de crédibilité autour des estimations individuelles. La barre horizontale (Fertilité médiane=1) représente la médiane à priori.

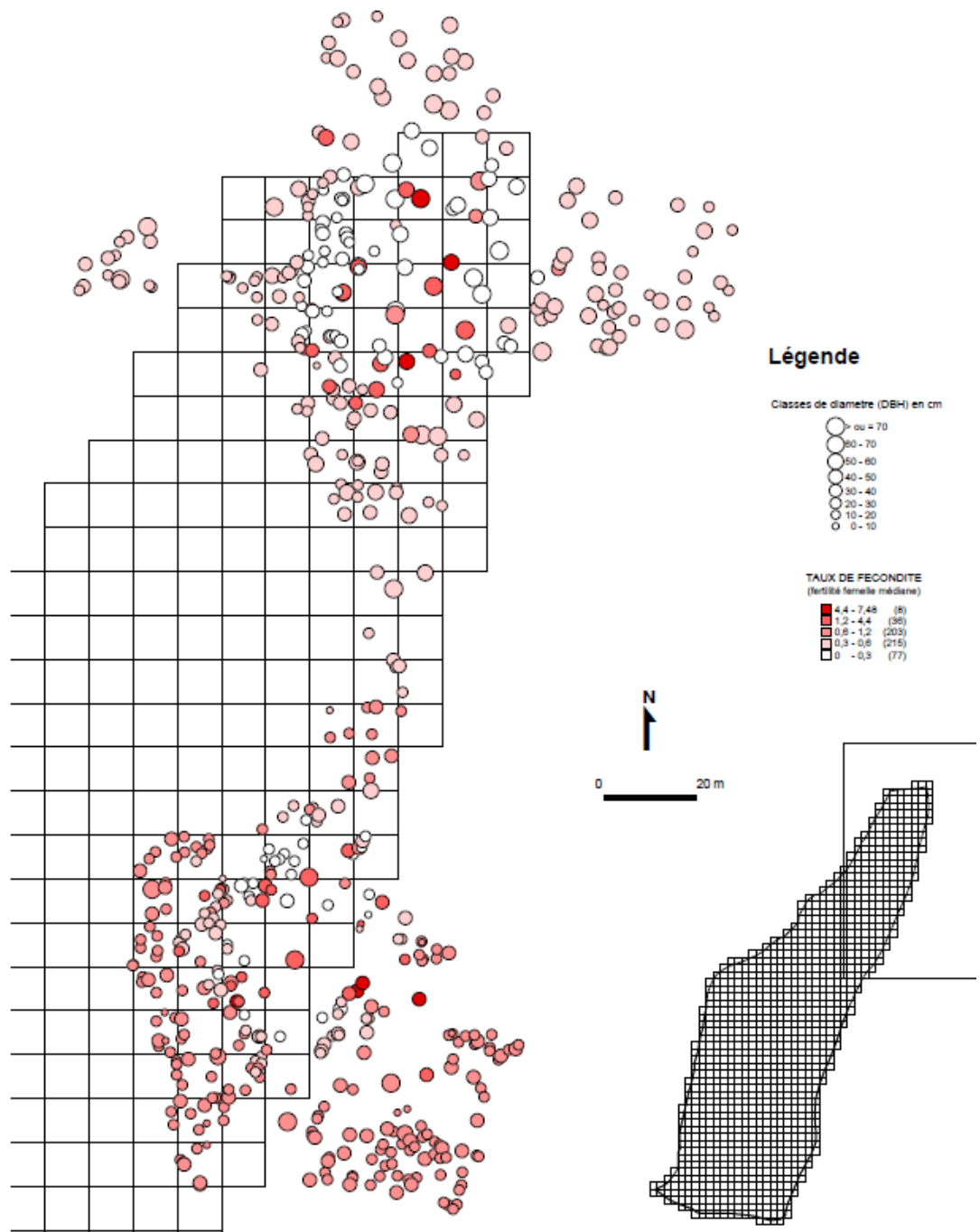




On observe également **un effet négligeable de la prise en compte des erreurs de géotypage sur l'estimation des fertilités intrinsèques femelles et mâles** (résultats non montrés). Il semble en particulier que les estimations de fertilité mâle intrinsèque du SEMM soient moins affectées par les erreurs de géotypage que les estimations de succès reproducteur mâle des assignations catégoriques.

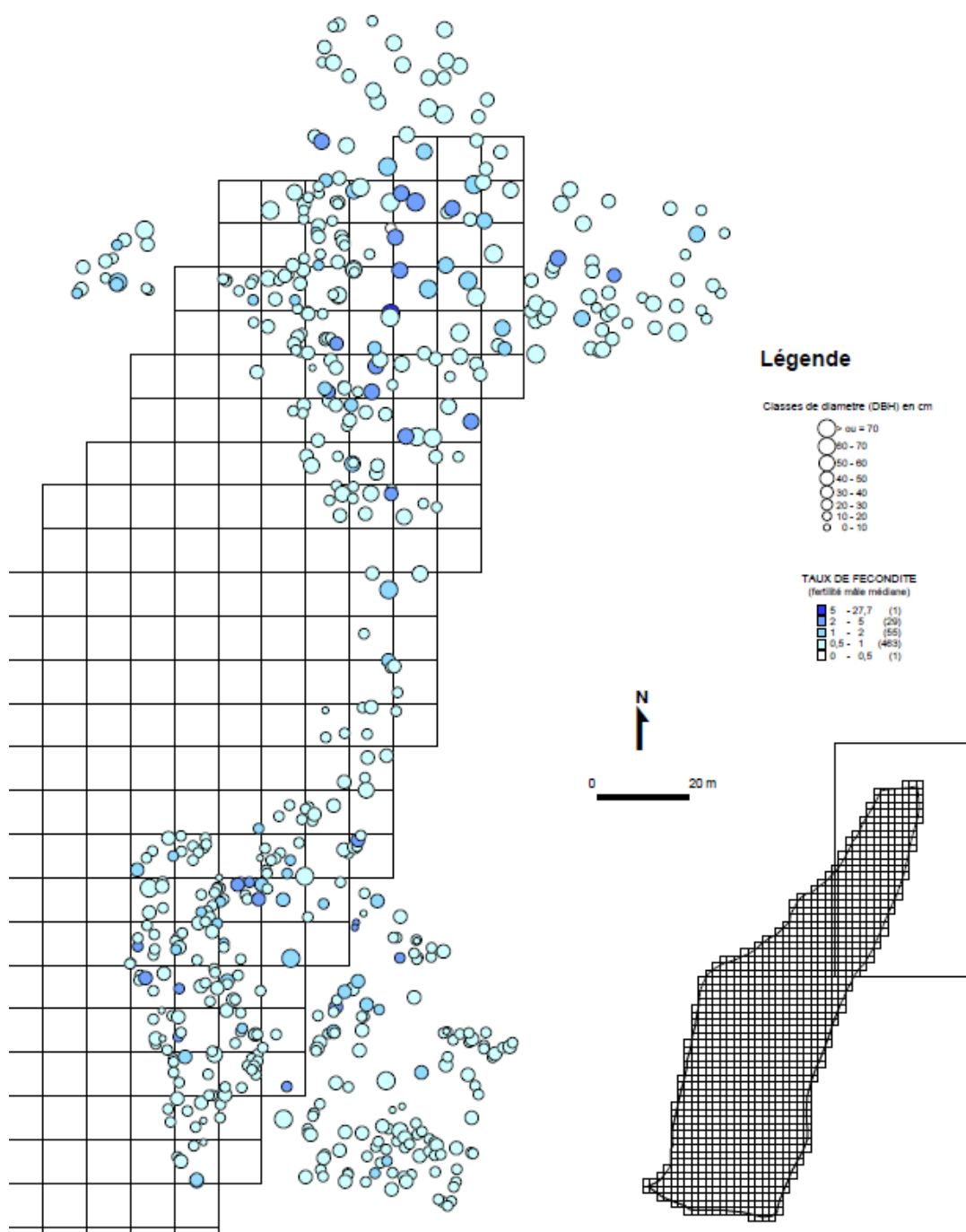
La distribution spatiale des fertilités intrinsèques estimées par le SEMM (Fig. 14) peut être comparée à celles des succès reproducteurs efficaces estimés par l'analyse de parenté/maternité (Fig. 9). Les fertilités intrinsèques sont les plus fortes pour les arbres les plus gros.

Figure 14 : Distribution spatiale des fertilités intrinsèques individuelles A- femelles et B- males estimées par le SEMM.



L'avantage en termes de fertilité intrinsèque des arbres du centre de chaque placette (les plus proches de semis) est beaucoup moins prononcé que ne l'est l'avantage en termes de succès reproducteur réalisé.

Enfin, nous avons étudié la corrélation entre fertilité mâle et femelle (Append. 6). Sur la placette 1, fertilité mâles et femelles sont fortement et positivement corrélées ( $cor = 0.60$ ,  $p$ -valeur  $< 0.001$ ). Sur la placette 2, fertilité mâles et femelles sont aussi corrélées positivement, mais plus faiblement ( $cor = 0.16$ ,  $p$ -valeur  $= 0.007$ ).



# EFFETS DE LA PHÉNOLOGIE ET DU DÉPÉRISSEMENT SUR LES FERTILITÉS

---

Nous avons considéré les fertilités femelles et mâles estimées par le SEMM pour les 327 adultes de diamètre supérieur à 10 cm (placette 1 : 196 ; placette 2 : 131) pour lesquels le dépérissement foliaire (DEF) et la présence du champignon lignicole *O. mucida* (OUD) a été suivis entre 2002 et 2012. Nous avons testé le modèle global incluant simultanément l'effet du dépérissement cumulé (DEFcum), de la présence cumulée du champignon (OUDcum), du diamètre (DBH) et de la production de faînes (FruitProd), et sélectionné le meilleur modèle sur la base de l'AIC.

Pour la fertilité femelle (Tab. 7), le meilleur modèle inclut un effet positif de DBH sur les deux placettes (Effet<sub>DBH-plac1</sub>=0.020, pval=0.008 ; Effet<sub>DBH-plac2</sub>=0.011, pval=0.002). Autrement dit, plus le diamètre de l'arbre augmente, plus la fertilité femelle est importante. Le déficit foliaire cumulé (DEFcum) a un effet négatif marginalement significatif sur la fertilité femelle et ce uniquement pour la placette 2 (Effet<sub>DEFcum-plac2</sub>= -0.032, pval=0.07) : le déficit foliaire est associé à des fertilités femelles plus basses pour cette placette. Si l'on considère seulement les arbres montrant un signal de dépérissement (ie DEFcum>0), l'effet négatif de DEFcum disparaît. FruitProd est corrélé positivement à la fertilité femelle de manière marginalement significative (Effet<sub>FruitProd-plac2</sub>= 0.15, pval=0.07).

Tableau 7 : Effet des facteurs écologiques cumulés sur A-les fertilités femelles et B- fertilités mâles estimées par le SEMM. Les fertilités estimées de 155 adultes (placette 1) et de 173 adultes (placette 2) suivis entre 2002 et 2012 ont été considérées pour estimer les effets conjoints du diamètre (DBH), du dépérissement foliaire cumulé sur 11 ans (DEFcum), de la présence de champignons cumulée sur 11 ans (OUDcum), et de la production de faines mesurées en 2013(FruitProd). Seuls les effets retenus dans le meilleur modèle sont affichés.

Facteur	Effet	SE	t	p-val.
<b>Placette 1</b>				
DBH	0.0203	0.0051	3.43	0.0008
Modèle	0.0656	1+152 DL		
<b>Placette 2</b>				
DEFcum	-0.0319	0.0168	-1.81	0.073
DBH	0.0109	0.0050	3.20	0.002
FruitProd	0.1492	0.0765	1.78	0.077
Modèle	0.0877	3+169 DL		

Facteur	Effet	SE	t	p-val.
<b>Placette 1</b>				
DBH	0.0100	0.0021	3.91	0.00014
FruitProd	-0.1910	0.0804	-1.44	0.15161
Modèle	0.0878	2+151 DL		
<b>Placette 2</b>				
OUDcum	-0.0200	0.0179	-1.45	0.148
FruitProd	0.0779	0.0342	2.57	0.011
Modèle	0.0361	2+170 DL		

Pour la fertilité mâle (Tab. 8) sur la placette 1, on observe un effet positif de DBH (Effet<sub>DBH-plac1</sub>=0.037, pval=0.0001 ); sur la placette 2, on observe seulement un effet positif de FruitProd (Effet<sub>FruitProd-plac2</sub> = -0.078, pval=0.011). On peut noter que sur la placette 1, à côté de l'effet positif et significatif de DBH, le meilleur modèle inclut aussi un effet négatif **non significatif** de FruitProd. Cela suggère qu'au sein d'un groupe d'arbres de même taille, l'investissement à la production de graine pourrait être associé à un investissement moins fort à la production de pollen, soit sous l'effet d'un compromis entre production de gamètes mâles et femelles, soit à cause des phénologies sensiblement décalées des floraisons mâles et femelles chez le Hêtre. Sur la placette 2, l'effet de FruitProd inclut également l'effet du diamètre (puisque la production de faînes augmente avec le diamètre, Append. 3).

Enfin, pour tenter d'estimer l'effet du dépérissement sur la production quantitative de graine, nous avons analysé l'effet de DEFcum sur la production de faînes (FruitProd), considéré comme une variable qualitative ordonnée, et après correction des effets du diamètre. On observe un effet marginalement significatif et positif de DEFcum sur FruitProd (Effet<sub>DEFcum</sub>=0.021, pval=0.059 ).



Tableau 8 : Comparaison du régime de reproduction et des paramètres de dispersion entre la Massane et le Ventoux. Les paramètres estimés sont le taux d'autofécondation ( $s$ ), les taux de migration par pollen et par graine ( $m_P$  et  $m_S$ ), les distances moyennes de dispersion par pollen et par graine ( $\delta_P$  et  $\delta_S$ ) et le paramètre de forme du noyau ( $b_P$  et  $b_S$ ), ainsi que le rapport de/dobs empirique (calculé à partir des variances de fertilité dans la population) et théorique (calculé d'après le paramètre de la distribution des fertilités, ici une loi Gamma). Pour chaque paramètre sont donnés la moyenne des valeurs médianes estimées, et l'intervalle de crédibilité à 95% (IC- et IC+) autour de cette moyenne.

Paramètre	La Massane (2 placettes)			Ventoux (3 placettes)		
	moyenne	IC-	IC+	moyenne	IC-	IC+
$\delta_S$	9.6	8.2	12.0	19.7	21.0	23.0
$b_S$	1.5	1.0	2.3	1.3	0.9	1.6
$m_S$	0.19	0.12	0.28	0.14	0.04	0.18
$\delta_P$	92.0	21.5	5624.0	162.0	45.0	2700.3
$b_P$	0.4	0.1	1.6	0.2	0.1	0.6
$m_P$	0.68	0.58	0.78	0.60	0.53	0.68
$s$	$3.3 \cdot 10^{-05}$	$1.0 \cdot 10^{-18}$	0.002	0.009	0.001	0.030
$d_{obs}/d_{e-theo} \text{ ♀}$	3.4	1.5	7.7	7.2	3.2	12.3
$d_{obs}/d_{e-emp} \text{ ♀}$	4.4	3.1	8.1	8.1	2.7	16.0
$d_{e-emp} \text{ ♀ (/ha)}$	50	27	70			
$d_{obs}/d_{e-theo} \text{ ♂}$	1.0	1.0	4.4	4.7	1.1	9.1
$d_{obs}/d_{e-emp} \text{ ♂}$	2.2	1.8	5.3	5.3	3.0	9.1
$d_{e-emp} \text{ ♂ (/ha)}$	98	41	120			







# Discussion

---





# RÉGIME DE REPRODUCTION, FLUX DE GÈNES ET TAILLE EFFICACE DE POPULATION DANS LA HÊTRAIE DE LA MASSANE

---

Cette étude nous a d'abord permis de caractériser le régime de reproduction, les capacités de dispersion primaires du pollen et des graines et la densité efficace de reproducteurs dans la hêtraie de La Massane (Tab. 8). **Le régime de reproduction** est essentiellement allogame, avec un taux d'autofécondation négligeable. **La distance moyenne de dispersion des graines** est de  $\delta_s=9.6$  m [10.15 ; 12m], c'est-à-dire  $\sim 10$  fois plus faible que **la distance moyenne de dispersion du pollen :  $\delta P = 92$  m** [21.5 ; 5624m]. Le noyau de dispersion\* du pollen, à queue lourde ( $b_p=0.4$ ), indique en outre des **capacités de dispersion de pollen à longue distance non négligeables**, alors que de tels événements sont négligeables par voie de graine ( $b_s=1.5$ ). Ces caractéristiques des noyaux de dispersion (estimées à partir des événements de dispersion intra-parcelle) sont cohérentes avec les taux de migration par pollen élevés ( $m_p\sim 68\%$ ), alors que les taux de migration par graines sont plus faibles ( $m_s\sim 19\%$ ). En outre, les capacités de dispersion primaires du pollen et des graines sont similaires entre les deux placettes, et donc peu affectées par la structure du peuplement (plus hétérogène sur la placette 1). Enfin, la dérive génétique due aux variations de fertilité entre individus apparaît relativement limitée, avec des **densités efficaces de reproducteurs par voie femelle de 50 individus/ha, et de 98 individus/ha par voie mâle** (alors que la densité observée de reproducteurs est de  $d_{obs}=216$  individus/ha).

Ces différents éléments permettent d'estimer approximativement le **nombre efficace de reproducteurs (Nb)** dans la hêtraie de la Massane. Nb est lié à la notion de taille efficace de population ( $N_e$ )\*, un paramètre majeur pour évaluer le statut de conservation et les menaces sur la santé génétique des populations. En effet,  $N_e$  détermine l'intensité de la dérive génétique\* : quand  $N_e$  est faible (typiquement quelques dizaines d'individus), la dérive génétique érode la variabilité génétique, augmente la probabilité de fixation des allèles délétères et réduit l'efficacité de la sélection, ce qui conduit globalement à réduire la valeur adaptative de la population et son potentiel d'adaptation (Hare et coll., 2011; Palstra & Ruzzante, 2008). Chez beaucoup d'espèces, l'effectif de recensement de la population ( $N_{obs}$ ) et la taille efficace  $N_e$  varient typiquement dans la gamme  $N_e\sim 0.5-0.10\times N_{obs}$  ; mais chez certaines espèces (par exemple marines), on observe des ratios beaucoup plus faible ( $N_e\sim 10^{-5}\times N_{obs}$ ). Il est donc crucial de pouvoir estimer  $N_e$ , en particulier dans des

populations concernées par des enjeux de conservation. À partir de nos estimations des densités efficaces de reproducteurs, et en considérant que la hêtraie de la Massane est très grande (336 ha pour la seule réserve intégrale), que les hêtres y sont distribués de façon continue dans l'espace et que leurs capacités de dispersion par pollen en particulier sont non négligeables, le nombre efficace de reproducteurs à l'échelle de la réserve pourrait être autour de  $N_b = 10,000$  individus (voir App. 11 pour les détails). On est donc clairement dans une situation où **la dérive génétique est suffisamment faible pour ne pas réduire la valeur adaptative de la population de la Massane et son potentiel d'adaptation.**

**Les capacités de dispersion primaires du pollen et des graines estimées dans la hêtraie de la Massane sont cohérentes avec celles estimées dans d'autres populations du sud de l'aire de distribution.**

Avec la même méthode (le SEMM) dans la hêtraie du Ventoux (Tab. 8), nous avons estimé des valeurs moyennes de dispersion similaires ( $\bar{\delta}_s \sim 19.7$  m ;  $\bar{\delta}_p \sim 162$  m, avec  $\bar{\delta}_p/\bar{\delta}_s \sim 8$ ), des formes des noyaux semblables ( $b_p = 0.2$  ;  $b_s = 1.3$ ), et des taux de migration par pollen et graines similaires ( $m_p \sim 78\%$  ;  $m_s \sim 14\%$ ). Les taux d'autofécondation estimés à la Massane sont particulièrement faibles ( $s < 0.01\%$ ), même en comparaison des taux d'autofécondation faibles estimés au Ventoux ( $s = 0.9\%$ ). Ce résultat suggère des conditions légèrement plus défavorables à la fécondation croisée au Ventoux qu'à la Massane, peut-être dues à la plus grande hétérogénéité en espèces des peuplements mélangés (Hêtre-Sapin-Pin) au Ventoux (Gauzère *et coll.*, 2013). **Malgré les conditions écologiques a priori marginales dans la réserve de la Massane, le régime de reproduction est donc quasiment un régime d'allogamie strict**, ce qui est rassurant pour le potentiel évolutif de la population. Par ailleurs, les capacités de dispersion du pollen du Hêtre à la Massane sont du même ordre de grandeur que celles estimées chez d'autres Fagacées anémophiles (Ashley, 2010), et confirment que la dispersion du pollen par le vent est un facteur majeur d'homogénéisation génétique à longue distance. En ce qui concerne les plus faibles capacités de dispersion des graines, elles sont également cohérentes avec l'écologie du Hêtre, et le fait que les faînes soient essentiellement dispersées par gravité, ou par de petits rongeurs (*Apodemus* sp. et dans une moindre mesure l'écureuil plus rare à la Massane) qui cachent les graines (Jensen, 1985). La sittelle (qui déplace les faînes avant de les manger et les loge dans un creux de l'écorce – figure Y ; qui serait la 12 et tout se décale ensuite), le geai, ou les mésanges sont également de bons disperseurs de graines. Des événements rares de dispersion de faînes à longue distance (1 km ou plus) sont ainsi possibles, mais ne peuvent pas être détectés par notre étude dont l'échelle spatiale est trop réduite, avec des placettes de moins d'un hectare. En parallèle des capacités de dispersion primaire du pollen et des graines, nous avons



aussi caractérisé **la dispersion efficace\* du pollen et des graines** à la Massane, à partir des assignations catégoriques de parenté/maternité. Les distances médianes de dispersion des graines ( $\delta_{S\text{-eff}}=15.6$  m) et du pollen ( $\delta_{S\text{-eff}}=42$  m) intègrent, en plus des capacités de dispersion primaire, l'effet de la disposition spatiale des parents et des semis, ainsi que les processus de recrutement post-dispersion. Il est délicat de comparer directement les distances de dispersion primaires et efficaces (car les méthodes utilisées pour les estimer font des hypothèses très différentes). Néanmoins, on note que comme pour les capacités de dispersion primaire, les capacités de dispersion efficace du pollen sont plus importantes que celles des graines (avec un rapport de 3). Par ailleurs, sur la placette 2, entre les 65 semis de l'année et les 100 semis plus vieux, on observe un signal d'augmentation de la distance de dispersion efficace au cours du recrutement. Ce signal est attendu notamment si la mortalité des semis est plus forte à courte distance de l'arbre mère, du fait de processus densité-dépendant (Bontemps *et coll.*, 2013).

Enfin, à côté des **capacités contemporaines de dispersion du pollen et des graines**, nous avons également caractérisé la structure génétique spatiale (SGS), qui reflète **l'action cumulée sur plusieurs générations des flux de pollen et de graines dits historiques**. En accord avec les capacités contemporaines limitées de dispersion du pollen et surtout des graines, on observe à La Massane une SGS significative, avec une diminution rapide de l'apparement entre adultes ou semis avec la distance les séparant. Une façon de comparer les flux de gènes contemporains\* et historiques\* consiste à convertir les distances contemporaines de dispersion du pollen et des graines estimées par le SEMM en une distance moyenne de dispersion des gènes ; dans notre cas,  $\sigma_{\text{contemp}} = 67$  m (équation 2). Cette valeur est cohérente avec l'estimation des capacités historiques de dispersion :  $\sigma_{\text{hist}} = 65.7$  m (Tab. 4), estimée à partir de la SGS en supposant un rapport  $d_{\text{obs}}/d_e = 10$  (par ailleurs cohérent avec les rapports  $d_{\text{obs}}/d_e$  contemporains estimés). **Cette similarité des flux de gènes contemporains et historiques suggère que les processus de dispersion du pollen et des graines n'ont pas changé de façon majeure au cours des dernières générations.**

# EFFET DES CARACTÈRES DES ARBRES ADULTES (TAILLE, PRODUCTION DE GRAINES ET DÉPÉRISSEMENT) ET CONSÉQUENCES SUR LEURS FERTILITÉS RELATIVES

---

Cette étude a également permis de caractériser **les variations relatives de fertilités mâles et femelles** entre individus, et **d'analyser l'effet de différents caractères des arbres adultes** (taille, production de fruit, dépérissement foliaire, présence de champignon lignicoles) **sur ces variations de fertilité**. De façon générale, nous observons une **distribution déséquilibrée des fertilités intrinsèques mâles ou femelles entre individus, avec peu d'individus présentant une fertilité forte, et beaucoup d'individus présentant une fertilité faible** (Fig. 12, 13). Cette hétérogénéité est la règle plutôt que l'exception, comme le montrent les études portant sur la variabilité individuelle des fertilités dans les populations d'arbres {Chybicki, 2013 #1732 ;Hampe, 2005 #1733}. Dans cette étude, l'hétérogénéité des fertilités résulte en partie de l'augmentation des fertilités femelles ou mâles avec la taille des individus (ici mesurée via le diamètre ; Tab. 7). Cet effet de la taille peut s'expliquer par un investissement plus important à la reproduction des individus les plus gros, et/ou par une dispersion/réception du pollen plus efficace pour les individus les plus hauts, ou ceux qui ont un houppier plus développé. Les fertilités femelles ont d'ailleurs tendance à être plus déséquilibrées sur la placette 1 (la plus hétérogène en diamètre) que sur la placette 2, comme le montrent les plus forts ratios de densité observée sur densité efficace de reproducteurs par voie femelle ( $d_{obs}/d_e$  ♀, 1.4 plus élevés sur la placette 1 en moyenne), mais cette tendance n'est pas significative. Le nombre de faînes produites est aussi un bon prédicteur de la fertilité femelle ou mâle (Tab. 7), même si seule la production de faînes en 2013 a pu être mesurée dans cette étude (alors que les semis étudiés ont été recrutés entre 1982 et 2012). Les fertilités femelles et mâles ont aussi tendance à être plus déséquilibrées au Ventoux qu'à La Massane, comme le montrent les ratios  $d_{obs}/d_e$  plus élevés (rapport de 2 pour les ratios ♀ et de 3.6 pour les ratios ♂), mais cette tendance n'est pas non plus significative. Les peuplements du Ventoux sont moins hétérogènes en diamètre que ceux de la Massane, donc cette comparaison suggère que le facteur taille seul n'est pas suffisant pour comprendre les variations de fertilités entre individus.

Sur une des deux placettes étudiées, **l'analyse des fertilités individuelles estimées par l'approche génétique suggère un effet négatif du dépérissement foliaire sur l'investissement individuel**

**à la reproduction, compatible avec la limitation des ressources carbonée et le stress hydrique subi par l'arbre dépérissant.** Cet effet négatif du stress est observé sur la fécondité femelle, mais pas sur la fécondité mâle : cela est cohérent avec le coût de production des organes reproducteurs, plus élevé chez le hêtre pour les graines et fruits que pour les fleurs et étamines. De façon intéressante, **l'analyse des productions individuelles de faîne en 2013 suggère au contraire un effet positif du dépérissement foliaire sur l'investissement individuel à la reproduction, qui lui serait compatible avec l'hypothèse alternative d'une augmentation possible de la production de graines en réponse au stress** (Lee, 1988). Ces deux résultats antagonistes soulignent l'intérêt de distinguer l'effet du stress sur d'une part les fertilités individuelles **réalisées**, c'est-à-dire estimées à partir de semis recrutés, et d'autre part les **fertilités individuelles potentielles**, estimées à partir de la quantité de fruits ou de graines produites. Les fertilités réalisées tiennent compte indirectement des différences de qualité des graines entre individus (parasitisme, aptitude à la germination, quantité de réserve), ce qui n'est pas le cas des **fertilités potentielles**. **Globalement, ces résultats pourraient indiquer que le stress induit une production de graine plus importante en quantité, mais de moins bonne qualité.** Nous n'observons pas en revanche d'effets clairs de la présence du champignon lignicole *O. mucida* sur la fécondité, probablement en lien avec la faible prévalence du champignon.

Pour mieux comprendre l'effet des facteurs écologiques sur la quantité de la régénération, il serait bien sûr encore plus intéressant de caractériser les variations de fertilité individuelles réalisées **absolues**; mais cela demanderait de compter exhaustivement le nombre de semis issus des arbres étudiés. Dans cette étude, nous n'avons recensé les semis que sur une surface restreinte au sein de chaque placette, ce qui permet de comparer les contributions relatives des adultes à ce sous-échantillon de la régénération. Les succès reproducteurs efficaces estimés avec les assignations catégoriques de parenté/maternité ne tiennent pas compte des effets spatiaux liés à la dispersion limitée par la distance, ce qui se traduit mécaniquement par des succès reproducteurs plus fort pour les arbres adultes les plus proches des semis (Fig. 9). Au contraire, les estimations de fertilité intrinsèque obtenues avec le SEMM tiennent compte des effets spatiaux, ce qui se traduit par le fait que des arbres loin du centre des placettes peuvent avoir des fertilités intrinsèques élevées (Fig. 14), malgré un succès reproducteur efficace faible.

De façon générale, **la taille et/ou la production de fruit expliquent moins de 10% de la variation des fertilités estimées** (Tab 7). De nombreux autres facteurs peuvent être à l'origine de ces variations de fertilités, notamment les coïncidences des phénologies reproductives

individuelles. La hêtraie de la Massane présente une intéressante originalité de ce point de vue : environ 10% des hêtres y présentent une phénologie végétative beaucoup plus précoce que les autres [REF]. Or phénologie végétative et reproductrice sont très bien corrélées chez le Hêtre. Néanmoins, dans cette étude, nous n'avons détecté aucun effet significatif de la précocité sur la variation des fertilités. D'autres analyses permettraient de prendre en compte l'effet de la compétition entre individus sur les fertilités individuelles. Par ailleurs, il serait intéressant de tenir compte de l'investissement relatif des arbres à la croissance par rapport à celui accordé à la reproduction, ce qui pourrait être fait par une analyse des cernes de croissance des individus (en projet).

## CONCLUSION

---

La hêtraie de la Massane se caractérise donc par un régime de reproduction allogame, une taille efficace de population importante et des flux de gènes en particulier par pollen non négligeables. Ces résultats sont intéressants en regard du statut de réserve de la Massane, et suggèrent que le potentiel adaptatif de cette population marginale de haute valeur patrimoniale n'est pas menacé par la dérive génétique.

En outre, nos résultats suggèrent un effet limité des dépérissements actuellement observés sur les processus reproductifs. Le dépérissement foliaire (mesuré par la présence de branches mortes dans le houppier), qui peut être l'indicateur d'une réponse au stress hydrique, induit une production de graine légèrement plus importante en quantité, mais de moins bonne qualité, ce qui se traduit par une diminution de la contribution individuelle aux semis recrutés. Il serait intéressant de poursuivre ces observations dans le temps, en particulier si les dépérissements foliaires s'accroissent ; il est en effet probable qu'il existe un seuil de dépérissement foliaire, non encore atteint actuellement, tel que l'investissement à la reproduction soit brutalement limité. De tels points de basculement dans les effets du stress hydrique pourraient compromettre la régénération. L'idéal pour suivre cette dynamique serait probablement de mettre en place un suivi individuel sur un échantillon d'arbres dépérissant et non dépérissant, avec prélèvement de graines et analyse fine de leur qualité (taux de graine vides ou parasitées, faculté de germination).





## Matériel et Méthodes

---



# SITE D'ÉTUDE ET SUIVI DES ARBRES ADULTES

---

Cette étude a été conduite dans la partie nord-est de la réserve intégrale, zone où le hêtre est l'essence dominante à plus de 90% (contre 70% en moyenne sur l'ensemble de la réserve intégrale). Nous avons délimité deux placettes (appelées Placette 1 et 2) couvrant respectivement 0.9 et 0.8 ha et distantes de ~80 m (Fig. 5). Sur chaque placette tous les adultes ont été cartographiés, et leur diamètre à 1.30 a été mesuré. Les deux placettes présentent une structure en âge du peuplement sensiblement différente : dans la placette 1, le diamètre moyen du peuplement est plus élevé que dans la placette 2 ( $DBH_{\text{placette1}} = 36.7$  vs  $DBH_{\text{placette2}} = 21.5$ ), ce qui résulte du déficit des diamètres supérieurs à 50cm sur la placette 2 par rapport à la placette 1 (Append. 7). Sur un transect de 20 m entre les deux placettes (appelé corridor), 21 hêtres adultes sur ont été également cartographiés et échantillonnés.

Une station météo est installée au cœur de la réserve intégrale depuis 1959 (pluviomètre totalisateur et thermomètre pour les minima et maxima). Une station automatique est présente depuis 1996. Des travaux de la Massane consacrés à la climatologie (analyse des données) paraissent annuellement depuis 1984. Ces travaux avaient mis évidence un effet particulièrement marqué sur le hêtre de la conjonction prolongée de sécheresse et de fortes chaleurs de l'année 2003, et suggéré de suivre dans le temps les impacts de ces effets, au travers en particulier de suivis individuels de différents symptômes.

Ainsi, sur un sous échantillon des arbres adultes (Placette 1 : 169 et Placette 2 : 248), **des symptômes visibles de dépérissement foliaire ont été suivis qualitativement pendant 11 ans (2002-2012)**. Au minimum deux contrôles annuels sont effectués, pendant lesquels les cas récents de mortalité de branches de l'année (feuilles desséchées encore visibles sur branches et rameaux) sont relevés. Cette présence de branches mortes dans la partie haute du houppier (descente de cime) a été retenue pour suivre le déficit foliaire. L'observation repose sur l'absence (0) /présence (1) du symptôme quand il est significatif. Le degré n'est pas précisé, mais le cumul des observations sur les dix dernières années permet une certaine finesse dans l'interprétation : intensité du stress subi, fréquence, progressivité du dépérissement et possible récupération de l'individu. Nous avons donc analysé la note individuelle qualitative DEFcum, obtenue en sommant les notes annuelles de dépérissement. DEFcum varie de 0 à 9, avec une moyenne de  $\mu_{\text{DEFcum}} = 1.43$  sur les deux placettes ( $\mu_{\text{DEFcum}} = 1.53$  sur la placette 1 et  $\mu_{\text{DEFcum}} = 1.3$  sur la placette 2 ; Append. 8 a et b).

**La présence du champignon lignicole *Oudemansiella mucida* a également été suivie** sur la base d'un examen macroscopique: absence (0) / présence (1) d'au moins un sporophore. Il y a donc des cas d'attaque fongique qui échappent aux contrôles. Pour le champignon lignicole, une note qualitative de 1 a été attribuée chaque année aux arbres sur lesquels *O. mucida* est présente, et de 0 sinon. Le cumul sur les 10 années de suivi de la présence de *O. mucida* permet donc d'obtenir une note individuelle qualitative OUDcum, qui varie de 0 à 10, avec une moyenne de  $\mu_{\text{OUDcum}} = 0.29$  sur les deux placettes ( $\mu_{\text{OUDcum}} = 0.07$  sur la placette 1 et  $\mu_{\text{OUDcum}} = 0.15$  sur la placette 2 ; Append. 9).



Figure 15 : Etapes de prélèvement des feuilles de hêtres chez les adultes et les plantules ou semis.

# ÉCHANTILLONNAGE DU MATÉRIEL BIOLOGIQUE ET MARQUAGE DES INDIVIDUS

---

Les campagnes de terrain ont été effectuées entre juin et juillet 2013, pour prélever du matériel biologique (feuille) pour le génotypage sur deux cohortes (parents et semis) et marquer les semis

Les 684 adultes échantillonnés (Placette 1 : 266 adultes ; Placette 2 : 397 adultes; Corridor : 21 adultes ; Tableau 1) sont situés : (1) soit à l'intérieur de la réserve intégrale et font partie du suivi forestier engagé depuis 2002 (438 individus), (2) soit hors de cette zone de suivi et ont été cartographiés pour les besoins de cette étude (246 individus). 21 hêtres adultes sur un transect de 20 m entre les deux placettes ont également été cartographiés et échantillonnés. Ainsi, le suivi forestier est actualisé avec ajout d'une entrée associée au génotypage pour chacun des adultes sélectionnés.

Au milieu de chaque placette et dans la réserve intégrale, 200 (Placette 1) et 165 (Placette 2) semis ont été échantillonnés non exhaustivement (Tab. 1). Sur chaque placette, ont été distingués les «semis» (qui ont germé au printemps 2012) et les «plantules» (d'âge variant entre 2 et 30 ans en 2012). L'âge des plantules est estimé par le comptage des cicatrices de croissance (Fig. 15). Semis et plantules ont été marqués à l'aide d'étiquettes plastiques à tenue permanente (Fig. 15) et leur position dans le quadrillage du suivi forestier est consignée.

Deux feuilles de chaque hêtre adulte sont prélevées pour les analyses de biologie moléculaire. Lorsque les feuilles des reproducteurs potentiels sont inaccessibles, elles sont prélevées en sectionnant de petits rameaux à l'aide d'un échenilloir ou bien à l'aide d'un sac de lancer (techniques de grimpeurs élagueurs). Le poids relié à une cordelette est lancé à travers la frondaison quand celle-ci n'est pas accessible. La cordelette passe au-dessus d'un rameau. Il suffit ensuite de la tirer pour casser le rameau et le récupérer au sol (Fig 15). Une seule feuille a été prélevée pour les semis, deux pour les plantules. Chaque échantillon reçoit un numéro d'identification propre à la présente étude de génotypage. Le matériel biologique de chaque individu est placé dans un sachet de thé numéroté, lui-même placé dans un sachet plastique contenant du silicagel et à fermeture étanche. Ce mode de stockage assure la dessiccation complète du tissu végétal et une préservation optimale de l'ADN (Fig. 15).

Tableau 9 : Diversité et qualité des marqueurs microsatellites utilisés.

N est le nombre total d'individus lus et %DM le pourcentage de données manquantes. Na est le nombre d'allèles différents, RA la richesse allélique pondérée (le nombre d'allèles dans un échantillon aléatoire de 36 individus), Ne le nombre efficace d'allèles. He (l'hétérozygotie attendue) est l'indice diversité génétique de Nei. Les fréquences des allèles nuls (fnul) ont été estimées dans chaque combinaison de placette/cohorte (P1R=Placette 1, reproducteur ; P2R=Placette 2, reproducteur ; P1S=Placette 1, semis ; P2S=Placette 2, semis) et seules les valeurs significatives sont données. Les probabilités d'exclusion en analyse de maternité (PE1P) et de parenté (PEPP) sont aussi données, avec leur valeur cumulée pour les 21 marqueurs sur la dernière ligne. En vert les 4 marqueurs présentant des fréquences d'allèles nuls non négligeables. Ils sont exclus de l'analyse.

Marqueur	N	%DM*	Na	RA	Ne	He	f(null)				PE-1P	PE-PP
							P1R	P2R	P1S	P2S		
Csolfagus_19	956	8.8%	12	8.07	5.69	0.82		0.01			0.52	0.17
Csolfagus_7	726	30.7%	6	5.04	4.50	0.78	0.03	0.02	0.07		0.63	0.27
F1_15	928	11.5%	18	9.83	4.77	0.79	0.02				0.56	0.18
Fs3_4	821	21.7%	4	2.69	2.02	0.50					0.87	0.69
Sfc0007	854	18.5%	7	5.21	4.20	0.76		0.01	0.03		0.65	0.29
Sfc1143	746	28.8%	10	7.05	3.17	0.68	0.01	0.05	0.01		0.70	0.31
Csolfagus_25	906	13.5%	6	4.73	2.17	0.54		0.01			0.85	0.53
Csolfagus_29	908	13.4%	5	4.04	2.68	0.63					0.79	0.45
Csolfagus_31	742	29.2%	11	6.65	3.17	0.68			0.03	0.05	0.72	0.36
Csolfagus_6	1025	2.2%	10	6.66	4.23	0.76	0.01				0.63	0.27
Fi05	297	71.7%	7	5.36	1.99	0.50	0.02	0.04		0.15	0.86	0.54
Mfc7	915	12.7%	7	5.76	1.98	0.49		0.04			0.86	0.50
sfc061	559	46.7%	13	9.72	5.19	0.81	0.04	0.02		0.01	0.54	0.18
concat14_A_0	689	34.3%	6	5.15	3.50	0.71	0.02				0.70	0.35
DE576	552	47.3%	6	4.55	3.16	0.68					0.74	0.41
DUKCT_A_0	920	12.2%	5	4.97	2.66	0.62	0.01			0.01	0.78	0.41
DZ447_A_0	1019	2.8%	6	5.46	3.70	0.73					0.68	0.31
EEU75	891	15.0%	9	5.26	1.91	0.48	0.07				0.88	0.55
EJV8T	1001	4.5%	8	4.53	2.61	0.62	0.05	0.08	0.05	0.02	0.80	0.50
EMILY_A	869	17.1%	6	5.50	3.06	0.67	0.01			0.02	0.74	0.39
ERHIBI_A_0	718	31.5%	6	3.58	2.75	0.64	0.02	0.10	0.02		0.80	0.49
Moyenne	811.5	22.6%	8.0	5.7	3.29	0.66						
Cumulée											0.001	5.8 10 <sup>-10</sup>

## Génotypage

L'ADN de l'ensemble des individus (n=1048) a été extrait à l'aide du kit DNeasy 96 Plant Kit (Qiagen). Deux témoins (un positif et un négatif) sont présents dans chacune des plaques. Un jeu de 21 marqueurs microsatellites (Append. 1) a été utilisé pour génotyper tous les individus. Ces marqueurs ont été développés ou déjà testés au laboratoire sur le Hêtre commun. Les 21 marqueurs microsatellites ont d'abord été amplifiés par PCR. Vu les conditions techniques (taille des fragments, séquenceur), 3 PCR sont nécessaires à l'amplification des 21 marqueurs (PCR multiplex). La révélation des marqueurs a été réalisée à l'aide d'un séquenceur ABI 3730xl (LifeTechnologies) à 48 capillaires. La lecture des génotypes s'est faite sous le logiciel GeneMapper (LifeTechnologies). Afin de minimiser les erreurs de lecture, une double lecture des génotypes a été réalisée par 2 personnes indépendamment, et ces deux lectures sont ensuite confrontées.



Le jeu de données final comporte **un taux de données manquantes (22.6%) pour l'ensemble des 1048 individus et des 21 marqueurs** (Tab. 9). Ce taux varie fortement entre marqueurs : 4 marqueurs ont un taux de données manquantes inférieur à 10%; inversement, le locus Fi05 a un taux de données manquantes supérieur à 70%. Les marqueurs sfc061 et DE576 ont été lus chez un individu sur 2. Un seul échantillon ne donne aucun résultat sur les 1048 individus analysés. Par ailleurs, 7 paires d'individus partagent le même génotype à tous les locus ; pour la plupart il s'agit d'arbres extrêmement proches, voire présentant des tiges fusionnées à leur base (correspondant à différents brins d'une même cépée).

L'analyse de la variabilité génétique globale sur l'ensemble des 1048 individus (Tab. 9) montre que tous les marqueurs sont polymorphes, avec un nombre d'allèles moyen  $N_a=8$  allèles/marqueur, correspondant à une richesse allélique pondérée de 5.7 allèles (pour 36 individus). Le marqueur FS3\_4 présente la plus faible richesse allélique avec  $N_a=4$  allèles et  $RA=2.7$ . Les marqueurs Fs1\_15 et sfc061 présentent la plus forte richesse allélique avec  $N_a= 18$  et 13 allèles, et  $RA=9.8$  et 9.7, respectivement. L'indice de diversité de Nei est en moyenne de  $He=0.66$ . Trois marqueurs présentent une valeur de  $He$  inférieure à 0.50 : EEU75 ( $He= 0.48$ ), Mfc7 ( $He=0.49$ ) et Fi05 ( $He = 0.50$ ). Cette faible diversité n'est pas forcément associée à une richesse allélique plus faible que la moyenne ( $RA=5.7$  pour Mfc7), mais elle résulte surtout du déséquilibre des fréquences des différents allèles. A l'opposé, le marqueur avec la plus forte diversité de Nei est CsolFagus\_19 ( $He=0.82$ ), avec  $N_a=12$ ,  $RA=8$ , et des fréquences alléliques relativement équilibrées (pas d'allèle majoritaire). Le nombre efficace d'allèles par locus ( $N_e$ ), qui intègre ces deux composantes de la diversité (richesse allélique et diversité de Nei) est en moyenne de  $N_e= 3.3$  allèles efficaces/marqueur. Il varie entre 1.91 (EEU75) et 5.7 (CsolFagus\_19).

**Les fréquences d'allèles nuls varient entre marqueurs, et aussi entre les combinaisons de placette ou de cohortes.** Quatre marqueurs (CsolFagus\_7, EEU75, EJV8T et ERHIBI\_A\_0) présentent des fréquences d'allèle nuls non négligeables ( $f_{null}>0.05$ ) dans au moins une des combinaisons placette/cohorte considérées (Tab. 5). Mais ces fréquences ne sont jamais très élevées ( $f_{null}<0.10$ ), sauf pour le marqueur Fi05 ( $f_{null}=0.15$ ), à cause du grand nombre de données manquantes.

Par ailleurs, la puissance théorique du jeu de 21 marqueurs pour les analyses de maternité/parenté est très bonne, avec une probabilité multilocus  $PE_{1P}=0.001$  d'exclure un individu non apparenté dans l'analyse de maternité et une probabilité  $PE_{PP} = 5.8 \cdot 10^{-10}$  d'exclure un couple d'individus non apparentés dans l'analyse de parenté. Si l'on considère le jeu de 17 marqueurs (en excluant CsolFagus\_7, EEU75, EJV8T et ERHIBI\_A\_0), la qualité reste élevée, avec  $PE_{1P}=0.0028$  et  $PE_{PP} = 2 \cdot 10^{-8}$ .



Enfin, 7 paires d'individus de la placette 2 présentent un génotype identique (clones issus de recépage). Le nombre d'individus uniques considérés pour la plupart des analyses sur la placette 2 est donc 390 au lieu de 397.

## ANALYSES DE DIVERSITÉ ET DE STRUCTURE GÉNÉTIQUE SPATIALE

---

Le logiciel GenAlex a été utilisé pour estimer les indices classiques de diversité génétique : hétérozygotie observée ( $H_{obs}$ ) et attendue en panmixie ( $H_e$ , aussi appelée indice de diversité de Nei), et le nombre d'allèles observé ( $N_a$ ) et efficace ( $N_e$ ). La différenciation génétique entre placettes (1 et 2) et entre cohortes (adultes et semis) a été estimée par le  $F_{ST}$  de Weir et Cockerham (1984), calculé avec le logiciel FSTAT (Goudet, 2000). Les fréquences des allèles nuls ont été estimées par la méthode de maximum de vraisemblance proposée par Kalinowski et Taper (2006) et implémentées dans le ML-Null.

La structure génétique spatiale (SGS) a été caractérisée par l'analyse de l'apparentement génétique entre paires d'individus  $i$  et  $j$  (estimée par le coefficient  $F_{ij}$  de Loiselle et coll. (1995) en fonction de la distance géographique  $d_{ij}$  les séparant. Pour visualiser la SGS, les valeurs moyennes des coefficients  $F_{ij}$  par classe de distance ont été représentées en fonction de la distance géographique (auto-corrélogramme). Pour tester la SGS, le coefficient de régression ( $b_{log}$ ) de  $F_{ij}$  sur le logarithme de  $d_{ij}$  a été calculé ; l'hypothèse d'absence de SGS a été testée par 5000 permutations des positions spatiales des individus, pour obtenir la distribution de  $b_{log}$  sous l'hypothèse nulle d'absence de corrélation entre  $F_{ij}$  et  $d_{ij}$ . Enfin, l'intensité de la SGS a été caractérisée par  $Sp = b_{log}(F1-1)$ , ou  $F1$  est la valeur moyenne du coefficient  $F_{ij}$  dans la première classe de distance. Tous ces calculs ont été effectués avec le logiciel SPAGED1 (Hardy & Vekemans, 2002).

# ESTIMATION INDIRECTE DES FLUX DE GÈNES

---

Le modèle théorique d'isolement par la distance montre que la pente de la régression ( $b_{log}$ ) peut être utilisée pour estimer le produit de la densité efficace de reproducteur ( $d_e$ ) et de la variance axiale de dispersion des gènes ( $\sigma^2$ ), à travers la relation suivante :

$$4pd_e\sigma^2 = \frac{1}{b_{log}} \cdot \left[ \frac{F_1}{F_0} \right] \quad \text{Eq (1)}$$

Le produit  $4pd_e\sigma^2$  est aussi appelé taille de voisinage. La variance axiale de dispersion des gènes ( $\sigma^2$ ) reflète l'intensité des flux de gènes\* historiques dans la population. La densité efficace de reproducteurs  $d_e$  reflète la force de la dérive génétique à l'œuvre dans la population. Pour estimer  $\sigma^2$ , nous avons utilisé la procédure itérative implémentée dans SPAGEDI (Hardy & Vekemans, 2002).

La valeur de  $\sigma$  estimé reflète le flux de gènes\* totaux, qui peut être décomposé en flux de gènes par graine et flux de gènes par pollen selon la relation suivante : (Crawford, 1984)

$$s^2 = s_s^2 + \frac{1}{2}s_p^2 \quad \text{Eq (2)}$$

Avec  $s^2$  la distance moyenne de dispersion des gènes,  $s_s^2$  la variance axiale de dispersion des graines et  $s_p^2$ , la variance axiale de dispersion du pollen.

# ANALYSE DE PARENTÉ/MATERNITÉ

---

Nous avons tenté de retrouver les parents de chaque semis échantillonné, afin de caractériser d'une part les flux de pollen et de graines efficaces, et d'autre part les succès reproducteurs mâles et femelles réalisés. Les assignations de parenté et maternité ont été réalisées par une méthode de maximum de vraisemblance implémentée dans le logiciel CERVUS (Marshall *et coll.*, 1998).

## Principe

Le logiciel CERVUS propose deux analyses utilisant les génotypes de descendants et de leurs parents potentiels pour rechercher pour chaque descendant soit le meilleur parent compatible (analyse de maternité), soit le meilleur couple de parents compatibles (analyse de parenté). Dans l'analyse de maternité, la **probabilité que chaque parent potentiel considéré soit le vrai parent** par rapport à un individu pris au hasard dans la population est calculée à partir des génotypes du descendant, du parent potentiel et des fréquences alléliques dans la population. Dans l'analyse de parenté, la **probabilité que chaque couple de parents potentiels considéré soit le vrai couple de parents** par rapport à un couple d'individus pris au hasard dans la population est calculée à partir des génotypes du descendant, du couple de parents potentiels et des fréquences alléliques dans la population. La décision d'assignation de maternité/parenté à un parent/couple de parent donné est basée sur le logarithme de cette probabilité relative (appelée LOD Score), ou sur la différence de LOD-score (appelée D-score) entre le premier et le deuxième parent (respectivement couple de parents) les plus probables.

L'originalité de l'approche de Marshall et coll. (1998) est de **déterminer, par simulations en amont de l'assignation, la valeur critique du LOD-score ou du D-score au-delà de laquelle la maternité/parenté peut être attribuée à un parent/couple de parents**. La procédure de simulation consiste : (1) à générer des génotypes pour des parents et leurs descendants à partir des fréquences alléliques du jeu de données ; puis (2) à calculer pour chaque descendant simulé le LOD-score et le D-score de tous les parents potentiels (analyse de maternité) ou de tous les couples de parents potentiels (analyse de parenté). Le pourcentage de « vrais » parents ou couples de parents retrouvés par cette analyse à partir de données simulées permet de déterminer la valeur critique du D-score (appelée  $\Delta_{crit}$ ) au-delà de laquelle la maternité/parenté peut être attribuée à un parent/couple de parents. Ainsi, cette valeur critique peut être directement associée à un niveau de confiance donné; typiquement, le  $\Delta_{crit}$  à 80 % (respectivement 95%) est la valeur de D-score seuil tel qu'au moins 80% (respectivement 95%) des descendants assignés ont

bien été assignés à leur vrai parent/couple de parents. Cela signifie que l'erreur de Type I (le risque d'assigner un descendant à un parent/couple de parent non apparenté) est de 20% (respectivement 5%). Les simulations déterminant les valeurs de  $\Delta_{crit}$  peuvent tenir compte de différents paramètres impactant l'analyse de maternité/parenté : un taux d'erreur de génotypage (TErr), la taille totale de la population des parents potentiels et le pourcentage des parents génotypés (PropEch).

Après cette étape de simulation, **CERVUS tente d'assigner chaque descendant du vrai jeu de donnée au meilleur parent (assignation de maternité) ou au meilleur couple de parents (assignation de parenté)**. Pour un descendant donné, l'analyse de maternité consiste à calculer le LOD-score de tous les parents potentiels, et à comparer le D-score du meilleur parent à la valeur seuil  $\Delta_{crit-MAT}$  déterminée par simulations ; il y a trois conclusions possibles à cette comparaison (Tab. 10A). Il faut noter que la dénomination « analyse de maternité » est arbitraire, dans la mesure où l'on ne peut généralement pas savoir si le parent assigné est la mère ou le père du descendant (à moins d'utiliser par exemple des marqueurs génétiques à hérédité maternelle

Tableau 10A : Décisions possibles de l'analyse de maternité pour chacun des descendants considérés.

Décision	Résultat de l'analyse de maternité	Conclusion
<b>M1</b>	Aucun parent compatible n'est trouvé (aucun LOD-score >0)	Le descendant est considéré comme provenant d'une mère à l'extérieur de la placette (événement de migration)
<b>M2</b>	Un ou plusieurs parents compatibles sont trouvés, mais le D-score du meilleur parent est inférieur à $\Delta_{crit-MAT}$ . Cela peut être soit parce que 2a) son LOD-score est faible, soit parce que 2b) les LOD-scores des deux meilleurs parents sont élevés.	Le descendant est considéré comme provenant d'une mère indéterminée à l'extérieur ou à l'intérieur de la placette ( <b>non-assignation</b> ).
<b>M3</b>	Le meilleur (ou le seul) parent présente un LOD-score suffisamment fort pour que le D-score soit supérieur à $\Delta_{crit}$	Le descendant est assigné à ce parent ( <b>assignation catégorique</b> ).

ou paternelle).

De même, l'analyse de parenté consiste à calculer le LOD-score de tous les couples de parents potentiels, et à comparer le D-score du meilleur couple à la valeur seuil  $\Delta_{\text{crit-PP}}$  déterminée par simulations ; il y a trois conclusions possibles à cette comparaison (Tab. 10B). Il faut noter que la différence fondamentale entre analyse de maternité et de parenté est que dans la première, on teste chaque couple {descendant ; parent potentiel}, alors que dans la deuxième, on teste chaque trio {descendant ; parent potentiel1 ; parent potentiel2} ; si l'analyse de parenté assigne au descendant  $i$  un couple de parents  $\{P1_i ; P2_i\}$ , et que l'analyse de maternité assigne à ce même descendant la mère  $\{P\text{Seul}_i\}$ , on s'attend donc à ce que  $\{P\text{Seul}_i\}$  soit inclut dans le couple  $\{P1_i ; P2_i\}$ .

**Tableau 10B : Décisions possibles de l'analyse de parenté pour chacun des descendants considérés**

Décision	Résultat de l'analyse de parenté	Conclusion
<b>P1</b>	Aucun couple de parents compatible n'est trouvé (aucun LOD-score >0)	Au moins un des deux parents du descendant est considéré comme provenant de l'extérieur de la placette (événement de migration)
<b>P2</b>	Un ou plusieurs couples de parents compatibles sont trouvés, mais le D-score du meilleur couple est inférieur à $\Delta_{\text{crit-PP}}$ . Cela peut être soit parce que 2a) son LOD-score est faible, soit parce que 2b) les LOD-scores des deux meilleurs couples sont élevés.	Le descendant est considéré comme provenant d'un couple indéterminé à l'extérieur ou à l'intérieur de la placette (non-assignation).
<b>P3</b>	Le meilleur (ou le seul) couple de parents présente un LOD-score suffisamment fort pour que le D-score soit supérieur à $\Delta_{\text{crit-PP}}$ .	Le descendant est assigné à ce couple de parents (assignation catégorique).

## Procédure suivie et calcul des distances de dispersion du pollen et des graines

Les 365 descendants échantillonnés sur les 2 placettes ont été analysés conjointement, en cherchant leurs parents dans l'ensemble des 677 adultes des deux placettes. Nous avons combiné analyse de maternité et de parenté pour arriver pour chaque descendant à une des 4 décisions suivantes (Tableau 10C):

1. Le descendant provient de parents à l'extérieur de la placette (événement de migration par graine) ;
2. Le descendant provient d'une mère  $P_m$  à l'intérieur des placettes (**graine locale**), et d'un père non échantillonné, à l'extérieur des placettes (événement de migration par pollen) ;
3. Le descendant provient d'un couple de parents à l'intérieur des placettes (**graine et pollen locaux**).
4. Aucune décision ne peut être prise concernant l'origine du pollen et de la graine (**non-assignation**).

Tableau 10C : Décisions possibles de la combinaison des deux analyses effectuées dans cette étude

Décision	Résultat de l'analyse de maternité		Résultat de l'analyse de parenté	Conclusion
<b>D1</b>	M1 : Aucun parent compatible n'est trouvé (aucun LOD-score >0)	ET	P1 : Aucun couple de parents compatible n'est trouvé (aucun LOD-score >0)	Le descendant provient de 2 parents à l'extérieur de la placette (événement de migration par graine)
<b>D2</b>	M3 : Le meilleur (ou le seul) parent $P_m$ présente un LOD-score suffisamment fort pour que son D-score soit supérieur à $\Delta_{crit}$	ET	P1 :Aucun couple de parents compatible n'est trouvé (aucun LOD-score >0)	Le descendant provient de la mère $P_m$ ( <b>graine locale</b> ) et d'un père non échantillonné, à l'extérieur de la placette (événement de migration par pollen)
<b>D3</b>	Non considéré		P3 : Le meilleur (ou le seul) couple de parents présente un LOD-score suffisamment fort pour que son D-score soit supérieur à $\Delta_{crit-PP}$	Le descendant provient de ce couple, l'arbre le plus proche étant la mère $P_m$ , et le plus éloigné le père $P_p$ ( <b>graine et pollen locaux</b> ).
<b>D4</b>	M2 : Un ou plusieurs parents compatibles sont trouvés, mais le D-score du meilleur parent est inférieur à $\Delta_{crit-MAT}$ .	OU	P2 : Un ou plusieurs couples de parents compatibles sont trouvés, mais le D-score du meilleur couple est inférieur à $\Delta_{crit-PP}$ .	Aucune décision ne peut être prise concernant l'origine du pollen et de la graine ( <b>non-assignation</b> ).



Notons que dans le cas D3, l'analyse de maternité peut avoir elle aussi identifié un parent présentant un  $D_{\text{-score}} > \Delta_{\text{crit-MAT}}$  et que ce parent peut appartenir ou non au couple de parents identifié par l'analyse de parenté. Nous avons compté le nombre  $N_{\text{conflit}}$  de cas où analyses de parenté et de maternité n'étaient pas cohérentes, et nous avons donné la priorité à l'analyse de parenté (car le risque d'assigner un descendant à un faux couple de parents compatibles est plus faible que le risque d'assigner un descendant à un faux parent compatible).

À partir de tous les descendants pour lesquels la décision D2 ou D3 a été prise, la distance efficace de dispersion des graines ( $d_{\text{eff-graine}}$ ) a été calculée comme la moyenne des distances entre chaque descendant et son parent le plus proche. À partir de tous les descendants pour lesquels la décision D3 a été prise, la distance de dispersion efficace du pollen ( $d_{\text{eff-pollen}}$ ) a été calculée comme la moyenne des distances entre les deux parents assignés.

### Analyse de sensibilité des assignations

Les analyses de parenté/maternité sont notoirement sensibles à différents facteurs, notamment: le sous-échantillonnage de la population des parents potentiels, la présence d'erreur de génotypage, le mauvais comportement de certains marqueurs (présence d'allèles nulles). Nous avons adopté une démarche d'analyse de sensibilité pour étudier comment l'incertitude sur ces paramètres affectait les décisions d'assignations et les paramètres d'intérêt (notamment la distance de dispersion des graines et du pollen).

D'une part, les simulations et les assignations de maternité et de parenté ont été réalisées en jouant sur deux paramètres : le pourcentage de parents génotypés (PropEch= 100,95 ou 80%) et le taux d'erreur de génotypage (TErr = 0 ou 3 ou 5%). Supposer que PropEch =100% revient à faire l'hypothèse que tous les parents potentiels ont été échantillonnés, et à accorder une confiance maximale aux assignations à un parent/couple de parent échantillonné. A l'inverse, en considérant PropEch =80% (c-a-d 20% des parents extérieurs à la Placette étudiée), on tient compte du fait qu'un parent/couple de parents génétiquement compatibles avec un semis, mais partageant avec lui des allèles très fréquents, puisse ne pas être le vrai parent/couple de parents, parce que celui-ci n'a pas été échantillonné. De la même façon, supposer un taux d'erreur de génotypage nul (TErr = 0%) revient à accorder une confiance maximale aux lectures de génotypes effectuées ; nous avons considéré par ailleurs TErr =5% (correspondant au taux d'erreur calculé au laboratoire sur un projet similaire), et TErr =3% (intermédiaire). Pour chaque combinaison de paramètres (PropEch & TErr), des simulations d'analyse de maternité et de parenté ont été réalisées (1000 descendants simulés), pour déterminer les différentes valeurs de  $\Delta_{\text{crit-MAT}}$  (maternité) et de  $\Delta_{\text{crit-PAR}}$  (parenté), au niveau de confiance de 80 et 95%.

Outre les erreurs de génotypage, la présence d'allèles nuls\* à certains marqueurs microsatellites peut biaiser les analyses de maternité ou de parenté, en conduisant à rejeter le vrai parent (quand il partage l'allèle nul avec le descendant). Les allèles nuls sont fréquents avec les marqueurs microsatellites, et s'expliquent généralement par une mutation sur le site où l'amorce s'hybride qui empêche l'hybridation et donc l'amplification du microsatellite (Kalinowski et Taper, 2006). Les individus portant un allèle nul sont hétérozygotes  $A_iA_n$  ( $A_n$  étant l'allèle nul et donc non détecté), mais sont détectés comme homozygotes  $A_iA_i$ . En plus de la sous-estimation de la fréquence des hétérozygotes, la présence d'allèle nul est susceptible de biaiser les analyses de paternité. En effet, le croisement entre une mère de génotype  $A_iA_i$  avec un père de génotype  $A_jA_n$  a 50% de chance de produire un descendant de génotype  $A_iA_n$ , apparemment incompatible avec son père. Une étude par simulations (Dakin & Avise, 2004) a montré que ce risque d'erreur de Type II (rejeter le vrai parent) augmente de façon quasi-linéaire avec la fréquence  $v_n$  des allèles nuls quand  $0 < v_n < 0.4$  (Dakin et Avise, 2004). L'erreur de Type II (rejeter le vrai parent) peut ainsi avoir une probabilité supérieure à 0.10 quand la fréquence des allèles nuls dépasse 0.20. Nous avons étudié l'effet des allèles nuls sur les assignations et leurs résultats en réalisant les simulations et les assignations avec un jeu de marqueurs réduit, en excluant de l'analyse les marqueurs les plus problématiques (ceux présentant des fréquences d'allèles nuls > 5%).

Pour juger de la sensibilité des résultats des assignations à ces différents paramètres, nous avons calculé dans chaque cas la proportion de descendants assignés à un parent/couple de parents, le nombre  $N_{\text{conflit}}$  de cas où analyses de parenté et de maternité n'étaient pas cohérentes (Tab. 10C), ainsi que la distance moyenne de dispersion efficace des graines et du pollen ( $d_{\text{eff-graine}}$ ,  $d_{\text{eff-pollen}}$ ). En considérant comme référence (*ref*) les résultats des assignations avec le jeu de paramètres [PropEch=100% ; TErr=0% ; 21 marqueurs], nous avons calculé l'homologie de chacun des autres jeux de résultats des assignations (*comp*) avec le jeu de résultats *ref* de la façon suivante :

$$\text{Eq(3)}$$

Où  $N_{\text{desc}}$  est le nombre de descendants dans le fichier *ref*,  $u=1$  (resp.  $v=1$ ) si le parent  $P1_i$  (resp.  $P2_i$ ) du descendant  $i$  dans le jeu *ref* est retrouvé dans le jeu *comp* et 0 sinon ;  $n_{\text{com}[0-P1]}$  est le nombre de descendants (avec au moins un parent assigné  $P1$ ) communs aux jeux de données *ref* et *comp* (quel que soit  $P1$ ), et  $n_{\text{com}[0-P2]}$  est le nombre de descendants (avec un deuxième parent  $P2$  assigné) commun aux jeux de données *ref* et *comp* (quel que soit  $P2$ ).

**Tableau 11 : Analyse de sensibilité des résultats des assignations de maternité/parenté aux paramètres de l'analyse (PropEch = le taux d'échantillonnage des parents ; Terr= le taux d'erreur de géotypage ; Nm<sub>q</sub>= nombre de marqueurs)**  
 Pour chaque niveau de confiance (NC), sont donnés le seuil critique de D-score (Δ<sub>crit</sub>) et le nombre de descendants assignés par l'analyse de maternité (Nm) et de parenté (Np), ainsi que le nombre de conflits (conf) entre ces deux analyses (cas où la mère assignée ne fait pas partie du couple de parents assignés). À partir des Nut assignations (ie de la somme Np + Nm\*, avec Nm\* le nombre de descendants assignés seulement à une mère, mais pas à un couple de parents), les distributions des distances de dispersion réalisées du pollen et des graines ont été calculées, et sont résumées par les moyennes (μ) et médiane (med) de ces distributions. Enfin, le pourcentage d'homologie de chaque jeu de résultats d'assignation par rapport au jeu de données de référence (1ère ligne) est donné : % glob est le pourcentage global de parents communs, % Mère est le pourcentage de mères communes et % Père est le pourcentage de pères communs aux deux jeux de résultats comparés.

Paramètres			Résultats des assignations					Distance de dispersion					% homologie					
Nmq	PropEch	Terr	NC	Δ <sub>crit</sub>	Nm	Δ <sub>crit</sub>	Np	conf	Nm*	Nut	Med	μ	Med	μ	Pollen (d <sub>eff-pollen</sub> )	% glob	% Mère	% Père
21	100%	0%	95%	0	361	0	274	83	87	361	15.6	30.0	42.8	70.9		Réf.	100%	
21	95%	0%	80%	0	361	0	274	«	«	«	«	«	«	«	«	100%	100%	«
21	80%	0%	80%	0	361	0	274	«	«	«	«	«	«	«	«	«	«	«
21	80%	0%	95%	1.13	193	0	274	24	64	338	15.6	30.0	42.8	70.9		100%	100%	100%
21	100%	3%	80%	0	361	0	274	«	«	«	«	«	«	«	«	100%	100%	«
21	100%	5%	80%	1.55	222	0	365	13	0	365	14.5	26.7	46.8	75.06		74%	66%	73%
21	100%	3%	80%	0	365	0	365	?	«	«	«	«	«	«	«	«	«	«
21	100%	5%	80%	1.8	114	0.63	216	4	39	255	12.8	25.2	44.6	44.6		74%	67%	73%
21	95%	3%	80%	0	365	0	365	90	0	365	14.3	26.0	47.2	75.0		70%	63%	68%
21	95%	3%	80%	1.97	115	0.16	315	9	11	326	13.7	25.8	44.6	73.8		75%	68%	75%
21	80%	3%	80%	0	365	0	365	95	0	365	14.5	26.6	56.8	75.1		74%	66%	73%
21	80%	3%	80%	3.2	45	2.67	48	0	37	85	10.1	18.2	25.5	46.9		85%	82%	92%
17	100%	0%	80%	1.25	173	0.77	197	8	74	271	13.0	25.7	41.6	71.3		77%	71%	78%
17	100%	0%	80%	0.54	234	0	317	54	28	345	16.5	30.8	52.4	80.3		61%	60%	52%
17	100%	0%	80%	0	364	0	317	107	47	364	16.6	30.6	«	«		61%	59%	52%

## Résultats de l'analyse de sensibilité des assignations.

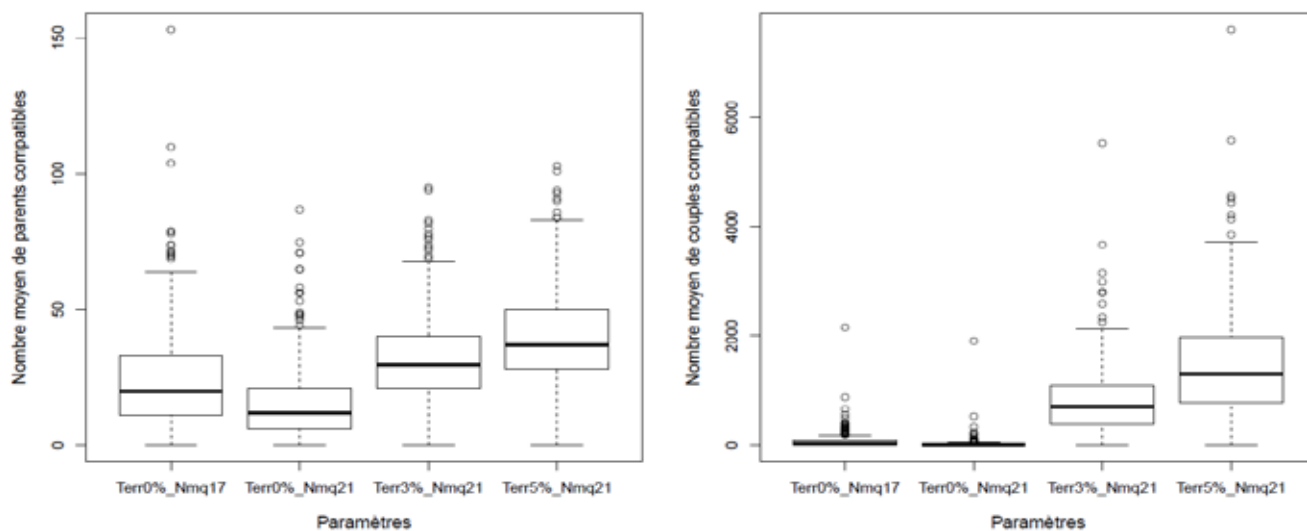
Le tableau 11 résume les résultats de l'analyse de sensibilité des assignations de maternité/parenté à trois paramètres d'intérêt : le nombre de marqueurs (Nm<sub>q</sub>), le taux d'échantillonnage des parents (PropEch) et le taux d'erreur de génotypage (Terr).

Le premier résultat à considérer est **la valeur du seuil critique de D-score** ( $\Delta_{crit}$ , déterminée par simulations) au-delà de laquelle la maternité/parenté peut être attribuée à un parent/couple de parents. Un  $\Delta_{crit} = 0$  signifie que le parent/couple de parents avec le meilleur LOD-score est le vrai parent/couple de parents pour au moins 80% ou 95% des descendants simulés (selon le niveau de confiance), quel que soit le LOD-score du second meilleur parent/couple de parents. Cela indique donc généralement une bonne puissance théorique de l'analyse de maternité/parenté (typiquement, seul le vrai parent/couple de parent a un LOD-score positif dans les simulations). Nos résultats montrent qu'au niveau de confiance de 80%,  $\Delta_{crit} = 0$  pour la plupart des paramètres de l'analyse de maternité ou parenté (sauf en combinant PropEch=80% avec Terr=3%). Notre jeu de marqueurs permet donc en théorie de retrouver le vrai parent/couple de parents pour au moins 80% des descendants (simulés), et ce vrai parent/couple de parents est simplement celui avec le LOD-score le plus positif. Cette bonne puissance théorique des analyses de maternité/parenté est cohérente avec les faibles probabilités d'exclusion  $PE_{1P}$  et  $PE_{PP}$ .

Le deuxième résultat à considérer est **la proportion de descendants du vrai jeu de donnée effectivement assignés à une mère ( $N_m$ ) ou à un couple de parents ( $N_p$ )**. Comme  $\Delta_{crit}$  prend la même valeur (0) au niveau de confiance de 80% pour différentes combinaisons de paramètres considérés,  $N_m$  et  $N_p$  sont aussi relativement similaires (sauf en combinant PropEch=80% avec Terr=3%). Pour l'analyse de maternité, entre  $N_m=361$  (Terr=0%) et  $N_m=365$  (Terr>0%) descendants peuvent être assignés à un parent échantillonné (décision 1 dans Tab 3A). En particulier, dès que l'on considère la possibilité d'erreur de génotypage (Terr>0%), tous les descendants peuvent être assignés à un parent échantillonné (avec au maximum 3 locus/21 qui ne sont pas compatibles entre le descendant et un parent). Néanmoins, chaque descendant a en général un nombre élevé de parents compatibles (c'est-à-dire avec un LOD-score >0) : le nombre moyen de parents compatibles/ descendants varie entre 16.2 (Terr=0%, Nm<sub>q</sub>=21) et 40.1 (Terr=5%, Nm<sub>q</sub>=21) (Fig. 8). Les descendants non assignés (Terr=0%) n'ont aucun parent échantillonné compatible (décision 2 dans Tab 3A). Pour l'analyse de parenté, entre  $N_p=274$  (Terr=0%) et  $N_p=365$

(Terr>0%) descendants peuvent être assignés à un couple de parents échantillonnés (et ceux non assignés n'ont pas de couple échantillonné compatible). Comme en analyse de maternité, pour Terr>0%), tous les descendants peuvent être assignés à un parent échantillonné (avec au maximum 5 locus/21 non compatibles entre le descendant et son couple de parents). Chaque descendant a aussi un nombre élevé de couples de parents compatibles (c'est-à-dire avec un LOD-score >0) : le nombre moyen de couples de parents compatibles/ descendant varie entre 90.4 (Terr=0%, Nm<sub>q</sub>=21) et 1858.2 (Terr=5%, Nm<sub>q</sub>=21) (Fig. 16). Enfin, il y a un nombre relativement élevé de conflits (jusqu'à 107) entre analyse de maternité et de parenté, c'est-à-dire de cas où la mère assignée par l'analyse de maternité ne fait pas partie du couple de parents assignés par l'analyse de parenté.

Figure 16 : Distribution du nombre de parents (analyse de maternité) ou de couples de parents (analyse de parenté) compatibles en moyenne avec un descendant (c'est-à-dire avec un LOD-score positif pour ce descendant). Les valeurs moyennes sont données dans le tableau sous les graphiques.

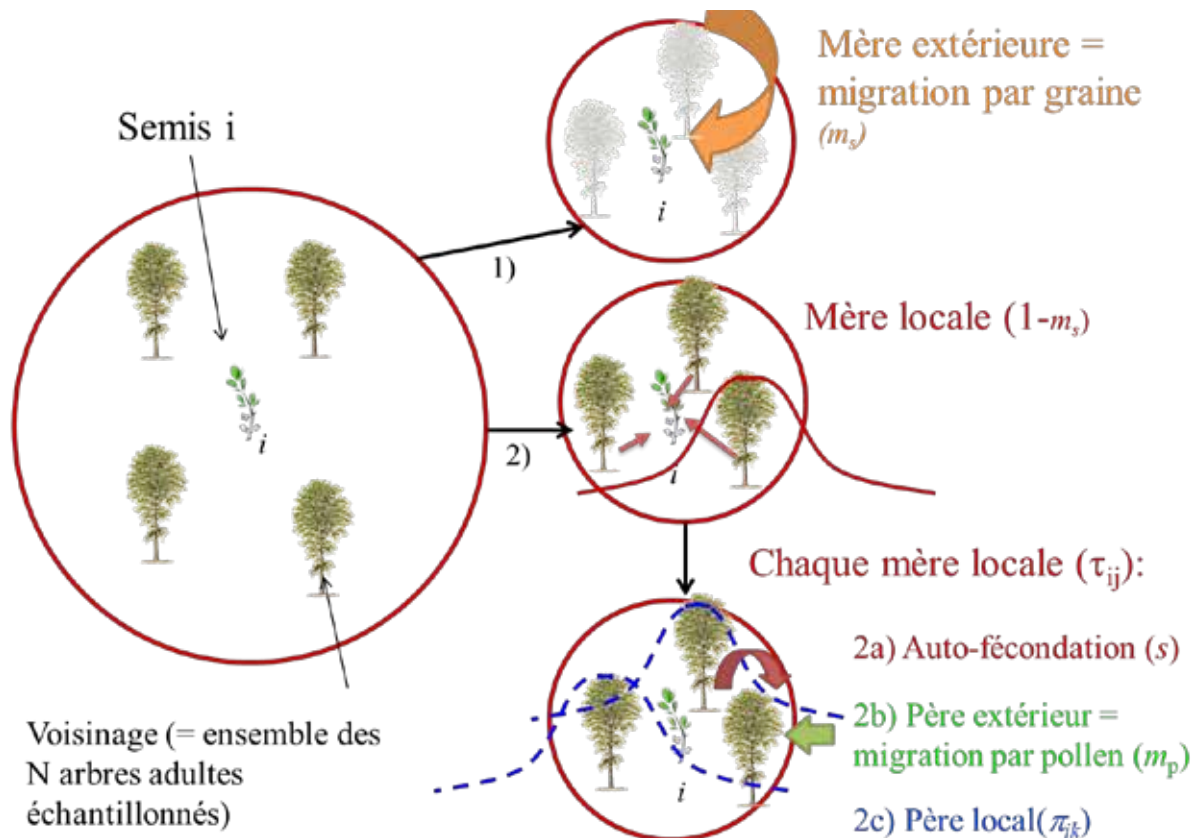


Terr	Nm <sub>q</sub>	Ncompat-Mat	Ncompat-PP
0	17	24.64	172.51
0	21	16.27	90.42
3	21	32.51	1101.21
5	21	40.11	1858.24

Au niveau de confiance de 95%, des différences substantielles apparaissent dans les valeurs de  $\Delta_{crit}$ ,  $N_m$  et  $N_p$  en fonction des paramètres PropEch, Terr et  $N_{mq}$ . Globalement,  $\Delta_{crit}$  augmente peu avec PropEch, plus fortement par Terr, et de façon encore plus marquée quand on combine PropEch<100% et Terr>0%. Cela se traduit par une réduction du nombre de descendants assignés à une mère (au minimum  $N_m=45$ ) ou à un couple de parents (au minimum  $N_p=48$ ), due à l'augmentation du nombre de décisions de non-assignation (i.e. décision 2 dans Tab 3 A et B). Ces résultats illustrent un dilemme classique du test d'assignation de parenté : si l'on veut minimiser le taux d'erreur de Type 1 (le risque d'assigner un descendant à un adulte non apparenté ; ici 5% pour CL=95%), le risque d'erreur de type 2 (celui de ne pas assigner un descendant à son parent échantillonné) augmente.

**Les distances efficaces de dispersion du pollen et des graines sont néanmoins peu sensibles aux erreurs éventuelles d'assignation, au nombre de marqueurs utilisés, ou au taux d'échantillonnage de la population de parents.** Pour les graines, la distance médiane de dispersion varie entre 10.1 et 16.5 m (et la distance moyenne entre 18.2 et 30 m). Pour le pollen, la distance médiane de dispersion varie entre 25.5 et 52.4 m (et la distance moyenne entre 46.92 et 80.3 m). Globalement, il n'y a pas de différence significative dans les distributions des distances de dispersion entre chaque jeu de résultats et le jeu de résultats de référence.

Figure 17 : Schéma du modèle de croisement dans un voisinage spatialement explicite (SEMM).





En revanche, les jeux de résultats d'assignation peuvent notablement différer dans le détail des filiations reconstruites pour chaque descendant. La similarité des couples {descendant; parent} retrouvés est de 75% entre les analyses qui considèrent un taux d'erreur de génotypage de 3% ou 5% et les analyses qui ignorent ce risque d'erreur. Cela signifie que pour 25% des descendants, les LOD-score calculés en tenant compte de l'erreur de génotypage changent suffisamment pour que les rangs des parents/couples de parents soient modifiés. Le taux d'échantillonnage des parents joue sur  $N_m$  et  $N_p$ , mais ne change jamais les LOD-score des parents, et donc les parentés/maternités reconstituées. **L'homologie la plus faible avec le jeu de résultat de référence (%glob=61%) est obtenue en enlevant de l'analyse les 4 marqueurs suspectés de présenter des allèles nuls.** La proportion importante de filiations ainsi affectées peut indiquer que la présence d'allèle nul augmente le taux d'erreur de type 2 (un vrai parent ou couple de parent n'est pas assigné, car il partage avec le descendant un allèle nul), et qu'enlever les quatre locus avec des allèles nuls permet de limiter ces erreurs. Mais cela peut aussi indiquer que la puissance des assignations est plus faible avec 17 marqueurs qu'avec 21, comme le montre l'augmentation du nombre de parents ou de couples de parents compatibles avec chaque descendant quand on enlève des 4 marqueurs de l'analyse (Fig. 16).

Tableau 12 : Paramètres du modèle SEMM estimés par l'analyse.

Symbole	Signification	Équation
$\bar{d}_s$	Distance moyenne de dispersion des graines	6,7
$b_s$	Paramètres de forme du noyau de dispersion des graines	6
$m_s$	Taux de migration des graines	8
$\bar{d}_p$	Distance moyenne de dispersion du pollen	6,7
$b_p$	Paramètres de forme du noyau de dispersion du pollen	6
$m_p$	Taux de migration du pollen	8
$s$	Taux d'autofécondation	8
$d_{obs}/d_{e-theo} \text{ ♀}$	Rapport de la densité observée de reproducteurs par voie maternelle à la densité efficace, estimé d'après la variance théorique des fertilités femelles ( $\sigma_F^2$ )	9
$d_{obs}/d_{e-theo} \text{ ♂}$	Rapport de la densité observée de reproducteurs par voie paternelle à la densité efficace, estimé d'après la variance théorique des fertilités mâles ( $\sigma_M^2$ )	9
$d_{obs}/d_{e-emp.} \text{ ♀}$	Rapport de la densité observée de reproducteurs par voie maternelle à la densité efficace, estimé d'après la variance des fertilités femelles individuelles (FF)	9
$d_{obs}/d_{e-emp.} \text{ ♂}$	Rapport de la densité observée de reproducteurs par voie maternelle à la densité efficace, estimé d'après la variance des fertilités male individuelles (FM)	9

# ESTIMATION DE LA FONCTION DE DISPERSION DU POLLEN ET DES GRAINES ET DES FÉCONDITÉS INDIVIDUELLES

Nous avons ensuite estimé les fonctions primaires de dispersion du pollen et des graines, et la distribution des fécondités individuelles. Pour cela nous avons utilisé le modèle de croisement dans un voisinage (Spatially Explicit Mating Model, SEMM), développé par Burczyk et al (2006) et adapté par Oddou-Muratorio et Klein (2008).

## Principe du SEMM

Le SEMM (Fig. 16) considère que chaque semis  $i$  échantillonné peut être issu d'une mère (1) hors du voisinage N avec la probabilité  $m_s$ , ou (2) dans le voisinage N avec la probabilité  $(1-m_s)$ . Dans ce deuxième cas, on considère que le semis  $i$  peut être issu (2a) d'autofécondation avec la probabilité  $s$ ; (2b) d'un père à l'extérieur du voisinage N avec la probabilité  $m_p$ , ou (2c) d'un père échantillonné du voisinage N avec la probabilité  $1-s-m_p$ . Dans le cas d'une dispersion locale des graines (cas 2), la probabilité de contribution de la mère  $j$  à la pluie de graines à la position du semis  $i$ ,  $\tau_{ij}$ , est explicitement modélisée en fonction de la distance entre  $i$  et  $j$ ,  $d_{ij}$ , (à travers le noyau de dispersion des graines) et de la fertilité individuelle de la femelle  $j$ ,  $FF_j$ :

$$\text{Eq(4)}$$

où  $p_g$  est la densité de probabilité du noyau de dispersion des graines.

Dans le cas d'un croisement par allo-fécondation (2c), la probabilité d'appariement entre la mère  $j$  et le père  $k$ ,  $\pi_{jk}$ , est explicitement modélisée en fonction de la distance entre  $j$  et  $k$ ,  $d_{jk}$ , (à travers le noyau de dispersion du pollen) et de la fertilité individuelle du mâle  $k$ ,  $FM_k$ :

$$\text{Eq(5)}$$

où  $p_p$  est la densité de probabilité du noyau de dispersion du pollen.

Le noyau de dispersion des graines est la fonction de probabilité qu'un arbre-mère situé à la position (0,0) ait émis une graine retrouvée à la position (x,y). Le noyau de dispersion du pollen est la fonction de probabilité qu'un arbre-père situé à la position (0,0) ait émis le grain de pollen ayant fertilisé une mère située à la position (x,y) (Oddou-Muratorio et Klein, 2008). Pour modéliser ces noyaux, nous avons utilisé la famille des exponentielles puissances :

$$p_{ep}(a,b;x,y) = \frac{b}{2pa^2} \exp\left(-\frac{r}{a}\right) \frac{r^b}{\Gamma(b)} \quad \text{Eq(6)}$$

avec  $r =$

Le paramètre  $a$  est un paramètre d'échelle, lié à la distance moyenne de dispersion  $\bar{\delta}$  par :

$$\bar{\delta} = a \frac{\tilde{A}^b}{\tilde{A}} \quad \text{Eq(7)}$$

Le paramètre  $b$  est un paramètre de forme. Ainsi, si  $b=1$ , la fonction de dispersion est une exponentielle et si  $b=2$ , alors il s'agit d'une fonction gaussienne. Pour  $b>1$ , on considère que le noyau est à queue légère, c'est-à-dire que la dispersion se fait majoritairement près de l'arbre émetteur. Pour  $b<1$ , on a un noyau à queue lourde, avec des événements de dispersion à moyenne et longue distance non négligeable.

Les fertilités femelle  $\mathbf{FF} = \{FF_j\}_{j=1,\dots,N}$ , (avec  $N$  : le nombre total de parents) et mâles  $\mathbf{FM} = \{F_k\}_{k=1,\dots,N}$ , suivent une loi Gamma de variance  $\sigma_F^2$  ou  $\sigma_M^2$ , cette distribution en  $L$  permet de bien prendre en compte les rares individus à forte fertilité.

La vraisemblance d'obtenir un semis de génotype  $G_i$  s'écrit de la manière suivante :

$$L(G_i | \mathbf{FF}, \mathbf{FM}, \sigma_F^2, \sigma_M^2, \bar{\delta}_s, b_s, \bar{\delta}_p, b_p, s, m_s, m_p) = m_s \times T_1(G_i | AF) + (1-m_s) \times \sum_{j,i} [s \times T_2(G_i | G_j) + m_p \times T_3(G_i | M_j, AF) + (1-s-m_p) \sum_{j,k} \pi_{j,k} \times T_4(G_i | M_j, F_k)]$$

Eq(8)

où  $AF$  sont les fréquences alléliques à l'extérieur de la population,  $T$  représente les probabilités mendéliennes d'obtenir le génotype  $G_i$  ( $T_1$ ) par le croisement entre une mère et un père tous les deux à l'extérieur, ( $T_2$ ) par l'autofécondation de la mère  $j$  au génotype  $g_j$ , ( $T_3$ ) par le croisement de la mère  $j$  avec un père extérieur ( $AF$ ), et ( $T_4$ ) pour le croisement de la mère  $j$  avec un père  $k$  au génotype  $g_k$ . Il est possible de tenir compte des erreurs de génotypage en considérant une probabilité mendélienne  $T_2$ ,  $T_3$  ou  $T_4$  non nulle d'obtenir le génotype  $G_i$  à partir d'un génotype parental a priori incompatible.

Cette approche permet d'estimer directement les différents paramètres d'intérêt (Tab. 12), c'est-à-dire les paramètres du régime de reproduction ( $s$ ,  $m_s$  et  $m_p$ ), les noyaux de dispersion des graines ( $\bar{\delta}_s$ ,  $b_s$ ) et du pollen ( $\bar{\delta}_p$ ,  $b_p$ ), les vecteurs des fertilités femelles ( $\mathbf{FF}$ ) et mâles ( $\mathbf{FM}$ ) individuelles, et leur variance ( $\sigma_F^2$ ,  $\sigma_M^2$ ). Cette estimation ne requiert pas d'assignation de paternité catégorique, mais passe par une analyse de parenté fractionnelle, où chaque père  $k$  a une probabilité d'avoir fécondé la mère  $j$  pour donner un descendant  $o$ . Plutôt que de donner les  $N$  estimations des fertilités individuelles (les vecteurs  $\mathbf{FF}$  et  $\mathbf{FM}$ ),

ou la variance de ces fertilités ( $\sigma_F^2$ ,  $\sigma_M^2$ ), qui dépend du choix de la fertilité moyenne, on peut résumer l'information sur ces paramètres via le rapport de la densité observée de reproducteurs à la densité efficace ( $d_{obs}/d_e$ ) ; un rapport proche de 1 traduit que les fertilités intrinsèques sont équilibrées. Le rapport  $d_{obs}/d_e$  peut être estimé à partir de la variance des fertilités mesurée par le paramètre de leur distribution théorique ( $\sigma_F^2$ ,  $\sigma_M^2$ ) de la façon suivante :

Eq(9)

Le rapport  $d_{obs}/d_e$  peut aussi être estimé à partir de la distribution empirique des fertilités individuelles. Nous avons estimé ces deux rapports ( $d_{obs}/d_{e-theo}$  et  $d_{obs}/d_{e-emp}$ ) à la fois sur la base des fertilités femelles et des fertilités males sur chaque placette.

### Estimation dans le cadre Bayésien

La méthode standard basée sur le maximum de vraisemblance ne serait pas adaptée à l'estimation de toutes les valeurs de  $\mathbf{FF} = \{FF_j\}_{j=1,\dots,N}$  ou  $\mathbf{FM} = \{F_k\}_{k=1,\dots,N}$  comme paramètres, ni à l'estimation directe des valeurs de  $\sigma_F^2$  et  $\sigma_M^2$  par intégration sur toutes les valeurs possibles de  $\mathbf{FF}/\mathbf{FM}$ . Une extension Bayésienne du modèle SEMM a été proposée (Klein *et coll.*, 2008) pour pouvoir estimer les fertilités individuelles. Le cadre Bayésien consiste à utiliser une distribution *a priori* pour les paramètres du modèle à estimer, puis à combiner ces distribution *a priori* avec les observations réalisées pour obtenir une distribution *a posteriori* pour ces paramètres. Ce cadre est résumé par le théorème de Bayes, qui s'écrit dans ce cas :

$$\mathbf{P}(\text{paramètres}|\text{observation}) = \frac{\mathbf{P}(\text{observation}|\text{paramètres})}{\mathbf{P}(\text{paramètres})}$$

[ Distribution *a posteriori* = Vraisemblance x Distribution *prior* des paramètres / Constante de normalisation ]

L'analyse Bayésienne, développée sous le logiciel MEMM, fonctionne à partir de distributions *a priori* des paramètres  $\sigma_F^2$ ,  $\sigma_M^2$ ,  $a_S$ ,  $b_S$ ,  $a_P$ ,  $b_P$ ,  $s$ ,  $m_S$ ,  $m_P$ . Ensuite, une chaîne de Markov - Monte Carlo donne aléatoirement de nouvelles valeurs aux paramètres et aux variables non-observées (c'est-à-dire les fertilités  $\mathbf{FF}$  et  $\mathbf{FM}$  des individus dans notre cas), et accepte ces nouvelles valeurs avec une probabilité d'autant plus grande qu'elles (i) fournissent une vraisemblance plus élevées et (ii) sont des valeurs ayant un poids élevé dans la loi *a priori*. Les valeurs visitées par la chaîne fournissent des distributions *a posteriori* des paramètres et des fécondités individuelles. En prenant la médiane de ces distributions, on obtient de bonnes estimations de nos paramètres (Klein *et coll.*, 2011). À partir des données écologiques (coordonnées spatiales) et des données génétiques, le modèle va pouvoir estimer au mieux les différents paramètres d'intérêt.

Pour traiter ce modèle, le logiciel MEMM a été utilisé sous forme d'un exécutable développé en C++. Nous avons utilisé MEMM avec une distribution Gamma pour les fertilités, des lois *a priori* sur les intervalles  $[0, 1]$ ,  $[0, 1]$ ,  $[0, 10000]$ ,  $[0, 10]$  et  $[0, 1000]$  pour les paramètres  $m$ ,  $s$ , delta,  $b$  et  $dobs/de$  respectivement. Nous avons utilisé des chaînes MCMC de 50 000 itérations après avoir enlevé les 500 premières itérations pour que la chaîne atteigne un état stationnaire. L'approche a été appliquée à chacune des deux placettes successivement. Les résultats ont ensuite été visualisés sous R.

# ANALYSE DE L'EFFET DES VARIABLES ÉCOLOGIQUES SUR LES FÉCONDITÉS INDIVIDUELLES

---

Nous avons étudié l'effet de différentes variables écologiques (ou facteurs) mesurées sur les arbres reproducteurs sur les fécondités femelles et mâles intrinsèques (FF et FM, estimées avec le SEMM). Les facteurs considérés sont : le diamètre (DBH), la production de faînes (FruitProd, mesurées en 2013), le dépérissement foliaire cumulé sur les 10 années d'observation (DEFcum) et la présence du champignon lignicole *O. mucida* cumulée sur les 10 années d'observation (OUDcum). Les analyses ont été faites sur l'ensemble des arbres adultes suivi de la réserve intégrale et de diamètre >10cm (155 sur la placette 1, 174 sur la placette 2), en excluant tous ceux qui avaient au moins une donnée manquante (327 arbres au total). Les fertilités relatives femelles et mâles ont été log-transformées. Nous avons utilisé le modèle linéaire suivant, appliqué sur chacune des placettes, ou sur les deux placettes ensemble (après normalisation des fertilités relatives par la moyenne intra-placette) :

$$\text{Log(FF) ou Log(FM)} = \text{DBH} + \text{DEFcum} + \text{OUDcum} + \text{FruitProd}.$$



Ces analyses ont été réalisées avec les procédures *lm* et *step* (sélection de variables) dans le logiciel R.

Nous n'avons pas considéré de modèles avec les mesures annuelles de dépérissement ou de présence du champignon. En effet, une analyse préalable en composante principale (Append. 10) révèle que :

- les différentes notes de dépérissement foliaires sont fortement et positivement corrélées (et que les notes de 2002 et 2003 sont peu informatives)
- les différentes notes de présence du champignon sont fortement et positivement corrélées (et que la note de 2002 est peu informative)

En revanche, les notes de dépérissement foliaire d'une part et de présence du champignon d'autre part sont indépendantes.



# Annexes

---





# Liste des Figures, Tableaux, et Appendices

---

**Figure 1** : Boucle de rétroaction négative attendue entre dépérissement et qualité génétique de la régénération.

**Figure 2** : Distribution européenne du Hêtre commun (*Fagus sylvatica*)

**Figure 3** : Dépérissement foliaire chez le Hêtre

**Figure 4** : Le champignon lignicole *Oudemansiella mucida*.

**Encadré 1** : Notion de noyau de dispersion primaire et efficace

**Figure 5A.** Vue d'ensemble de la réserve naturelle de la Massane, et des deux placettes étudiées.

**Figure 5B.** Zoom sur la placette 1 et la placette 2.

**Tableau 1** : Nombre d'adultes reproducteurs et de descendants échantillonnés.

**Tableau 2** : Diversité génétique moyenne des combinaisons cohortes/placettes étudiées, et structure génétique spatiale (SGS).

**Tableau 3** : Différentiation génétique moyenne ( $F_{ST}$ ) entre paires de cohortes/placettes.

**Figure 6** : Structure génétique spatiale des cohortes de semis (A) et de reproducteurs (B).

**Tableau 4** : Distance moyenne de dispersion des gènes ( $\sigma$ , en m) estimée par l'approche indirecte sur la cohorte des semis ou des adultes. NC indique que la procédure itérative ne converge pas.

**Tableau 5** : Nombre (et pourcentage) de descendants issus de dispersion de graine et de pollen inter- ou intra-placette.

**Figure 7** : Distribution des distances de dispersion efficace A/C : des graines et B/D : du pollen sur l'ensemble des deux placettes.

**Figure 8** : Distribution conjointe des distances de dispersion des graines et du pollen par placette et par stratégie.

**Figure 9** : Distribution spatiale des succès reproducteurs individuels A-femelles et B- mâles estimés par l'analyse de parenté/maternité avec la stratégie 1.

**Figure 10** : Comparaison des succès reproducteurs efficaces femelle et mâle estimés avec les deux stratégies

**Figure 11** : Distances de dispersion efficace des graines A. pour les plantules sur chaque placette ; B. pour les semis de l'année sur chaque placette

**Tableau 6** : Régime de reproduction, noyaux de dispersion primaires du pollen et des graines, et variance des fertilités intrinsèques mâles et femelles pour chaque placette

Figure Y : qui décale la suite donc : Exemple de dispersion par zoochorie d'une faîne probablement logée dans l'écorce d'un chêne par une sitelle.

**Figure 12** : Distribution empirique des fertilités femelles et mâles sur la placette 1

**Figure 13** : Distribution empirique des fertilités femelles et mâles sur la placette 2.

**Figure 14** : Distribution spatiale des fertilités intrinsèques individuelles A- femelles et B- mâles estimées par le SEMM.

**Tableau 7** : Effet des facteurs écologiques cumulés sur A-les fertilités femelles et B- fertilités mâles estimées par le SEMM.

**Tableau 8** : Comparaison du régime de reproduction et des paramètres de dispersion entre la Massane et le Ventoux.

**Figure 15** : Étapes de prélèvement des feuilles de hêtres chez les adultes et les plantules ou semis.

**Tableau 9** : Diversité et qualité des marqueurs microsatellites utilisés.

**Tableau 10A** : Décisions possibles de l'analyse de maternité pour chacun des descendants considérés B) l'analyse de parenté et C) la combinaison des deux analyses effectuées dans cette étude

**Tableau 11** : Analyse de sensibilité des résultats des assignations de maternité/parenté aux paramètres de l'analyse (PropEch = le taux d'échantillonnage des parents ; Terr= le taux d'erreur de génotypage ; Nm<sub>q</sub>= nombre de marqueurs)

**Figure 16** : Distribution du nombre de parents (analyse de maternité) ou de couples de parents (analyse de parenté) compatibles en moyenne avec un descendant (c'est-à-dire avec un LOD-score positif pour ce descendant).

**Figure 17** : Schéma du modèle de croisement dans un voisinage spatialement explicite (SEMM).

**Tableau 12** : Paramètres du modèle SEMM estimés par l'analyse.

**Append. 1** : Liste des marqueurs microsatellites utilisés

**Append. 2** : Distribution spatiale des succès reproducteurs individuels A- femelles et B- mâles estimés par l'analyse de parenté/maternité avec la stratégie 2.

**Append. 3** : Distribution des diamètres à 1.30m en fonction de la classe de production de faînes (0 : pas de faînes ; 1 : très peu ; 2 : production moyenne ; 3 : production importante).

**Append. 4** : Corrélation entre fertilités femelles et mâles estimées sur chaque placette. Les fertilités ont été au préalable ajustées des effets du diamètre.

**Append. 5** : Distribution des diamètres mesurés à 1.30 m (DBH) par placette.

**Append. 6** : Distribution des dépérissements foliaires cumulés sur 11 ans (DEFcum) par placette.

**Append. 7** : Distribution des présences du champignon *O. mucida* cumulés sur 11 ans (OUDcum) par placette.

**Append. 8** : Analyse en composante principale des variables individuelles suivies sur 327 arbres de la réserve intégrale de diamètre > 10cm

**Append. 9** : Calcul du nombre efficace de reproducteurs dans la hêtraie de la Massane



# Références bibliographiques

---

- Allen, C.D., Macalady, A.K., Chenchouni, H., Bachelet, D., McDowell, N., Vennetier, M., Kitzberger, T., Rigling, A., Breshears, D.D., Hogg, E.H.T., Gonzalez, P., Fensham, R., Zhang, Z., Castro, J., Demidova, N., Lim, J.-H., Allard, G., Running, S.W., Semerci, A., & Cobb, N. (2010) A global overview of drought and heat-induced tree mortality reveals emerging climate change risks for forests. *Forest Ecology and Management*, **259**, 660-684.
- Ashley, M. (2010) Plant Parentage, Pollination, and Dispersal: How DNA Microsatellites Have Altered the Landscape. *Critical Reviews in Plant Sciences*, **29**, 148–161.
- Asuka, Y., Tani, N., Tsumura, Y., & Tomaru, N. (2004) Development and characterization of microsatellite markers for *Fagus crenata* Blume. *Molecular Ecology Notes*, **4**, 101-103.
- Bontemps, A., Klein, E.K., & Oddou-Muratorio, S. (2013) Shift of spatial patterns during early recruitment in *Fagus sylvatica*: evidence from seed dispersal estimates based on genotypic data. *Forest Ecology and Management*, **305**, 67-76.
- Broquet, T. & Petit, E.J. (2009) Molecular Estimation of Dispersal for Ecology and Population Genetics. *Annual Review of Ecology, Evolution, and Systematics*, **40**, 193-216.
- Burczyk, J., Adams, W.T., Birkes, D.S., & Chybicki, I.J. (2006) Using genetic markers to directly estimate gene flow and reproductive success parameters in plants on the basis of naturally regenerated seedlings. *Genetics*, **173**, 363-372.
- Bykova, O., Chuine, I., Morin, X., & Higgins, S.I. (2012) Temperature dependence of the reproduction niche and its relevance for plant species distributions. *Journal of Biogeography*, **39**, 2191-2200.
- Chybicki, I.J. & Burczyk, J. (2013) Seeing the forest through the trees: comprehensive inference on individual mating patterns in a mixed stand of *Quercus robur* and *Q. petraea*. *Annals of Botany* **112** 561-574
- Crawford, T.J. (1984) The estimation of neighborhood parameters for plant populations. *Heredity*, **52**, 273-283.
- Dajoz, R. (1998) *Les insectes et la forêt: Rôle et diversité des insectes dans le milieu forestier* Lavoisier Tec&Doc.
- Dakin, E.E. & Avise, J.C. (2004) Microsatellite null alleles in parentage analysis. *Heredity*, **93**, 504-509.
- Gauzère, J., Klein, E.K., & Oddou-Muratorio, S. (2013) Ecological determinants of mating system within and between three *Fagus sylvatica* populations along an elevational gradient. *Molecular Ecology*, **22**, 5001-15.
- Goudet, J. (2000) FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.1). Available from <http://www.unil.ch/izea/software/fstat.html>.
- Hampe, A. (2005) Fecundity limits in *Frangula alnus* (Rhamnaceae) relict populations at the species southern range margin. *Oecologia*, **143**, 377-386 LA - English.
- Hare, M.P., Nunney, L., Schwartz, M.K., Ruzzante, D.E., Burford, M., Waples, R.S., Ruegg, K., & Palstra, F. (2011) Understanding and Estimating Effective Population Size for Practical Application in Marine Species Management. *Conservation Biology*, **25**, 438-449.

Hardy, O. & Vekemans, X. (2002) SPAGeDi: a versatile compute program to analyse spatial genetic structure at the individual or population levels. *Molecular Ecology Notes*, **2**, 618.

Janzen, D. (1970) Herbivores and the number of tree species in tropical forests. *The American Naturalist*, **104**, 501-528.

Jensen, T.S. (1985) Seed-seed predator interactions of European beech, *Fagus sylvatica* and forest rodents, *Clethrionomys glareolus* and *Apodemus flavicollis*. *Oikos*, **44**, 149-156.

Jones, A.G., Small, C.M., Paczolt, K.A., & Ratterman, N.L. (2010) A practical guide to methods of parentage analysis. *Molecular Ecology Resources*, **10**, 6-30.

Jump, A.S. & Woodward, F.I. (2003) Seed production and population density decline approaching the range-edge of *Cirsium* species. *New Phytologist*, **160**, 349-358.

Kalinowski, S. & Taper, M. (2006) Maximum likelihood estimation of the frequency of null alleles at microsatellite loci. *Conservation Genetics*, **7**, 991-995.

Kelly, D. & Sork, V.L. (2002) Mast seeding in perennial plants: Why, How, Where? *Annual Review of Ecology and Systematics*, **33**, 427-447.

Klein, E.K., Carpentier, F., & Oddou-Muratorio, S. (2011) Estimating the variance of male fecundity from genotypes of progeny arrays: evaluation of the bayesian forward approach. *Methods in ecology and Evolution*, **2**, 349-361.

Klein, E.K., Desassis, N., & Oddou-Muratorio, S. (2008) Pollen flow in the wildservice tree, *Sorbus torminalis* (L.) Crantz. IV. Whole inter-individual variance of male fecundity estimated jointly with the dispersal kernel. *Molecular Ecology*, **17**, 3323-3336.

Lee, T.D. (1988). Patterns of fruit and seed production. In *Plant reproductive ecology: patterns and strategies* (eds J. Lovett-Doust & L. Lovett-Doust), pp. 179–202. Oxford University Press, New York, NY2.

Lefevre, S., Wagner, S., Petit, R.J., & De Lafontaine, G. (2012) Multiplexed microsatellite markers for genetic studies of beech. *Molecular Ecology Resources*, **12**, 484-491.

Loiselle, B.A., Sork, V.L., Nason, J., & Graham, C. (1995) Spatial genetic structure of a tropical understory shrub, *Psychotria officinalis* (Rubiaceae). *American Journal of Botany*, **82**, 1420-1425.

Marshall, T.C., Slate, J., Kruuk, L.E.B., & Pemberton, J.M. (1998) Statistical confidence for likelihood-based paternity inference in natural populations. *Molecular Ecology*, **7**, 639-655.

Michelot, A., Bréda, N., Damesin, C., & Dufrêne, E. (2012) Differing growth responses to climatic variations and soil water deficits of *Fagus sylvatica*, *Quercus petraea* and *Pinus sylvestris* in a temperate forest. *Forest Ecology and Management*, **265**, 161-171.

Mimura, M. & Aitken, S.N. (2007) Increased selfing and decreased effective pollen donor number in peripheral relative to central populations in *Picea sitchensis* (Pinaceae). *American Journal of Botany*, **94**, 991-998.

Moran, E.V. & Clark, J.S. (2011) Estimating seed and pollen movement in a monoecious plant: a hierarchical Bayesian approach integrating genetic and

ecological data. *Molecular Ecology*, **20**, 1248-1262.

Morin, X., Viner, D., & Chuine, I. (2008) Tree species range shifts at a continental scale: new predictive insights from a process-based model. *Journal of Ecology*, **96**, 784-794.

Oddou-Muratorio, S. & Klein, E.K. (2008) Comparing direct vs. indirect estimates of gene flow within a population of a scattered tree species. *Molecular Ecology*, **17**, 2743-2754.

Palstra, F.P. & Ruzzante, D.E. (2008) Genetic estimates of contemporary effective population size: what can they tell us about the importance of genetic stochasticity for wild population persistence? *Molecular Ecology*, **17**, 3428-3447

Pastorelli, R., Smulders, M.J.M., Van't Westende, W.P.C., Vosman, B., Giannini, R., Vettori, C., & Vendramin, G.G. (2003) Characterization of microsatellite markers in *Fagus sylvatica* L. and *Fagus orientalis* Lipsky. *Molecular Ecology Notes*, **3**, 76-78.

Rasmussen, K.K. & Kollmann, J. (2004) Poor sexual reproduction on the distribution limit of the rare tree *Sorbus torminalis*. *Acta Oecologica*, **25**, 211-218.

Sork, V., Nason, J., Campbell, D., & Fernandez, J. (1999) Landscape approaches to historical and contemporary gene flow in plants. *Trends in Ecology & Evolution*, **14**, 219-224.

Talon & al (in prep), Dynamique holocène d'une vieille hêtraie montagnarde : la forêt de la Massane, Argelès-sur-Mer, Pyrénées-Orientales.

von Wuehlisch, G. (2008). EUFORGEN Technical Guidelines for genetic conservation and use for European beech (*Fagus sylvatica*), Rome, Italy.

Vornam, B., Decarli, N., & Gailing, O. (2004) Spatial distribution of genetic variation in a natural beech stand (*Fagus sylvatica* L.) based on microsatellite markers. *Conservation Genetics*, **5**, 561-570.

Weir, B.S. & Cockerham, C.C. (1984) Estimating F-statistics for the analysis of population structure. *Evolution*, **38**, 1358-1370.

Zhu, K., Woodall, C.W., Ghosh, S., Gelfand, A.E., & Clark, J.S. (2014) Dual impacts of climate change: forest migration and turnover through life history. *Global Change Biology*, **20**, 251-264.

# Appendices

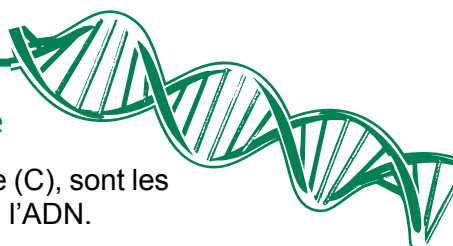
## Appendice 1 : Liste des marqueurs microsatellites utilisés

Multiplex	Marqueur	Motif répété *	Taille lue	Référence	Fluoro-chromme
1	Csolfagus_19	(TC)13	150-190	Vendramin (non publié)	Tam
	Csolfagus_7	(AT)11	100-135	Vendramin (non publié)	Tam
	Fs1_15	(GA)26	90-138	Pastorelli et coll. (2003)	Hex
	Fs3_4	(GCT)6	190-215	Pastorelli et coll.(2003)	Hex
	Sfc0007	(AG)24	139-170	Asuka et coll.(2004)	Hex
	Sfc1143	(AG)21	100-140	Asuka et coll.(2004)	Fam
2	Csolfagus_25	NA	150-195	Vendramin (non publié)	Hex
	Csolfagus_29	(ST)11	135-165	Vendramin (non publié)	Tam
	Csolfagus_31	(AG)12	100-135	Vendramin (non publié)	Hex
	Csolfagus_6	(AG)13	200-230	Vendramin (non publié)	Tam
	Fi05	NA	140-170	Vendramin (non publié)	Fam
	Mfc7	(GA)9	105-35	Vornam et coll. (2004)	Tam
	sfc061	(AG)22	100-140	Asuka et coll.(2004)	Fam
3	concat14_A_0	(TC)9	165-205	Lefevre et coll. (2012)	Hex
	DE576	(CAA)10	200-240	Lefevre et coll. (2012)	Fam
	DUKCT_A_0	(AC)14	65-100	Lefevre et coll. (2012)	ATTO56
	DZ447_A_0	(TC)10	180-200	Lefevre et coll. (2012)	Tam
	EJV8T_A_0	(TC)10	136-165	Lefevre et coll. (2012)	Tam
	EEU75	(CT)10	80-120	Lefevre et coll. (2012)	Hex
	EMILY_A	(GA)11	130-160	Lefevre et coll. (2012)	Fam
	ERHIBI_A_0	(AG)9	150-175	Lefevre et coll. (2012)	ATTO56

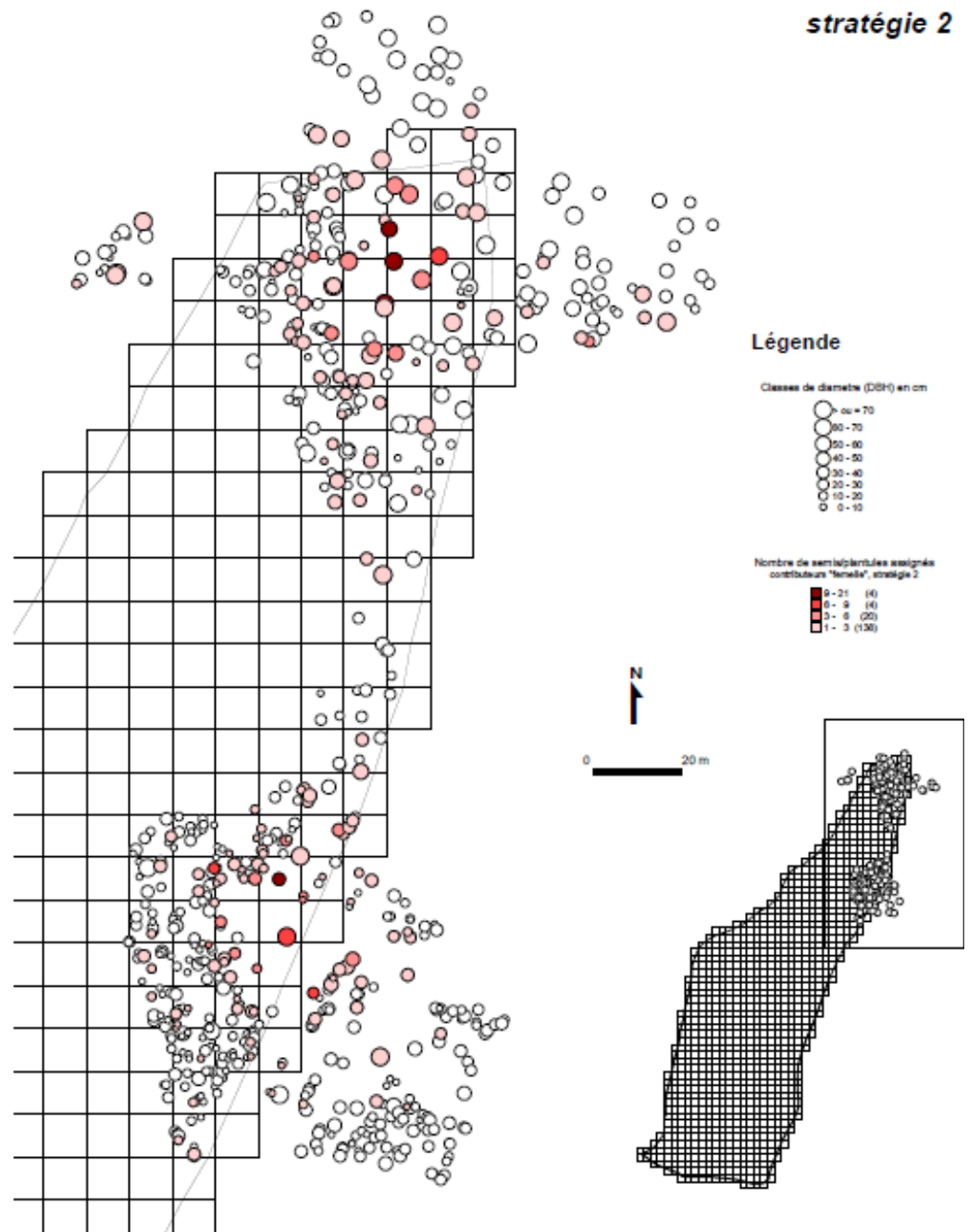
### \* Séquence d'ADN et marqueur microsatellite

L'adénine (A), la guanine (G), la thymine (T), et la cytosine (C), sont les quatre molécules organiques à la base de la structure de l'ADN.

Une séquence microsatellite est une portion d'ADN formée par une répétition continue de motifs composés de 2 à 10 de ces nucléotides. Ces « motifs » sont très abondants dans tout le génome de tous les eucaryotes. Le nombre de ces répétitions va de quelques-unes à une centaine et varie d'un individu à l'autre et d'un allèle (variant d'un locus) à l'autre chez un même individu.

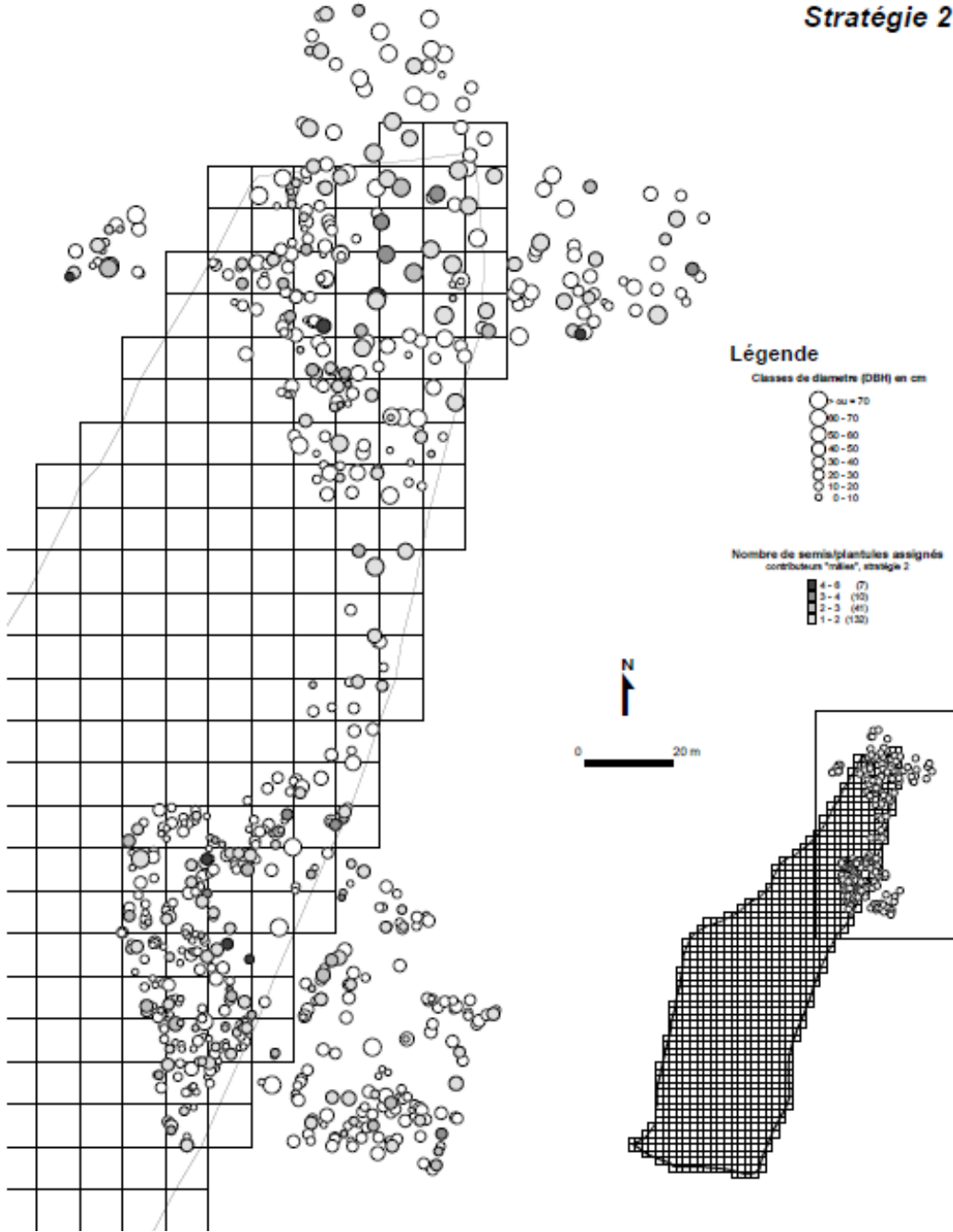


**Appendice 2** : Distribution spatiale des succès reproducteurs individuels  
A- femelles et B- mâles estimés par l'analyse de parenté/maternité avec  
la stratégie 2.

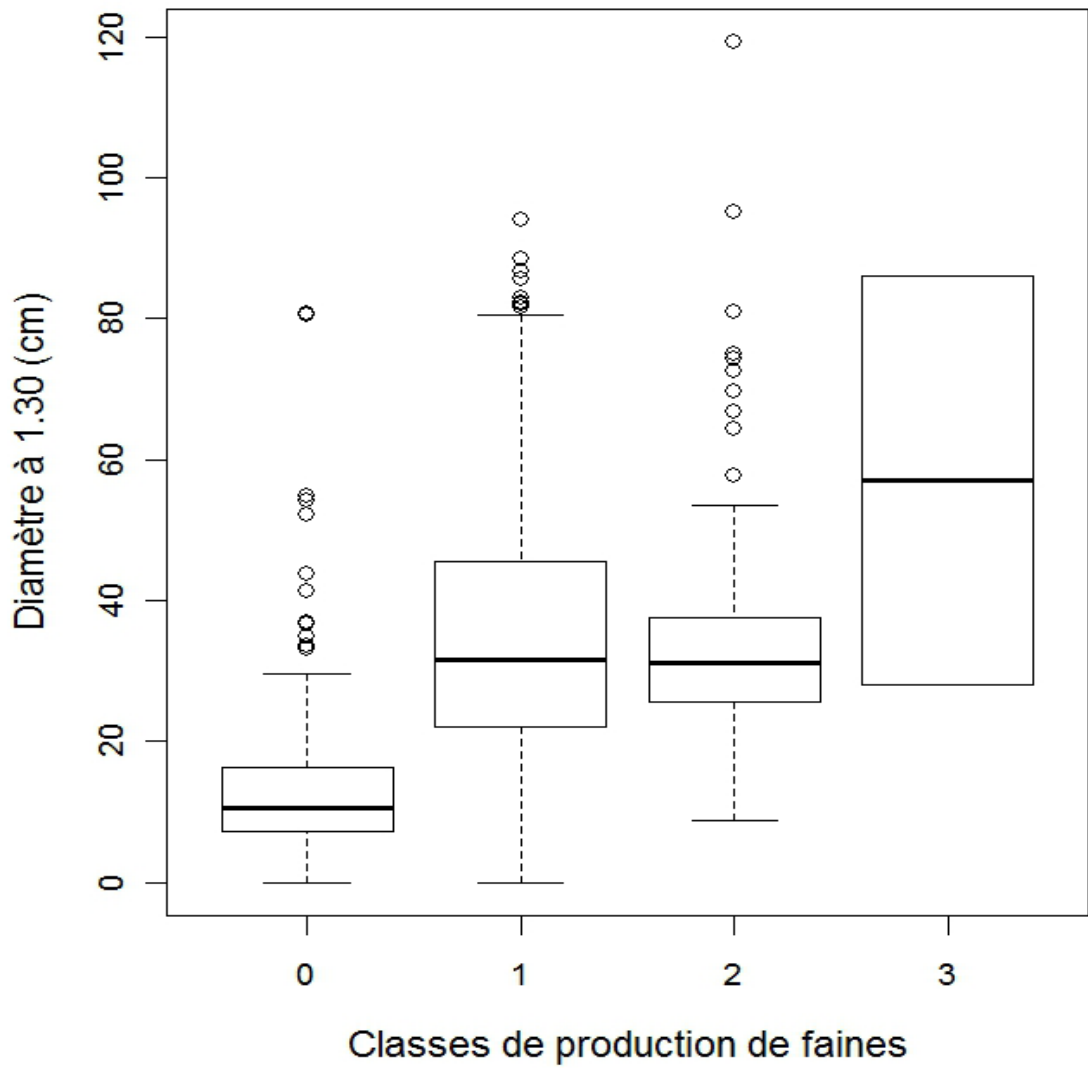




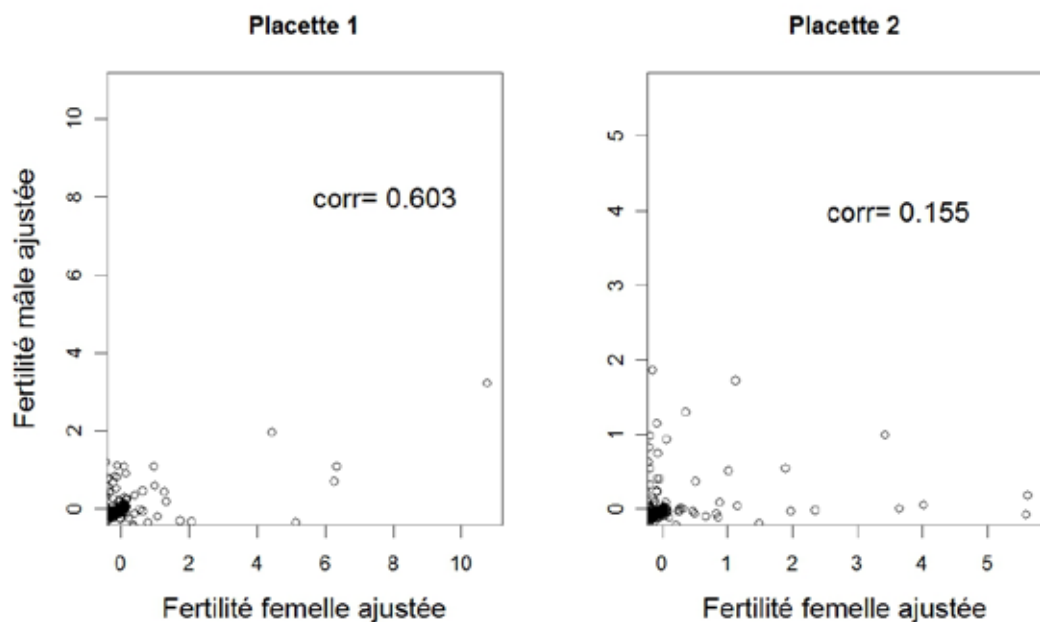
# Stratégie 2



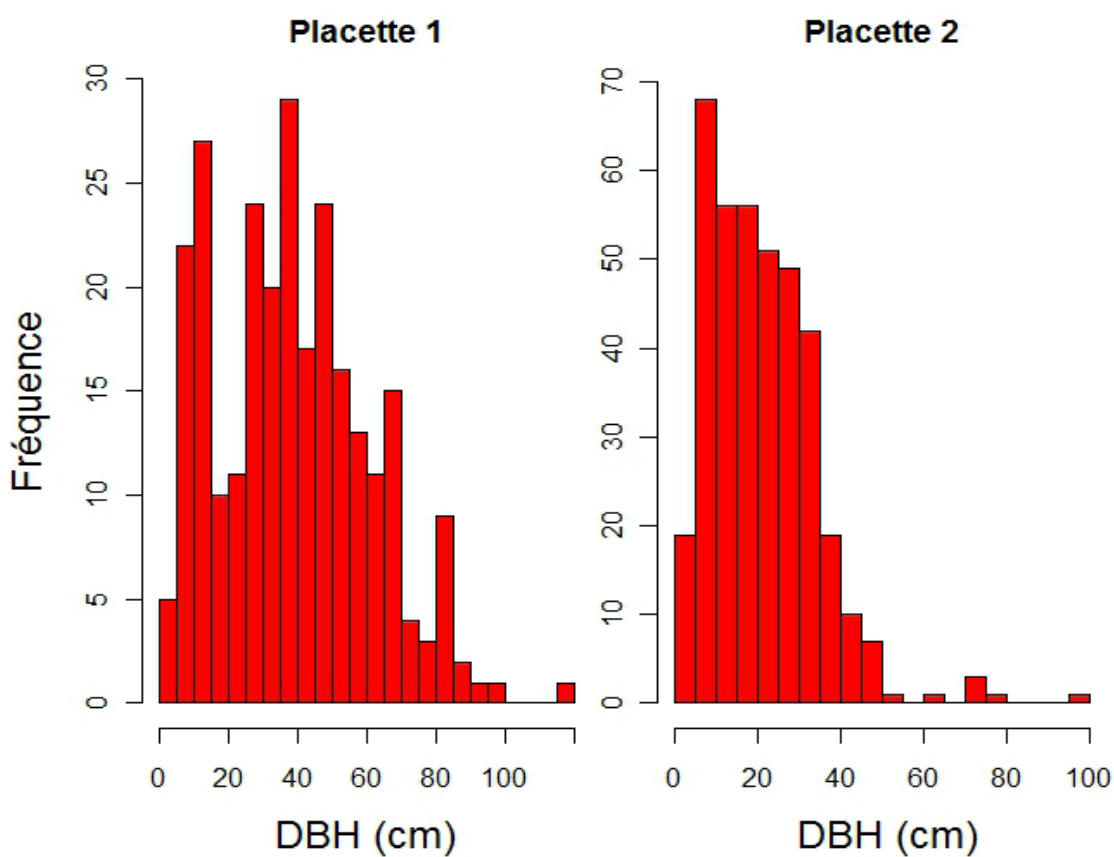
**Appendice 3** : Distribution des diamètres à 1.30m en fonction de la classe de production de faînes (0 : pas de faînes ; 1 : très peu ; 2 : production moyenne ; 3 : production importante). Données de production de faînes mesurée en 2013 pour 670 arbres.



**Appendice 4** : Corrélation entre fertilités femelles et mâles estimées sur chaque placette. Les fertilités ont été au préalable ajustées des effets du diamètre.

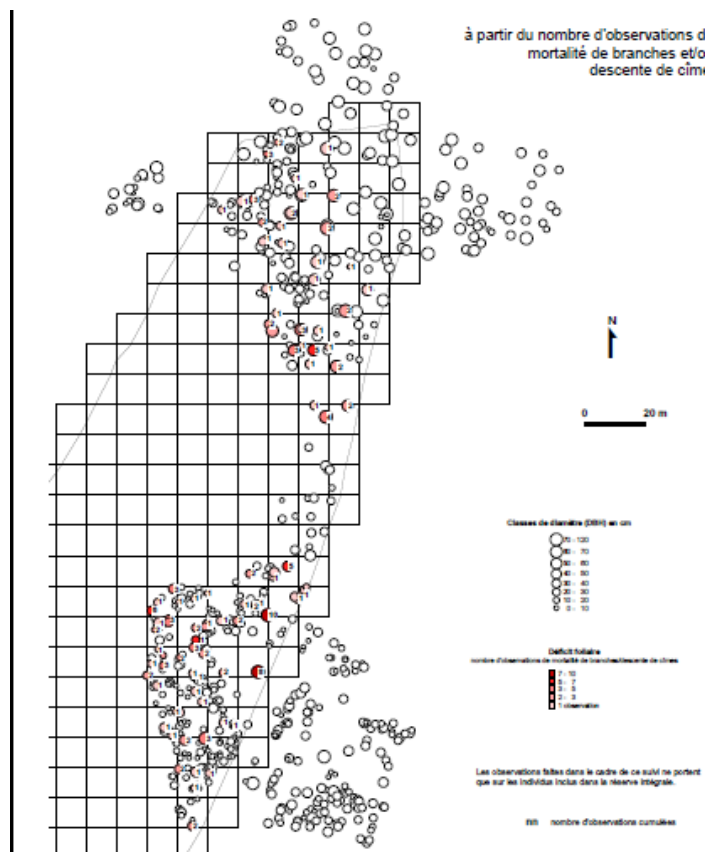
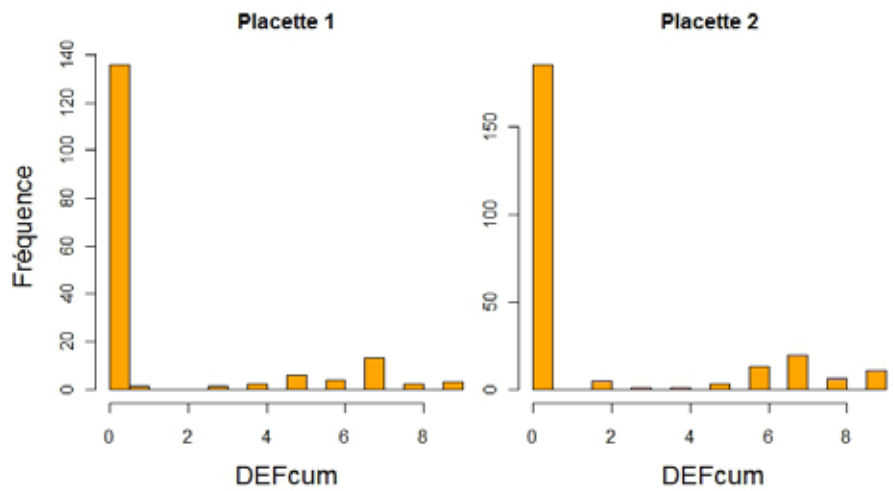


**Appendice 5** : Distribution des diamètres mesurés à 1.30 m (DBH) par placette.

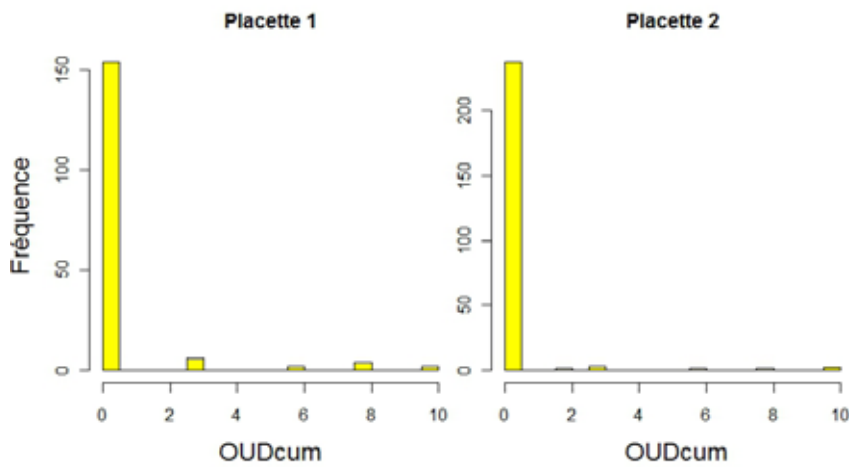


**Appendice 6** : Dépérissement foliaire cumulé sur 11 ans de suivi (DEFcum) et basé sur des observations de mortalité de branches et/ou de descente de cime.

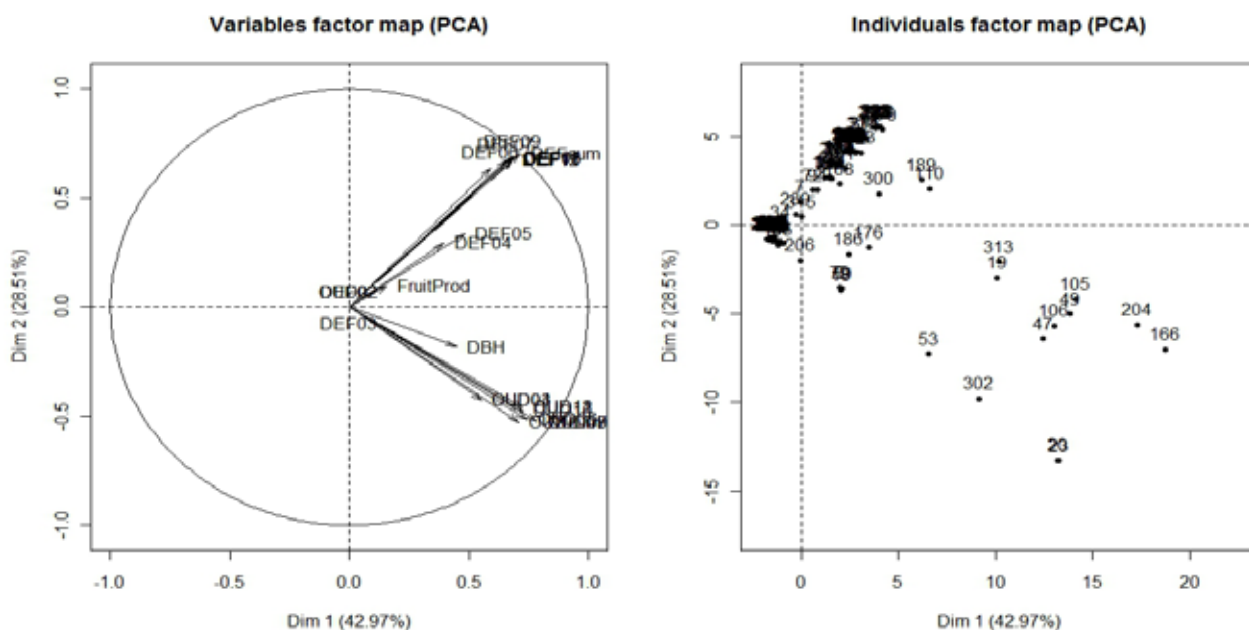
- a) Distribution des DEFcum par placette et par classe de diamètre
- b) Cartographie des DEFcum pour la zone suivie pour les placettes 1 et 2



**Appendice 7** : Distribution des présences du champignon *O. mucida* cumulées sur 11 ans (OUDcum) par placette.



**Appendice 8** : Analyse en composante principale des variables individuelles suivies sur 327 arbres de la réserve intégrale de diamètre > 10cm. Les variables sont les 11 notes de dépérissements foliaires (DEFn, avec 2002<n< 2012), les 11 notes de présence du champignon *O. mucida* (OUDn, avec 2002<n< 2012), le diamètre (DBH) et la production de faînes en 2013 (FruitProd). Le graphe de gauche montre la projection des variables, celui de droite celle des individus. Les deux premiers axes expliquent 71.5% de la variation.





**Appendice 9** : Calcul du nombre efficace de reproducteurs dans la hêtraie de la Massane

Cet appendice décrit comment le nombre efficace de reproducteurs a été estimé à partir des ratios  $d_{obs}/d_e$  empiriques par voie femelle ( $R_{\text{♀}}$ ) mâle ( $R_{\text{♂}}$ ) estimés par le SEMM, et des densités observées de reproducteurs sur les placettes 1 et 2.

Nous avons considéré la densité observée  $d_{obs}$  de reproducteur moyenne sur les deux placettes, estimée à partir du nombre d'individus de diamètre DBH>10 cm (Nrep=116 et 243 respectivement sur les placettes 1 et 2) et de leur surface (0.9 et 0.8 ha respectivement). On obtient  $d_{obs}=216$  individus/ha.

Nous avons ensuite dérivé les valeurs des densités efficaces de reproducteurs par voie femelle et male comme  $d_e = d_{obs}/R$ , comme estimé dans le tableau ci-dessous (avec les intervalles de confiance autour de  $d_e$ ):

Voie mâle ou femelle	R moyen	$d_e$ moyen	R IC+	$d_e$ IC-	R IC-	$d_e$ IC +
femelle	4.4	<b>49.2</b>	8.1	<b>26.7</b>	3.1	<b>69.8</b>
mâle	2.2	<b>98.3</b>	5.3	<b>40.8</b>	1.8	<b>120.2</b>

Les nombres de reproducteurs par voie femelle et male ont été calculés en considérant la totalité de la surface de la réserve intégrale (336ha), ce qui repose sur l'hypothèse que malgré la limitation par la distance de la dispersion du pollen et surtout des graines, tous les individus sont potentiellement connectés entre eux et peuvent contribuer à la pluie de graine en tout point de la réserve. On obtient  $Nb_{\text{♀}} = 16520$  individus et  $Nb_{\text{♂}} = 33\ 037$  individus. Il est probable que ces valeurs de Nb soient surestimées du fait de notre hypothèse.

Enfin, pour obtenir de nombre efficace de reproducteurs total, nous avons pris la moyenne harmonique de  $Nb_{\text{♀}}$  et  $Nb_{\text{♂}}$ , **soit Nb~11000 individus.**

Il faut noter que chez les espèces sémelpares comme les arbres,  $Ne \sim Nb \times TG$ , où TG, le temps de génération, est le temps moyen de graine à graine (typiquement entre l'âge à la première reproduction et la durée de vie) ; on considère généralement que TG~60 ans chez des arbres comme le hêtre.



## Allèles\*

Formes alternatives d'un gène en un locus donné. Les allèles peuvent se substituer les uns aux autres au même locus. Ils dérivent les uns des autres par mutation et assurent une même fonction.

## Allèle nul

La technique de caractérisation des marqueurs microsatellites\* fait intervenir (lors de la PCR\*) l'hybridation d'une amorce sur le site cible que l'on souhaite caractériser. Il est possible que le site en question ait subi des mutations qui empêchent l'hybridation de l'amorce et donc l'amplification du microsatellite. Les individus portant un allèle nul sont hétérozygotes  $A_iA_n$  ( $A_n$  étant l'allèle nul et donc non détecté), mais seront détectés comme homozygotes  $A_iA_i$ . La présence d'allèles nuls sous-estime ainsi la fréquence des hétérozygotes, et est susceptible de biaiser les analyses de maternité ou de parenté.

## Allfécondation\*

Fusion entre un gamète mâle et un gamète femelle produits par deux individus distincts.

## Allogamie\*

Mode de reproduction sexuée dans lequel la fécondation est croisée. Les gamètes qui fusionnent lors de la fécondation sont issus de deux individus distincts.

## Autofécondation\*

Fusion entre un gamète mâle et un gamète femelle produits par le même individu. L'autofécondation est source de consanguinité et est évitée chez de nombreuses espèces.

## Autogamie\*

Mode de reproduction sexuée dans lequel les fécondations ont lieu entre gamètes provenant d'un même individu (une fleur hermaphrodite effectuant une autopolinisation par exemple). L'autogamie favorise l'homozygotie sur l'ensemble du génome.

## Cohorte

Désigne un ensemble d'individus ayant vécu un même événement au cours d'une même période. Plus précisément chez les arbres, organismes à long cycle de vie où des individus d'âge différents coexistent dans les populations naturelles, on peut théoriquement distinguer  $N$  cohortes d'âges différents si l'âge maximum d'un arbre est  $N$ . Dans cette étude, on distingue seulement deux cohortes : celle des adultes (diamètre  $>5\text{cm}$  à 1.30) et les semis.

## Consanguinité

Situation obtenue lors du croisement d'individus apparentés. Chez les organismes diploïdes, le coefficient (ou indice) de consanguinité est la probabilité pour que deux allèles que possède un individu en un locus quelconque soient identiques par descendance. Il est noté  $F$  et est compris entre 0 et 1. Ceci suppose donc qu'il existe un ou plusieurs ancêtres communs aux deux parents de l'individu étudié.

### Dérive génétique\*

Variations des fréquences alléliques entre générations successives dues à la limitation de la taille de la population. Dans les populations de petite taille (une dizaine d'individus), la variance d'échantillonnage lors du tirage au sort des gamètes induit des fluctuations importantes des fréquences alléliques qui peuvent conduire à la fixation des allèles.

### Diploïde\*

État d'une cellule ou d'un organisme dont les noyaux comportent deux exemplaires de chaque type de chromosome.

### Flux de gènes\*

Transfert d'allèle entre individus et populations. Chez les plantes, les flux de gènes sont liés au transport du pollen et des graines. Les flux de gènes sont responsables des migrations et des introgressions. Les flux de gènes historiques correspondent aux taux de migration cumulés sur un grand nombre de générations, par opposition aux flux de gènes contemporains, qui correspondent aux mouvements des gènes entre parents et enfants sur un événement de reproduction.

### Gamétophyte\*

Chez les Angiospermes, le gamétophyte femelle est le sac embryonnaire, qui contient l'oosphère (gamète femelle) et le gamétophyte mâle est réduit au grain de pollen, structure bi- ou tricellulaire donnant des gamètes mâles.

### Gène\*

Séquence codante caractérisée par une fonction et une localisation en un locus précis dans le génome. Il est transcrit en ARN messager puis traduit en protéine, ou seulement transcrit en ARN de transfert ou ribosomique.

### Hétérozygote\*

Un locus d'un individu diploïde est dit hétérozygote lorsque les deux allèles portés par les chromosomes homologues à ce locus sont différents. L'hétérozygotie est la proportion de locus hétérozygotes chez un individu diploïde.

### Homozygote\*

Un locus d'un individu diploïde est dit homozygote lorsque les deux allèles portés par les chromosomes homologues à ce locus ne sont pas différents. L'homozygotie est la proportion de locus homozygotes chez un individu diploïde

### Indice de diversité ou hétérozygotie de Nei (He)

Cet indice mesure le polymorphisme pour un locus donné en fonction des fréquences de chaque allèle dans une population. Plus précisément, c'est la fréquence théorique des hétérozygotes dans une population panmictique\* (he). S'il y a n allèles avec les fréquences  $f_1, f_2, \dots, f_n$ , alors  $he = 1 - (f_1^2 + f_2^2 + \dots + f_n^2)$ . he varie entre 0 (s'il n'existe qu'un seul allèle pour le locus en question) et 1 (cas théorique d'un nombre infini d'allèles en fréquences égales). Si plusieurs loci sont

considérés, l'hétérozygotie moyenne ( $H_e$ ), représentant la moyenne du taux d'individus hétérozygotes par population sera la moyenne arithmétique de toutes les valeurs de  $h_e$ . Plus le nombre d'allèles est grand et plus leurs fréquences sont équilibrées, plus  $H_e$  augmente.

### **Locus\***

Emplacement occupé par un gène sur le génome. Les différentes formes possibles de ce gène, les allèles, sont portées à ce locus. Par extension, les séquences non codantes caractérisées par un polymorphisme sont aussi localisées à des locus.

### **Marqueur moléculaire\***

Séquence ou fragment d'ADN dont le polymorphisme sert à caractériser le génotype des individus. Un marqueur moléculaire peut être un gène entier, mais plus souvent (comme les marqueurs microsatellites utilisés dans cette étude), un marqueur moléculaire est anonyme, c'est-à-dire que c'est un fragment d'ADN polymorphe, amplifié par PCR\* à l'aide d'une amorce ou paire d'amorces de séquence aléatoire, ne permettant pas son identification directe par une fonction ou une propriété autre que sa taille.

### **Masting (Terme anglais)**

Chez les arbres, désigne une forte variation interannuelle dans la production individuelle de graines, accompagnée d'une synchronisation entre arbres des productions de graine à l'échelle du peuplement (avec des bonnes et mauvaises années de production). Ce phénomène peut être très marqué chez certaines espèces, comme le chêne (glandées massives tous les 5-6 ans), ou plus discret (chez le hêtre, les bonnes fênaies semblent plus régulières), voire complètement absent (exemple du cèdre).

### **Microsatellite\***

Motif observé dans le génome, constitué de 2 à 10 nucléotides et répété en tandem un nombre variable de fois. Les microsatellites fournissent les marqueurs moléculaires\* intéressants, car très polymorphes\*.

### **Modèle en île (ou en archipel)**

Représentation simplifiée de la structure de populations utilisée pour caractériser les flux de gènes\* en génétique des populations. Ce modèle considère  $n$  populations de taille  $N$  chacune (la taille de population totale est donc  $n \times N$ ), entre lesquelles la migration est aléatoire; ainsi, la composition génétique des immigrants (entrant) dans une population est la composition génétique moyenne dans l'archipel (population totale). On appelle taux de migration ( $m$ ) la proportion d'individus immigrant dans une population par génération; le nombre d'immigrants dans une population est  $Nm$ .

### **Modèle d'isolement par la distance (IBD)**

L'isolement par la distance désigne la structuration spatiale de la variation génétique locale sous l'effet de la dispersion des gènes limitée dans l'espace. Le modèle d'IBD est utile pour déterminer la distribution des fréquences de gènes sur une région géographique. Contrairement au modèle en îles, la variance dispersion et les probabilités de migration



sont variables dans ce modèle et les deux contribuent à la différenciation génétique locale. L'isolement par la distance est généralement le modèle le plus simple pour expliquer l'isolement génétique entre les populations. A l'intérieur d'une population, l'isolement par la distance\* génère une forme de Structure Génétique Spatiale telle que l'apparentement entre individus diminue quand la distance géographique qui les sépare augmente.

### **Noyau de dispersion**

Fonction de probabilité qu'une plante source (ici un arbre) située à la position (0,0) ait émis une propagule (ici, pollen ou graine) retrouvée à la position (x,y)

### **Noyau de dispersion primaire**

Noyau de dispersion considérant uniquement le mouvement et la survie des graines pendant leur dispersion.

### **Noyau de dispersion efficace**

Noyau de dispersion considérant le mouvement et la survie des graines pendant leur dispersion ainsi que la survie des graines post-dispersion, conduisant ainsi au recrutement pour la génération suivante.

### **Panmixie**

Contexte où les individus seraient répartis de manière homogène au sein de la population et se reproduiraient tous aléatoirement. En d'autres termes, cette situation postule qu'un individu appartenant à une population panmictique a autant de chances de se reproduire avec n'importe quel autre individu de cette population

### **PCR, Polymerase Chaine reaction\***

Réaction de polymérisation en chaîne. C'est la réaction fondamentale de l'amplification génétique, qui permet après plusieurs cycles de dénaturation, appariement et élongation et l'action d'une ADN-polymérase (enzyme permettant de synthétiser l'ADN), l'obtention de nombreuses copies d'un fragment d'ADN.

### **Polymorphisme d'un caractère**

Coexistence de plusieurs types reconnaissables de ce caractère dans une population. Les caractères utilisés sont soit phénotypiques (par exemple la date de débourrement végétatif dans une population une année donnée, qui peut varier entre arbres) soit génétiques (par exemple un marqueur microsatellites qui peut présenter différents allèles dans une population donnée. Un caractère est dit polymorphe s'il présente effectivement plusieurs types reconnaissables dans la ou les populations étudiées.

### **SEMM**

Méthode statistique basée sur modèle de croisement dans un voisinage spatialement explicite (en anglais Spatially Explicit Mating Model) permettant d'estimer les flux de pollen et de graines, les paramètres du régime de reproduction, et la variance des fertilités mâles et femelles à partir de données génétiques (analyse de parenté) et spatiales (données du suivi forestier par SIG).

### **Structure génétique spatiale (SGS)**

Mode d'organisation de la diversité génétique locale dans lequel l'apparement entre individus dépend de la distance géographique qui les sépare. L'isolement par la distance\* génère une forme de SGS particulière, dans laquelle l'apparement entre individus diminue quand la distance géographique qui les sépare augmente.

### **Taille efficace d'une population**

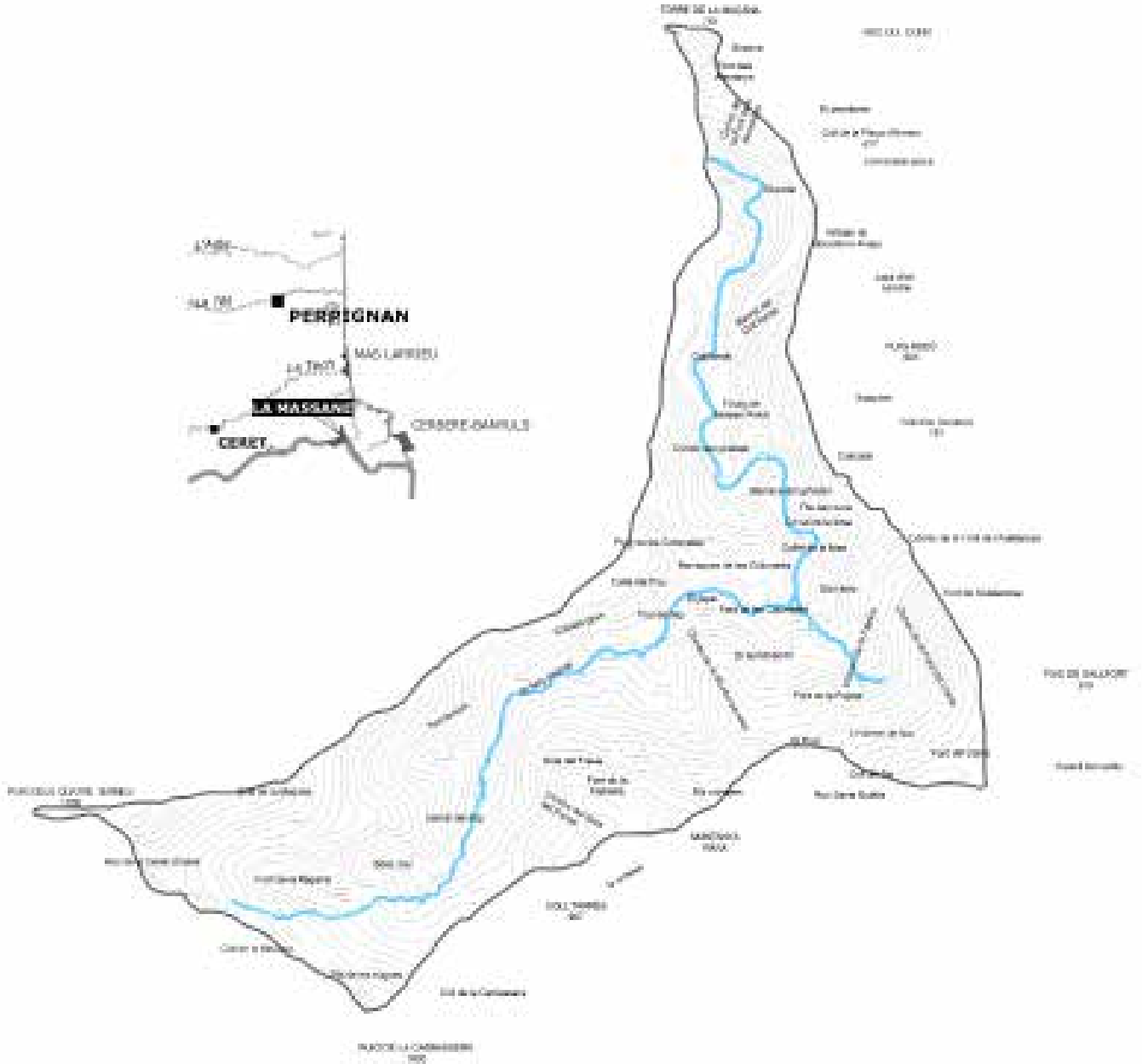
Dans les populations naturelles, tous les individus ne participent pas forcément au processus reproductif, si bien qu'en général l'effectif de la population qui détermine le rythme de la dérive génétique n'est pas égal à l'effectif de recensement de la population ( $N$ ). On définit donc l'effectif efficace de la population (ou taille efficace) comme l'effectif d'une population idéale (de type Wright-Fisher) pour laquelle on aurait une fluctuation du polymorphisme équivalente à celle de la population naturelle. C'est donc le nombre d'individus d'une population idéale pour lequel on aurait un degré de dérive génétique équivalent à celui de la population réelle. On note ce nombre comme  $N_e$ .

\*Définition tirée de Prat, D., Faivre-Rampant, P., & Prado, E. (2006) *Analyse du génome et gestion des ressources génétiques forestières*, Editions INRA QUAE, Paris.



# Réserve Naturelle FORÊT DE LA MASSANE

Laboratoire Arago, 66650 Banyuls-sur-Mer  
www.rnmassane.fr - 04.68.88.12.18



**Crédits iconographiques**  
Toutes les photos sont de Élodie Magnanou, sauf RNN  
Massane J. Garrigue : page 11 ; D. Sorel : pages 20-21.

## Résumé

---

Cette étude investigate le régime de reproduction de la hêtraie de la Massane grâce à une approche démogénétique. Pour cela, nous avons bénéficié du suivi cartographique de la forêt et procédé au génotypage des semis établis et des adultes reproducteurs sur deux placettes de la réserve intégrale. Le régime de reproduction est essentiellement allogame, avec un taux d'autofécondation négligeable, contrairement à ce qui est souvent observé pour des populations marginales. Les contributions des adultes à la reproduction ne sont pas équilibrées. Ces variations interindividuelles des fécondités mâles et femelles sont largement expliquées par le diamètre des reproducteurs. Nous avons identifié en outre un effet négatif du dépérissement foliaire (mesuré par la présence de branches mortes) sur la fécondité femelle, mais pas sur la fécondité mâle. La présence du champignon lignicole *Oudemansiella mucida* n'a en revanche pas d'effet systématique sur la fécondité, mâle ou femelle.

## Abstract

---

We coupled genotyping approaches to the ongoing GIS forest survey to characterize the reproductive mode of the Massane beech forest. For this purpose, we genotyped adult trees and seeding from two study sites within the fenced plot of the reserve. The reproductive mode is dominated cross-pollination, with a negligible autofecundation rate. This pattern is quite unusual for marginal forests. The relative contribution of males and females adult trees to the regeneration (fecundity) mainly varies according to their diameter (an age proxy). There is a negative effect of tree dieback (based on the observation of dead branches) on female fecundity but not on male one. We found no clear correlation between the presence of *Oudemansiella mucida* – a wood fungi and tree male or female fecundity.

