HAL open science

Variabilité spatio-temporelle de la reproduction chez une espèce pérenne monoïque: Le cas du sapin pectiné, Abies alba Miller, en limite sud de son aire de répartition

Gwendal Restoux

▶ **To cite this version:**

Gwendal Restoux. Variabilité spatio-temporelle de la reproduction chez une espèce pérenne monoïque: Le cas du sapin pectiné, Abies alba Miller, en limite sud de son aire de répartition. Sciences de l'environnement. Université Paul Cézanne (Aix Marseille 3), Aix-en-Provence, FRA., 2009. Français. NNT : . tel-03906575

HAL Id: tel-03906575

https://hal.inrae.fr/tel-03906575v1

Submitted on 19 Dec 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITE PAUL CEZANNE - AIX MARSEILLE III

N° attribué par la bibliothèque

THÈSE

pour obtenir le grade de

Docteur de L'université Paul Cézanne - Aix-Marseille III

Spécialité : **Ecologie, Génétique des populations**

préparée au laboratoire **Ecologie des Forêts Méditéranéennes - UR629**

INRA Avignon

dans le cadre de l'École Doctorale **Sciences de l'environnement**

présentée et soutenue publiquement

par

Gwendal Restoux

le 12 Juin 2009

Titre:

Variabilité spatio-temporelle de la reproduction chez une espèce pérenne monoïque: Le cas du sapin pectiné, *Abies alba* Miller, en limite sud de son aire de répartition

Directeur de thèse: **Bruno FADY - INRA, Avignon**

Co-directeur de thèse: **Etienne KLEIN - INRA, Avignon**

Jury

Dr. Catherine BASTIEN - INRA, Orléans,	Rapporteur
Dr. Nathalie FRASCARIA-LACOSTE - Université Paris Sud, Paris,	Rapporteur
Dr. Alex BAUMEL - Université Paul Cézanne, Marseille,	Examinateur
Pr. Frederic MEDAIL - Université Paul Cézanne, Marseille,	Examinateur
Dr. Emmanuelle PORCHER - MNHN, Paris,	Examinateur
Dr. Bruno FADY - INRA, Avignon,	Directeur de thèse
Dr. Etienne KLEIN - INRA, Avignon,	Co-directeur de thèse

Résumé

Cette thèse s'intéresse aux variations spatio-temporelles de la reproduction du sapin pectiné (*Abies alba*), en limite sud de son aire de répartition dans les populations du mont Ventoux et de la montagne de Lure. En général, les populations marginales présentent des densités faibles et variables. Pour nos populations de sapins, nous avons montré que les densités faibles entraînent une augmentation des taux d'autofécondation par loi d'action de masse, à l'échelle de la population mais également individuelle. Les populations moins denses présentent également une réduction du taux de germination par expression de la dépression de consanguinité, mais qui est compensée par une meilleure survie des semis pouvant provenir d'une purge plus efficace du fardeau génétique. Cependant, la diminution de la densité entraîne dans nos populations une baisse de la production efficace de graines par augmentation du taux de graines vides, probablement à cause d'une dispersion du pollen limitée. En effet, nos analyses génétiques conduisent à une distance de dispersion moyenne estimée de 20 à 50m. Cette faible dispersion induit également une forte variabilité individuelle du régime de reproduction liée aux variations locales de densité et de productions individuelles de pollen. Finalement, nous avons trouvé des variations inter-annuelles importantes de la structure génétique spatiale du nuage de pollen avec une différenciation entre arbres-mères de 8% et une différenciation entre années au sein des mères de 13%. Les années avec les plus fortes productions de pollen étaient également celles avec la plus faible diversité génétique globale. Cette étude confirme donc la nécessité de considérer de multiples années quand on s'intéresse à la reproduction d'une espèce pérenne comme le sapin, à la fois pour des études théoriques et pour des questions centrées sur la gestion de la diversité.

Mots-clefs : Variabilité spatio-temporelle, succès reproducteur, populations marginales, régime de reproduction, dépresion de consanguinité, dispersion pollinique.

Abstract

This study focuses on the spatio-temporal varibility of reproduction in silver fir(*Abies alba*), in its French rear edge populations in the Mont Ventoux and Lure mountain. Usually, marginal populations contrast from core populations with lower and more variable densities. In our populations of firs, we found that lower densitiy leads to higher selfing rates through a mass-action law at both population and individual levels. As a consequence, germinative abilities were lower for seeds from lower density populations, mainly because of inbreeeding depression. But this was balanced by a higher survival rates at the seedling stage. A more efficient genetic purge could explain this. However, in our populations, a lower density also induced a lower effective seed production, through an increased rate of empty seeds. This can result from pollen limitation. Indeed, genetic analyses revealed a rather high spatial genetic structure of pollen pool, indicating a limited dispersal with a mean dispersal distance between 20 and 50m. This limited dispersal also induced a high inter-individual variability of the mating system mainly due to variations of local density and individual pollen production. Finally, we found strong inter-annual variations of the pollen pools : giving an inter-mothers differentiation of 8%, we obtained an inter-annual differenciation of 13% within mothers. The year with the higher pollen productions also presented the lower overall genetic diversity. Thus this study confirms the necessity to consider multiple years in studies on the reproduction of a perennial tree such as fir, either for theoretical concerns of for practical questions focused on the management of genetic diversity.

Keywords : Spatio-temporal variability, reproductive successes, marginal populations, mating system, inbreeding depression, pollen dispersal.

Table des matières

Résumé iii

Abstract iv

Table des matières v

Introduction **1**

1 Avant propos 1

1.1 Contexte général 1

1.2 But de l'étude 4

1.3 Organisation de la thèse 5

2 La reproduction et les régimes mixtes 7

2.1 L'autofécondation et les régimes mixtes de reproduction . . . 7

2.2 Considérations évolutives concernant les régimes mixtes de reproduction 9

2.2.1 Modèles écologiques 10

2.2.2 Modèles génétiques 12

2.2.3 Les régimes mixtes de reproduction, une énigme insoluble ? 14

3 Démo-génétique et reproduction 18

3.1 La démographie 18

3.1.1 La théorie : approche matricielle 18

3.1.2 Phénomènes démographiques 21

3.2 Approche génétique 23

3.3 La démo-génétique et la biologie de la conservation 24

4 Les arbres : un modèle biologique particulier 28

4.1 Extinction et recolonisation : Le paradoxe de la diversité génétique . . . 28
4.2 A quelle échelle intégrer la production de graines ? . . . 30
4.3 Les flux de pollen : Un déterminant essentiel ? . . . 31
5 Objectif de la thèse . . . 34

1 Sapin pectiné et sites d'études 35

1 Le sapin pectiné, *Abies alba* Miller . . . 35
2 Les sites d'études . . . 41
2.1 Le Mont Ventoux . . . 41
2.2 La montagne de Lure . . . 43
3 Les dispositifs expérimentaux . . . 44
3.1 Localisation et échantillonnage . . . 44
3.2 Estimation de la fertilité . . . 49
3.2.1 Mesure des fertilités femelles et mâles . . . 49
3.2.2 Estimation quantitative de la fertilité mâle . . . 49
3.3 Estimation de la densité locale . . . 50
3.4 Estimation du succès reproducteur femelle . . . 51
3.5 Génotypage . . . 53

2 Aire de répartition et régime de reproduction 55

1 Synthèse . . . 55
2 Life at the margin : the mating system of Mediterranean conifers . . . 58

3 Succès reproducteur femelle 69

1 Synthèse . . . 69
2 Variability of female reproductive success in the conifer tree *Abies alba* : pollen quantity, pollen quality and climate . . . 71

4 Détection de la dépression de consanguinité 109

1 Synthèse . . . 109
2 Variability of individual genetic load : consequences for the detection of inbreeding depression through a regression approach . . . 111

5 Succès reproducteur mâle **141**

1 Synthèse 141

2 Introduction 143

2.1 L'importance des flux de gènes et son lien avec le succès reproducteur 143

2.2 Estimation du régime de reproduction et des flux de gènes . . 144

2.2.1 Flux de gènes "historiques" 144

2.2.2 Les flux de gènes instantanés 146

2.3 Variabilité spatio-temporelle du régime de reproduction et des flux de gènes 149

2.3.1 Variabilité spatiale 149

2.3.2 Variabilité temporelle 150

2.4 Objectifs de cette étude 151

3 Matériel & Méthode 154

3.1 Matériel biologique 154

3.2 Méthode 154

3.2.1 Caractérisation de la dispersion du pollen par l'approche TwoGener 155

3.2.2 Estimation du régime de reproduction par les modèles d'appariement 165

4 Résultats 173

4.1 Estimation de la structure génétique spatiale du nuage de pollen et de la dispersion pollinique 173

4.1.1 Approche globale 173

4.1.2 Approche par couples de mères, Φ_{ft-pw} 174

4.2 Estimation du régime de reproduction 183

4.2.1 Taux d'autofécondation individuels 183

4.2.2 Estimation des corrélations de paternité 183

4.2.3 Comparaison des modèles 186

4.3 Variabilité spatio-temporelle du régime de reproduction 186

4.3.1 Le taux d'autofécondation 188

4.3.2 La corrélation de paternité 190

5 Discussion 194
5.1 Structure génétique spatiale du nuage de pollen et dispersion . 194
5.1.1 Structure génétique spatiale du nuage pollinique . . . 194
5.1.2 Dispersion du pollen 195
5.1.3 Impact de l'autofécondation 197
5.1.4 Descripteurs de la fonction de dispersion 199
5.2 Variabilité du régime de reproduction 201
5.2.1 La corrélation de paternité globale, $Q_0^{s-r_p}$ 201
5.2.2 Le taux d'autofécondation, s 201
5.2.3 La corrélation de paternité au sein du nuage de pollen allofécondant, r_p 203
5.3 Variabilité temporelle de la structure génétique spatiale du nuage pollinique 206
6 Conclusion 212

Conclusion **217**

Bibliographie **227**

Introduction

1 Avant propos

1.1 Contexte général

Le but de cette thèse est de décrire et d'expliquer la variabilité dans l'espace et le temps de la reproduction chez le sapin pectiné *Abies alba* Miller. Les populations étudiées se trouvent sur le mont Ventoux et sur la montagne de Lure. Elles sont en situation marginale et constituent la limite sud de l'aire française de répartition du sapin dans les Alpes (figure 1.6). Cette situation de limite écologique les rend extrêmement vulnérables au changement climatique prévu pour les décennies à venir (Parmesan 2006, Mimura et Aitken 2007*b*, Kawecki 2008). Or le sapin pectiné est une espèce importante pour le maintien de la biodiversité, en particulier dans nos sites d'études, puisqu'il constitue avec le hêtre (*Fagus sylvatica*) les bases d'un écosystème auquel sont inféodées de nombreuses espèces tant animales que végétales (Blondel 2007). La conservation de ces populations est donc d'un intérêt majeur. Une politique de restauration a été mise en place entre la seconde moitié du XIXeme et le début du XXeme siècle (voir chapitre 1) afin de lutter contre la disparition de ces forêts causée par la surexploitation (production de bois) et un pastoralisme intensif (Cheikh Al Bassatneh et al. 2007, Simon et al. 2007). Ainsi, les sapinières du Mont Ventoux et de la montagne de Lure sont un modèle de choix pour comprendre comment l'homme peut modifier son environnement, en fragilisant les populations (surexploitation hier et changement climatique aujourd'hui) ou au contraire en les conservant (politiques de restauration, gestion forestière).

Le maintien des populations face à un changement de leur environnement peut

s'effectuer selon trois modalités, combinées ou non : **i**) migrer, **ii**) s'accommoder en utilisant la plasticité phénotypique, ou **iii**) s'adapter génétiquement au nouvel environnement. L'objet de cette thèse porte sur cette dernière hypothèse, notre but étant de comprendre le potentiel évolutif de nos populations et donc leur capacité à se maintenir dans le futur.

La durée de vie d'un arbre forestier excède fréquemment le siècle, leur conservation implique ainsi certaines spécificités (Petit et Hampe 2006). Le sapin pectiné possède une durée de vie moyenne de l'ordre de 300 ans et un âge moyen de première reproduction assez élevé (40 ans en moyenne) ce qui conduit à un temps de génération très long. Or les climatologues prédisent un réchauffement climatique important et rapide avec $1,1$ à $6,4$°c d'augmentation en moyenne d'ici à 2100 (IPCC 2007), ce qui pourra avoir des conséquences dramatiques notamment pour le maintien des populations à la limite sud de leurs aires de répartition (Parmesan 2006, Hampe et Petit 2005). Le potentiel évolutif de nos populations marginales de sapin devra donc être particulièrement élevé pour pouvoir répondre très rapidement au changement climatique qui est prévu à une échelle de temps au moins aussi courte que celle de leur temps de génération. Les paramètres clefs du potentiel évolutif et du maintien des populations sont **i**) la démographie à travers la production de descendants en nombre suffisant et **ii**) la conservation d'une diversité génétique élevée, gage de capacité adaptative (Lande 1988, Spielman et al. 2004, Coulson et al. 2006). L'approche à adopter pour étudier le devenir des populations est donc celle de la démo-génétique qui va s'intéresser à ces deux aspects et à leur interaction avec la stochasticité environnementale et démographique induite respectivement par le réchauffement climatique d'une part, et la taille relativement réduite des populations (i.e. populations marginales) d'autre part (figure 1).

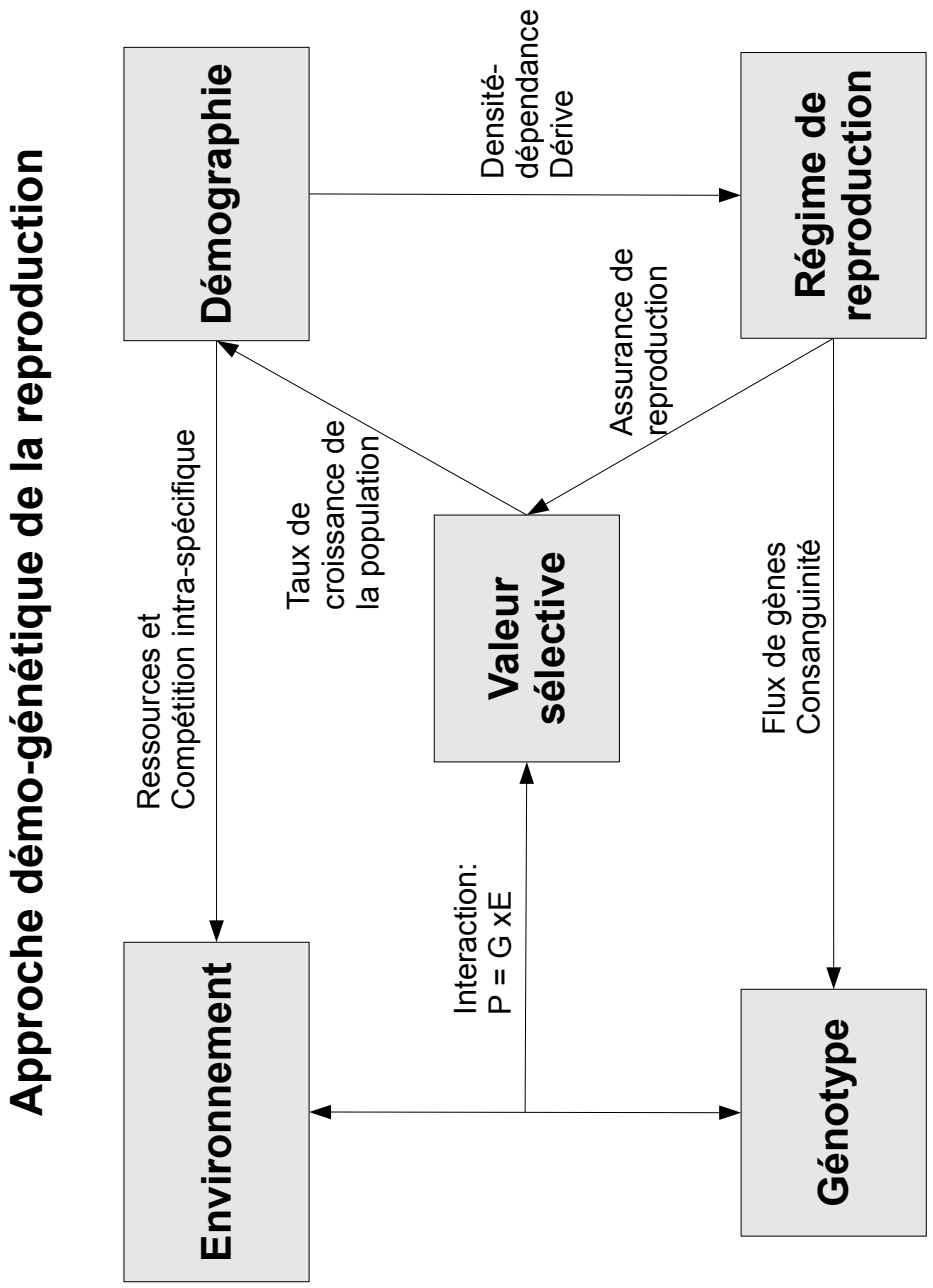


Fig. 1: **Démo-génétique et reproduction.** Schéma fonctionnel des liens existants entre la démographie et la génétique dans le cadre de la reproduction et plus particulièrement l'effet des flux de gènes et du régime de reproduction.

1.2 But de l'étude

Dans cette thèse, nous avons étudié le succès reproducteur des individus au stade graine. Dans le cas des espèces monoïques comme le sapin pectiné, le succès reproducteur individuel se décompose en succès reproducteur mâle et succès reproducteur femelle. Au stade graine, le premier est représenté par le nombre d'ovules fécondés par le pollen de l'individu étudié qui donneront une graine viable. Le second, quant à lui, est représenté par le nombre de graines viables produites par l'individu étudié. L'étude de ces paramètres à une échelle individuelle nous permet de quantifier le déséquilibre de contribution des individus à la reproduction, ce qui détermine l'effectif efficace de la population. Pour une année donnée, la fertilité individuelle chez les arbres est très variable notamment en fonction de la taille des individus (Petit et Hampe 2006, Kang et al. 2003). Tous les individus participent donc inégalement à chaque épisode de reproduction en fonction notamment de leur âge. Nous avons caractérisé ce déséquilibre de contributions des individus à la fois pour la voie mâle et pour la voie femelle. Ce déséquilibre a été étudié de manière spatiale à l'échelle d'un épisode de reproduction, et de manière temporelle en considérant les quatre années qu'a duré le suivi de ces populations.

Les arbres ont la capacité de participer à de nombreux épisodes de reproduction au cours de leur vie, leur fertilité et donc leur succès reproducteur peuvent varier d'une année à l'autre en fonction des conditions climatiques, de leur âge ou du *masting* (Kang et al. 2003, Kelly et Sork 2002, Satake et Iwasa 2002). Il est donc essentiel d'intégrer la variabilité temporelle pour comprendre comment évolue le déséquilibre de contribution des individus à la reproduction. Notre suivi pendant quatres années est probablement insuffisant pour estimer le succès reproducteur réel des arbres dont la durée de vie est bien supérieure, mais il nous permet d'estimer au moins partiellement les variabilités inter-annuelle et inter-individuelle. Ce n'est qu'à cette condition que l'on pourra estimer les paramètres clefs de la conservation des populations et notamment le taux d'accroissement moyen, l'effectif efficace mais aussi les stochasticités environnementales et démographiques (Lande 1988, Coulson et al. 2006).

1.3 Organisation de la thèse

Afin de faciliter la lecture de ce document de thèse nous allons décrire brièvement son organisation. L'introduction présente quelques notions utiles à la compréhension de la suite de la thèse :

- Dans un premier temps nous présenterons la reproduction et plus particulièrement les régimes mixtes de reproduction et les causes évolutives expliquant le maintien d'une telle stratégie.
- Dans une seconde partie, nous présenterons l'approche démo-génétique dans le cadre de la biologie de la conservation et en particulier celle que nous avons choisi d'adopter.
- Enfin, nous soulignerons les caractéristiques des arbres qui font d'eux des modèles biologiques particuliers et pour lesquels une approche spatio-temporelle est particulièrement appropriée.

Le corps de la thèse est, quant à lui, composé de cinq chapitres :

- Le premier chapitre présente le matériel biologique, le sapin pectiné, et les sites d'étude du Mont Ventoux et de la montagne de Lure. C'est également dans cette partie que sont présentés les différents échantillonnages et les différentes mesures effectués.
- Le second chapitre traite des régimes mixtes de reproduction et discute de la stratégie évolutivement stable qu'ils pourraient constituer pour les populations en marge de l'aire de distribution de l'espèce, là où les stochasticités environnementales et démographiques sont les plus fortes (i.e. impact sur la densité).
- Le troisième chapitre présente une étude du succès reproducteur femelle effectuée sur les sites du Mont Ventoux et de la montagne de Lure pendant quatre ans. Nous y avons étudié la variabilité inter-annuelle et inter-individuelle de la production de graines viables en relation avec des facteurs génétiques (dépression de consanguinité) et démographiques (limitation pollinique).

- Le quatrième chapitre pose la question de la détectabilité de la dépression de consanguinité lorsque le fardeau génétique est hétérogène entre les individus d'une population. Cette hypothèse fait écho au chapitre précédent dans lequel nous n'avons pas pu prouver d'effet de la dépression de consanguinité sur la production de graines viables, ce qui est étonnant chez un conifère. Cette étude repose sur des simulations et présente également quelques recommandations concernant les plans expérimentaux permettant de maximiser la détection de la dépression de consanguinité.
- Le cinquième et dernier chapitre est, quant à lui, consacré à la variabilité inter-individuelle et inter-annuelle des flux de pollen et du régime de reproduction dans la population de sapins de la montagne de Lure entre 2002 et 2005. A partir de marqueurs moléculaires, nous avons estimé les paramètres de structuration génétique spatiale des nuages polliniques à partir desquels nous avons estimé la fonction de dispersion du pollen. Nous avons également estimé les paramètres individuels du régime de reproduction. Enfin, nous avons mis en relation ces paramètres avec la densité ou la fertilité mâle, afin de comprendre quels étaient les déterminants du déséquilibre de contributions des individus à la reproduction *via* la voie mâle, dans l'espace et le temps.

Les chapitres deux à quatre sont structurés de manière identique, composés d'une synthèse en français suivie d'un article en anglais. Il y a donc deux niveaux de lectures possibles : synthétique ou approfondi en fonction de l'intérêt du lecteur pour chacune des parties traitées.

Finalement une synthèse générale fait le bilan des résultats et présente les applications et les perspectives envisageables pour donner suite à ce travail de thèse notamment dans le cadre de la biologie de la conservation. Cette conclusion discute également de manière plus générale de l'intérêt de tenir compte des variabilités spatiales et temporelles lors de l'étude de la reproduction des arbres et des espèces pérennes en général. Elle souligne également quelques aspects des populations marginales et l'intérêt de protéger ces dernières dans le cadre de la biologie de la conservation.

2 La reproduction et les régimes mixtes

Le cycle de vie des individus est finalisé lorsque ceux-ci produisent des descendants, c'est la reproduction. Ce compartiment du cycle de vie est donc essentiel au maintien des populations et *de facto* à celui des espèces. La reproduction revêt un intérêt particulier lorsque l'on s'intéresse à l'évolution des populations d'un point de vue écologique ou génétique, puisque la notion de valeur sélective (ou *fitness*) repose en partie sur le nombre de descendants produits par individu ou encore sur le nombre de copies de son génome transmises à sa descendance. Chez les organismes sessiles et les plantes en particulier, c'est la phase de reproduction qui va permettre la colonisation de nouveaux milieux par le biais de la dispersion de propagules. La reproduction va donc modifier la démographie et la structure des populations (Wang et Smith 2002, Wright 1943; 1946). Il convient cependant de distinguer deux types de reproduction : i) la reproduction asexuée qui va produire des clones des parents à quelques exceptions près (mutations somatiques) et ii) la reproduction sexuée, qui, quant à elle, fait intervenir deux gamètes (obtenus par meïose) lors d'un épisode de fécondation afin de produire un ou plusieurs descendants. C'est uniquement à ce dernier type de reproduction que nous nous intéresserons dans la suite de cette thèse. La reproduction sexuée implique donc, en théorie, la présence de deux individus dont l'union des gamètes aboutira à la formation d'un ou plusieurs nouveaux individus. Bien que la reproduction sexuée repose sur l'union de deux gamètes, en réalité elle ne nécessite pas toujours l'implication de deux individus adultes.

2.1 L'autofécondation et les régimes mixtes de reproduction

L'autofécondation est une forme de reproduction sexuée particulière dans laquelle les deux gamètes proviennent d'un unique individu et conduisent à la formation d'un descendant qui n'est pas nécessairement génétiquement identique à son unique parent. L'autofécondation est un phénomène répandu chez les plantes et qui a suscité l'intérêt des biologistes dès les prémices de la théorie de l'évolution (Darwin 1876). Ceci peut avoir lieu lorsque les organes sexuels (fleurs) remplissent les fonctions mâles et femelles, on parle alors d'hermaphrodisme au sens strict, ou bien lorsque

les organes mâles et femelles sont distincts mais présents sur le même individu, dans ce dernier cas on parle alors de monoécie. L'autofécondation est beaucoup plus rare chez les animaux (Tsitrone et al. 2003*b*, les mollusques en sont l'exemple le plus connu). L'autofécondation peut être obligatoire pour les espèces dont la morphologie ne permet pas l'import de gamètes extérieurs. C'est notamment le cas des espèces cléïstogames dont les fleurs hermaphrodites restent closes lors de la période de pollinisation, cette dernière est alors interne (Oakley et Winn 2008). A l'inverse beaucoup d'espèces ont développé des mécanismes pour éviter l'autofécondation : les systèmes d'auto-incompatibilité (Castric et Vekemans 2004, par exemple), les décalages phénologiques (McCarthy et Quinn 1990, par exemple) ou l'hétérostylie (Shimono et Washitani 2002, par exemple). Finalement pour d'autres espèces la capacité de s'autoféconder est facultative, leur régime de reproduction est dit mixte. Ainsi, la descendance des espèces à régime de reproduction mixte est constituée, ou peut potentiellement être constituée, d'individus issus d'autofécondation et d'allofécondation. On distingue chez les espèces à régime mixte de reproduction trois "types" d'autofécondation (Lloyd 1979) :

- L'autofécondation dite décalée (*delayed selfing*), dans ce cas l'autofécondation intervient seulement lorsque qu'il n'y a plus d'opportunités de s'alloféconder, c'est-à-dire à la fin de la saison de pollinisation. Ceci est le cas chez certaines espèces où à la fin de la période d'émission du pollen, la déhiscence des anthères va les faire entrer en contact avec les étamines situées à la base de celles-ci (Goodwillie et al. 2005). Alors, si des ovules demeuraient encore non fécondés à la fin de la période de pollinisation, ils pourront être autofécondés. Cette stratégie de reproduction semble être, pour des raisons que nous détaillerons dans la suite de cette introduction, avantageuse d'un point de vue écologique et génétique (Tsitrone et al. 2003*a*).
- L'autofécondation "primaire" (*primary selfing*), dans ce cas l'autofécondation intervient avant toute possibilité d'allofécondation. Ceci est par exemple le cas lorsque les fleurs hermaphrodites ne s'ouvrent que tardivement, ainsi la majorité des ovules sont fécondés par le pollen de la plante avant d'être exposés au pollen extérieur et donc d'être potentiellement allofécondés. Cela peut également résulter d'un décalage phénologique, où la précocité de floraison va

virtuellement réduire la densité en conspécifique, et donc limiter les possibilités d'allofécondation, conduisant à un taux d'autofécondation élevé (Naito et al. 2008). Dans ce dernier cas, s'il y a un coût ou un bénéfice à une telle stratégie, la sélection sur la date de floraison sera fréquence-dépendante.

– L'autofécondation compétitrice (*competing selfing*), dans ce cas l'autopollen (celui conduisant à l'autofécondation) entre en compétition avec l'allopollen (Leclerc-Potvin et Ritland 1994).

C'est à ce dernier type d'autofécondation que nous allons principalement nous intéresser dans cette thèse. C'est de loin le cas le plus courant de régime mixte de reproduction au sein du règne végétal (Goodwillie et al. 2005), bien que l'autofécondation décalée semble être une stratégie optimale dans bien des cas (Tsitrone et al. 2003*a*).

2.2 Considérations évolutives concernant les régimes mixtes de reproduction

Les régimes mixtes de reproduction soulèvent de nombreuses questions en terme d'intérêt évolutif. Leur existence et surtout leur maintien a été et est encore au coeur de nombreux débats, en écologie et en génétique (Holsinger 1991, Goodwillie et al. 2005, Porcher et Lande 2005*a*, Cheptou et Schoen 2007). Les régimes mixtes de reproduction sont connus et étudiés depuis les prémices de la botanique et des sciences l'évolution (Darwin 1876). Pourtant leur importance relative au sein du règne végétal a été sous-estimée vers le milieu des années 80 par les généticiens. En effet, Schemske et Lande (1985) observent une distribution bimodale des régimes de reproduction, majoritairement allogames ou autogames (figure 2), arguant alors que les régimes mixtes de reproduction ne sont pas une stratégie évolutivement stable. Vingt ans plus tard Goodwillie et al. (2005) observent, grâce à une étude bibliographique mise à jour, que les espèces majoritairement autogames sont en réalité moins fréquentes que dans l'étude précédente (figure 2). Les régimes mixtes de reproduction demeurent un mystère évolutif car différentes causes peuvent être invoquées d'un point de vue écologique et/ou génétique pour justifier un aspect transitoire ou

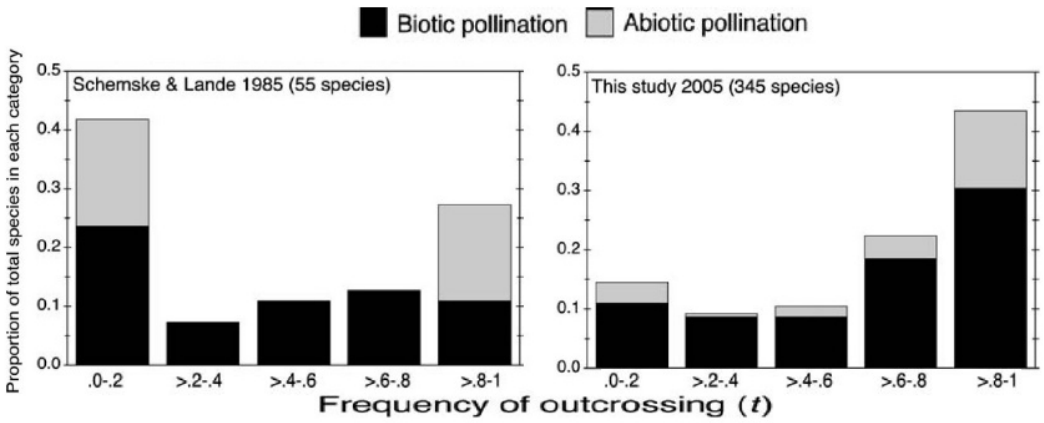


FIG. 2: Distribution des régimes de reproduction chez les végétaux (i.e. taux d'allofécondation par classes en abscisse), avec en gris les espèces dont la pollinisation est abiotique (anémophiles principalement) et en noir lorsqu'elle est biotique (entomophiles, oiseaux, mammifères). Le graphique de gauche représente l'étude bibliographique effectuée par (Schemske et Lande 1985), alors que le graphique de droite représente celle mise à jour par (Goodwillie et al. 2005). Figure extraite de Goodwillie et al. (2005).

stable de ce mode de reproduction. Nous allons décrire ici les principaux avantages ou désavantages imputés à l'autofécondation et leurs conséquences évolutives sur les régimes mixtes de reproduction.

2.2.1 Modèles écologiques

D'un point de vue écologique deux théories s'opposent quant à l'avantage sélectif de l'autofécondation sur l'allofécondation. Elles aboutissent à des prévisions différentes concernant la stabilité évolutive des régimes mixtes de reproduction :

- L'autofécondation confère un avantage aux individus la pratiquant en leur permettant de se reproduire sexuellement, indépendamment de l'accès aux gamètes des autres partenaires (*reproductive assurance*). Cette caractéristique peut s'avérer particulièrement avantageuse lorsque la densité d'une population est faible (Kalisz et Vogler 2003, Cheptou 2004, Kennedy et Elle 2008). C'est d'ailleurs sur cette idée que Baker (1955; 1967) énonce sa loi liant les capacités de dispersion à longue distance (i.e. événement rare) et l'absence de mécanismes d'auto-incompatibilités. En effet, ces deux caractères permettent

la colonisation d'un nouveau milieu même avec une densité très faible (ex : les îles).

- Cependant l'autofécondation a un coût non négligeable, simplement parce que le nombre de gamètes produits par un individu est un nombre fini. Ainsi, le pollen utilisé pour s'autoféconder est perdu pour l'allofécondation et occasionne donc un coût dès lors que l'on exporte moins de pollen que les autres individus de la population, ce concept est connu sous le terme anglo-saxon de *Pollen discounting* (Harder et Wilson 1998).

Holsinger (1991) a étudié l'évolution du régime de reproduction en ne considérant que des aspects écologiques. Il a défini un modèle de loi d'action de masse dans lequel le taux d'autofécondation est le résultat de la dilution sur les stigmates de son propre pollen parmi l'ensemble du pollen fécondant exporté, suivant la formule suivante :

$$s = \frac{\alpha.Q_s}{Q_o + \alpha.Q_s}, \tag{1}$$

avec Q_s et Q_o, les quantités respectives de pollen autofécondant (*selfed*) et allofécondant (*outcrossed*) arrivant sur les stigmates et α la capacité fécondante du pollen autofécondant relativement au pollen allofécondant. Ainsi, le taux d'autofécondation dépend à la fois de la quantité de pollen produit mais également de la quantité de pollen exporté par les autres individus de la population, il est donc densité (i.e. quantité de pollen reçu, Q_o) et fréquence-dépendant (i.e. quantité de pollen produit, Q_s). En utilisant ce modèle de loi d'action de masse et en considérant le coût de la perte des gamètes mâles (*pollen discounting*), Holsinger (1991) montre qu'il est possible de conserver des régimes mixtes de reproduction à deux conditions :

- La densité de la population ne doit pas être trop forte de sorte que la probabilité d'exporter son pollen (i.e. qu'il féconde d'autres plantes) ne soit pas trop élevée d'après la loi d'action de masse, sinon l'allofécondation devient la stratégie évolutivement stable.
- La quantité de pollen ne doit pas être limitante de sorte que tous les ovules doivent être fécondés (concept de *pollen limitation*) sinon l'autofécondation devient la stratégie évolutivement stable en procurant une assurance de reproduction.

Ainsi l'avantage de l'autofécondation devient dépendant de la densité des individus (assurance de reproduction et export de pollen). Il est également dépendant de la fréquence des différents morphes autofécondant, de sorte qu'il ne faut pas exporter moins de pollen que les autres individus de la population (*pollen discounting*). Les régimes mixtes de reproduction demeurent donc évolutivement stables dès lors que le taux d'autofécondation conserve la possibilité de fluctuer.

2.2.2 Modèles génétiques

Lorsque l'on considère la stabilité évolutive des régimes mixtes de reproduction du point de vue génétique, deux modèles se confrontent quant à l'avantage évolutif de l'autofécondation :

- L'avantage automatique de l'autofécondation, basé sur le fait qu'elle permet la transmission de 50% de matériel génétique supplémentaire à la descendance par rapport à l'allofécondation, avec en moyenne deux copies par descendance autofécondée (voies mâle et femelle) et une par allofécondation (par la voie mâle en tant que parent d'ovules allofécondés) contre deux copies pour les individus hermaphrodites strictement allogames (Fisher 1941). Si l'autofécondation est déterminée génétiquement, elle est automatiquement sélectionnée au sein d'une population strictement allogame puisqu'elle favorise sa propre transmission.
- La dépression de consanguinité (δ) qui constitue l'argument génétique le plus couramment opposé à l'avantage automatique de l'autofécondation. Elle résulte d'une baisse de valeur sélective des individus consanguins par rapport aux individus non consanguins (Charlesworth et Charlesworth 1987). La dépression de consanguinité est définie de la façon suivante :

$$\delta = 1 - \frac{\omega_s}{\omega_o}, \tag{2}$$

avec ω_s et ω_o, les valeurs sélectives des descendances autofécondées et allofécondées respectivement. La consanguinité résulte de croisements entre individus apparentés, l'autofécondation conduisant à l'apparentement maximum entre les parents. La consanguinité se traduit par une augmentation de l'i-

dentité par descendance des deux gènes homologues d'un individu qui induit de fait une augmentation de l'homozygotie au niveau de la population. La baisse de valeur sélective des individus consanguins peut être expliquée par deux modèles :

- Le modèle de dominance partielle, où la baisse de la valeur sélective est due à un état homozygote d'un allèle délétère récessif par rapport à un allèle "sain". Dans le cas simple d'un locus à deux allèles (A et a) les valeurs sélectives s'écrivent de la manière suivante :

Génotype	**AA**	**Aa**	**aa**
Valeur sélective	1	1-hs	1-s

  où h représente le coefficient de dominance de l'allèle **a** sur l'allèle **A** (de 0 si **a** est strictement récessif à 1 s'il est strictement dominant) et s représente le coefficient de sélection associé à l'allèle délétère **a** soit la baisse de valeur sélective de **a** par rapport à **A** (de 0 si **a** procure une valeur sélective identique à **A** jusqu'à 1 s'il est létal).
- Le modèle de superdominance ou hétérosis. Dans ce cas, la valeur sélective de l'hétérozygote est toujours supérieure aux valeurs sélectives des homozygotes pour les deux allèles considérés (dans le cas d'espèces diploïdes). Cependant ce modèle ne semble pas nécessaire pour expliquer le maintien des stratégies mixtes de reproduction (Charlesworth et Charlesworth 1987), de plus les cas où il a pu être mis en évidence dans la nature sont assez rares (Roff 2002).

Lande et Schemske (1985) ont avancé l'hypothèse de purge génétique permettant d'appréhender le fardeau génétique de manière dynamique, ce dernier évoluant conjointement avec le taux d'autofécondation. Ils considèrent à la fois l'avantage automatique de l'autofécondation et la baisse de valeur sélective induite par la dépression de consanguinité dans le cadre d'un modèle de dominance partielle. Leur modèle conduit à deux états stables qui ne dépendent que de la valeur initiale de la dépression de consanguinité qui peut ou non contrebalancer l'avantage de 50% donné à l'autofécondation :

- si $\delta < \frac{1}{2}$ alors la population doit évoluer vers une autogamie stricte. En effet

dans ce cas la perte de valeur sélective engendrée par la dépression de consanguinité ($< \frac{1}{2}$) est inférieure à l'avantage procuré par l'autofécondation ($\frac{1}{2}$). L'autofécondation et donc la consanguinité vont donc augmenter en fréquence et le fardeau génétique pourra alors être purgé.

- si $\delta > \frac{1}{2}$ alors la population doit évoluer vers une allogamie stricte. Dans ce cas, l'autofécondation est désavantagée et ne peut contrebalancer le coût de la dépression de consanguinité. Alors l'allofécondation va augmenter en fréquence et le fardeau génétique ne sera pas purgé (i.e. il sera stable ou augmentera suivant le taux de mutation). En revanche la consanguinité devrait diminuer (baisse limitée par l'effet de la consanguinité biparentale en population finie).

Ce modèle théorique était illustré par une analyse bibliographique effectuée chez plusieurs espèces végétales (Schemske et Lande 1985). Elle montrait une distribution bimodale des régimes de reproduction avec une minorité d'espèces présentant des régimes mixtes de reproduction, considérées alors comme étant dans un état évolutif transitoire. Cependant, ce modèle de fardeau génétique dynamique ne semble plus suffisant aujourd'hui. En effet, Goodwillie et al. (2005) montrent par une étude bibliographique récente que les régimes mixtes de reproduction sont en réalité relativement fréquents au sein du règne végétal. Ceci soulève donc à nouveau la question de l'état transitoire des régimes mixtes de reproduction.

2.2.3 Les régimes mixtes de reproduction, une énigme insoluble ?

Les modèles écologiques et génétiques semblent avoir identifié tous les processus sélectifs associés à l'autofécondation. Cependant, il existe des situations où des phénomènes peuvent modifier l'effet de ces processus sélectifs. Par exemple, le modèle décrit par Lande et Schemske (1985) considère que l'effet de la dépression de consanguinité est constant dans le temps et pour tous les individus. Pourtant plusieurs études montrent que l'effet de la dépression de consanguinité peut être variable en fonction du niveau de stress auquel sont exposés les individus (Cheptou et al. 2000*b*). Le stress pourra être induit par un accès difficile aux ressources (lumière, eau...) ou par la compétition, facteurs variables à l'échelle du micro-environnment (ex : densité) ou dans le temps (ex : années plus ou moins sèches), ce qui va modifier la sélection

sur le taux d'autofécondation. Un autre exemple est celui des végétaux pratiquant la polyembryonie, c'est-à-dire la production de plusieurs embryons au sein d'une même graine dont un seul survivra, généralement celui avec la meilleure valeur sélective (Lindgren 1975). Dans ce cas l'effet de la dépression de consanguinité est d'autant plus faible que le nombre d'ovules qu'il est possible de féconder au sein d'une graine est élevé (Sorensen 1982), ce qui va modifier les équilibres prédits par Lande et Schemske (1985) (Porcher et Lande 2005*b*). Finalement, l'étude bibliographique de Byers et Waller (1999) montre que la taille des populations et leur histoire démographique peuvent également modifier l'efficacité de la purge génétique (voir aussi Kärkkäinen et al. (1996), Hedrick et al. (1999)). Les facteurs ayant un effet sélectif sur le régime de reproduction n'agissent donc pas de manière simple. En effet, l'environnement en terme d'histoire, de compétition ou de ressources mais également les caractéristiques biologiques des espèces modifient l'évolution des régimes mixtes de reproducion.

Les processus sélectifs impliqués dans les modèles d'évolution du régime de reproduction peuvent interagir et donc en changer les conclusions. Des approches intégratives considérant à la fois les aspects écologiques (*pollen discounting* et assurance de reproduction) et génétiques (avantage automatique de l'autofécondation et dépression de consanguinité) sont donc certainement plus représentatives de la réalité. Par exemple, le modèle de loi d'action de masse décrit par Holsinger (1991) peut sembler être une explication suffisante au maintien des stratégies mixtes de reproduction. Cependant la dépression de consanguinité est un phénomène répandu chez les plantes (Kärkkäinen et al. 1996, Goodwillie et Knight 2006, Lazaro et Traveset 2006, à titre d'exemples) et elle ne peut donc pas être ignorée puisque l'on sait qu'elle va modifier la sélection opérant sur l'autofécondation (Lande et Schemske 1985, Schemske et Lande 1985, Charlesworth et Charlesworth 1987). Harder et Wilson (1998) ont abordé cette question, et concluent que les processus de pollinisation doivent être mieux étudiés afin de comprendre l'évolution des régimes de reproduction, notamment comment le *pollen discounting* se met en place et sur quelle fraction du pollen il se produit. Un autre exemple est la limitation pollinique qui est également un phénomène couramment observé chez les plantes, notamment du

fait de la fragmentation des milieux (de Jong et al. 2005, Knight et al. 2005, Lazaro et Traveset 2006). Dans ce cas, les individus pratiquant l'autofécondation possèdent un avantage sélectif grâce à l'assurance de reproduction (Lloyd 1979, He et Smouse 2002, Ashman et al. 2004). Holsinger (1991) prédit que ce type de situation devrait conduire à la fixation de l'autofécondation dans la population mais ce n'est pas toujours le cas notamment lorsque l'on observe des mécanismes d'auto-incompatibilité partielle (Gigord et al. 1998). Cheptou et Schoen (2007) montrent également qu'en combinant une disponibilité en pollen variable (i.e. limitation pollinique variable dans le temps), la dépression de consanguinité et l'avantage automatique de l'autofécondation, il est possible d'obtenir des stratégies mixtes stables. Malgré tous ces exemples, peu d'études empiriques abordent à la fois des considérations écologiques et génétiques pour comprendre l'évolution des régimes mixtes (Cheptou et Schoen 2007, où ils observent un faible nombre d'articles citant à la fois Baker (1955) et Fisher (1941), articles fondateurs en écologie et en génétique sur ce sujet). L'étude des régimes mixtes de reproduction doit donc utiliser le formalisme développé par les généticiens et l'intégrer dans les divers scénarios décrits par les écologues. Ces modèles complexes sont plus réalistes quant aux conditions nécessaires à l'obtention d'équilibres stables (Porcher et Lande 2005*a*, qui intègrent une dépression de consanguinité dynamique, le *pollen discounting* et la limitation pollinique).

Harder et Wilson (1998) considèrent que la fertilité de tous les individus est identique et que l'intensité du *pollen discounting* est fixe, ce qu'ils admettent être une limite de leur modèle. En effet, la fertilité mâle individuelle est souvent variable au sein d'une population (Kang et al. 2003, Klein et al. 2008). Cependant, les modèles théoriques font toujours, à ma connaissance, l'hypothèse d'une fertilité identique des individus (même nombre d'ovules) et d'un rapport pollen/ovule constant, ce qui n'est pas toujours vérifié. Les modèles d'évolution des régimes de reproduction sont très informatifs car ils offrent un cadre de réflexion générale, mais obligent également à faire quelques simplifications. Ainsi ils soulèvent de nombreuses questions quant à leur robustesse dans le cas de populations dont l'environnement comportera une certaine combinaison de facteurs plus ou moins en accord avec les hypothèses théoriques pour lesquelles les modèles sont valides.

Nous nous interrogeons ici sur la stabilité évolutive des régimes mixtes de reproduction chez les conifères. Chez ces espèces, l'autofécondation ne semble pas être avantageuse car la limitation pollinique est probablement absente, en raison d'une dispersion du pollen relativement élevée (Robledo-Arnuncio et al. 2004*a*, Schwendemann et al. 2007) et des populations généralement denses. De plus, la dépression de consanguinité est importante chez les conifères (Williams et Savolainen 1996). On sait également que la fertilité peut être très variable dans le temps et l'espace, ce qui risque donc de modifier les conditions d'équilibre. Toutes ces inconnues nécessitent donc des mesures et des estimations *in natura* afin de pouvoir développer des modèles réalistes se basant sur une identification des paramètres influents avec des gammes de valeurs cohérentes. Ainsi l'étude du régime de reproduction nécessite en général et dans notre cas en particulier, une approche démographique/écologique afin de mesurer le succès reproducteur au travers de la quantité de graines viables produites sous différentes conditions de densité (dans ce cas on s'intéresse à la valeur sélective absolue et non relative comme précédemment). Cependant une étude génétique est également nécessaire afin de caractériser le régime de reproduction, le niveau de consanguinité et le déséquilibre de contribution des individus à la reproduction par voie paternelle.

3 La démo-génétique : une approche nécessaire à la conservation des populations

3.1 La démographie

3.1.1 La théorie : approche matricielle

La biologie de la conservation consiste à évaluer les capacités de maintien d'une population et à apporter des solutions pour les sauvegarder (Soulé 1985). Il faut donc étudier les capacités d'accroissement des populations mais aussi les expliquer à partir des données environnementales. La démographie est une approche particulièrement adaptée à ce type de questions, puisqu'elle consiste à étudier les variations d'effectif et leur structuration au sein d'une population. On peut dans ce cadre représenter de façon simple l'évolution de l'effectif d'une population par :

$$N_{t+1} = R.N_t, \tag{3}$$

où N_t représente la taille de la population au temps t et R représente le taux de croissance net de la population. Ainsi la taille d'une population à la génération suivante ne dépend que de sa taille à la génération précédente et de son taux de croissance net qui inclut la mortalité et la reproduction. Ce taux de croissance net peut être assimilé à la valeur sélective absolue moyenne de la population, $\overline{\omega}$, qui correspond au nombre moyen de descendants reproducteurs produits par individu. Elle s'écrit :

$$\overline{\omega} = \overline{f} \times \overline{sv}, \tag{4}$$

avec $\overline{f}$ et $\overline{sv}$, les fécondités et survies moyennes des individus de la population. On peut donc en reprenant l'équation 3 faire l'analogie entre le taux de croissance net de la population (R) et la valeur sélective absolue moyenne de la population ($\overline{\omega}$) de telle sorte que :

$$N_{t+1} = \overline{\omega}.N_t, \tag{5}$$

toujours avec N_t la taille de la population à la génération t. Ainsi la biologie de la conservation consiste à décrire et à expliquer les variations d'effectifs ou de valeurs sélectives absolues moyennes (i.e. le taux net d'accroissement).

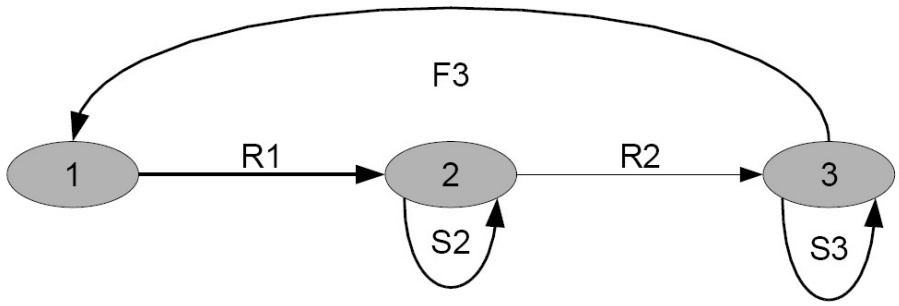


FIG. 3: Dynamique démographique simplifiée d'une population de sapin pectiné. La classe 1 représente les graines, la classe 2 représente les semis et la classe 3 les adultes reproducteurs. Les R représentent les taux de recrutement d'une classe à la suivante, et les S représentent quant à eux les taux de survie de chacune de ces classes. Le sapin ne présente pas de banque de graines donc S_1 est nul, en revanche on observe souvent une banque de semis représentée par S_2. Seul les adultes peuvent produire des graines avec un taux F_3.

Cependant le taux net d'accroissement de la population ne représente que de manière très synthétique le fonctionnement de la population. De même l'effectif total de la population ne représente pas la structure en âge de la population dont chaque classe diffère des autres par sa contribution à l'accroissement de la population. C'est pourquoi des modèles plus complexes de démographie formalisent la dynamique d'une population de façon matricielle. Soit la matrice $N(t)$ des effectifs de chaque classe (age, statut reproducteur...) au temps t telle que :

$$N(t) = \begin{bmatrix} N_1(t) \\ N_2(t) \\ N_3(t) \end{bmatrix}, \tag{6}$$

avec $N_x(t)$ l'effectif de la classe x au temps t.

Leslie (1945) définit une matrice de transition, L, qui intègre les probabilités de passages d'une classe à la suivante et aussi la fertilité de chacune de ces classes, cette matrice représente de manière plus détaillée la valeur sélective absolue moyenne de la population ($\overline{\omega}$). Soit L la matrice représentant la dynamique de la population

représentée dans la figure 3 :

$$L = \begin{bmatrix} 0 & 0 & F_3 \\ R_1 & S_2 & 0 \\ 0 & R_2 & S_3 \end{bmatrix}, \tag{7}$$

avec F_x et S_x les fécondités et les taux de survies de la classe x ; et R_x le taux de recrutement de la classe x vers la classe $x+1$. Ainsi cette matrice permet de calculer la valeur sélective absolue moyenne de la population puisqu'elle intègre à la fois la fécondité et la survie des différentes classes qui composent cette population.

On peut alors écrire la croissance de la population de manière matricielle :

$$\begin{bmatrix} N_1(t+1) \\ N_2(t+1) \\ N_3(t+1) \end{bmatrix} = \begin{bmatrix} 0 & 0 & F_3 \\ R_1 & S_2 & 0 \\ 0 & R_2 & S_3 \end{bmatrix} \times \begin{bmatrix} N_1(t) \\ N_2(t) \\ N_3(t) \end{bmatrix}. \tag{8}$$

On obtient alors les effectifs par classes d'ages à la génération $t+1$. Ce type de modèles matriciels est largement employé dans le domaine de la biologie de la conservation ou de la gestion des populations (Caswell 2001). Les propriétés mathématiques de ces matrices permettent d'estimer les paramètres clefs de l'évolution de la population considérée. En effet, il existe trois matrices, λ, V et W telles que :

$$LV = \lambda V \tag{9}$$

et

$$WL = W\lambda, \tag{10}$$

où λ est la plus grande valeur propre de la matrice L ; V est le vecteur propre à droite normalisé et W est le vecteur propre à gauche normalisé. λ représente la croissance asymptotique de la population, c'est-à-dire la croissance de la population quand la structure en classes est stabilisée (i.e. la population est à l'équilibre), alors :

$$N_{t+1} = \lambda N_t. \tag{11}$$

Le vecteur V représente la structure de la population à l'équilibre, avec les effectifs de chaque classe. Le vecteur W représente la contribution relative de chaque classe au taux d'accroissement de la population. Cette dernière information est très utile

aux gestionnaires puisqu'elle permet de savoir sur quelle classe de la population concentrer les efforts de conservation. Ce type de modèle est donc utilisé pour établir des plans de conservation pour des réintroduction par exemple. A l'inverse cela sert aussi à définir des quotas de chasse ou de pêche sans risquer de nuire à la survie de la population (Jensen 1974, un exemple appliqué aux pêcheries). Cependant afin de définir ces matrices et donc de prédire l'évolution des populations, il est nécessaire d'estimer *in natura* les valeurs des matrices de transition. Dans cette étude nous allons nous focaliser sur la production de graines viables ($F3$ dans la figure 3) dont nous allons caractériser la moyenne et la variabilité inter-individuelle. Le but ultime de cette étude sera d'intégrer ces valeurs ainsi que celles issues de différentes thèses sur le sapin pectiné actuellement en cours dans notre laboratoire au sein de modèles démographiques complets qui serviront pour les gestionnaires d'outils d'aide à la décision (en reprenant la notation de la figure 3 : Annabelle Amm paramètres $R1$ et $S2$, Aurore Bontemps paramètre $R2$ et Maxime Cailleret paramètre $S3$).

3.1.2 Phénomènes démographiques

Lande (1988) décrit deux principaux phénomènes démographiques pouvant affecter la survie des populations.

L'effet Allee selon lequel le taux de croissance est positivement corrélé avec la taille ou la densité de la population. Dans le cas des plantes on peut expliquer cela par une limitation pollinique, dont on sait qu'elle est concomitante avec une faible densité (Baker 1955, Allison 1990, Hirayama et al. 2007, voir également la partie précédente). Ainsi le taux d'accroissement de la population ne peut que diminuer dès lors que la taille de la population est passé sous seuil critique.

La stochasticité est dans le cadre de la démographie constituée de deux composantes :

- La stochasticité démographique : dans ce cas la valeur des paramètres de valeur sélective absolue de chaque individu de la population (survie, S_i et fertilité, F_i) est tirée aléatoirement. La variance inter-individuelle d'échantillonnage est

d'autant plus forte que la taille de la population est faible, et ceci a donc une influence sur la croissance globale de cette dernière. Ce phénomène pourrait être qualifié de "dérive démographique" par analogie à la dérive génétique. On peut représenter cela dans le cadre des matrices de Leslie en définissant deux matrices, une pour les moyennes des paramètres de valeurs sélective et une autre pour leurs variances (Schwartz et al. 2000), que l'on peut faire dépendre de la densité (Jensen 1974).

- La stochasticité environnementale : qui correspond aux fluctuations environnementales et modifie la valeur sélective absolue de tous les individus, mais de manière identique (Lande 1988). Dans ce cas l'impact de la stochasticité ne dépend pas de la taille de la population (Lande 1988).

Deux autres facteurs doivent également être considérés lorsque la démographie des individus est modélisée dans des modèles plus ou moins explicitement spatialisés, puisqu'ils vont conditionner l'efficacité de la migration :

- Les effets de bords qui modifient l'installation des individus situés à la limite de milieux favorables et défavorables. En effet, la coupure d'un milieu va entraîner un changement des conditions environnementales, qui seront bénéfiques ou au contraire désavantageuses, ce qui rendra la migration plus ou moins efficace. Par exemple la coupe d'une forêt rendra la lisière de la forêt attractive pour les espèces colonisatrices alors qu'elle sera répulsive pour les espèces d'ombres comme le sapin pectiné (Fagan et al. 1999, pour une revue assez complète concernant les effets de bord).
- Les probabilités d'extinction et de colonisation locales au sein d'une métapopulation peuvent influencer la dynamique globale de la méta-population. Nous n'aborderons que très peu ce dernier type de facteur qui ne correspond pas vraiment à notre cas particulier, mais le lecteur pourra en savoir plus sur ce type de modèles en lisant les articles de Hanski et Ovaskainen (2003), Ovaskainen et Hanski (2004).

La démographie implique donc une description détaillée des populations et surtout de leur milieu et de leur fonctionnement. Ainsi dans un premier temps il est nécessaire de décrire une structure en classe pertinente pour décrire la survie des populations considérées en réponse aux facteurs étudiés (structure en âge, en taille...).

Il faut également estimer la variabilité des paramètres de survie et de fertilité des individus sur de larges zones afin de tenir compte des différentes stochasticités dans les prédictions, qui seront alors plus précices. Dans cette thèse nous avons donc étudié la production de graines de nombreux individus afin d'intégrer la stochasticité démographique. De plus, le suivi a été effectué sur une large zone, contrasté en terme de densité, et pendant plusieurs années afin d'apréhender la stochasticité environnementale.

3.2 Approche génétique

Indépendamment des paramètres démographiques certains facteurs génétiques sont d'un intérêt particulier, puisqu'ils peuvent modifier les valeurs sélectives absolues des individus.

La dépression de consanguinité est le facteur génétique le plus étudié dans le domaine de la biologie de la conservation. En effet la taille finie des populations implique une augmentation de la consanguinité de la population, par simple dérive. Cette consanguinité peut se traduire par une baisse de vigueur des individus et donc entraîner une baisse du taux d'accroissement de la population (Williams et Savolainen 1996, voir également le chapitre précédent).

La diversité génétique permet aux populations de conserver leurs capacités évolutives. Au sein d'un environnement donné, les traits phénotypiques héritables les plus adaptés sont sélectionnés. Cependant lorsque l'environnement est hétérogène (dans le temps ou l'espace) il faut garder cette capacité d'adaptation. Or, cette évolution des individus vers les caractéristiques optimales ne peut se faire qu'à la condition que cette population renferme une diversité génétique suffisante pour exprimer une grande variété de phénotypes (Shaberg et al. 2008, Reed et Frankham 2003).

La diversité génétique peut être représentée par l'effectif efficace de la popula-

tion, Ne, qui représente la taille d'une population idéale (selon Wright-Fisher) pour laquelle la dérive génétique permettrait d'obtenir la même variation de polymorphisme que celle observée. De plus, la dérive génétique va induire une augmentation de la consanguinité et une baisse de l'hétérozygotie, d'autant plus rapidement que la taille de la population est faible (Vergeer et al. 2003). L'effet des facteurs génétiques étant devenu l'un des points de préoccupation majeure des acteurs de la conservation biologique, ils emploient régulièrement le rapport $\frac{Ne}{N}$ (avec N la taille *stricto sensu* de la population) comme indicateur de viabilité des populations (Nunney et Campbell 1993). Pourtant une faible diversité génétique ou un faible effectif ne signifie pas obligatoirement que la population est menacée (Johnson et al. 2009).

3.3 La démo-génétique et la biologie de la conservation

Les facteurs génétiques et démographiques peuvent interagir et ainsi s'auto-entretenir. Ce constat a été mis en évidence dans le cas des populations en déclin par Gilpin et Soulé (1986), qui définissent le concept de vortex d'extinction pour la première fois. Ce concept implique qu'une population dont l'effectif est passé sous un certain seuil ne peut que se diriger vers l'extinction. En effet, la petite taille de la population induit une augmentation de la consanguinité et une diminution de la diversité génétique, ces deux facteurs diminuant les capacités de survie, de reproduction et d'adaptabilité des individus. Ainsi la taille de la population diminue, ce qui la rend plus sensible à la stochasticité démographique et environnementale, le cycle continue alors jusqu'à l'extinction complète de la population (figure 3.3).

Cependant Lande (1988) remet en question l'importance des facteurs génétiques dans l'extinction d'une population. Il avance l'idée que l'influence de la diversité génétique ou même l'augmentation de la consanguinité et les conséquences qu'on lui connaît, ne peuvent agir sur les échelles de temps aussi courtes que celles de l'extinction des populations menacées. Il émet donc l'hypothèse que seuls les paramètres démographiques sont d'intérêt pour la survie de ces populations, puisqu'ils agissent à des échelles de temps compatibles avec le phénomène d'extinction. Cette affirma-

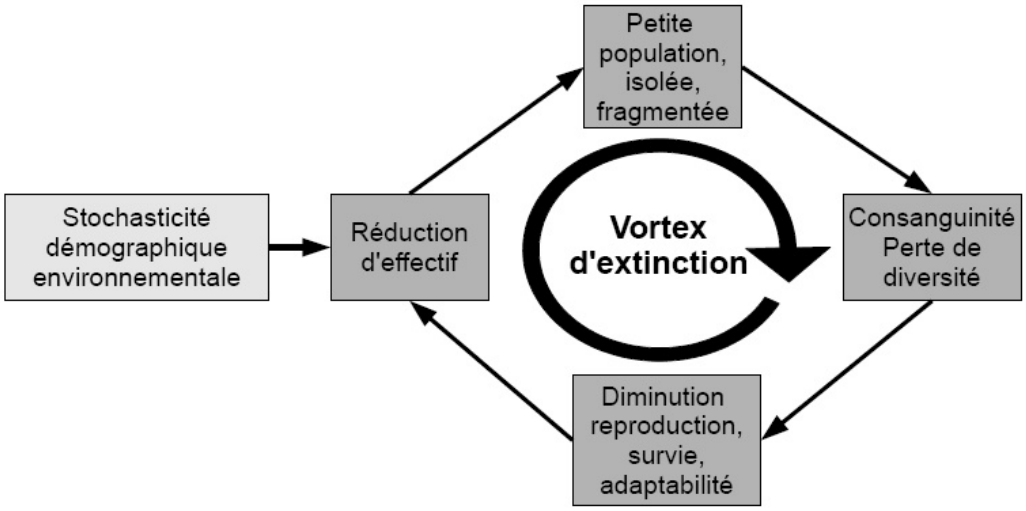


FIG. 4: Représentation schématique du vortex d'extinction défini par Gilpin et Soulé (1986). Une petite population dans ce cas ne peut que diminuer par le biais de facteurs génétiques associés à une plus grande sensibilité à la stochasticité démographique et environnementale.

tion est remise en question par de nombreuses études notamment celle de Spielman et al. (2004), qui montre une nette réduction de l'hétérozygotie chez les taxons menacées en comparaison à celle des taxons similaires non menacées (35% plus faible en moyenne). Ils en déduisent que cette différence notable d'hétérozygotie reflète de manière certaine que les facteurs génétiques ont le temps de s'exprimer lors du processus d'extinction. De plus la dépression de consanguinité semble nettement augmenter les probabilités d'extinction des populations (Byers et Waller 1999, Brook et al. 2002, Reed et al. 2003). Il est donc nécessaire d'intégrer les facteurs génétiques et démographiques dans une approche commune, en considérant toutes les interactions possibles entre les différents niveaux d'action dans un cadre théorique bien défini (With 1997). Ce cadre n'a été apporté que récemment par Coulson et al. (2006), qui identifie clairement le besoin de "cartographier" les liens qu'il existe entre les différents compartiments de la dynamique des populations et d'identifier à quel niveau d'organisation ils opèrent (figure 3.3).

La biologie de la conservation est donc une science pluridisciplinaire intégrant de nombreux domaines comme l'écologie, la physiologie, la génétique des populations,

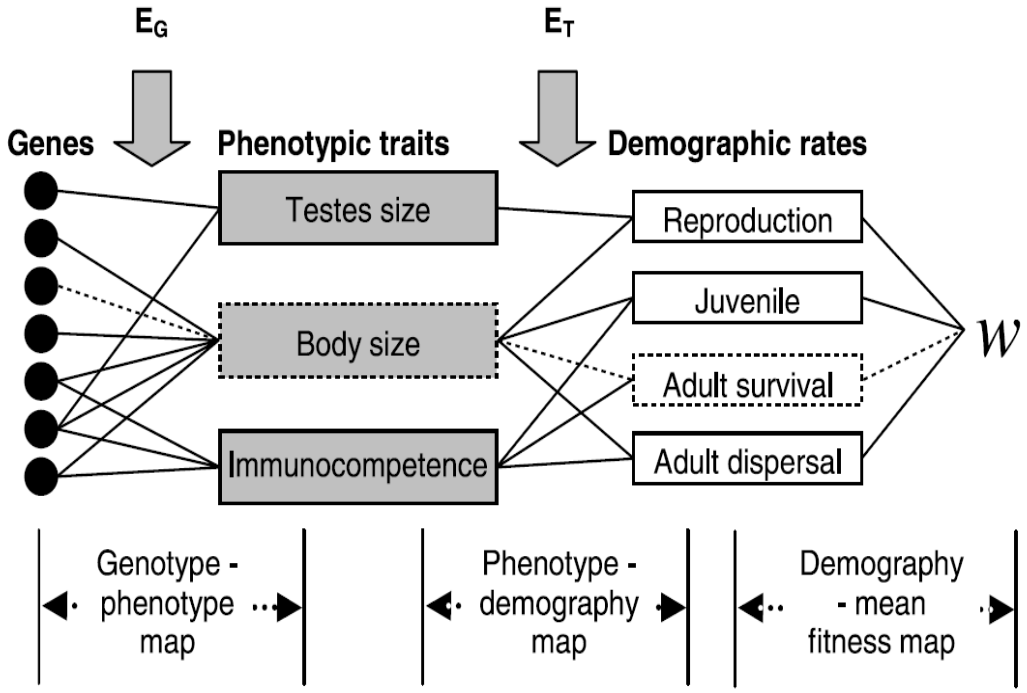


FIG. 5: Représentation schématique de différents compartiments de la démographie et comment ils peuvent interagir avec les facteurs génétiques. Eg et Et représentent respectivement l'effet de l'environnement sur l'expression des gènes et l'expression de ces traits en paramètres démographiques. Cette figure est extraite de Coulson et al. (2006).

la génomique ou la démographie mais dont le contour est encore assez mal défini (Soulé 1985, With 1997). Ce domaine de la science ne doit plus être une "science de l'urgence", ne s'intéressant aux causes d'extinction uniquement lorsque les populations sont déjà en déclin. En effet, les causes réelles d'extinction peuvent être très différentes de celles observées et il est alors difficile de savoir lequel des facteurs a engendré l'autre. Cela a pu conduire Lande (1988) à mettre en avant les facteurs démographiques ou Spielman et al. (2004) à insister sur l'effet des facteurs génétiques. Les mécanismes mis en jeu peuvent également différer selon qu'ils opèrent dans des populations de grande ou petite taille comme cela a été montré pour l'effet de la dépression de consanguinité par exemple (Brook et al. 2002, Reed et al. 2003). Il est donc important d'étudier des populations stables démographiquement pour comprendre leur dynamique et ne pas se focaliser uniquement sur les populations menacées. Par exemple les stochasticités démographiques ou environnementales ne peuvent être correctement estimées que sur un grand nombre d'individus ou de générations, ce qui rend leur estimation difficile dans le cas des petites populations en voie d'extinction (Lande et al. 2002; 2006, un exemple pour l'estimation de la densité dépendance). La calibration des modèles complexes de biologie de la conservation doit donc se faire sur des populations non menacées à court terme afin de comprendre les relations qui existent entre les différents compartiments de sa dynamique (génétique vs. démographique). De plus, les espèces pérennes et les arbres en particulier présentent des caractéristiques particulières renforçant l'importance de la variabilité spatio-temporelle lors de cette calibration.

4 Les arbres : un modèle biologique particulier

4.1 Extinction et recolonisation : Le paradoxe de la diversité génétique

Les forêts constituent un écosystème clef dont la biomasse représente 90% de la biomasse totale et qui couvrent près de 27% de la surface des terres émergées (Petit et Hampe 2006). Les arbres ont pourtant connu de fortes contractions de leurs aires de répartition lors de la dernière période glaciaire (Papageorgiou et al. 2008, Tollefsrud et al. 2008, Liepelt et al. 2009). Leur aire actuelle de répartition, la taille supposée des refuges ainsi que les données paléo-palynologiques, indiquent des vitesses de colonisation de l'ordre de 100 à 1000m par an, ce qui dépasse largement les capacités supposées de dispersion des graines (McLachlan et al. 2005, Sagnard et al. 2007). Ainsi seuls des événements rares de dispersion de graines à longue distance peuvent expliquer une telle vitesse d'expansion des populations (Petit et al. 1997, Clark et al. 2001; 2003). Cependant une colonisation de ce type implique deux phénomènes :

- Une production de graines suffisamment importante (ou un grand nombre d'épisodes de reproduction) pour que les événements rares de dispersion à longue distance se produisent.
- Des effets de fondation forts qui devraient engendrer une baisse de la diversité génétique.

Or nous observons une diversité génétique extrêmement élevée dans la majorité des populations d'arbres (Petit et Hampe 2006). De plus la production de graines est très variable entre les années et les individus (Herrera et al. 1998, Kelly et Sork 2002, voir figure 6). Pourtant la dispersion à longue distance semble être le seul scénario crédible sauf dans certains cas où l'on a pu démontrer la présence de refuges cryptiques (McLachlan et al. 2005).

Ainsi deux questions se posent :

- Dans quelle mesure varie la production de graines ?
- Comment est conservée la diversité génétique ?

Advantage	Drawback
Great potential of biomass gain	High maintenance costs
High competition after successful establishment	Extremely high recruit mortality
Endurance to short-term resource depletion	Increased probability of suffering catastrophic events
Escape from disturbances at ground level (e.g., grazing, fire)	Exposure to physical disturbances above ground (in particular wind)
Life-long increase in storage capacity and fecundity	High investment in supporting tissues and defense mechanisms reduces overall allocation to reproduction
Great lifetime fecundity	Delayed maturity
Little dependence on particular reproductive events	Trade-off between present reproductive output and future growth, survival, and reproduction
Attraction of mutualists (e.g., pollinators, seed dispersers, herbivore predators)	Attraction of antagonists (e.g., herbivores, pathogens)
Satiation of enemies (e.g., mast-fruiting)	Satiation of mutualists (e.g., geitonogamy, disperser satiation)
Large pollen and seed production and release height facilitate gene dispersal	Low plant density complicates mating and increases pollen limitation
Relatively little seed limitation of recruitment	Strong limitation of suitable sites and time windows for recruitment
Effective population size close to adult population size	Large differences in life spans exacerbate inequality in individual lifetime fecundity
Local adaptation favored by strong selection during early life stages	Local adaptation hindered by high gene flow
Reduced accumulation of mutations per unit of time	Increased mutation rate per generation
Strong inbreeding depression increases outcrossing rate and maintains genetic diversity	Lifelong accumulation of somatic mutations results in susceptibility to inbreeding depression
Long life span reduces extinction risk	Long generation time reduces speciation rate
Extensive intra- and interspecific gene flow reduces extinction risk	Extensive intra- and interspecific gene flow reduces speciation rate
Slower evolution than mutualists results in greater share of resources for host trees (Red King effect)	Slower evolution than antagonists results in host trees lagging behind in arms races (Red Queen effect)

FIG. 6: Tableau récapitulant les principaux avantages et inconvénients des caractéristiques des arbres. Tiré de Petit et Hampe (2006).

Les réponses à ces questions vont nous permettre de comprendre l'histoire passée des populations d'arbres mais vont également pouvoir nous informer sur les capacités de migration des populations actuelles face au changement climatique (IPCC 2007, Morin et al. 2008).

4.2 A quelle échelle intégrer la production de graines ?

Les arbres participent à de nombreux épisodes de reproduction au cours de leur vie. Il est donc difficile d'intégrer la production individuelle de graines sur l'ensemble de la vie d'un arbre afin d'obtenir une estimation correcte du succès reproducteur femelle individuel et de sa variabilité. En effet, la production de graines n'est pas constante entre les individus et entre les différents épisodes de reproduction (Krouchi et al. 2004). Plusieurs raisons peuvent être invoquées pour expliquer ces variations : l'âge, les conditions environnementales (climat, nutriments, quantité de pollen efficace, compétition) et la dépression de consanguinité (Agren 1996, Sorensen et Webber 1997, Herrera et al. 1998, Krouchi et al. 2004, de Jong et al. 2005). Un phénomène répandu chez les arbres est la variabilité temporelle de la production de graines connu sous le nom de *masting* (Kelly et Sork 2002). Il consiste en une production intermittente et synchronisée de graines résultant en une alternance d'années de fortes et de faibles productions à l'échelle du peuplement ou de la population. Il existe trois principales explications à ce phénomène :

- Les fluctuations temporelles, avec une alternance de bonnes ou mauvaises conditions, qui seraient donc bénéfiques pour la croissance et la reproduction. Cependant l'observation de nombreuses corrélations négatives entre l'allocation à la croissance et à la reproduction ont réfuté cette hypothèse (Kelly et Sork 2002).
- Les modèles d'allocation des ressources. Dans ce cas, les individus accumulent des ressources les bonnes et les mauvaises années qu'ils utilisent pour croître mais qu'ils investissent aussi dans la reproduction si la quantité de ressources accumulées dépasse un certain seuil. Ainsi les mauvaises années le seuil n'est pas atteint mais le surplus de ressources est conservé pour l'année suivante

qui donc nécessitera moins de ressources pour que la reproduction ait lieu, conduisant à des productions de graines intermittentes (Isagi et al. 1997). Ce modèle explique l'alternance mais ne peut expliquer la synchronisation entre les individus d'une population présentant une hétérogénéité fine et donc des conditions différentes entre individus.

– Les modèles d'allocation des ressources couplés à la disponibilité en pollen. Ce modèle couple le modèle précédent pour expliquer l'alternance de bonnes et de mauvaises années de reproduction, mais considère que la synchronie entre les individus s'effectue par le biais de la production de pollen, qui est nécessaire à la formation des graines (Satake et Iwasa 2002, Satake 2004).

Toutes les espèces ne présentent pas une production de graines suivant un phénomène de *masting*, même si cette production peut varier dans le temps et entre les individus (Herrera et al. 1998). En effet deux phénomènes peuvent influencer le succès reproducteur femelle, la dépression de consanguinité (Williams et Savolainen 1996) et la limitation pollinique (Ashman et al. 2004). Ces deux facteurs peuvent varier dans le temps et dans l'espace et même interagir (Byers et Waller 1999, Knight et al. 2005, Aizen et Harder 2007). Il est donc primordial d'évaluer la contribution de ces facteurs et de comprendre à quelle échelle ils agissent afin d'estimer de manière précise la production de graines et sa variabilité dans le temps et l'espace pour prédire le devenir des populations, notamment face au changement climatique.

4.3 Les flux de pollen : Un déterminant essentiel ?

La capacité des arbres à recoloniser rapidement un environnement ne dépend que d'une production de graines en quantité suffisante (Clark et al. 2001; 2003) et de leur capacité à s'y adapter en conservant une grande diversité génétique (Petit et Hampe 2006). La production de graines dépend des conditions environnementales (Isagi et al. 1997), de la disponibilité en pollen (Sorensen et Webber 1997, Satake et Iwasa 2002) et de la dépression de consanguinité (Williams et Savolainen 1996, Wang 2003), or ces deux derniers points dépendent de la dispersion du pollen. De plus, les flux polliniques extensifs observés chez les arbres ainsi que leur phase

juvénile prolongée (i.e. cumul de plusieurs épisodes de reproduction) peuvent également expliquer la forte diversité génétique que l'on observe au sein des populations d'arbres malgré les probables effets de fondation (Austerlitz et al. 2000, Petit et Hampe 2006). Les flux de pollen semblent donc être un déterminant essentiel de la production de graines et du maintien de la diversité génétique, qui sont les bases de l'importante capacité des arbres à recoloniser efficacement de nouveaux milieux.

Chez les arbres la dispersion du pollen s'effectue à grande distance en comparaison avec d'autres végétaux (Adams et Birkes 1991). Ces grandes capacités de dispersion peuvent s'expliquer par la hauteur à laquelle le pollen est libéré (Okubo et Levin 1989, voir aussi la figure 6). De plus, les conifères montrent une adaptation marquée à l'anémophilie avec la présence de deux ballonnets permettant une baisse de densité des grains de pollen et donc une dispersion importante (Schwendemann et al. 2007). Cependant la dispersion du pollen peut varier au sein des populations et même si elle permet de conserver une grande diversité au sein d'une aire de répartition, elle structure la diversité génétique au niveau local (Savolainen et al. 2007). Si l'on considère que le régime de reproduction suit une loi d'action de masse, cela va également modifier le régime de reproduction (Robledo-Arnuncio et al. 2004*a*). Plusieurs facteurs peuvent influencer la structuration génétique spatiale du nuage pollinique :

- La densité en conspécifiques qui va augmenter le nombre de contributeurs au nuage de pollen (Allison 1990, Morgante et al. 1991, Smouse et Sork 2004, Fukue et al. 2007, Naito et al. 2008).
- La densité globale qui peut augmenter la rugosité du milieu et donc "freiner" le déplacement du pollen, par exemple dans les milieux où la canopée est fermée (Smouse et Sork 2004).
- L'isolement des individus qui permet un mélange relativement homogène des contributions des individus situés à longue distance (Robledo-Arnuncio et al. 2004*a*, Klein et al. 2006, hypothèse de mélange des propagules à longue distance).
- La configuration spatiale notamment l'agrégation dont les effets peuvent être variables en fonction de l'échelle à laquelle se fait la dispersion et celle à laquelle

se fait l'agrégation (Fagan et al. 1999, Meagher et Vassiliadis 2003, Robledo-Arnuncio et Austerlitz 2006).

– La densité en pollinisateurs chez les espèces entomophiles qui va affecter indirectement la densité en conspécifiques et qui dépend du comportement des animaux (Franceshinelli et Bawa 2000).

De plus ces différents facteurs peuvent évoluer dans le temps et donc potentiellement modifier la structure génétique spatiale du nuage pollinique (Irwin et al. 2003, Kenta et al. 2004, Smouse et Sork 2004). Ainsi pour les arbres plus que pour toute autre espèce végétale la diversité génétique s'organise dans le temps et l'espace et ces deux sources de variabilité doivent être intégrées dans les analyses afin de comprendre la dynamique à long terme des peuplements forestiers.

5 Objectif de la thèse

Dans cette thèse nous avons caractérisé le succès reproducteur chez le sapin pectiné, *Abies alba*, dans une population en expansion mais qui subit déjà l'impact du réchauffement climatique (thèse de Maxime Cailleret en cours à l'URFM). En effet, la reproduction détermine le taux d'accroissement des populations et peut être influencée par de nombreux facteurs. La dispersion du pollen notamment, peut modifier le nombre d'ovules fécondés par limitation pollinique et donc la production de graines (Sorensen et Webber 1997, de Jong et al. 2005, Knight et al. 2005). Une dispersion limitée du pollen peut également réduire la diversité génétique localement, en diminuant le nombre de contributeurs au nuage pollinique fécondant (Smouse et Sork 2004). Cela peut alors modifier les capacités évolutives des populations confrontées à des environnements changeants (Mimura et Aitken 2007*b*, Bittencourt et Sebbenn 2008, de Lucas et al. 2008). Lorsqu'ils sont limités, les flux de pollen peuvent augmenter la consanguinité en induisant une structuration génétique spatiale (Robledo-Arnuncio et al. 2004*a*, Hirayama et al. 2005). Ils peuvent également induire de la mal-adaptation quand la dispersion s'effectue à des échelles supérieures à celle des limites écologiques du milieu (Ellstrand 1992*b*, Lazaro et Traveset 2006, Lopez et al. 2008). De plus les flux de pollen vont modifier le régime de reproduction dès lors que ce dernier est déterminé par une loi d'action de masse (Burczyk et al. 1996, Robledo-Arnuncio et al. 2004*a*, Geng et al. 2008, voir aussi partie précédente) ; ceci peut à son tour accentuer la consanguinité des populations et les exposer ainsi à la dépression de consanguinité. Nous avons donc abordé la reproduction du sapin en intégrant les succès reproducteurs mâles et femelles en relation avec l'environnement (ex : densité), les flux des gènes et le régime de reproduction (figure 1) ainsi que leurs variabilités dans le temps et l'espace (stochasticité), c'est ainsi que l'on pourra tirer des conclusions solides quand à la conservation des populations.

Chapitre 1

Matériel biologique, sites d'études et dispositifs expérimentaux

1 Le sapin pectiné, *Abies alba* Miller

Le sapin pectiné, *Abies alba* Miller, est un gymnosperme appartenant à l'ordre des coniférales et à la famille des Pinaceae (d'après la classification phylogénétique donnée par le National Center for Biotechnology Information voir aussi Liu (1971)). En Europe cet arbre est le plus grand des conifères et l'un des plus grands arbres, pouvant atteindre jusqu'à 60-65m de hauteur (figure 1.1). C'est un arbre à très longue durée de vie, jusqu'à 500 à 600 ans. Il atteint sa maturité sexuelle au bout de 20 à 25 ans lorsque les conditions sont favorables, mais il peut rester à l'état de semis jusqu'à 60 à 70 ans lorsque les conditions environnementales lui sont défavorables, notamment lorsque la compétition est forte en peuplement dense (Feurdean et Willis 2008). Les forêts de sapin pectiné présentent donc des banques de semis, dont les âges peuvent être très hétérogènes (figure 1.2). Cette particularité combinée à leur grande durée de vie et à leur long temps de génération produit des forêts à générations chevauchantes. Comme tous les représentants de sa famille, le sapin pectiné est monoïque, c'est à dire que les strobili (i.e. inflorescences mâles) et les cônes (i.e. inflorescences femelles) sont présents sur les mêmes individus (figure 1.3). Comme la plupart des conifères, le sapin pectiné ne présente pas de mécanisme connu d'auto-incompatibilité (Kormutak et Lindgren 1996) même si de récentes recherches

FIG. 1.1: Sapin pectiné sur le Mont Ventoux. *Restoux* **2006**

FIG. 1.2: Semis de sapin pectiné sur le Mont Ventoux poussant sous couvert dans une hétraie-sapinière. Cet individu est âgé de plus de 20 ans. *Restoux* **2006**

FIG. 1.3: Organes reproducteurs males et femelle du sapin pectiné. A gauche les strobili, ici après avoir émis leur pollen. A droite une fleur femelle au début de sa période de réceptivité, qui se transformera en cône par la suite. *Restoux* **2006-2007**

FIG. 1.4: Photographie microscopique d'un grain de pollen d' *Abies alba.* On y distingue clairement les deux ballonnets. *J. Benes - http ://www.lape.bf.jcu.cz/img/pyl-abies.jpg*

FIG. 1.5: Graines de sapin pectiné, *Abies alba* Miller. *Source inconnue*

suggèrent qu'ils pourraient exister dans certains cas mais ces résultats doivent être confirmés (Owens et al. 1998). Cependant les organes mâles et femelles ne sont pas présents dans la même zone du houppier : les premiers sont situés dans la partie basale et les derniers dans la partie apicale. Cette architecture particulière pourrait constituer une adaptation permettant l'évitement de l'autofécondation. Cette espèce est anémophile (i.e. la dispersion du pollen se fait par le vent). Son pollen est caractéristique des genres *Abies*, *Pinus* et *Picea* avec deux ballonnets diminuant sa densité et améliorant ainsi ses capacités de dispersion (Adams et Birkes 1991, figure 1.4). Même si le sapin pectiné reste majoritairement allogame, il présente donc un régime de reproduction mixte avec un taux moyen d'allofécondation de l'ordre de 80 % en peuplement mature (Fady et al. 1999, Feurdean et Willis 2008). Cependant l'autofécondation peut être conséquente avec des valeurs moyennes au niveau de la population de l'ordre de 95% dès lors que la densité en individus reproducteurs est faible ou que la production de pollen est faible (Feurdean et Willis 2008).

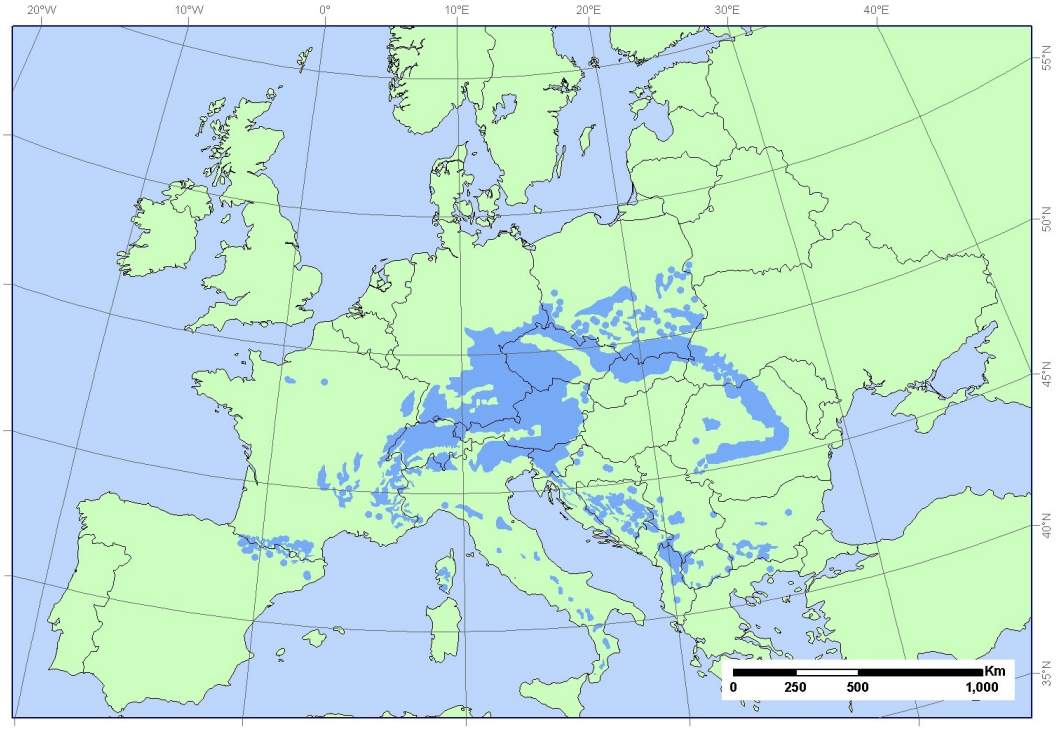


FIG. 1.6: Carte de répartition du sapin pectiné (*Abies alba*) en Europe. *Source réseau EUFORGEN - http ://www.bioversityinternational.org/networks/euforgen/Euf_Distribution_Maps.asp* **2008**

Les graines présentent des ailes conférant au sapin une dispersion anémochore (figure 1.5). Le vol des graines est caractérisé par un mouvement de tourbillon nommé *spining behavior* qui confère aux propagules des capacités de dispersion accrues (Minami et Azuma 2003, voir aussi la thèse de Julien Fayard au sein du laboratoire portant sur la modélisation physique de la dispersion des graines). Cependant la distance médiane de dispersion des graines de sapin pectiné est relativement faible entre 13 et 20 m (Sagnard et al. 2007).

Le sapin pectiné est une espèce présente à l'état naturel uniquement en Europe. On le trouve dans la plupart des massifs montagneux européens, généralement à une altitude comprise entre 400m et 1800m, c'est à dire de l'étage montagnard à l'étage subalpin inférieur (figure 1.6). C'est une espèce intolérante à la sécheresse et sciaphile au moins pendant les premiers stades de son développement, on le trouve donc principalement sur les ubacs (Konnert et Bergmann 1995, Carcaillet et Muller 2005). Le sapin pectiné est une espèce de fin de succession qui se développe sous le

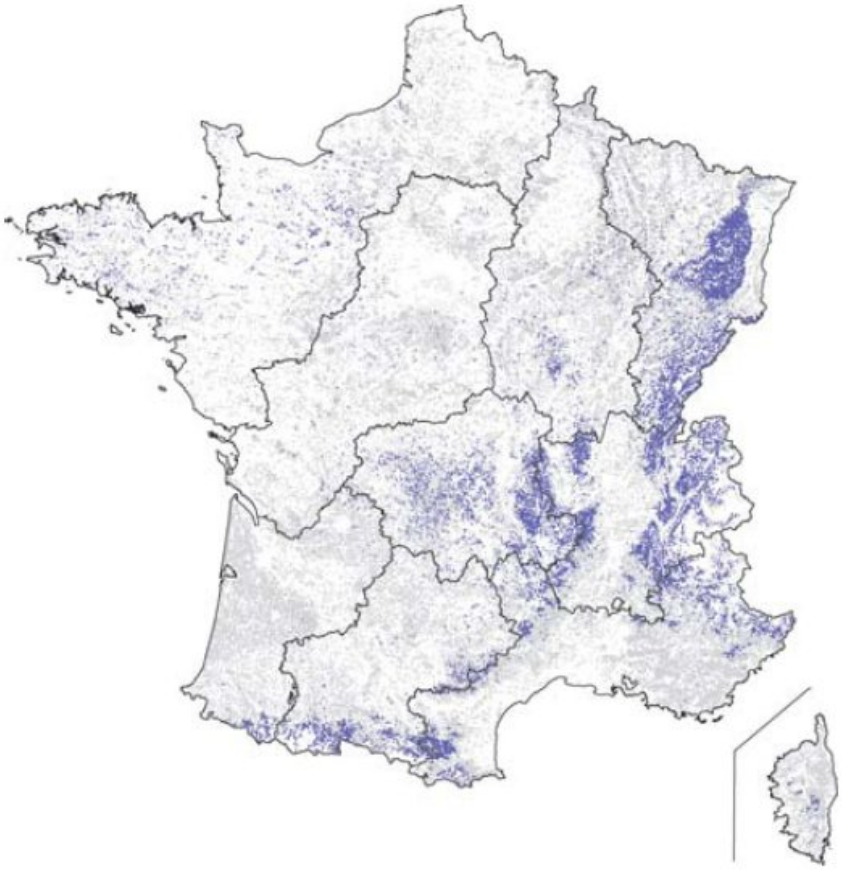


FIG. 1.7: Carte de répartition du sapin pectiné (*Abies alba* Miller) en France (en bleu). *Source Inventaire Forestier National* **2008**

couvert d'autres essences notamment les pins. Il est souvent associé au hêtre commun (*Fagus sylvatica*), avec qui il partage une partie de sa niche écologique ; on le trouve également en peuplements purs. *A. alba* a connu une forte contraction de son aire de répartition lors du dernier maximum glaciaire et n'était présent que dans 5 sites refuges. Cependant l'analyse de la diversité génétique des populations Européennes montre que la répartition actuelle du sapin pectiné serait le résultat d'une expansion à partir de 3 sites refuges (Konnert et Bergmann 1995, Liepelt et al. 2009) :

- le nord des Apennins,
- le sud Balkans, l'est des Carpates et les Alpes Dinariques,
- le nord ou l'ouest des Balkans.

Il occupe actuellement une surface de 580 000 hectares en France (figure 1.7). De

plus son aire de répartition s'élargit depuis environ 150 ans notamment dans les Alpes du Sud. Ceci serait principalement le résultat d'un changement d'usage des terres (Carcaillet et Muller 2005). Il représente également un fort intérêt économique avec une production annuelle de 7,1 millions de m^3 de bois soit environ 7% de la production annuelle française totale (Inventaire Forestier National 2008). De plus dans la région PACA, il représente un fort intérêt patrimonial et une forte valeur de biodiversité (Cheikh Al Bassatneh et al. 2007).

2 Les sites d'études

Cette étude s'est déroulée sur deux sites situés en limite sud de l'aire de répartition française dans les Alpes du sapin pectiné, le Mont Ventoux et la montagne de Lure. Ecologiquement en marge, elles sont géographiquement isolées du "coeur" de la population centrale (figure 1.6) et présentent de faibles effectifs, conséquence courante de la marginalité (Kawecki 2008, pour une définition de la marginalité). Ces deux massifs possédent des caractéristiques écologiques très proches, mais diffèrent tout de même sur quelques points notamment leur histoire, le climat et la nature des peuplements forestiers qui en résultent.

2.1 Le Mont Ventoux

Il est situé à la limite sud-occidentale des Préalpes méridionales et s'étale sur 24 km d'Ouest en Est et sur 15 km du Nord au Sud. Son sommet culmine à 1912 m et est localisé à 44°10′ de latitude Nord et 5°16′ de longitude Est. Le climat y est en grande majorité typique de l'étage montagnard médio-européen, avec une pluviométrie abondante entre septembre et janvier. La pluviométrie moyenne annuelle y est de 1457 mm et la température moyenne annuelle de $5,5$°C (Sagnard 2001). L'influence méditerranéenne est plus prononcée sur le versant sud, alors que le versant nord subit (notamment au-dessus de 1000 m) les influences du climat alpin.

Le Mont Ventoux est issu de la tectogénèse pyrénéo-provençale, expliquant son

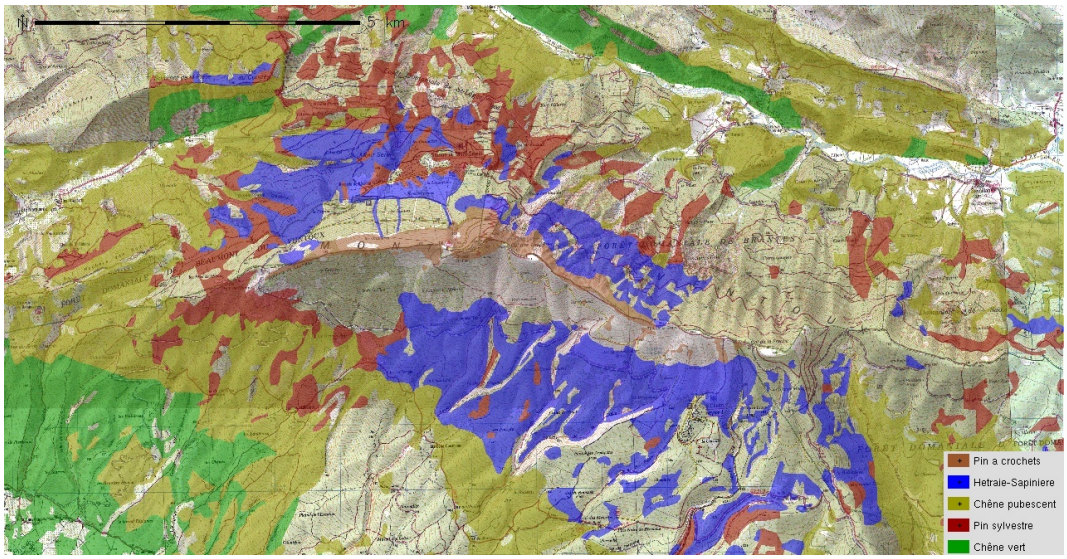


FIG. 1.8: Carte de végétation du Mont Ventoux d'après Guende (1978) et l'Institut National Géographique. *Restoux* **2008**

orientation est-ouest. La tectogénèse alpine a ensuite modifié la forme du massif en lui conférant une orientation secondaire selon un axe nord-sud. Les pentes du versant sud sont modérées alors que sur le versant nord les pentes très fortes sont fréquentes (Sagnard 2001). On observe une roche mère principalement calcaire, cependant l'érosion et l'activité biologique sur les parties recouvertes par de la forêt ont permis la formation de sols bruns forestiers. La qualité des sols et les ressources, notamment hydriques, dont ils disposent peuvent être très variables à l'échelle du massif (cf. thèse au sein du laboratoire de Maxime Cailleret, traitant de l'écophysiologie des sapins et des hêtres dans le cadre du dépérissement).

Au XIXeme siècle, le Mont Ventoux était en très grande partie déboisé, avec une absence quasi-totale de forêt en dessous de 1100m d'altitude. Ce défrichement était dû à des activités humaines intenses, nécessitant notamment du bois de chauffage, mais aussi à une absence quasi-totale de régénération du fait d'un pastoralisme intensif (Guende 1978). En effet les jeunes semis sont particulièrement appétants pour un grand nombre de mammifères. Afin de contrecarrer l'érosion qui découlait de ces pratiques, une politique de reboisement a été initiée par décret le 22 mars 1861, afin de restaurer les terrains de montagne (politique RTM). Cette politique dura

jusqu'au début de la première guerre mondiale. Elle a consisté à modifier l'usage des terres et notamment à limiter le pastoralisme. Dans une deuxième phase, des reboisements ont été effectués à partir d'essences indigènes (chênes, pin sylvestre, hêtre) ou non (pin noir, épicéa, mélèze...).

De nombreuses études tentent de comprendre quelles sont les politiques de restauration les plus efficaces, notamment dans le domaine forestier. Mais elles manquent de recul pour pouvoir évaluer correctement ces politiques (observations personnelles lors du colloque Seed Ecology II, Perth, Australia, 2007). Nous disposons avec le Mont Ventoux d'un outil particulièrement adapté pour évaluer sur le long terme l'efficacité des méthodes de gestion. En effet, cette recolonisation est un véritable succès car en près de 150 ans le Mont Ventoux a complètement changé de faciès. Il est devenu une des forêts naturelles les plus importantes du Vaucluse (Sagnard 2001, Carcaillet et Muller 2005, Blondel 2007). Un des exemples les plus frappants de cette restauration est la régénération naturelle des sapinières dont il ne restait que quelques peuplements et arbres "reliques". Celle-ci s'est opérée sous le couvert des autres essences plantées et notamment du pin à crochets (*Pinus uncinata*), qui désormais disparaît progressivement au niveau de la zone naturelle de la hêtraie-sapinière (Sagnard 2001). Le Mont Ventoux présente maintenant une forêt diversifiée, mélangeant espèces indigènes et introduites (figure 1.8) et qui fait également preuve d'une biodiversité remarquable (Blondel 2007).

2.2 La montagne de Lure

La montagne de Lure est située dans les Alpes de Haute Provence, avec un sommet situé à 44°7′ de latitude Nord et 5°58′ de longitude Est et culminant à 1826m d'altitude. Elle résulte du soulèvement à l'ère tertiaire du bassin sédimentaire de Forcalquier. Elle est donc principalement constituée de roches calcaires. Elle s'étale sur 7 km du Nord au Sud et 20 km d'Est en Ouest. Le climat est largement sous influence médio-européenne et méditerranéenne à basse altitude. Comme le Mont Ventoux, la montagne de Lure possède un passé riche en activités humaines, avec notamment la présence de charbonniers, qui utilisaient massivement le bois, jusqu'à

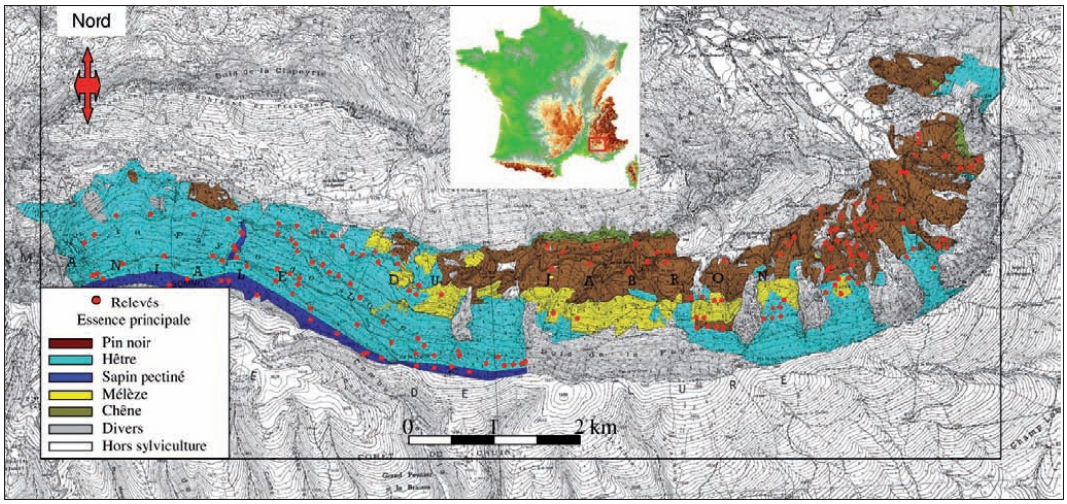


FIG. 1.9: Carte de végétation de la partie nord de la montagne de Lure extraite de Cheikh Al Bassatneh et al. (2007).

la moitié du XXeme siècle, et un pastoralisme relativement important (Simon et al. 2007). Dès la moitié du XIXeme siècle le massif a également fait l'objet de la politique RTM, mais principalement sur la partie basse et notamment pour éviter l'érosion des sols qui devenait critique (Cheikh Al Bassatneh et al. 2007). On y trouve donc une diversité d'essences forestières très proche de celle observée sur le Mont Ventoux, avec notamment des pins noirs, des hêtres, des sapins, des mélèzes et des chênes (figure 1.9). Cependant sur la montagne de Lure, la zone à sapin pectiné située principalement dans la partie haute et relativement inaccesible du massif n'a pas subi de gestion intensive, et s'est donc développée de manière quasi-naturelle (Cheikh Al Bassatneh et al. 2007, Simon et al. 2007).

3 Les dispositifs expérimentaux

3.1 Localisation et échantillonnage

Nous avons échantillonné 146 sapins pectinés, 95 sur le Mont Ventoux et 51 sur la montagne de Lure (figures 1.10 et 1.11). Ces individus ont été choisis pour leur forte production de graines lors de la première campagne d'échantillonnage en 2002, leur situation spatiale (altitudes, densités) et leur facilité d'accès. Le Mont

Ventoux est composé de deux sous-sites d'échantillonnage, le Mont Serein où la densité est moyenne et le Contrat où la densité est relativement forte. Les différences de densité de ces peuplements peuvent être expliquées par leur âge : le Contrat étant plus ancien avec notamment un plus grand nombre de "semenciers reliques" à l'origine de la régénération naturelle du Ventoux. Le Mont Serein, quant à lui, résulte majoritairement de la politique RTM et présente donc des arbres plus jeunes, parmi lesquels la deuxième "cohorte" de régénération est désormais mâture. De plus la partie du Contrat se trouve plus au centre de la population de sapin et dans des conditions plus propices à l'installation des semis de sapins avec des combes plus humides et ombragées que sur le Mont Serein. Le peuplement étant continu, on peut regrouper, lors de notre étude, ces deux sous-populations. Néanmoins elles diffèrent en termes de densité locale et de fertilité individuelle, qui seront donc considérées en tant que covariables dans nos analyses.

Les caractéristiques de reproduction (production de cônes et de pollen) des individus ont été suivies entre 2002 et 2005. Les effectifs par année, par site et par caractéristiques sont présentés dans le tableau 1.1. Ceux-ci dépendent essentiellement des floraisons et fructifications individuelles, mais aussi des moyens disponibles pour le suivi.

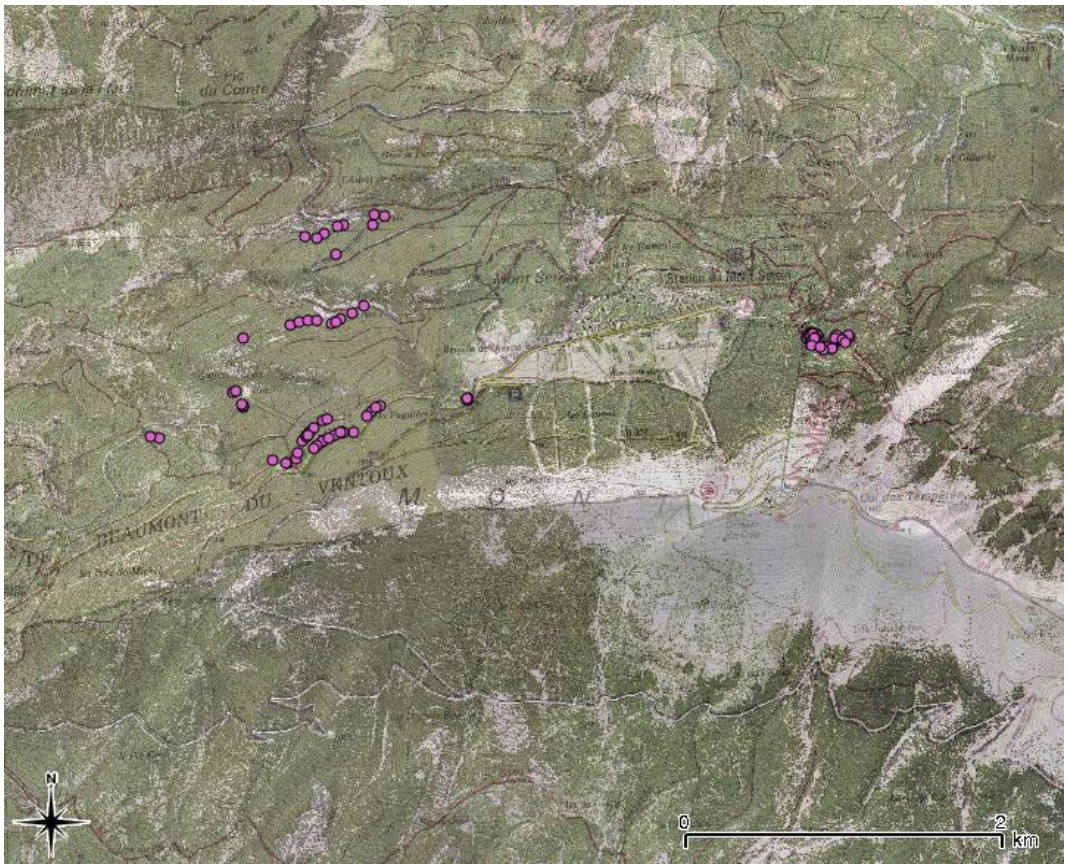


FIG. 1.10: Localisation géographique des sapins pectinés étudiés sur le mont Ventoux (cercles roses). Le fond de carte est une superposition de photographie aérienne et de cartographie au 1/25000 issu de l'IGN.

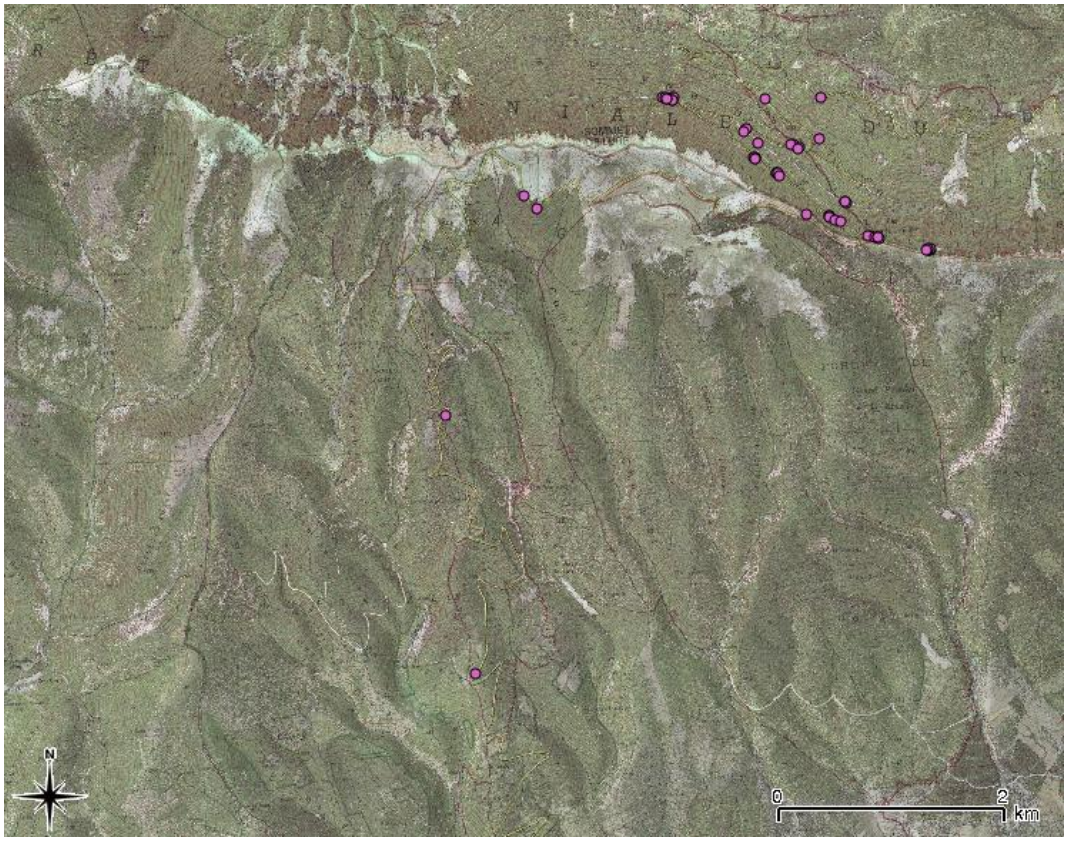


FIG. 1.11: Localisation géographique des sapins pectinés étudiés sur la montagne de Lure (cercles roses). Le fond de carte est une superposition de photographie aérienne et de cartographie au 1/25000 issu de l'IGN.

Site	Année	2002	2003	2004	2005	Ttes les années
Lure	Cones	**43** ; 50.7 ; [54.42]	**27** ; 28.41 ; [21.39]	**42** ; 14 ; [13.07]	**32** ; 34 ; [32.93]	**144** ; 32.1 ; [37.96]
	Pollen	**42** ; 105.8 ; [124.71]	**26** ; 86.92 ; [99.07]	**41** ; 143.9 ; [143.72]	**31** ; 83.1 ; [94.41]	**140** ; 108.4 ; [121.75]
	Tx autof.	**43** ; 0.29 ; [0.31]	**23** ; 0.2 ; [0.2]	**40** ; 0.33 ; [0.32]	**32** ; 0.15 ; [0.16]	**138** ; 0.26 ; [0.27]
	% vides	**43** ; 0.52 ; [0.29]	**26** ; 0.73 ; [0.24]	**42** ; 0.6 ; [0.2]	**32** ; 0.56 ; [0.27]	**143** ; 0.59 ; [0.26]
Mt Ventoux	Cones	**39** ; 36.77 ; [32.13]	**46** ; 24.57 ; [20.18]	**87** ; 60.32 ; [39.13]	**15** ; 44.13 ; [56.18]	**187** ; 45.32 ; [38.5]
	Pollen	**38** ; 196.1 ; [151.62]	**44** ; 80.85 ; [119.16]	**87** ; 234.6 ; [142.95]	**15** ; 74.01 ; [83.44]	**184** ; 176.8 ; [151.21]
	Tx autof.	**12** ; 0.27 ; [0.32]	**19** ; 0.14 ; [0.2]	**19** ; 0.33 ; [0.24]	**14** ; 0.2 ; [0.27]	**64** ; 0.23 ; [0.26]
	% vides	**39** ; 0.39 ; [0.25]	**46** ; 0.54 ; [0.23]	**88** ; 0.44 ; [0.17]	**15** ; 0.34 ; [0.2]	**188** ; 0.45 ; [0.21]
Deux Sites	Cones	**82** ; 44.07 ; [45.49]	**73** ; 25.99 ; [20.57]	**129** ; 45.24 ; [39.48]	**47** ; 37.23 ; [41.4]	**331** ; 39.57 ; [38.77]
	Pollen	**80** ; 148.7 ; [144.57]	**70** ; 83.1 ; [111.41]	**128** ; 205.5 ; [148.82]	**46** ; 80.13 ; [90.15]	**324** ; 147.2 ; [143.12]
	Tx autof.	**55** ; 0.29 ; [0.31]	**42** ; 0.17 ; [0.2]	**59** ; 0.33 ; [0.29]	**46** ; 0.17 ; [0.2]	**202** ; 0.25 ; [0.27]
	% vides	**82** ; 0.45 ; [0.28]	**72** ; 0.61 ; [0.25]	**130** ; 0.5 ; [0.19]	**47** ; 0.49 ; [0.27]	**331** ; 0.51 ; [0.25]

TAB. 1.1: Récapitulatif des moyennes individuelles par semencier des nombres de cônes, de la fertilité mâle, du taux d'autofécondation et de la proportion de graines vides par site et par année. Les valeurs indiquées représentent la taille de l'échantillon (en gras), la valeur moyenne de la variable mesurée ou estimée (souligné) et l'écart type pour cette même variable (entre parenthèses). **Note :** Les effectifs de la dernière colonne représentent le nombre de descendances et non le nombre de mères (i.e. plusieurs descendances par mère à différentes années).

3.2 Estimation de la fertilité

3.2.1 Mesure des fertilités femelles et mâles

La fertilité des individus a été suivie chaque année pendant 4 ans :

- La fertilité femelle a été mesurée en comptant exhaustivement le nombre de cônes portés par chacun des sapins. Ces observations ont été effectuées aux jumelles, ou bien dans certains cas directement en grimpant dans les arbres. La fertilité femelle est donc représentée par une variable quantitative (tableau 1.1).
- La fertilité mâle a été mesurée en attribuant à chaque individu une note de floraison, correspondant à la quantité de strobili portée par la partie du houppier susceptible d'en posséder. Nous avons défini 4 classes qualitatives de la façon suivante :
  - **0** : 0 à 25% des branches portent des strobili
  - **1** : 25 à 50% des branches portent des strobili
  - **2** : 50 à 75% des branches portent des strobili
  - **3** : 75 à 100% des branches portent des strobili

  L'estimation a été faite sur l'ensemble de la partie basse du houppier où l'on peut potentiellement trouver des organes reproducteurs mâles, et ce selon toutes les orientations (nord, sud, est, ouest). En effet, la production de strobili peut varier, notamment en fonction de la densité ou de l'exposition à la lumière (observations personnelles), la note globale est une moyenne sur l'ensemble du houppier. La fertilité mâle est donc représentée par une variable qualitative.

3.2.2 Estimation quantitative de la fertilité mâle

La note de production de pollen est relative au nombre de branches portées par un individu et ne reflète pas correctement la quantité de pollen réellement émise. En effet un arbre de petite taille peut avoir la même note qu'un arbre de grande taille bien que ce dernier produise bien plus de pollen que le premier. Nous avons donc transformé la note de fertilité mâle en une variable pseudo-quantitative en rapportant cette note à la taille du houppier. Les dimensions du houppier ont été

mesurées pour tous les individus pendant l'hiver 2006-2007 à partir des paramètres suivants :

- Hauteur totale de l'arbre, h_t
- Hauteur de la première branche basale du houppier, aussi appelée premier verticille vert, h_{pvv}
- Largeur du houppier selon les 4 directions cardinales, n,s,e et o.

A partir de ces données nous avons estimé la surface du houppier, S_h, et non son volume, car les fleurs mâles sont généralement situées à la périphérie. Pour calculer S_h nous avons assimilé le houppier à une pyramide à base quadrilatère, ainsi :

$$S_h = (h_t - h_{pvv}) \left(\frac{\sqrt{n^2+e^2}}{2} + \frac{\sqrt{e^2+s^2}}{2} + \frac{\sqrt{s^2+o^2}}{2} + \frac{\sqrt{o^2+n^2}}{2} \right). \tag{1.1}$$

Ensuite nous avons multiplié l'aire S_h, considérée constante pendant toute la durée de l'étude (4 ans), par la note qualitative de fertilité mâle, fql, pour obtenir la note pseudo-quantitative, fqt :

$$fqt_{ij} = fql_{ij} \times S_{h_i} \tag{1.2}$$

avec i l'individu et j l'année pendant laquelle la mesure a été effectuée. Les fertilités mâles ainsi estimées par site et par année représentent une mesure de la quantité de pollen produit, et sont présentées dans le tableau 1.1.

3.3 Estimation de la densité locale

Pour chaque arbre nous avons mesuré la densité locale en conspécifiques producteurs de pollen. Cette densité locale correspond au nombre de sapins potentiellement fécondants dans un rayon de 12,5m autour de chaque individu. Ce rayon correspond à une surface de $490m^2$ et semble cohérent avec des estimations de la dispersion du pollen effectuée sur la montagne de Lure qui montraient une distance moyenne de dispersion de cet ordre (Schossig 2004). Cette mesure a été effectuée pendant l'hiver 2006-2007, nous n'avons tenu compte que des individus producteurs de pollen présentant des cicatrices dus à la présence de strobili sur les rameaux datant d'au moins 2002. En effet, nous tenions à mesurer une densité d'arbres émetteurs de pollen pendant la durée de l'étude Nous avons ainsi ignoré les jeunes arbres produisant

pour la première fois du pollen en 2006. La densité locale moyenne sur le site de Lure est d'environ 186 $sapins.ha^{-1}$ (écart-type=150,8) et de 190 $sapins.ha^{-1}$ (écart-type=131,05) sur le Mont Ventoux. Au sein de ce dernier site on peut distinguer les deux zones de densités différentes :

- La partie du Contrat représentée par le groupe d'arbres échantillonnés à l'ouest (figure 1.10) dont la densité locale moyenne est de 266 $sapins.ha^{-1}$ (écart-type=145,8).
- La partie du Mont Serein qui est représentée dans notre échantillonnage par les arbres situés à l'est (figure 1.10), la densité locale moyenne estimée y est de 140 $arbres.ha^{-1}$ (écart-type=92,5).

3.4 Estimation du succès reproducteur femelle

Afin d'estimer le succès reproducteur femelle nous avons quantifié la viabilité des graines produites par chaque individu. Nous avons donc estimé la proportion de graines vides, c'est à dire ne comportant pas d'embryon viable. Chaque année ces analyses ont été effectuées sur 1 à 10 cônes en fonction de la fertilité femelle de l'arbre l'année en question. Nous avons employé deux méthodes différentes :

- La dissection des graines, qui était également nécessaire au génotypage. Dans ce cas, nous avons disséqué en général un minimum de 100 graines par arbre afin d'avoir des proportions de graines vides interprétables statistiquement (i.e. pour estimer correctement la proportion réelle de graines vides). Cependant nous avons dans certains cas disséqué beaucoup plus de graines (i.e. jusqu'à obtenir 10 embryons pour les génotypages, voir la partie suivante). Dans d'autres cas le nombre de graines récoltées était insuffisant (i.e. inférieur à 100), toutes les graines du lot étaient alors disséquées. Le nombre de graines disséquées varie entre 73 et 1754 par an et par arbre avec une moyenne de 183 $graines.arbre^{-1}$.
- La radiographie des graines par le biais de deux instruments : un appareil à radiographie (INRA Orléans) nécessitant de développer des clichés ou un scanner à rayon X (INRA-URFM Avignon). Dans le dernier cas, les graines

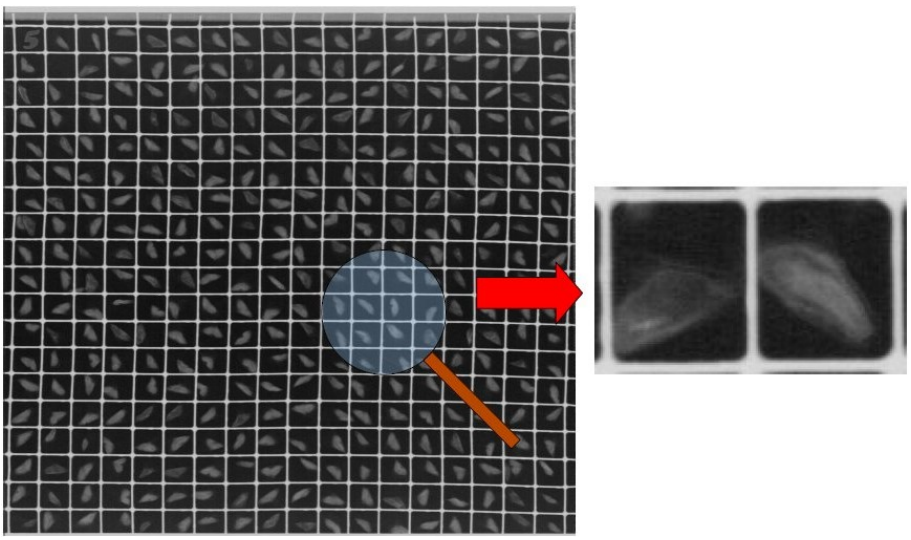


FIG. 1.12: Image obtenue grâce au scanner à rayons X. L'encart de droite représente en détail, une graine vide (à gauche) et une graine pleine (à droite).

ont été radiographiées sous une intensité de radiation de 20kV délivrée par une source Faxitron X-ray Mx-20 pendant 3 min 43 s. L'enregistrement de l'image a été faite avec un scanner EZ240 sous un balayage de 150 ms et une résolution moyenne de 2 pixel-binning. Ces réglages ont permis d'obtenir une qualité d'image suffisante tout en préservant la viabilité des graines pour d'éventuelles expériences ultérieures (notamment la germination). Dans les deux cas l'embryon était repéré par contraste : le vide apparaît noir à l'image, alors que les parties pleines apparaissent plus ou moins blanches en fonction de leur densité (i.e. notamment différence entre l'embryon et l'endosperme), ce que l'on voit clairement sur la figure 1.12.

Pour 12 descendances en 2003 et 22 descendances en 2004 nous avons estimé la proportion de graines vides en disséquant les graines et en les radiographiant. Nous avons donc estimé la proportion de graines vides que l'on considérait comme étant pleines avec la radiographie, ce taux d'erreur était de 8%. En considérant ce taux d'erreur indépendant de la descendance considérée, nous avons pu corriger les estimations basées uniquement sur des radiographies. Toutes les estimations et les effectifs considérés sont présentés dans le tableau 1.1. Il est à noter que de très faibles quantité de graines étaient parasitées par un insecte séminiphage, *Megastigmus spp.*. Cette

proportion de graines a été écartée des analyses car elle était toujours inférieure à 5% de la descendance échantillonnée. De plus, il est difficile de savoir si les graines en question ont ou non été fécondées. En effet, von Aderkas et al. (2005) ont montré que l'insecte pouvait déclencher la formation d'un endosperme en mimant la fécondation.

3.5 Génotypage

Nous avons génotypé certaines des descendances des arbres échantillonnées (voir les effectifs dans le tableau 1.1). Nous avons employé un jeu de 4 marqueurs microsatellites chloroplastiques (donc haploïdes) développés par Vendramin et al. (1996) et Liepelt et al. (2001). Les chloroplastes ont la particularité d'être hérités paternellement chez les conifères, comme le sapin pectiné (Vendramin et Ziegenhagen 1997). Nous avons extrait l'ADN en utilisant les kits DNEasy Plant Mini Kit (Qiagen©) et amplifié les microsatellites par P.C.R. avec le Qiagen© multiplex amplification kit sur des thermocycleurs Eppendorf© Mastercycler EP Gadient et un MJ Research© PTC200 (en suivant les protocoles standards et Vendramin et al. (1996) et Liepelt et al. (2001)). Les lectures ont été faîtes sur deux types de séquenceurs : un séquenceur sur LiCor© ReadIR 4200 et un séquenceur à capillaires Molecular Dynamics MegaBace© 1000. De plus par souci de précision, toutes les lectures ont été faites au moins trois fois par deux personnes différentes. Nous avons obtenu les génotypes de 1 à 55 graines par arbre, avec une moyenne de 15 $graines.arbres^{-1}.an^{-1}$ (i.e. dans le cas où nous n'avions qu'une graine elle n'a bien sûr pas servi aux analyses). Seules les familles représentées par au moins deux descendants ont été considérés pour les analyses. Les données génétiques sont présentées dans les figures 1.2.

Montagne de Lure

Années	Nb d'individus	Nb d'haplotypes	Probabilité d'exclusion
2002	452	72	0.94
2003	417	53	0.93
2004	410	37	0.92
2005	337	62	0.96
Total	1529	138	0.97

Mont Ventoux

Années	Nb d'individus	Nb d'haplotypes	Probabilité d'exclusion
2002	127	23	0.87
2003	837	69	0.94
2004	599	36	0.91
2005	142	24	0.86
Total	1661	95	0.93

TAB. 1.2: **Tableau récapitulatif des données génétiques utilisées** *"Nb individus"* représente le nombre total d'individus échantillonnés (mères et graines). *"Nb d'haplotypes"* correspond au nombre d'haplotypes complets observés. Les probabilités d'exclusion ont été calculées à partir des fréquences des haplotypes issues de l'estimation des probabilités d'autofécondation et de copaternité. Les graines dont les haplotypes présentent des données manquantes pour au moins trois des quatres loci ont été retirées du jeu de données.

Chapitre 2

Les espèces en marge de leur aire de répartition : Le régime mixte de reproduction chez les conifères méditerranéens

1 Synthèse

Dans ce chapitre nous nous sommes intéressés au régime mixte de reproduction des conifères et à l'avantage sélectif qu'il pouvait ou non procurer, et plus particulièrement dans les populations en marge de leur aire de répartition. La région méditerranéenne naturellement hétérogène de par sa géographie subit diverses pressions anthropiques : cultures, aménagement, incendies ; et présente donc une forte hétérogénéité de milieux. Dans un premier temps une analyse bibliographique nous a permis de faire un bilan du régime de reproduction de trois genres de conifères présents en région Méditerranéenne : *Pinus*, *Picea* et *Abies*. Cette analyse est basée sur 46 études étudiant le régime de reproduction dans une ou plusieurs populations et pour lesquelles nous avions accès à la densité de manière qualitative (d'après les auteurs). Nous en avons déduit des variations de régimes de reproduction et de densité. Nous avons également classé les populations suivant trois catégories : **M**, méditerranéennes, **nM**, non méditerranéennes et **nM-r** pour les populations non-

méditérranéennes mais présentes sous ce climat en limite de leur aire de répartition.

La densité au sein de ces milieux est positivement corrélée avec le taux d'allofécondation estimé au sein de plusieurs populations de conifères (voir tableau 1 et figure 3 de l'article). Il existe également une corrélation positive entre la gamme de variation du taux d'allofécondation et la gamme de variation de densité (figure 5 de l'article), avec généralement une plus grande variabilité au sein des populations en marge de leur aire de répartition (figure 2). Ceci suggère que le régime de reproduction est régi par une loi d'action de masse (Holsinger 1991), le taux d'autofécondation est alors le résultat de la dilution de l'autopollen au sein du nuage de pollen fécondant reçu des voisins. Les populations marginales présentent une large gamme de densités qui peut alors expliquer la large variation du régime de reproduction observée. Potentiellement, ces populations marginale peuvent donc également présenter une forte variabilité de l'expression de la dépression de consanguinité. Le régime mixte de reproduction semble donc *a priori* défavorable, sauf s'il procure une assurance de reproduction efficace, c'est à dire si l'avantage procuré par ce dernier peut compenser le désavantage lié à la dépression de consanguinité.

Nous avons testé cette hypothèse dans une population de sapin pectiné (*Abies alba* Miller) sur le Mont Ventoux. Nous y avons échantillonné 4 parcelles (figure 1 de l'article) qui diffèrent en termes d'âge et de densité, pour lesquelles nous avons procédé à un échantillonnage de graines. Les deux parcelles les plus denses présentent des taux d'allofécondation estimés plus forts que les parcelles à moyenne ou à faible densité. Ces résultats sont en accord avec la tendance observée sur les données bibliographiques, avec un régime de reproduction qui semble répondre à une loi d'action de masse. Afin de tester si les graines produites dans les parcelles peu denses, et donc avec un plus fort taux d'autofécondation, présentent une valeur sélective différente des graines issues des parcelles plus denses, nous avons effectué des tests de germination et de croissance sous trois niveaux de stress hydriques (normal, stressant et sécheresse). Les graines issues des parcelles dont la densité est faible possèdent un taux de germination significativement plus faible que les graines issues des peuplements plus denses (33,8% vs. 52,3%). En revanche les semis issus des

graines majoritairement autofécondées présentent une meilleure survie sur 4 ans que les semis issus des graines allofécondées lorsque cette survie est moyennée sur les 3 traitements de stress hydrique (81% vs. 63%). Cette différence s'explique principalement par la différence observée dans le traitement le plus stressant (voir tableau 2 de l'article).

Cette expérience montre qu'il existe une forte dépression de consanguinité au stade germination, et donc une purge importante du fardeau génétique. Ce phénomène de purge génétique confirme une étude parallèle effectuée dans les mêmes populations de sapin pectiné par Pichot et al. (2006), dans laquelle les taux d'autofécondation individuels (estimés à partir de marqueurs génétiques) sont globalement plus faibles pour les graines ayant germées que pour les graines n'ayant pas germées. Cependant, nos expériences de survie menées sous différentes conditions de stress hydrique montrent une survie globale (i.e. germination et survie sur 4 ans) équivalente entre les semis issus des peuplements denses et peu denses (33% vs. 27%). Ainsi la purge précoce des individus possédant le fardeau génétique le plus élevé pourrait permettre de sélectionner les individus les plus performants. En effet, les différences de valeurs sélectives entre individus issus de peuplements denses et peu denses n'apparaissent pas lorsque le niveau de stress est faible. Cependant lorsque le niveau de stress augmente, le fardeau génétique s'exprime par le biais d'une dépression de consanguinité élevée, ce qui est cohérent avec de précédentes études (Byers et Waller 1999, Cheptou et al. 2000*a*). Les régimes mixtes de reproduction pourraient donc être bénéfiques pour les individus en marge de leur aire de répartition, comme les populations de sapin sur le Mont Ventoux. En effet, ils permettraient à la fois une assurance de reproduction dans les endroits où la densité en partenaires potentiels est faible mais également une purge efficace du fardeau génétique dans des conditions de stress exacerbées en marge de l'aire de répartition.

Afin de vérifier ce résultat, nous allons dans les chapitres suivants préciser l'effet de la marginalité sur différents aspects du cyle de vie des arbres. La grande variabilité des taux de survie et de germination entre les individus (voir le tableau 2 de l'article), implique qu'il nous faut également estimer le régime de reproduction

de manière individuelle. De plus, nous souhaitons vérifier comment la densité locale influence la viabilité des graines au travers de la dépression de consanguinité ou de la disponibilité en pollen, à l'échelle individuelle et non plus à l'échelle de la population ou du peuplement. Ainsi afin de comprendre les potentialités évolutives et de survie des populations de sapin du Mont Ventoux, et du sapin en marge de son aire de répartition en général, il nous faut répondre aux trois questions suivantes :

- Les individus se comportent-ils de manière identique (chapitre 3 et 5) ?
- Comment est modifié le succès reproducteur femelle par la densité locale ? Dépression de consanguinité ou limitation pollinique (chapitre 3) ?
- Comment la diversité génétique spatiale du nuage pollinique est-elle structurée (chapitre 5) ?

Nous développerons ces questions dans les chapitres suivants.

2 Life at the margin : the mating system of Mediterranean conifers

Cet article a été publié dans le numéro 8 de Web Ecology en Juillet 2008, il fait suite à une conférence donnée en septembre 2007 à Perth en Australie dans le cadre du colloque MEDECOS XI, une conférence consacrée aux écosystèmes méditerranéens. Certains des résultats présentés ici proviennent de la thèse de Fabrice Sagnard (soutenue en 2001) et du stage de Master 2 de Daniel E. Silva (soutenu en 2006).

Web Ecology 8: 94–102.

Life at the margin: the mating system of Mediterranean conifers

Gwendal Restoux, Daniel E. Silva, Fabrice Sagnard, Franck Torre, Etienne Klein and Bruno Fady

Restoux, G., Silva, D. E., Sagnard, F., Torre, F., Klein, E. and Fady, B. 2008. Life at the margin: the mating system of Mediterranean conifers. – Web Ecol. 8: 94–102.

Mixed mating, where a single tree progeny results from a mixture of selfing and outcrossing, is widespread in conifers and could be an evolutionary advantage at ecological margins when mating partners become scarce. This study analyzes how the mating system responds to bioclimate and density variations. We surveyed published data on the mating system of *Abies*, *Picea* and *Pinus* species when information on bioclimate and stand density was available. Our survey revealed that Mediterranean species demonstrate a lower selfing rate than other species and that the proportion of selfed versus outcrossed progeny is not fixed within species. The highest variability in mating types within populations was found when stand density was the most variable.

To show how density affects the proportion of selfed versus outcrossed progeny, we used isozymes to genotype single tree seeds from a marginal *Abies alba* forest in Mediterranean France (Mont Ventoux) where low-to high-density stands are found. We then tested the adaptive potential of the different high and low density progenies by sowing them under controlled nursery conditions and measuring germination rate and seedling survival after 4 years under 3 different water regimes. Although the mean value of outcrossing rate was typical for mixed mating conifers (t_m = 0.85), individual outcrossing rates varied from 0.05 to 0.99 and were strongly correlated with stand type and density (t_m from 0.87 in high-density to 0.43 in low-density marginal stands). Significantly fewer seeds from the low density marginal stand germinated (32% vs 53% in the high density mature stand), although seedlings from the marginal stand had a better 4-year survival (81%) than seedlings from the high-density mature stand (63%) when the three water regimes (from least to most stressful) were averaged.

Trees from low density stands may be at a selective disadvantage because they produce more selfed progeny than trees from high-density stands. However, selfed seeds may be purged early on (at the embryo stage) and the remaining seeds tend to produce seedlings with better fitness. Variability of the selfing rate might thus be an adaptive strategy for species in highly contrasted environments where selection of both reproductive assurance and avoidance of inbreeding depression may be density dependent.

G. Restoux, F Sagnard and B. Fady (fady@avignon.inra.fr), INRA, UR629, Ecologie des Forêts Méditerranéennes (URFM), Domaine St Paul, FR–84914 Avignon cedex, France. – D. E. Silva and F. Torre, IMEP, UMR CNRS 6116, Univ. Paul Cézanne Aix-Marseille III, Europôle Méditerranéen de l'Arbois, B.P. 80, FR–13545 Aix-en-Provence cedex 04, France. – E. Klein, INRA, UR546, Biostatistique et Processus Spatiaux, Domaine St Paul, FR–84914 Avignon cedex, France. – D. E. Silva also at: INRA, UMR Écologie et écophysiologie forestières, FR–54280 Champenoux, France. – F. Sagnard also at: ICRISAT/CIRAD, c/o Campus ICRAF, PO Box 39063, Nairobi, Kenya.

Accepted 23 June 2008
Copyright © EEF
ISSN 1399-1183

The study of reproduction in plants encompasses a wide range of different processes, from flower and seed fecundity to mating barriers, from the scale and shape of pollen and seed flow to the frequency of long-distance dispersal (Griffin and Sedgley 1989). These processes are traditionally analyzed either from a population genetics or an ecological perspective (Cheptou and Schoen 2007). From the population genetics perspective, the mating system is a complex trait that affects the expression of selective forces. By increasing heterozygosity, outcrossing increases the effective rate of recombination by mixing non-related genomes (promoting the spread of beneficial alleles and consequently the purge of deleterious ones) while selfing increases the expression of genetic load (and subsequently the purge of deleterious recessive mutations). Mixed mating, in which a single tree progeny results from a mixture of selfing and outcrossing, is widespread in plants, although its maintenance remains an evolutionary puzzle (Goodwillie et al. 2005).

On the one hand, if selfing is advantageous (because two copies of the same genes are transmitted to the progeny instead of just one, Fisher 1941), alleles responsible for selfing should spread and become fixed in the population. On the other hand, if selfing is disadvantageous (because of inbreeding depression following the appearance of deleterious alleles at a homozygous state), selfing alleles should disappear from the population. However, because the genetic load is purged by selfing, either complete outcrossing or complete selfing should be evolutionary stable strategies depending on the relative fitness of selfed versus outcrossed progeny (Lande and Schemske 1985).

Several theories have been proposed to explain the widespread co-existence of both selfing and outcrossing within individuals, populations and species: spatial and temporal variability of inbreeding depression, pollen quantity-dependant outcrossing rate ('mass action' and 'pollen discounting') and reproductive assurance (Holsinger 1991, Porcher and Lande 2005; reviewed by Goodwillie et al. 2005, Cheptou and Schoen 2007).

The reproductive assurance hypothesis explains why mixed mating may be conserved states where there is a need to balance the cost of producing selfed seeds (suffering from inbreeding depression) with the cost of producing no seeds at all when partners (or pollinators) become scarce, while still maintaining the potential to produce outcrossed seeds (with limited inbreeding depression) when partners (or pollinators) are plentiful (reviewed by Cheptou and Schoen 2007). Empirical evidence demonstrating larger seed-sets for mixed mating versus strictly outcrossing individuals in low density populations is still very limited (Cheptou and Dieckmann 2002, Kalisz et al. 2004).

Mixed mating is widespread in trees, particularly in conifers (Mitton 1992). Although conifers have various life history traits, most are wind-pollinated and monoecious, which can lead to automatic selfing due to geitonogamy. Furthermore, they colonize a wide range of (often severe) environments (from boreal to tropical) to which they have adapted since their emergence in the Carboniferous (DiMichele et al. 2001) and where their population densities have varied considerably. Therefore, they are a good model to test whether or not demography (and thus mate availability) can affect reproductive output and whether selfing contributes to reproductive assurance at low densities. For example, Robledo-Arnuncio et al. (2004) showed that low-stand density significantly increased both the proportion of selfed progeny and the rate of empty seeds due to inbreeding depression in the mixed mating conifer *Pinus sylvestris*.

Here, we test and discuss the hypothesis that reproductive assurance can be an evolutionary advantage at a species' ecological margin when and where potential mates become scarce (low census and/or low density). First, we analyze how bioclimatic and density variations can affect variability in the mating system using a set of published data compiled for conifers. We focused our attention on the Mediterranean area, where we expected to find strong local variations in population spatial structure due to the high geographical, climatic and environmental heterogeneity of this biome (Blondel and Aronson 1999). Second, we used original data to show how stand density affects the selfing rate, inbreeding depression and survival potential of *Abies alba*, a mixed mating European conifer that has numerous marginal populations under the Mediterranean climate (Quézel and Médail 2003).

Material and methods

We first conducted a survey of published data on the mating system of conifers, restricting our research to the three genera *Abies*, *Picea* and *Pinus* (where all species are mixed mating), where information on bioclimate (using 3 classes defined as Mediterranean, M, non-Mediterranean, nM, and non-Mediterranean but at the rear-edge of the natural distribution area, nM-r) and stand density were available. We collected 6, 30 and 10 studies dealing with M, nM and nM-r species, respectively. From these data, we computed the mean, minimum and maximum values of the average outcrossing rate (t_m) within each study (Table 1). We considered that t_m values greater than 1 were due to aberrations (MLTR or MLTF, Ritland 2002), and thus made them equal to 1 for all computations of outcrossing ranges and means. We attributed a density level to each population and classified studies according to species (Table 1). We separated density into 3 levels: high (3), medium (2) and low (1). We relied on authors' estimates of what they considered low, medium or high density for the species studied. For *Abies alba*, we considered that stand density was (i) low when adult trees were fewer than 25 per hectare, (ii) medium when adult trees were between 25 and 100 per hectare and (iii) high when density was over 100. The definition of these density classes is based on our own

Table 1. Comparison of mean outcrossing rates (t_m) in *Abies*, *Picea* and *Pinus* species in relation to bioclimate and mean population densities. Values of t_m in parenthesis are population extremes. An absence of population extremes indicate that only a single population was investigated. Values of t_m are estimated using isozyme or DNA markers and seed material, except Sorensen (1994) who used phenotypic markers. M = Mediterranean, nM = non-Mediterranean, nM-r = 'rear edge' Mediterranean distribution of a typically non-Mediterranean species,* = seed orchard.

Species	Bioclimate	Stand density	t_m	Geographic origin	Reference
Abies alba	nM-r	low to high	0.72 (0.43–0.87)	Medit. France	This study
Abies alba	nM	medium to high	0.89 (0.68–1.00)	German Alps	Schroeder 1989
Abies amabilis	nM	medium to high	0.82 (0.80–0.84)	W. Canada	El-Kassaby et al. 2003
Abies balsamea	nM	medium to high	0.89 (0.78–0.99)	E. USA	Neale and Adams 1985
Abies borisii regis	M	high	0.94	Greece	Fady and Westfall 1997
Abies lasiocarpa	nM	medium to high	0.89 (0.65–0.94)	Colorado, USA	Shea 1987
Abies procera	nM	high	0.94 (0.90–0.98)	Denmark (*)	Siegismund and Kjaer 1997
Picea abies	nM	medium to high	0.95 (0.95–0.96)	Italy	Morgante et al. 1991
Picea abies	nM	medium	0.83 (0.82–0.84)	Italy	Boscherini et al. 1993
Picea breweriana	nM-r	low to high	0.87 (0.75–1.00)	California, USA	Ledig et al. 2005
Picea chihuahuana	nM-r	low	0.08 (0.00–0.15)	Mexico	Ledig et al. 1997
Picea engelmanii	nM	medium to high	0.86 (0.78–0.96)	Colorado, USA	Shea 1987
Picea glauca	nM	low to high	0.90 (0.89–0.92)	Ontario, Canada	O'Connell et al. 2006
Picea glauca	nM	medium	0.73 (0.70–0.76)	E Canada	Innes and Ringius 1990
Picea mariana	nM	medium to high	0.72 (0.63–0.83)	Alberta, Canada	Sproule and Dancik 1996
Picea mariana	nM	high	0.93 (0.89–0.98)	Eastern Canada	Boyle and Morgenstern 1986
Picea martinezii	nM-r	low to medium	0.56 (0.40–0.69)	Mexico	Ledig et al. 2000
Picea mexicana	nM-r	low to high	0.68 (0.59–0.81)	Mexico	Ledig et al. 2002
Picea omorika	nM-r	medium to high	0.95 (0.84–1.02)	Serbia	Kuittinen and Savolainen 1992
Picea rubens	nM	low to high	0.60 (0.53–0.65)	E Canada	Rajora et al. 2000
Picea sitchensis	nM	medium to high	0.86 (0.65–0.97)	W Canada and USA	Mimura and Aitken 2007
Pinus albicaulis	nM	medium to high	0.86 (0.73–0.98)	W Canada and USA	Bower and Aitken 2007
Pinus attenuata	M	high	0.96	California, USA	Burczyk et al. 1996
Pinus banksiana	nM	high ?	0.89 (0.83–0.95)	Ontario, Canada	Fu et al. 1992
Pinus brutia	M	high	0.95 (*)	Turkey	Kaya et al. 2006
Pinus cembra	nM	low	0.69	Ukraine	Krutovskii et al. 1995
Pinus cembra	nM	high	0.81	Italy	Lewandowski and Burczyk 2000
Pinus chiapensis	nM	medium to high	0.87 (0.49–1.20)	Mexico	Del Castillo and Trujillo 2008
Pinus contorta	nM	high	0.95 (0.93–0.98)	Alberta, Canada	Perry and Dancik 1986
Pinus greggii	nM	high	0.97 (0.92–1.00)	Mexico	Parraguirre-Lezama et al. 2004
Pinus jeffreyi	nM	medium to high	0.94 (0.88–0.97)	Oregon, USA	Furnier and Adams 1986
Pinus kesyia	nM	low to high	0.83 (0.68–0.97)	Thailand	Boyle et al. 1991
Pinus koraiensis	nM	high	0.97 (0.92–1.03)	Far east Russia	Krutovskii et al. 1995
Pinus leucodermis	M	medium to high	0.77 (0.72–0.82)	Italy	Morgante et al. 1993
Pinus maximartinezii	nM-r	low to medium	0.82 (0.44–0.97)	Mexico	Ledig et al. 1999
Pinus merkusii	nM	high	0.87 (*)	Indonesia	Siregar and Hattemer 2001
Pinus merkusii	nM	low to high	0.47 (0.02–0.84)	Indonesia	Changtragoon and Finkeldey 1995
Pinus pinaster	M	high	0.98 (0.95–1.00)	Spain	De Lucas et al. 2008
Pinus pinceana	nM-r	medium	0.87 (0.84–0.90)	Mexico	Ledig et al. 2001
Pinus ponderosa	nM	high	0.91 (0.8–0.92)	NW USA	Sorensen 1994
Pinus radiata	M	medium to high	0.96 (0.9–1.00)	California and Mexico	Vogl et al. 2002
Pinus sibirica	nM	high	0.89 (0.82–0.98)	Siberia	Krutovskii et al. 1995
Pinus strobus	nM	medium to high	0.93 (0.87–0.99)	Canada	Rajora et al. 2002
Pinus sylvestris	nM-r	low to high	0.93 (0.75–0.99)	Spain	Robledo et al. 2004
Pinus sylvestris	nM	high	0.98 (0.96–0.99) *	Poland	Burczyk 1998
Pinus washoensis	nM-r	medium to high	0.86 (0.76–0.92)	Nevada, USA	Mitton et al. 1997

observations of natural stands in our study site (Mont Ventoux), where a significant spatial genetic structure was observed on a progeny when density was below 25 trees per ha (Sagnard 2001). From this data set, we estimated the variability of outcrossing rates between stands for each study by subtracting the minimum t_m value observed in the study from the maximum one. We also estimated the variability of stand density for each study: from one, when there was only one density class, to three, when density ranged from low to high. We tested the effect of bioclimate (M vs nM and nM-r either individually or grouped in a single non-Mediterranean class) on average outcrossing rate for each study, using a non-parametric ANOVA (Kruskal–Wallis rank sum test), because we could not test for variance equality. We also tested the relationship between stand density and the average outcrossing rates of stands using the same Kruskall–Wallis rank sum test.

Secondly, we used isozymes to investigate the genetic composition of individual progeny from a marginal *Abies alba* population growing under Mediterranean climate in a southeastern French mountain (Mont Ventoux, 44°10´N, 5°15´E). Mont Ventoux is characterized by stands of different ages, from old-growth relictual stands to young stands where *Abies alba* has recolonized grazing areas abandoned in the late 19th century. Stands are also characterized by contrasting mature tree densities, from low (fewer than 25 trees per hectare in recolonized areas) to high (more than 100 trees per hectare in old, mature stands). We collected seeds from 52 trees from 4 different stands (Fig. 1): two high-density old-growth stands (HD1 and HD2), one medium-density old-growth stand at the periphery of an old-growth high-density stand, MD, and one low-density recolonizing stand (LD). A set of 7 polymorphic Mendelian isozymes developed for *Abies alba* was used on embryos, following the standard extraction, electrophoresis and scoring protocols of Fady et al. (1999). A total of 842 seeds were used to assess selfing rate and pollen pool composition using the MLTR software (Ritland 2002). We computed selfing rates for each sampled individual in LD and MD stands but only for stand level in HD1 and HD2 stands.

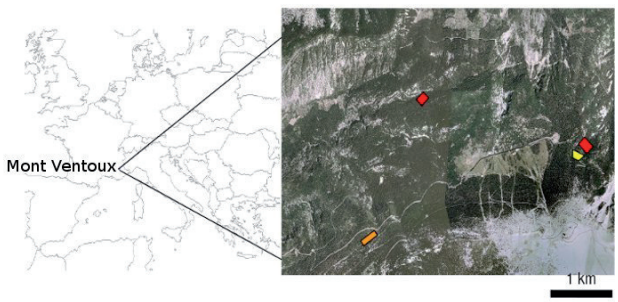


Fig.1. Localization of Mont Ventoux at the European scale and localization of the stands analyzed within Mont Ventoux: high density mature stands HD1 and HD2 (in red), medium density peripheral stand MD (in orange) and low density recolonizing stand LD (in yellow).

Finally, we tested early inbreeding depression and survival potential of the different progenies by sowing them under controlled nursery conditions and measuring germination rate and seedling survival after a 4-year period. A subset of 32 seed trees originating either from the HD1 or LD stands was used for this experiment. Germination rate, i.e. the proportion of seeds in each progeny which produced a 5 mm long emerging radicle after stratification (6 weeks at +4°C in moist peat), was measured for all collected seeds (from 215 to 1790 seeds per tree). From these seed lots, 2400 germinated seeds (i.e. 75 seeds per seed tree) were planted in the nursery, using 9 complete blocks and a randomized progeny design within blocks. Three groups of 3 blocks each received different water regimes: drought (Wilting point, Wp = 3–3.5), light water stress (Wp = 2.7) and normal watering (Wp = 1.5). Progeny survival was measured each year for 4 years. Significance of germination rate and of survival was tested using a logistic regression within an analysis of deviance ('glm' function in R).

All statistical analyses were carried out using the R software, R Development Core Team.

Results

Our literature survey of conifer trees showed that the variability in the mating system of genera *Abies*, *Picea* and *Pinus* were highly variable among populations and species. The average t_m for each bioclimatic class was 0.92, 0.84 and 0.70 for Mediterranean (M), non-Mediterranean (nM) and non-Mediterranean rear-edge (nM-r) species or populations, respectively. There was a significant difference of t_m between the three bioclimatic classes ($Chi^2 = 9.13$, DF = 2, $p = 0.01$, Fig. 2). Furthermore, there was a significant difference between the average t_m over Mediterranean type species ($t_m = 0.92$) and over non-Mediterranean species, including both nM and nM-r species, ($t_m = 0.81$) ($Chi^2 = 5.68$, DF = 1, $p = 0.017$). A significantly higher outcrossing rate was found when stand density was higher ($Chi^2 = 23.81$, DF = 4, $p < 0.0001$, Fig. 3). Further, we found a general trend toward a higher t_m range in marginal populations (Fig. 4), although this could not be tested as this range was mainly dependent on the number of stands sampled in each study. Finally, a higher t_m range was observed when the density range was higher ($Chi^2 = 13.88$, DF = 2, $p = 0.001$, Fig. 5).

Our isozyme data showed that the mean multi-locus outcrossing rate (t_m) of the Mont Ventoux *Abies alba* population over the 4 stands was 0.72 (Table 1). However, although trees were in the same meso-environment, there were considerable differences in mean stand and individual outcrossing rates. Within the old growth stands HD1 and HD2 (tree density > 100 ha^{-1}), average t_m were 0.85

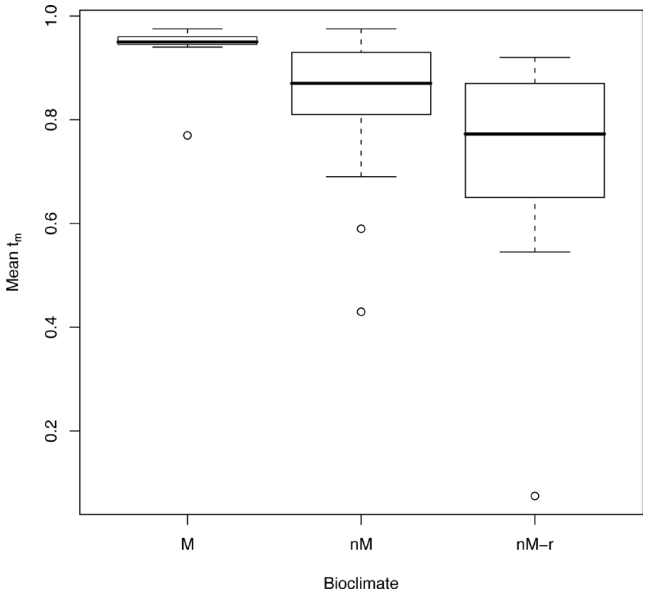


Fig. 2. Outcrossing rates over three bioclimates (Mediterranean, M, non-Mediterranean, nM, and rear edge non-Mediterranean nM-r) in *Abies*, *Picea* and *Pinus* species worldwide.

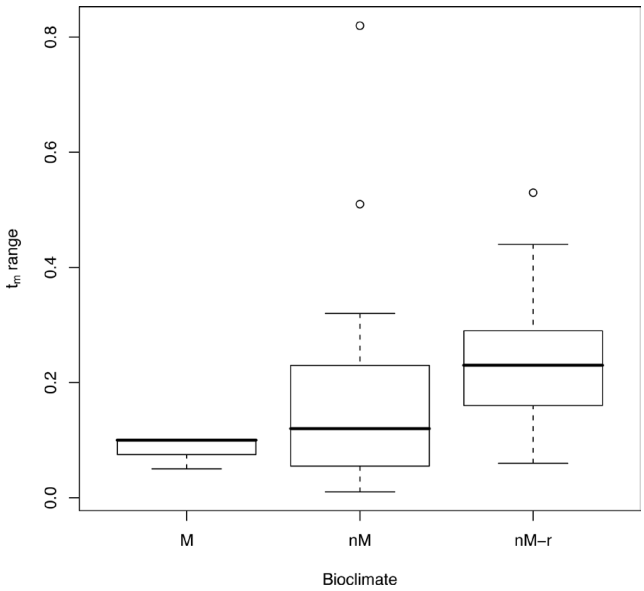


Fig. 4. Outcrossing rate variability among populations over three bioclimates (Mediterranean, M, non-Mediterranean, nM, and rear edge non-Mediterranean, nM-r) in *Abies*, *Picea* and *Pinus* species worldwide.

and 0.87, respectively. At stand edge, t_m decreased to 0.71 within the medium density mature stand (MD), with individual values ranging from 0.45 to 0.99. In the marginal low density recolonizing stand LD (density < 30 ha^{-1}), t_m was 0.43 (individual values from 0.05 to 1.18). We found no difference in pollen pool genetic diversity between the different stands (H_e from 0.141 to 0.162).

Significantly fewer seeds from the low density marginal stand LD were able to germinate: 33.8% vs 52.3% in the old-growth high-density stand HD1 (F = 18.09, DF = 30, p = 0.0002, Table 2). However, seedlings from LD had an overall better survival (81%) than seedlings from HD1 (63%) after 4 years (p <0.0001). The analysis of deviance showed that water regime, stand density type and their interaction with progeny could explain differences in survival (p <0.0001). Survival decreased with water stress for seeds originating from both LD and HD1 (survival rates: 82% and 73% for LD and HD1, respectively). Averaging

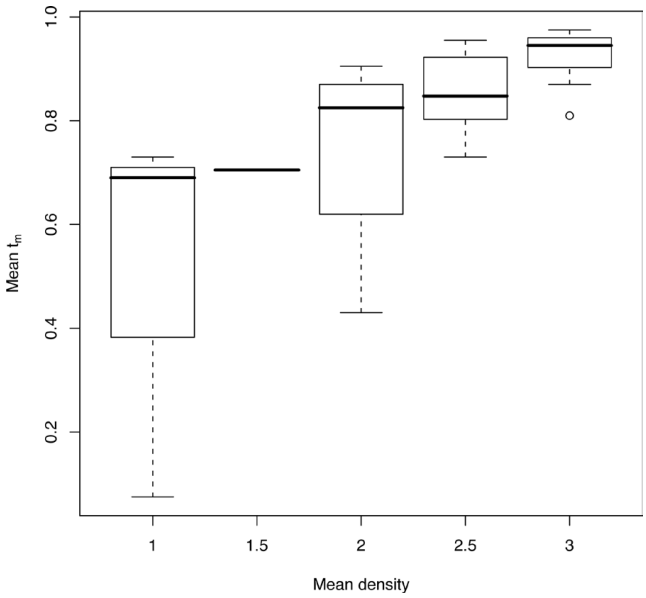


Fig. 3. Effect of density on outcrossing rate at the population level over three genera *Abies*, *Picea* and *Pinus*.

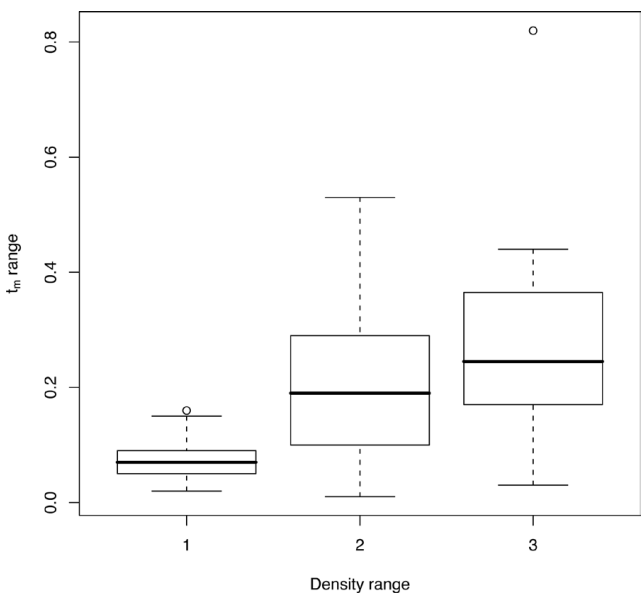


Fig. 5. Outcrossing rate variability versus density variability in *Abies*, *Picea* and *Pinus* species worldwide.

all three water treatments, LD seedlings had a better survival than HD1 ones. However, this difference is mainly due to the driest water regime (Table 2) and no significant difference between LD and HD1 seedlings was observed under the normal water regime.

Discussion

For reasons ranging from highly heterogeneous topography to the heterogeneity of land uses due to urbanization and agriculture in highly anthropogenic environments,

Table 2. Germination rate of the 32 seed lots tested according to stand density. Density type: HD = high density, LD = low density.

Name of seed tree	Density type	Total number of germinated seeds	Total number of seeds tested	Germination rate in %	Mortality in % after 4 years under high water stress
s 23 hp	HD	580	995	58.29	21.9
s 24 hp	HD	954	1790	53.30	46.7
s 25 hp	HD	838	1568	53.44	16.7
s 26 hp	HD	663	1251	53.00	39.3
s 27 hp	HD	323	577	55.98	29.6
s 28 hp	HD	561	1296	43.29	18.2
s 29 hp	HD	446	826	54.00	18.5
s 30 hp	HD	202	697	28.98	18.4
s 31 hp	HD	251	359	69.92	25.0
s 32 hp	HD	414	901	45.95	25.6
s 33 hp	HD	622	1147	54.23	41.5
s 34 hp	HD	355	1089	32.60	15.7
s 35 hp	HD	444	972	45.68	15.4
s 36 hp	HD	467	897	52.06	9.7
sem 21 hp	HD	796	1118	71.20	70.4
sp 22 hp	HD	857	1280	66.95	18.2
parc 5	LD	361	1367	26.41	2.5
parc 661	LD	134	698	19.20	11.8
parc 813	LD	158	627	25.20	23.8
s 18 hp	LD	338	1048	32.25	12.8
s 19 hp	LD	227	513	44.25	21.2
s 4 newhp	LD	162	460	35.22	7.1
sem 1 hp	LD	123	509	24.17	32.0
sem 12 hp	LD	152	435	34.94	14.8
sem 13 hp	LD	101	351	28.77	38.9
sem 14 hp	LD	463	907	51.05	9.5
sem 17 hp	LD	458	1212	37.79	23.1
sem 2 hp	LD	65	823	7.90	50.0
sem 20 hp	LD	304	794	38.29	4.3
sem 3 hp	LD	215	610	35.25	7.7
sem 6 hp	LD	591	1049	56.34	8.7
sem 8 hp	LD	208	595	34.96	20.0

species distribution areas are typically fragmented in the Mediterranean, thus leading to stands of highly contrasted densities, census and effective sizes (Blondel and Aronson 1999, Fady and Médail 2004). Out of the 9 *Abies* and 9 *Pinus* species found in Europe, 13 are Mediterranean endemics and an additional 3 have the 'rear-edge' of their distribution within the Mediterranean bioclimatic zone (Quézel and Médail 2003). Thirteen of these 16 species have patchy distribution areas under Mediterranean conditions. In such heterogeneous environments in terms of density and population size, pollen limitation may be a critical factor for reproduction at fine spatial scale (Restoux et al. unpubl.). Thus, a mixed mating system may be at an evolutionary advantage in such environments, since selfing ensures reproductive assurance when pollen is limited (Knight et al. 2005). On the other hand, pollen discounting, i.e. the cost of selfing due to pollen loss which is not dedicated to outcrossing, favors outcrossing when pollen is not limited (Holsinger 1991).

However, it is not because species live under Mediterranean conditions, characterized by heterogeneity in terms of density, that higher selfing or more variable mating types among and within populations are observed. Mediterranean-type conifers are typically major outcrossers, just like other conifers elsewhere (Table 1). The mean multi-locus outcrossing rate (t_m) found for *Pinus attenuata* in California by Burczyk et al. (1996), for example, is typical for that of mixed-mating conifers, with a low overall level of selfing. What leads to high selfing rates and a high variability in selfing rate is that populations and individuals within populations may experience low and highly variable levels of density, often because they are at the margin of the species distribution area. In *Abies alba* on the Mont Ventoux, not only does average selfing rate within a stand increase when stand density decreases, but also individual selfing rate variability increases when local density decreases (i.e. when spatial structure becomes more heterogeneous) as shown by Robledo-Arnuncio et al. (2004) for *Pinus sylvestris*.

Trees from low-density stands may face an uncertain demographic future because they produce both more selfed progeny, subject to inbreeding depression, and more empty seeds, due to pollen limitation and early expression of inbreeding depression, than trees from higher-density stands. However, selfed seeds may be purged early on (at the embryo stage, reviewed by Williams and Savolainen 1996) and the remaining seeds may produce seedlings with better fitness, possibly because of counter-selection of individuals with a higher genetic load (Pichot et al. 2006). This is just what we observed for *Abies alba* in the Mont Ventoux. Low-density stands (the younger stands) produced fewer viable seeds than high-density (mature) stands, but their fitness, as measured by seedling survival after 4 years in the nursery, was higher, especially under stressful conditions. Also, the fitness reduction at low density caused by inbreeding depression and pollen limitation could be compensated by the demographic advantage of being in an open area where density dependence is weaker (e.g. fate of a gene from the colonization front, Edmonds et al. 2004).

Overall, the pattern observed here is typical of competing self-fertilization (Lloyd 1979), where self-pollen competes with outcross-pollen for pollination, leading to selfing rates governed by a 'mass-action law' (Holsinger 1991): selfing rate is proportional to the ratio of the amount of self-pollen over the total amount of pollen received by stigmas. As a consequence, individuals in low density stands might have a low outcrossing rate because their self-pollen is less diluted within the overall outcrossed pollen. The variability in the mating system in our study is due to obligate selfing in geitonogamous organisms, but we can wonder why such a mating system is conserved if pollen limitation does not occur, since it is costly in terms of fitness (Porcher and Lande 2005). One explanation could be the position of flowers on each tree which may decrease selfing rate because female flowers are located at the top of the crown and exposed to wind-dispersed pollen, whereas male flowers are much below, at the bottom of the crown. On steep slopes, such as those where mountain conifers grow, they are thus much closer to the female flowers of a neighboring seed tree than to their own female flowers. When density decreases, the benefit of selfing as a form of reproductive assurance is thus still possible.

We suggest that the mixed mating system has been conserved as a consequence of the contrasted environments where plants grow, particularly in Mediterranean environments. In such heterogeneous environments, a tradeoff between reproductive assurance, avoidance of inbreeding depression and possible pollen discounting may be selected at different times and places, leading to the maintenance of mixed mating as a viable reproductive strategy.

Acknowledgments – We wish to thank Patrice Brahic and the staff at the State Forest Nursery of Aix en Provence for preparing, installing and measuring the nursery test. This research was made possible by the grant 'Conservation des Ressources Génétiques Forestières' from the French Ministry of Agriculture to INRA. We are very grateful to Julia Fady-Welterlen for her English revision of the whole manuscript.

References

Blondel, J. and Aronson, J. 1999. Biology and wildlife of the Mediterranean region. – Oxford Univ. Press.

Bower, A. D. and Aitken, S. N. 2007. Mating system and inbreeding depression in whitebark pine (*Pinus albicaulis* Engelm.). – Tree Genet. Gen. 3: 379–388.

Boyle, T. J. B. et al. 1991. Genetic studies in a tropical pine, *Pinus kesiya*: III. The mating system in four populations from northern Thailand. – J. Trop. For. Sci. 4: 37–44.

Boyle, T. J. B. and Morgenstern, E. K. 1986. Estimates of outcrossing rates in six populations of black spruce in central New Brunswick. – Silvae Genet. 35: 102–106.

Boscherini, G. et al. 1993. Mating system analysis in two Italian populations of Norway spruce. – J. Genet. Breed. 47: 45–48.
Burczyk, J. et al. 1996. Mating patterns and pollen dispersal in a natural knob-cone pine (*Pinus attenuata* Lemmon.) stand. – Heredity 77: 251–260.
Burczyk, J. 1998. Mating system variation in a Scots pine clonal seed orchard. – Silvae Genet. 47: 155–158.
Changtragoon, S. and Finkeldey, R. 1995. Patterns of genetic variation and characterization of the mating system of *Pinus merkusii* in Thailand. – For. Genet. 2: 87–97.
Cheptou, P. O. and Dieckmann, U. 2002. The evolution of self-fertilization in density-regulated populations. – Proc. R. Soc. Lond. B 269: 1177–1186.
Cheptou, P. O. and Schoen, D. J. 2007. Combining population genetics and demographical approaches in evolutionary studies of plant mating systems. – Oikos 116: 271–279.
Del Castillo, R. F. and Trujillo, S. 2008. Effect of inbreeding depression on outcrossing rates among populations of a tropical pine. – New Phytol. 177: 517–524.
De-Lucas, A. I. et al. 2008. Mating system and pollen gene flow in Mediterranean maritime pine. – Heredity 100: 390–399.
DiMichele, W. A. et al. 2001. Response of late Carboniferous and early Permian plant communities to climate change. – Annu. Rev. Earth Plan. Sci. 29: 461–487.
Edmonds, C. A. et al. 2004. Mutations arising in the wave front of an expanding population. – Proc. Natl Acad. Sci. USA 101: 975–979.
El-Kassaby, Y. A. et al. 2003. Genetic evaluation of alternative silvicultural systems in coastal montane forests: western hemlock and amabilis fir. – Theor. Appl. Genet. 107: 598–610.
Fady, B. et al. 1999. Genetic differentiation in *Abies alba* populations from south-eastern France. – For. Genet. 6: 129–138.
Fady, B. and Médail, F. 2004. Mediterranean forest ecosystems. – In: Burley, J. et al. (eds), Encyclopedia of forest science. Elsevier, pp. 1403–1414.
Fady, B. and Westfall, R. D. 1997. Mating system parameters in a natural population of *Abies borisii regis* Mattfeld. – Ann. For. Sci. 54: 643–647.
Fisher, R. A. 1941. Average excess and average effect of a gene substitution. – Ann. Eugenics 11:53–63.
Fu, Y. B. et al. 1992. Pollen pool heterogeneity in Jack pine (*Pinus banksiana* Lamb) – a problem for estimating outcrossing rates. – Theor. Appl. Genet. 83: 500–508.
Furnier, G. R. and Adams, W. T. 1986. Mating system in natural populations of Jeffrey pine. – Am. J. Bot. 73: 1002–1008.
Goodwillie, C. et al. 2005. The evolutionary enigma of mixed mating systems in plants: occurrence, theoretical explanations, and empirical evidence. – Annu. Rev. Ecol. Evol. Syst. 36: 47–79.
Griffin, A. R. and Sedgley, M. 1989. Sexual reproduction of tree crops. – London Acad. Press.
Holsinger, K. E. 1991. Mass-action models of plant mating systems: the evolutionary stability of mixed mating systems. – Am. Nat. 138: 606–622.
Innes, D. J. and Ringius, G. G. 1990. Mating system and genetic structure of two populations of white spruce (*Picea glauca*) in eastern Newfoundland. – Can. J. Bot. 68: 1661–1666.
Kalisz, S. et al. 2004. Context-dependent autonomous self-fertilization yields reproductive assurance and mixed mating. – Nature 430: 884–887.
Kaya, N. et al. 2006. Mating system and pollen contamination in a *Pinus brutia* seed orchard. – New For. 31: 409–416.
Knight, T. M. et al. 2005. Pollen limitation of plant reproduction: pattern and process. – Annu. Rev. Ecol. Evol. Syst. 36: 467–497.
Krutovskii, K. V. et al. 1995. Isozyme study of population genetic structure, mating system and phylogenetic relationships of the five stone pine species (subsection Cembrae, section Strobi, subgenus Strobus). – In: Baradat, P. et al. (eds), Population genetics and conservation of forest trees. SPB Acad. Publ., pp. 280–304.
Kuittinen, H. and Savolainen, O. 1992. *Picea omorika* is a self fertile but outcrossing conifer. – Heredity 68: 183–187.
Lande, R. and Schemske, D. W. 1985. The evolution of self-fertilization and inbreeding depression in plants. I. Genetic models. – Evolution 39: 24–40.
Ledig, F. T. et al. 1997. Recent evolution and divergence among populations of a rare Mexican endemic, Chihuahua spruce, following Holocene climatic warming. – Evolution 51: 1815–1827.
Ledig, F. T. et al. 1999. Evidence for an extreme bottleneck in a rare Mexican pinyon: genetic diversity, disequilibrium and the mating system in *Pinus maximartinezii*. – Evolution 53: 91–99.
Ledig, F. T. et al. 2000. The mating system and genetic diversity in Martinez spruce, an extremely rare endemic of Mexico's Sierra Madre Oriental: an example of facultative selfing and survival in interglacial refugia. – Can. J. For. Res. 30: 1156–1164.
Ledig, F. T. et al. 2001. Genetic diversity and the mating system of a rare Mexican pinon, *Pinus pinceana*, and a comparison with *Pinus maximartinezii* (Pinaceae). – Am. J. Bot. 88: 1977–1987.
Ledig, F. T. et al. 2002. Genetic diversity, mating system and conservation of a Mexican subalpine relict *Picea mexicana*. – Conserv. Genet. 3: 113–122.
Ledig, F. T. et al. 2005. Genic diversity, genetic structure, and mating system of Brewer spruce (Pinaceae), a relict of the Arcto-Tertiary forest. – Am. J. Bot. 92: 1975–1986.
Lewandowski, A. and Burczyk, J. 2000. Mating system and genetic diversity in natural populations of European larch (*Larix decidua*) and stone pine (*Pinus cembra*) located at higher elevations. – Silvae Genet. 49: 158–161.
Lloyd, D. G. 1979. Some reproductive factors affecting the selection of self-fertilization in plants. – Am. Nat. 113: 67–79.
Mimura, M. and Aitken, S. N. 2007. Increased selfing and decreased effective pollen donor number in peripheral relative to central populations in *Picea sitchensis* (Pinaceae). – Am. J. Bot. 94: 991–998.
Mitton, J. B. 1992. The dynamic mating systems of conifers. – New For. 6: 197–216.
Mitton, J. B. et al. 1997. The pattern of inbreeding in Washoe pine and survival of inbred progeny under optimal environmental conditions. – Silvae Genet. 46: 215–219.
Morgante, M. et al. 1991. Effects of stand density on outcrossing rate in two Norway spruce (*Picea abies*) populations. – Can. J. Bot. 69: 2704–2708.
Morgante, M. et al. 1993. Selection against inbreds in early life-cycle phases in *Pinus leucodermis* – Ant. Heredety 70: 622–627.
Neale, D. B. and Adams, W. T. 1985. Allozyme and mating system variation in balsam fir (*Abies balsamea*) across a continuous elevational transect. – Can. J. Bot. 63: 2448–2453.

O'Connell, L. M. et al. 2006. Impacts of forest fragmentation on the mating system and genetic diversity of white spruce (*Picea glauca*) at the landscape level. – Heredity 97: 418–426.

Parraguirre-Lezama, C. et al. 2004. Mating system in four natural populations of *Pinus greggii* Engelm. – Agrociencia 38: 107–119.

Perry, D. J. and Dancik, B. P. 1986. Mating system dynamics of lodgepole pine in Alberta, Canada. – Silvae Genet. 35: 190–195.

Pichot, C. et al. 2006. Déterminants et conséquences de la qualité génétique des graines et semis lors de la phase initiale de régénération naturelle des peuplements forestiers. – Les actes du BRG 6: 277–297.

Porcher, E. and Lande, R. 2005. The evolution of self-fertilization and inbreeding depression under pollen discounting and pollen limitation. – J. Evol. Biol. 18: 497–508.

Quézel, P. and Médail, F. 2003. Ecologie et biogéographie des forêts du bassin méditerranéen. – Elsevier.

Rajora, O. P. et al. 2000. Indicators of population viability in red spruce, *Picea rubens*. II. Genetic diversity, population structure, and mating behavior. – Can. J. Bot. 78: 941–956.

Rajora, O. P. et al. 2002. Mating system and reproductive fitness traits of eastern white pine (*Pinus strobus*) in large, central versus small, isolated, marginal populations. – Can. J. Bot. 80: 1173–1184.

Ritland, K. 2002. Extensions of models for the estimation of mating systems using n independent loci. – Heredity 88: 221–228.

Robledo-Arnuncio, J. J. et al. 2004. Increased selfing and correlated paternity in a small population of a predominantly outcrossing conifer, *Pinus sylvestris*. – Mol. Ecol. 13: 2567–2577.

Sagnard, F. 2001. Dynamique de recolonisation des pinèdes pionnières par la hêtraie-sapinière: Etude de la structure des peuplements forestiers et de l'évolution génétique des populations de sapin pectiné (*Abies alba* Mill.) sur le Mont Ventoux. PhD thesis. – Paul Cézanne Univ., Aix-Marseille III, France.

Schroeder, S. 1989. Outcrossing rates and seed characteristics in damaged natural populations of *Abies alba* Mill. – Silvae Genet. 38: 185–189.

Shea, K. L. 1987. Effects of population structure and cone production on outcrossing rates in Engelmann spruce and subalpine fir. – Evolution 41: 124–136.

Siegismund, H. R. and Kjaer, E. D. 1997. Outcrossing rates in two stands of noble fir (*Abies procera* Rehd.) in Denmark. – Silvae Genet. 46: 144–146.

Siregar, I. Z. and Hattemer, H. H. 2001. Gene flow and mating system in a seedling seed orchard and a natural stand of *Pinus merkusii* (Jungh. et de Vriese) in Indonesia. – In: Muller-Starck, G. and Schubert, R. (eds), Genetic response of forest systems to changing environmental conditions. Kluwer, pp. 281–292.

Sorensen, F. C. 1994. Frequency of seedlings from natural self-fertilization in Pacific Northwest ponderosa pine (*Pinus ponderosa* Dougl. ex Laws.). – Silvae Genet. 43: 100–108.

Sproule, A. T. and Dancik, B. P. 1996. The mating system of black spruce in north-central Alberta, Canada. – Silvae Genet. 45: 159–164.

Vogl, C. et al. 2002. High resolution analysis of mating systems: inbreeding in natural populations of *Pinus radiata*. – J. Evol. Biol. 15: 433–439.

Williams, C. G. and Savolainen, O. 1996. Inbreeding depression in conifers: Implications for breeding strategy. – For. Sci. 42: 102–117.

Chapitre 3

Variation de fertilité et conséquence sur le succès reproducteur femelle

1 Synthèse

La production de graines est une composante essentielle du succès reproducteur femelle, et donc une production en quantité suffisante est nécessaire au maintien et/ou à la croissance des populations (Clark et al. 2007). Or chez les conifères la production de graines vides (i.e. ne comportant pas d'embryons viables) est un phénomène courant (Burczyk et al. 1991, Kormutak et Lindgren 1996). Ainsi la seule production de graines n'est pas une information suffisante chez les conifères, il est également nécessaire d'estimer la production efficace de graines, c'est-à-dire la production de graines viables. Plusieurs mécanismes peuvent modifier la production de graines et notamment la quantité et la qualité génétique du pollen (Aizen et Harder 2007). En effet la limitation pollinique est un phénomène répandu chez les plantes qui conduit à l'avortement des ovules non fécondés, ce qui chez les conifères se traduit par des graines vides (Knight et al. 2005, Ashman et al. 2004). D'autre part, la dépression de consanguinité est également commune chez les conifères et se traduit souvent par un avortement précoce de l'embryon qui conduit à la formation d'une graine vide (Kormutak et Lindgren 1996, Williams et Savolainen 1996). Nous avons donc choisi d'étudier, à l'échelle individuelle, l'effet de la densité sur la production de graines vides. En effet, une diminution de la densité locale peut potentiellement

réduire la production efficace de graines en modifiant la qualité et la quantité de pollen reçu :

- Par dépression de consanguinité, en augmentant le taux d'autofécondation par loi d'action de masse.
- Par limitation pollinique en diminuant le nombre de contributeurs au nuage de pollen fécondant.

Nous avons étudié ces deux facteurs au sein de nos deux populations marginales de sapin pectiné : Le Mont Ventoux et la montagne de Lure. Afin de comprendre si les effets observés pouvaient varier dans le temps, nous avons suivi ces populations pendant une durée de 4 ans.

Tout d'abord nous avons estimé les taux d'autofécondation individuels des arbres échantillonnés. Nous avons donc spécialement développée une méthode que pour traiter les données obtenues à partir de marqueurs génétiques hérités paternellement et en tenant compte des données manquantes. Nous avons trouvé une corrélation négative et significative du taux d'autofécondation avec la densité locale en arbres producteurs de pollen ($p = 0.03$) et une corrélation positive avec la production individuelle de pollen ($p = 6.10^{-4}$). Ceci confirme l'hypothèse émise dans le second chapitre selon laquelle le régime de reproduction serait régi par une loi d'action de masse.

Nous avons ensuite étudié la variabilité de la proportion de graines vides. La production efficace de graines n'est pas affectée par la production de pollen ($p = 0.35$) ni par le taux d'autofécondation. En revanche nous avons trouvé un très fort effet de l'année ($p < 10^{-4}$), sans doute expliqué par l'année 2003, présentant une proportion de graines vides très élevée. Nous avons également trouvé un effet négatif et significatif de la densité locale sur la production efficace de graines ($p = 0,045$).

Nous avons donc conclu que la production efficace de graines était principalement influencée par la quantité de pollen reçue. Ce résultat est assez étonnant car même si les densités locales peuvent être très variables, la densité globale est relativement élevée. De plus la dispersion du pollen des conifères s'effectue en général sur des

distances relativement grandes. Nous ne nous attendions donc pas à observer une limitation pollinique dans notre population à une si faible échelle spatiale. De plus nous n'avons pas détecté d'expression de la dépression de consanguinité au stade graine, là encore cela est inattendu puisque ce phénomène est connu pour largement expliquer la présence de graines vides chez les conifères (Williams et Savolainen 1996). Ainsi la variabilité de l'environnement (densité et climat) dans ce cas semble expliquer à elle seule la production de graines vides, et la dépression de consanguinité au contraire ne semble pas, au stade graine, avoir d'influence sur le succès reproducteur femelle.

2 Variability of female reproductive success in the conifer tree *Abies alba* : pollen quantity, pollen quality and climate

Cet article sera prochainement soumis à Heredity.

Variability of female reproductive success in the conifer tree *Abies alba*: pollen quantity, pollen quality and climate

Gwendal Restoux[1,3,*], Catherine Bonnet[1], Christian Pichot[1] ,
Giovanni G. Vendramin[2], Etienne K. Klein[3] & Bruno Fady[1]

Address:

1 INRA, UR 629, Ecologie des Forêts Méditerranéennes, 84914 Avignon, France

2 Istituto di Genetica Vegetale, CNR,Via Madonna del Piano 10, 50019 Sesto Fiorentino, Firenze, Italy

3 INRA, UR 546, Biostatistique et Processus Spatiaux, 84914 Avignon, France

* corresponding author : gwendal.restoux@avignon.inra.fr

Abstract

Female reproductive success is one of the components of fitness and thus of population persistence. To better understand forest dynamics for both conservation and evolutionary purposes, factors affecting fitness, and female reproductive success in particular, have to be known. This is what we proposed to examine in this study of silver fir, (*Abies alba* Mill.). Production of empty seed, selfing rate, pollen production, and local density were monitored during four years for 146 silver firs on two mountains in southern France. Selfing rate was found to be environment dependent and driven by a trade-off between local density (outcross pollen quantity) and individual pollen production (self pollen quantity), typical of the mass action law. The proportion of empty seed within a progeny was influenced by pollen quantity emitted in the neighborhood, indicating pollen limitation, and not by pollen quality, defined here as the proportion of self-fertilization. Our results suggest that inbreeding depression is either weak or undetectable, refuting several recent studies suggesting that inbreeding depression is mainly responsible for the production of empty seeds. Density, past history and year to year variations influence the production of viable seeds. These environmental factors largely explain the variability of viable-seed production.

Keywords: Seed-set; Mixed mating; Pollen limitation; Pollen quality; Local density; Inbreeding depression; Selfing; Microsatellites; Climate

Introduction

Individual fitness can be defined as the number of zygotes produced by a zygote of the previous generation or similarly by survival of the progeny until sexual maturity and its reproductive success. For hermaphroditic plant species, individual reproductive success is the sum of male reproductive success, i.e. pollen grains fertilizing ovules, and female reproductive success, i.e. ovules to be fertilized. Both reproductive successes are achieved when fertilization results in the formation of viable seeds. Ecological and biological conservation studies usually measure the number of seeds produced per individual (i.e. seed-set) as an estimate of reproductive success (Aigner 2004, Hirayama et al. 2005, Ledig et al. 2006). However, presence of empty or nonviable seeds is widespread, particularly in conifers (Lindgren 1975, Williams and Savolainen 1996, Williams 2008). An effective seed-set measure consisting in removing empty seeds from the total seed count would be more appropriate than standard seed-set to estimate reproductive success in the presence of empty seeds (Kormutak and Lindgren 1996, Ortiz et al. 1998, Wang and Smith 2002, Burczyk et al. 1991).

Many species in natural populations experience pollen limitation, i.e. seed-set is limited by the quantity of pollen available (Burd 1994, Knight et al. 2005). In this case, isolated individuals receive less pollen than individuals in denser populations, leading to a reduced seed-set (or effective seed-set). Historically, pollen limitation was tested by pollen supplementation experiments in which seed production reached a maximum value, indicating the number of potentially fertilizable ovules. The importance of pollen limitation was directly related to the difference between experimental maximum seed production and natural seed production (Ashman et al. 2004, Aizen and Harder 2007). The concept of pollen limita-

tion was enhanced (Haig and Westoby 1988, Burd 2008) by incorporating the influences of both pollen availability and female resources used to enhance pollination. This hypothesis is more appropriate to entomophilous species, for which many ressources have to be dedicated to attracting pollinators (Burd 2008), but is also suitable for wind-pollinated species when nutrient ressources limit seed development (Allison 1990).

Most approaches used to investigate pollen limitation deal with the quantity of available pollen and not its genetic quality. However, the importance of pollen quality in pollen limitation was discussed early on through the effect of self-incompatibility mechanisms that reduce seed-set. These studies concluded that self-incompatible species are more prone to pollen limitation than self-compatible ones (Burd 1994, Agren 1996, Ashman et al. 2004, Knight et al. 2005). Pollen genetic quality must also be considered from a wider perspective since early inbreeding depression resulting from mating between genetically related individuals can lead to a decrease in seed production through abortion of embryos (Kormutak and Lindgren 1996, Donohue 1998, Gigord et al. 1998, Goodwillie and Knight 2006, Williams 2008). Thus both the quantity and quality of pollen interfere in a pollen limitation context as underlined recently by Aizen and Harder (2007) in their empirical and theoretical study. From an inbreeding depression point of view, differences between low and high quality pollen grains can be accessed by the genetic relatedness between pollen grain and ovule: the higher the relatedness, the lower the quality. Thus a spatial genetic structure and a significant kinship coefficient can increase the relatedness probability between pollen and ovules, especially if mating occurs within a single individual (selfing) or between closely related individuals. The probability of abortion caused by deleterious mutations through inbreeding depression would then increase consequently (Williams and Savolainen 1996).

Density should affect both quality and quantity of pollen received by plants. Increasing density increases pollen quantity available for individuals within a population by increasing the number of potential pollen contributors, leading to an increase of effective seed production. At the same time, increasing density should affect pollen quality by decreasing both the proportion of self-fertilization and the genetic relatedness among progenies (Robledo-Arnuncio et al. 2004), which in turn can limit seed abortion (Aizen and Harder 2007). By studying if both pollen quality and ecological factors significantly modify effective seed production, we tried to identify to what extent pollen quality has to be considered in the pollen limitation framework of Aizen and Harder (2007).

From an evolutionary point of view, studying the effects of selfing on effective seed-set raises the question of what factors maintain a mixed mating system (i.e. combined production of outcrossed and selfed offspring by individual parents) (Goodwillie et al. 2005)? Selfers appear to be at an evolutionary advantage by twice transmitting copies of their (nuclear) genes to the next generation (through male and female functions) as opposed to strictly outcrossers (Fisher 1941, Lande and Schemske 1985, Schemske and Lande 1985). Furthermore, selfing provides reproductive assurance (Cheptou 2004, Latouche-Hallé et al. 2004), i.e. the guarantee for an individual to have its ovules fertilized even in the absence of cross-pollen, which is not possible for strict outcrossers, such as dioecious species. On the other hand self-fertilization would be counterselected because of i) inbreeding depression that can lead to production of offspring with low fitness (Lande and Schemske 1985, Schemske and Lande 1985, Cheptou 2004, Latouche-Hallé et al. 2004) and ii) pollen discounting that consists of a cost to selfing due to the loss of pollen for outcrossing (Holsinger 1991, Harder and Wilson 1998). Despite theoretical studies evaluating the joint

and relative importance of these factors (Maurice and Fleming 1995, Porcher and Lande 2005), few experimental studies attempt to answer this question under natural conditions (Goodwillie et al. 2005).

To understand which factors influence effective seed-set, we monitored male and female production (i.e. pollen and cone) of 146 silver firs (*Abies alba* Mill.) in the southern French Alps during four years. For each individual we measured the local density of conspecifics and harvested seeds that we analysed to estimate the proportion of empty seeds. For some of these trees we also genotyped seeds to estimate their selfing rates. Finally, using statistical analysis we estimated relationships between the proportion of empty seeds and other variables to understand which are the most influential factors on this component of reproductive success.

The aim of this study is to answer the three following questions: (i) What are the respective effects of inbreeding depression and pollen limitation on effective seed production ? (ii) Is effective seed-set only pollen limited or is there any influence of climate in our silver fir population ? (iii) Can we bring some new empirical elements from this study for understanding the maintenance of mixed mating systems ?

Material and methods

Biological material and study site

Silver fir (*Abies alba* Mill.) is a sciaphilous European conifer, distributed thoughout temperate mountain areas, such as the Alps, at moderate elevations between 300 m and 1800 m. This monoecious, wind-pollinated species is self-compatible and has a mixed mating system, although it passively avoids self-fertilization: female strobili (i.e. cones) are located within the top third of the crown, whereas male strobili are located within the bottom two thirds of the crown.

Our study site in southeastern France is divided into two mountainous locations: Lure (summit : 44°7'23.11"N ; 5°48'27.40"E) and Mont Ventoux (summit : 44°10'27.42"N ; 5°16'46.27"E) approximately 40 km apart. On Lure (hereafter noted **L**), silver fir is found at quite high density (mean local desnity =186 trees.ha^{-1}; sd=150.8) in monospecific stands or associated with beech, *Fagus sylvatica*. Silver fir density is highly variable on Lure. On Mont Ventoux (hereafter noted **MV**), *Abies alba* is found in stands of local densities as variable as on Lure (mean=190 trees.ha^{-1}; sd=131.1), but more often associated with beech.

Sampling

A total of 146 mature trees were monitored during four years between 2002 and 2005: 51 trees on Lure and 95 on Mont Ventoux (see table 1). These trees had been chosen in 2002 because they produced cones that year and were easily accessible for harvesting. Each of these trees was mapped using GPS equipment (Trimble© Ranger Pro XR). The number of

female cones was counted each year. Male strobili production was visually evaluated each year using a discrete score between 0 when there was no strobilus and 3 when more than 75% of the branches on the lower two-thirds of the crown were covered by male strobili. Crown area was estimated assuming a pyramidal shape using basal branch length in four directions and crown height. Because male strobili are located at the periphery of the crown, we estimated pollen production for each tree by multiplying its crown area by the discrete pollen production score. Thus, calculations on a short tree and a tall tree, with the same relative score of male strobilus abundance, resulted in different amounts of pollen.

Local density was defined as the number of potential pollen donors within 12.5m of each studied tree. Conspecific neighbors were mapped relative to the central mature tree under study using a compass and a laser telemeter (Zimmer© Vertex IV).

For each tree, 1 to 10 cones, depending on tree production, were harvested during late summer. Harvested cones were dried and their seeds collected and stored at -20°C. To estimate the proportion of empty seeds for each tree for a given reproductive event, 10 to 2887 seeds per tree per year were analyzed using dissection or X-ray scanning (whereby embryos contrast with maternal tissues). To analyze jointly the values from scanned and dissected seeds, we used both methods for 12 progenies in 2003 and 22 in 2004. We then estimated the fraction of false full seeds detected from X-ray scan to be 8% (assumed independent of the mother-tree). We corrected all scan-based empty seed proportions by adding the expected number of false full seeds (number of filled seeds multiplied by the fraction of false-full seeds) to the number of observed empty seeds.

Finally, 10 to 55 seeds per tree for each year were genotyped using 4 chloroplast mi-

crosatellites : Pt71936; Pt63718; Pt30241; Pt30249 (Vendramin et al. 1996, Liepelt et al. 2001). These markers are paternally inherited in conifers and in *Abies alba* in particular (Vendramin and Ziegenhagen 1997). DNA was extracted using the DNEasy Plant Mini Kit (Qiagen©) and was amplified using the Qiagen© multiplex amplification kit on an Eppendorf© Mastercycler EP Gadient and an MJ Research PTC200 Peltier thermal cyclers (following the standard protocols of manufacturers). As the microsatellites used are mononucleotide repeats and thus difficult to score, all amplified genotypes were scored twice using different sequencers: a LiCor© ReadIR 4200 DNA sequencer (INRA Avignon, France) and a Molecular Dynamics MegaBace© 1000 capillary sequencer at the GenExpress lab (CNR Firenze, Italy). All profiles obtained were read at least three times by two different people to avoid mistyping.

Genetic and statistical methods

Estimation of selfing rates

The annual individual selfing rate of 73 mature trees (50 in **L** and 23 in **MV**) was estimated from the haplotypic data obtained from the seeds using a maximum likelihood approach similar to that of Ritland and Jain (1981) and Ritland (2002). The likelihood was maximized using an E.M. algorithm (Cheliak et al. 1983), which is more robust for these estimations than classical Newton-Raphson procedure (Ritland 1986). We adapted this estimation method to the particular genetic markers we used, which were physically linked, haploid, and paternally inherited.

The following equation describes the probability of obtaining a seed with haplotype h_o,

knowing its maternal haplotype h_m. This seed can result from selfing with probability s_m. In this case, the probability to read $h_o = h_m$ is 1 and the probability of reading $h_o \neq h_m$ is 0. This seed can also result from outcrossing with probability $(1 - s_m)$. In this case the probability to read h_o is $f(h_o)$, the frequency of haplotype h_o in the background pollen pool, assumed similar for all trees considered for computation. The likelihood for this seed is

$$L(s_m, f) = P(h_o|h_m) = s_m \mathbb{1}_{\{h_o=h_m\}} + (1 - s_m) f(h_o), \tag{1}$$

where $\mathbb{1}_{\{h_o=h_m\}}$ has value 1 if $h_o = h_m$ and 0 otherwise. The algorithm maximises the sum of the log-likelihoods for all the seeds to estimate both the selfing rates s_m for all seed-trees and the frequency of each haplotype $f(h_o)$ in the background pollen pool,

$$log(L(s, f)) = \sum_m \sum_o log(P(h_o|h_m)) = \sum_m \sum_o log(s_m \mathbb{1}_{\{h_o=h_m\}} + (1 - s_m) f(h_o)), \tag{2}$$

where the sum is over all maternal-trees m and all seeds o from each maternal-tree.

The E.M. algorithm is an iterative procedure. All selfing rates and all haplotypes frequencies were first initialized at 0.1 and (number of different haplotypes)$^{-1}$, respectively. Then, for each iteration t, for which intermediate s_m and $f(h)$ are denoted s_m^t and $f(h)^t$, we computed for all mothers m:

$$s_m^{t+1} = \sum_o \frac{s_m^t \mathbb{1}_{\{h_o=h_m\}}}{N_s(m)[s_m^t \mathbb{1}_{\{h_o=h_m\}} + (1 - s_m^t) f(h_o)^t]}, \tag{3}$$

where $N_s(m)$ is the number of seeds analyzed for mother m.

We computed for all haplotypes h

$$f'(h)^t = \sum_m \sum_o \frac{(1-s_m^t)f(h)^t}{s_m^t \mathbb{1}_{\{h_o=h_m\}} + (1-s_m^t)f(h)^t} \mathbb{1}_{\{h_o=h\}}, \tag{4}$$

and

$$f(h)^{t+1} = \frac{f'(h)^t}{\sum_h f'(h)^t}. \tag{5}$$

The iterations continued until the log-likelihood difference between consecutive iterations was $< 10^{-4}$. Haplotypic frequencies were estimated separately for the **C**, **MS** and **L** populations, since background pollen pools were expected to differ over a few kilometers. Finally, missing data at some loci for some seeds were taken into account by distributing the contribution of incomplete haplotypes to all matching complete haplotypes relative to their estimated frequencies in the background pollen pool. All computations were conducted using R (R Development Core Team 2007). The source code is available upon request.

Individual selfing rates were estimated on a stratified sample including on average 20 germinated and 20 non-germinated seeds for 12 progenies in 2003 and 22 in 2004. The global selfing rate for each of these mother trees was then obtained by weighting selfing rates from its germinated and non-germinated seeds by the proportion of germinated and non-germinated seeds in the offspring estimated by germination tests (further details in Pichot et al. (2006)).

Statistical analysis of results

We investigated variation in cone production among individuals (qualitative), years (qualitative, 4 levels), sites (qualitative, 2 levels) and according individual pollen production (quantitative) and local density (quantitative) using a linear mixed-effects model. Indi-

viduals were considered as a random effect nested in the year (one variance estimated per year). We also tested for interactions among fixed effects up to the second order. Using a similar linear mixed-effects model, we investigated the individual pollen productionusing the same effects (except the pollen production). Because our measure of pollen production combines estimates of crown area and discrete pollen production score, we also used these two variables in two linear models to test which component was most influenced by year or local density.

We investigated inter-individual variations of selfing rate with a linear mixed-effects model that considered individuals (qualitative), year (qualitative with 4 levels), local density (quantitative), self-pollen production (quantitative), and second-order significant interactions. Individuals were considered as a random effect nested in the year (one variance estimated per year). Because the estimates of individual selfing rate were based on different numbers of seeds, we weighted each observation by the sample size.

The mixed-effects linear model used to assess variation in the proportion of empty seeds considered the fixed effects of year (qualitative with 4 levels), site (qualitative with 2 levels), local density (quantitative) and self-pollen production (quantitative). Individuals were considered as a random effect nested in the year (one variance estimated per year). We weighted each observation by the size of the sample on which the proportion of empty seeds was estimated.

All computations were also made using R (R Development Core Team 2007) and the nlme package (Pinheiro et al. 2008). We used a type III approach (marginal type ANOVA in R) to test fixed effects for all mixed-effects linear models with more than one effect.

Random effects of individual were tested by comparing models with or without this random effect through a likelihood ratio test. We also compared models for which among individuals variance was different according the year, and models for which we estimated only one variance component.

Results

Cone and pollen production varied significantly among individuals, and the among individuals variability is significantly different between the years (model comparision results not shown here). Both cone and pollen production were significantly different among years (tables 2, 3) and both presented a significant interaction effect of the year and the site (tables 1, 2 and 3, and figure 1). Cone production was significantly different between the Lure and the Mont Ventoux with a higher mean production in the latter one. The number of cones and the quantity of pollen produced were significantly correlated ($r = 14\%$, $F_{1,166} = 41.75$, $p < 10^{-4}$).

Crown area was significantly and negatively affected by local density ($F_{1,139} = 13.24$, $p = 4 \times 10^{-4}$). In contrast, the pollen production score was significantly influenced only by year ($F_{3,177} = 216.33$, $p < 10^{-4}$), but not by either local density ($F_{1,138} = 0.63$, $p = 0.43$), sites ($F_{1,138} = 0.28$, $p = 0.60$) or year x local density interaction ($F_{3,177} = 1.58$, $p = 0.20$). Furthermore the yearly variations of pollen production scores were significatly different between the two sites ($F_{3,177} = 11.87$, $p < 10^{-4}$).

Selfing rates differed considerably among individuals and this among individuals vari-

ability differed significantly among years (table 1, results from the models comparisions not shown here). There was no significant site effect thus we did not consider it for further analysis of selfing rate. Selfing rates differed significantly among years (table 4). Selfing rate and local density were significantly and negatively correlated ($r = 8.53 \times 10^3$) and this correlation differed among years (table 4). Selfing rate varied very negatively with an individual's local density and positively with an individual's estimated pollen production (figure 2 and table 4). Pollen production had a positive and significant effect on selfing rate in 2003 which resulted in a significant "year $\times$ pollen production" interaction effect (table 4). Density and pollen production did not have interacting effects on selfing rate (table 4).

The proportion of empty seeds differed significantly among individuals and this variability was significantly different among years (results of models comparisions not shown here). The mean proportions of empty seed were significantly different among years (with a largely higher value in 2003, figure 1 and table 1) and sites with a higher proportion of empty seeds in Lure than in Mont Ventoux (table 5). Neither individual pollen production nor any interaction involving it affected the proportion of empty seeds (table 5).

Discussion

Pollen limitation and mating system

The annual production of pollen varies yearly and is synchronous with annual cone production (figure 1). This synchrony is congruent with results observed in other firs such as

subalpine fir, *Abies lasiocarpa* (Shea 1987). Pollen production varies yearly and is negatively affected by individual tree local density. We can divide this variability between (i) variability of crown size mainly influenced by density, and (ii) variability of the relative discrete flowering score mainly influenced by year. This is due to the growth of the crown which is constrained by competition at high density, leading to a smaller part of the crown holding male strobili, independently of the average "year fecundity". Cone production was mainly influenced by the year and the sites, without any significant influence of density. The effect of the year on cone and pollen production differed among sites. Indeed we observed a large production of cone in 2004 in the **MV** site whereas it was particularly low in the **L** site. Furthermore there was a very pronounced bi-annual alternance of large and low amount of pollen produced in the **MV** site whereas such a pattern was less pronounced in the **L** site.

Individual local density and individual pollen production were found to influence selfing rate. The effect of pollen production was significant in 2003 and 2005, the two years with the lowest annual pollen production. The pattern we observed is typical of the the mass action law (Holsinger 1991): a higher local density implies that self pollen is more diluted into the pollen of neighbors leading to a decrease in selfing rate. Inversely, a higher individual pollen production, relative to pollen exported by other trees, increases the proportion of self-pollen sampled by the tree, and thus increases the level of self-fertilization. We can observe this effect only in 2003 and 2005, because the globally low level of pollen production permitted to increase the effect of self-pollen which was then less diluted into the total pollen cloud received by the seed-tree. This result is consistent with previous results on conifers in Robledo-Arnuncio et al. (2004) where the individual selfing rates ranged from 0% to 92% for Scots pine (*Pinus sylvestris*), and in Mimura and Aitken (2007) where rates

ranged from 2% to 35% at the population level (about the same magnitude at the individual level) for *Picea sitchensis*, depending on stand size and density.

The proportion of empty seeds in a progeny was found to vary between 0% and 99%. This range is broader than values reported by Kormutak and Lindgren (1996) for silver fir, which varied between 31.3% and 88.5%. This difference is probably explained by the broader range of local density that we investigated. Another explanation is that the population we studied has a particular history that could have led to a significant level of inbreeding due to the founder effect (i.e. establishment of a population from a small number of individuals), which could potentially lead to higher inbreeding depression. The silver fir populations of Mt Ventoux and Lure were indeed very small until the French government encouraged mountain forest restoration in the mid 19th century by planting pioneer species over large areas. This was followed by recolonization of *Abies alba* in about 2 generations over extended areas. Nevertheless the size reduction of the population was more pronounced in the **MV** site. This could explain the higher empty seed proportion in the **L** site than in the **MV** site (59% vs 45%), which could be the result of a more efficient but incomplete purge of deleterious mutation load in the former site due to a higher inbreeding. Evidences of rapid purging had been found for silver fir in the Mont Ventoux population by Sagnard (2001). However, the mean proportion of empty seeds in a progeny obtained in this study, 51%, is in accordance with the 55% obtained for European larch (*Larix decidua*) by Burczyk et al. (1991) and with a mean of 61% for a large number of different conifers (Kormutak and Lindgren 1996).

Pollen quantity versus pollen quality

The aim of this study was to assess the relative effects of pollen quantity and pollen quality on the proportion of empty seeds. This question is based on the hypothesis that self-pollenization can induce a fitness decrease leading to a lower seed-set (Aizen and Harder 2007), or in our case to a lower effective seed-set induced by pollen limitation. Unfortunately it is impossible to morphologically distinguish an aborted seed from a non-pollinated ovule for silver fir, unlike some other conifers (e.g. Scots pine where a non-pollinated ovule only leads to the formation of wings without seeds while a pollinated aborted seed is fully developed, (Robledo-Arnuncio et al. 2004)).

We found a significant and negative effect of individual local density on the proportion of empty seeds. As previously assumed, the local density should have an effect on the quantity of pollen received by the tree : individuals in lower density stands experience higher pollen limitation leading to production of more empty seeds than in moderate or high density stands. This result was unexpected since all trees we studied were in high mean density stands. Thus, local density variations can have strong effects on effective reproduction, and so can be relevant at the individual level in a pollen limitation context which is assumed to occur only when population size is very limited. This result highlights the limited pollen dispersal experienced in this case, which is often assumed to be extensive for wind-pollinated species, despite some new evidences suggesting that short-distance pollen dispersal is widespread and can influence seed production (Koenig and Ashley 2003).

We did not find any effect of pollen quality (through individual pollen production) on the proportion of empty seeds within a progeny at population level, neither a pos-

itive relationship indicating inbreeding depression nor a negative one synonymous with reproductive assurance in a pollen limited population. Furthermore we should stress that when only considering the effects of selfing rate, year and site, there was no significant effect of this proportion of self-fertilization ($F_{1,117} = 0.63$, $p = 0.43$) on empty seed proportion (even through interaction effects). This is in contradiction with the results of Kormutak and Lindgren (1996) on the same species and with Wang and Smith (2002) on beech, who found a significant effect of selfing rate on the proportion of empty seed using a one way ANOVA. However, both studies were controlled pollination experiments that compared a completely outcrossed progeny to a completely selfed one. As a consequence the effect of pollen quality was tested by comparing two extreme classes of progenies, while our study was in natural conditions and relied on a linear relationship between selfing rate and proportion of empty seed over a continuous range of selfing rates. Nevertheless our results are in accordance with those found by Burczyk et al. (1991), who did not find any relation between outcrossing rates and proportion of filled seed. To explain these results Burczyk et al. (1991) insisted on the fact that both the inter-individual variability of genetic load and its mean value could affect the proportion of empty seeds. Our results on determinants of empty seed proportion is also compatible with this hypothesis. The model developed by Aizen and Harder (2007) which stated that pollen limitation has to be considered with respect to both pollen quantity and quality is not verified in our study since it did not explain the effective seed-set better than pollen quantity alone.

In our study selfing rates were estimated from filled seeds. This measure can be unadapted to investigate the effect of inbreeding depression on the proportion of empty seeds, because it did not include the selfed seeds that aborted following early inbreeding depression. Since our analyses provided evidences that selfing rate is driven by the mass action

law, we rather considered individual pollen production as a surrogate of selfing rate in this analysis. Furthermore local density is a major component of selfing rate, so when included both effects in the anaylisis it resulted in non-significant effect of neither selfing rate nor local density by counfounding effect (not shown). Thus we used both local density and self-pollen production in analyis to disentangle between quantity and quality effect on the empty seed proportion.

Climate influence

In addition to the quantity of pollen received related to the local density and the quality of pollen received related to individual pollen production, a third determinant of effective seed-set appears to be climate (or at least a yearly effect). In 2003 the proportion of empty seeds reached 60% compared to 48% for the 3 other years combined. On our site as in most of Europe, 2003 was characterized by extremely high temperatures and drought during the summer (Rebetez et al. 2006). Seeds already fertilized (during late spring) were exposed to drought during their maturation, which would have led to a high level of seed abortions (Bennell et al. 2007), i.e. high proportion of empty seeds. Indeed the interaction between year and local density has significant effect on the proportion of empty seed at the 10% risk, with positive effect of local density in 2003 and 2004 (significant in 2004). These were the driest 2 years of the study (data kindly provided by M. Cailleret) with possibly limited resource availability (mainly access to water). Water access was probably lower or at least more heterogeneous between individuals in high density, and it could explain the significant interaction effect found due to these water stressful years. This hypothesis of climate influence (rather than year by itself) was confirmed by observations of effective

seed production for the same trees in 2007, also characterized by very low precipitation and high temperatures (personal observations), that showed an extremely high proportion of empty seeds (about 90%). These results point out another unexpected effect of global warming. It will not only reduce existing populations but it could also decrease the capacity of populations to grow or to maintain themselves, leading to a reduced migration capability which is necessary for escaping from stressful environments.

Evolutionary considerations

Finally we provide here another example of competing selfing, defined as simultaneous self and cross pollination leading to competition between self and cross pollen for fertilization (Leclerc-Potvin and Ritland 1994), which conserves the mixed mating system. Maintenance of such a mating strategy is often assumed to result from a trade-off between the advantage of self fertilization (reproductive assurance) and its drawbacks (inbreeding depression and/or gamete discounting) as reviewed by Goodwillie et al. (2005). In our case we did not find any effect of inbreeding depression at the seed stage since there was no positive relationship between the proportion of empty seeds and either individual pollen production (table 5 and regression coefficient not shown) or selfing rate (not shown). However Pichot et al. (2006) demonstrated a strong effect of inbreeding depression in the same population at the seedling stage. We did not find any evidence of reproductive assurance since we were unable to detect a significant negative relationship between the proportion of empty seed and either individual pollen production (table 5 and regression coefficient not shown) or selfing rate (not shown). Ecological factors alone are sufficient to explain the variations of selfing rate as described by Holsinger (1991) in the mass action law context.

We suggest that, in heterogeneous environments, the capacity to react to different pollination conditions can be conserved (without selection as a consequence of unavoidable geitonogamy). This was pointed out for low to moderate selfing rate populations (averaged selfing rate over the 4 years was 0.25, table 1) by Porcher and Lande (2005) using simulations. However, simulation models provided by these latter authors were made in an adaptative dynamics context and then only tested for the invasion probability of a selfer mutant (with a defined level of selfing) in an outcrosser population. As a consequence these kinds of models did not take into account the effect of among-individual variability of selfing rates, what we observed in natural conditions. We think that such a variability can have an unexpected effect on the predictions made by adaptative dynamics models. Furthermore we cannot exclude an effect of selection through pollen discounting leading to counter-selection of selfing. We observed an avoidance mechanism of selfing (i.e. location of male and female strobili within the crown) but we did not test for pollen discounting (no control of pollen export) (Holsinger 1991). Finally, to answer the question of mixed mating system maintenance, an integrative measure of all components of fitness (i.e. seed, seedling and adult stages) is necessary.

Conclusion

This study shows that inbreeding depression is not expressed in our population at the seed stage or is not strong enough to decrease significantly effective seed production. However, because this experiment was conducted in natural conditions, we cannot test independently the effects of individual genetic and environment components. Although many studies

proved that demographic or genetic considerations alone are not sufficient to understand both seed production and maintenance of a mixed mating system (Cheptou and Schoen 2007), in this particular case we proved that both demography and climate had a stronger effect than genetic factors. As a conclusion, to study female reproductive success and effective seed production in particular, one must consider both the interaction between environments at different scales (i.e. local for density and broader for climate) and the past history of the populations that are under consideration (Byers and Waller 1999).

Acknowledgments

This study is part of Gwendal Restoux's Ph.D. project, co-funded by The Provence Alpes Côtes d'Azur Regional Government and INRA. We would like to thank Daisy Jouaud, Olivier Gilg, Franck Rei and Norbert Turion for data collection and Anne Roig, Bruno Jouaud and Thierry Langlois for genotyping. We acknowledge the financial support of the Ministry of Agriculture National project on forest genetic resources. We also thank Ronald Daigle for his English revision of the manuscript, and Francois Lefevre and Pierre-Olivier Cheptou for their advices on the manuscript.

References

Agren, J. 1996. Population size, pollinator limitation, and seed set in the self-incompatible herb *Lythrum salicaria*. Ecology **77**:1779–1790.

Aigner, P. A. 2004. Ecological and genetic effects on demographic processes : pollination,

clonality and seed production in *Dithyrea maritima*. Biological Conservation **116**:27–34.

Aizen, M. A., and L. D. Harder. 2007. Expanding the limits of the pollen-limitation concept : effects of pollen quantity and quality. Ecology **88**:271–281.

Allison, T. D. 1990. Pollen production and plant density affect pollination and seed production in *Taxus canadensis*. Ecology **71**:516–522.

Ashman, T.-L., T. M. Knight, J. A. Steets, P. Amarasekare, M. Burd, D. R. Campbell, M. R. Dudash, M. O. Johnston, S. J. Mazer, R. J. Mitchell, M. T. Morgan, and W. G. Wilson. 2004. Pollen limitation of plant reproduction: Ecological and evolutionnary causes and consequences. Ecology **85**:2408–2421.

Bennell, M. R., H. Cleugh, J. Leys, and D. Heind. 2007. The effect of hot dry wind on the pod set of faba bean (*Vicia faba*) cv. Fiord: a preliminary wind tunnel study. Australian Journal of Experimental Agriculture **47**:1468–1475.

Burczyk, J., G. Kosiński, and A. Lewandowski. 1991. Mating pattern and empty seed formation in relation to crown level *Larix decidua* Mill. clones. Silva fennica **25**:201–205.

Burd, M. 1994. Bateman principle and plant reproduction - The role of pollen limitation in fruit and seed-set. Botanical Review **60**:83–139.

Burd, M. 2008. The Haig-Westoby model revisited. The American Naturalist **171**:400–404.

Byers, D. L., and D. M. Waller. 1999. Do plant populations purge their genetic load? Effects of population size and mating history on inbreeding depression. Annual Review of Ecology and Systematics **30**:479–513.

Cheliak, W. M., K. Morgan, C. Strobeck, F. C. H. Yeh, and B. P. Dancik. 1983. Estimation

of mating system parameters in plant populations using the EM algorithm. Theoretical and Applied Genetics **65**:157–161.

Cheptou, P.-O. 2004. Allee effect and self-fertilization in hermaphrodites: Reproductive assurance in demographically stable populations. Evolution **58**:2613–2621.

Cheptou, P.-O., and D. J. Schoen. 2007. Combining genetics and demographical approaches in evolutionary studies of plant mating systems. Oïkos **116**:271–279.

Donohue, K. 1998. Effects of inbreeding on traits that influence dispersal and progeny density in *Cakile edentula* var. *Lacustris* (*Brassicaceae*). American Journal of Botany **85**:661–668.

Fisher, R. A. 1941. Average excess and average effect of a gene substitution. Annals of Eugenics **11**:53–63.

Gigord, L., C. Lavigne, and J. A. Shykoff. 1998. Partial self-incompatibility and inbreeding depression in a native tree species of La Reunion (Indian Ocean). Oecologica **117**:342–352.

Goodwillie, C., S. Kalisz, and C. G. Eckert. 2005. The evolutionary enigma of mixed mating systems in plants : Occurence, theoretical explanations, and empirical evidence. Annual Review of Ecology, Evolution and Systematics **36**:47–79.

Goodwillie, C., and M. C. Knight. 2006. Inbreeding depression and mixed mating system in *Leptosiphon jepsonii* : A comparison of three populations. Annals of Botany **98**:351–360.

Haig, D., and M. Westoby. 1988. On limits to seed production. The American Naturalist **131**:757–759.

Harder, L., and W. Wilson. 1998. A Clarification of Pollen Discounting and Its Joint Effects with Inbreeding Depression on Mating System Evolution. The American Naturalist **152**:684–695.

Hirayama, K., K. Ishida, and N. Tomaru. 2005. Effects of pollen shortage and self-pollination on seed production of an endangened tree, *Magnolia stellata*. Annals of Botany **95**:1009–1015.

Holsinger, K. E. 1991. Mass-action models of plant mating systems: The evolutionary stability of mixed mating systems. The American Naturalist **138**:606–622.

Knight, T. M., J. A. Steets, J. C. Vamosi, S. J. Mazer, M. Burd, D. R. Campbell, M. R. Dudash, M. O. Johnston, R. J. Mitchell, and T.-L. Ashman. 2005. Pollen limitation of plant reproduction : Pattern and process. Annual Review of Ecology, Evolution and Systematics **36**:467–497.

Koenig, W. D., and M. V. Ashley. 2003. Is pollen limited ? The answer is blowin' in the wind. Trends in Ecology and Evolution **18**:157–159.

Kormutak, A., and D. Lindgren. 1996. Mating system and empty seeds in silver fir (*Abies alba* Mill.). Forest Genetics **3**:231–235.

Lande, R., and D. W. Schemske. 1985. The evolution of self-fertilization and inbreeding depression in plants. I.Genetic models. Evolution **39**:24–40.

Latouche-Hallé, C., A. Ramboer, E. Bandou, H. Caron, and A. Kremer. 2004. Long-distance pollen flow and tolerance to selfing in a neotropical tree species. Molecular Ecology **13**:1055–1064.

Leclerc-Potvin, C., and K. Ritland. 1994. Modes of self-fertilization in Mimulus guttatus (Scrophulariaceae): A field experiment. American Journal of Botany **81**:199–205.

Ledig, T. F., P. D. Hodgskiss, and D. R. Johnson. 2006. Genetic diverity and seed production in Santa Lucia fir (*Abies bracteata*), a relict of the Miocene Broadleaved Evergreen Forest. Conservation Genetics **7**:383–398.

Liepelt, S., V. Kuhlenkamp, M. Anzidei, G. G. Vendramin, and B. Ziegenhagen. 2001. Pitfalls in determining size homoplasy of microsatellites loci. Molecular Ecology Notes **1**:332–335.

Lindgren, D. 1975. The relationship between self-fertilization, empty seeds and seeds originating from selfing as a consequence of polyembryony. Studia Forestalia Suecica **126**:1–24.

Maurice, S., and T. H. Fleming. 1995. The effect of pollen limitation on plant reproductive systems and the maintenance of sexual polymorphisms. Oïkos **74**:55–60.

Mimura, M., and S. N. Aitken. 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.

Ortiz, P. L., M. Arista, and S. Talavera. 1998. Low reproductive success in two subspecies of *Juniperus oxycedrus* L. International Journal of Plant Science **159**:843–847.

Pichot, C., C. Bastien, F. Courbet, B. Demesure-Musch, P. Dreyfus, B. Fady, N. Frascaria-Lacoste, S. Gerber, F. Lefevre, M.-E. Morand-Prieur, S. Oddou, E. Tessier Du Cros, and A. Valadon. 2006. Déterminants et conséquences de la qualité génétique des graines et semis lors de la phase initiale de régénération naturelle des peuplements forestiers. Les actes du BRG **6**:277–297.

Pinheiro, J., D. Bates, S. DebRoy, D. Sarkar, and the R Core team, 2008. nlme: Linear and Nonlinear Mixed Effects Models.

Porcher, E., and R. Lande. 2005. The evolution of self-fertilization and inbreeding depression under pollen discounting and pollen limitation. Journal of Evolutionary Biology **18**:497–508.

R Development Core Team, 2007. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. URL `http://www.R-project.org`.

Rebetez, M., H. Mayer, O. Dupont, D. Schindler, K. Gartner, J. P. Kropp, and A. Menzel. 2006. Heat and drought 2003 in Europe: a climate synthesis. Annals of Forest Science **63**:569–577.

Ritland, K. 1986. Joint maximum likelihood estimation of genetic and mating structure using open-pollinated progenies. Biometrics **42**:25–43.

Ritland, K. 2002. Extensions of models for the estimation of mating systems using n independent loci. Heredity **88**:221–228.

Ritland, K., and S. Jain. 1981. A model for the estimation of outcrossing rate and gene frequencies using n independent loci. Heredity **47**:35–52.

Robledo-Arnuncio, J.-J., R. Alía, and L. Gil. 2004. Increased selfing and correlated paternity in a small population of predominantly outcrossing conifer, *Pinus sylvestris*. Molecular Ecology **13**:2567–2577.

Sagnard, F., 2001. Dynamique de recolonisation des pinèdes par la hêtraie-sapinière : Etude de la structure des peuplements forestiers et de l'évolution génétique des populations de

sapin pectiné (*Abies alba* Mill.) sur le Mont Ventoux. Ph.D. thesis, Université d'Aix-Marseille III.

Schemske, D. W., and R. Lande. 1985. The evolution of self-fertilization and inbreeding depression in plants. II. Empirical observations. Evolution **39**:41–52.

Shea, K. L. 1987. Effects of population structure and cone-production on outcrossing rates in Engelmann spruce an subalapine fir. Evolution **41**:124–136.

Vendramin, G., and B. Ziegenhagen. 1997. Characterisation and inheritance of polymorphic plastid microsatellites in *Abies*. Genome **40**:857–864.

Vendramin, G. G., L. Lelli, P. Rossi, and M. Morgante. 1996. A set of primers for the amplification of 20 chloroplast microsatellites in *Pinaceae*. Molecular Ecology **5**:595–598.

Wang, B. C., and T. B. Smith. 2002. Closing the seed dispersal loop. Trends in Ecology and Evolution **17**:379–385.

Williams, C. G. 2008. Selfed embryo death in *Pinus taeda*: a phenotypic profile. New Phytologist **178**:210–222.

Williams, C. G., and O. Savolainen. 1996. Inbreeding depression in conifers : Implications for breeding strategy. Forest Science **42**:102–117.

Tables & Figures

Site	Year	2002	2003	2004	2005	All the years
Lure	Cones	**43** ; 50.7 ; [54.42]	**27** ; 28.41 ; [21.39]	**42** ; 14 ; [13.07]	**32** ; 34 ; [32.93]	**144** ; 32.1 ; [37.96]
	Pollen	**42** ; 105.8 ; [124.71]	**26** ; 86.92 ; [99.07]	**41** ; 143.9 ; [143.72]	**31** ; 83.1 ; [94.41]	**140** ; 108.4 ; [121.75]
	Selfing rate	**43** ; 0.29 ; [0.31]	**23** ; 0.2 ; [0.2]	**40** ; 0.33 ; [0.32]	**32** ; 0.15 ; [0.16]	**138** ; 0.26 ; [0.27]
	Empty seeds	**43** ; 0.52 ; [0.29]	**26** ; 0.73 ; [0.24]	**42** ; 0.6 ; [0.2]	**32** ; 0.56 ; [0.27]	**143** ; 0.59 ; [0.26]
Mt Ventoux	Cones	**39** ; 36.77 ; [32.13]	**46** ; 24.57 ; [20.18]	**87** ; 60.32 ; [39.13]	**15** ; 44.13 ; [56.18]	**187** ; 45.32 ; [38.5]
	Pollen	**38** ; 196.1 ; [151.62]	**44** ; 80.85 ; [119.16]	**87** ; 234.6 ; [142.95]	**15** ; 74.01 ; [83.44]	**184** ; 176.8 ; [151.21]
	Selfing rate	**12** ; 0.27 ; [0.32]	**19** ; 0.14 ; [0.2]	**19** ; 0.33 ; [0.24]	**14** ; 0.2 ; [0.27]	**64** ; 0.23 ; [0.26]
	Empty seeds	**39** ; 0.39 ; [0.25]	**46** ; 0.54 ; [0.23]	**88** ; 0.44 ; [0.17]	**15** ; 0.34 ; [0.2]	**188** ; 0.45 ; [0.21]
Both sites	Cones	**82** ; 44.07 ; [45.49]	**73** ; 25.99 ; [20.57]	**129** ; 45.24 ; [39.48]	**47** ; 37.23 ; [41.4]	**331** ; 39.57 ; [38.77]
	Pollen	**80** ; 148.7 ; [144.57]	**70** ; 83.1 ; [111.41]	**128** ; 205.5 ; [148.82]	**46** ; 80.13 ; [90.15]	**324** ; 147.2 ; [143.12]
	Selfing rate	**55** ; 0.29 ; [0.31]	**42** ; 0.17 ; [0.2]	**59** ; 0.33 ; [0.29]	**46** ; 0.17 ; [0.2]	**202** ; 0.25 ; [0.27]
	Empty seeds	**82** ; 0.45 ; [0.28]	**72** ; 0.61 ; [0.25]	**130** ; 0.5 ; [0.19]	**47** ; 0.49 ; [0.27]	**331** ; 0.51 ; [0.25]

Table 1: Summary of the number of cones, pollen production, selfing rate and empty seed proportion at individual level. The values indicated are the sample sizes (bold), mean values (underlined) and standard deviations (between braces) for each year and each site (Lure and Mont Ventoux).

Effects on cone production	Df.1	Df.2	F	Pr(F)
Year	4	166	12.45	$< 10^{-4}$
Site	1	138	6.44	0.01
Local density	1	138	1.23	0.27
Pollen production	1	166	41.75	$< 10^{-4}$
Year×Site	3	166	20.99	$< 10^{-4}$
Year×Local density	3	166	1.8	0.15
Year×Pollen production	3	166	1.03	0.38
Site×Local density	1	138	2.75	0.1
Site×Pollen production	1	166	1.18	0.28
Local density×Pollen production	1	166	4.17	0.04

Table 2: Effects of year, site, local density and pollen production on indivudal cone production. The table presents the results from an ANCOVA on the fixed effects (type III approach) of a mixed-effects model including an additional significant random effect of individuals depending on the year (not shown here).

Effects on pollen production	Df.1	Df.2	F	Pr(F)
Year	4	173	39.21	$< 10^{-4}$
Site	1	138	2.13	0.15
Local density	1	138	7.29	0.01
Year×Site	3	173	9.33	$< 10^{-4}$
Year×Local density	3	173	3.01	0.03
Site×Local density	1	138	0.01	0.9

Table 3: Effects of year, site and local density on individual pollen production. The table presents the results from an ANCOVA on the fixed effects (type III approach) of the corresponding mixed-effects model including an additional significant random effect individuals depending on the year (not shown here).

Effects on selfing rate	Df.1	Df.2	F	Pr(F)
Year	4	115	12.37	$< 10^{-4}$
Local density	1	70	5.02	0.03
Pollen production	1	115	1.88	0.17
Year×Local density	3	115	6.28	6×10^{-4}
Year×Pollen production	3	115	3.5	0.02
Local density×Pollen production	1	115	0.31	0.58

Table 4: Effects of year, local density and pollen production on individual selfing rate. The table presents the results from an ANCOVA on the fixed effects (type III approach) of the corresponding mixed-effects model including an additional significant random effect of individuals depending on the year (not shown here).

Effects on empty seeds proportion	Df.1	Df.2	F	Pr(F)
Year	4	166	86.47	$< 10^{-4}$
Site	1	138	11.22	10^{-3}
Local density	1	138	4.08	4.53×10^{-2}
Pollen production	1	166	0.89	0.35
Year×Site	3	166	0.79	0.50
Year×Local density	3	166	1.51	0.21
Year×Pollen production	3	166	0.45	0.72
Site×Local density	1	138	0.21	0.64
Site×Pollen production	1	166	0.22	0.64
Local density×Pollen production	1	166	1.36	0.24

Table 5: Effects of year, site, local density and pollen production on empty seeds proportion. The table presents the results from an ANCOVA on the fixed effects (type III approach) of the corresponding mixed-effects model including an additional significant random effect of individuals depending on the year (not shown here).

Figure 1: Boxplot of individual cone and pollen production, selfing rate and empty seed proportion for each site and each year.

Figure 2: Selfing rate versus local density. The radius of each circle represents the number of seeds used for the estimation of the selfing rate. The solid line represents the linear regression of selfing rate by local density.

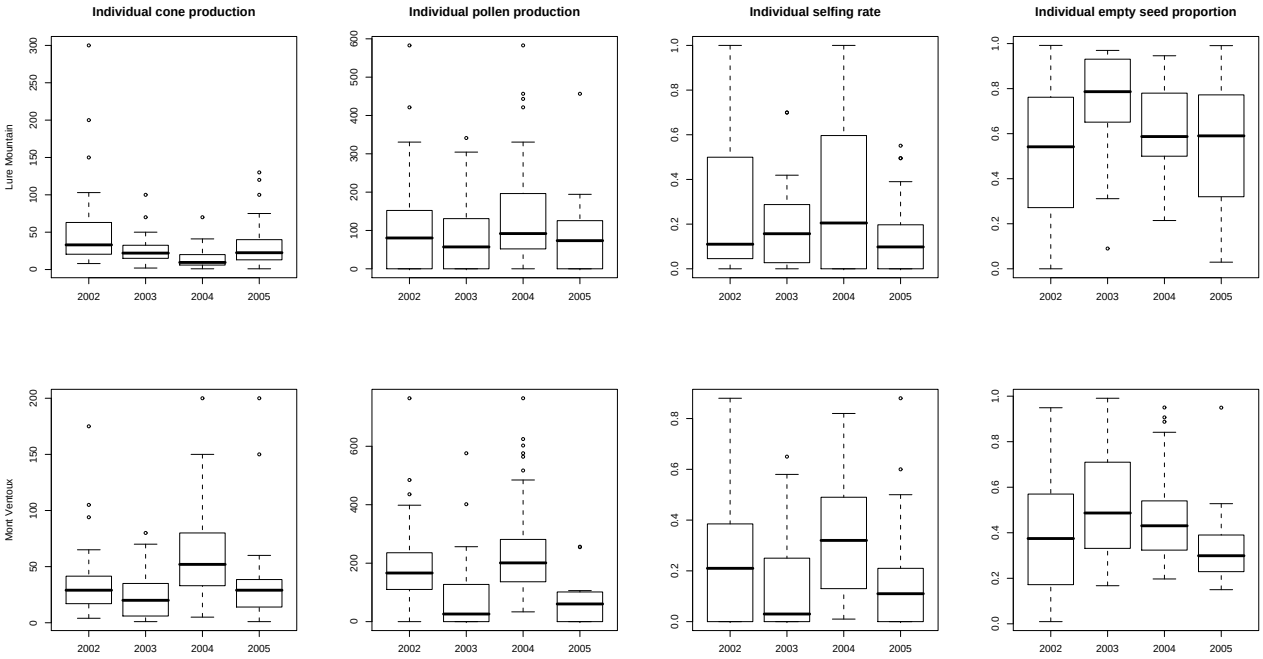

Individual cone production
Individual pollen production
Individual selfing rate
Individual empty seed proportion
Lure Mountain
Mont Ventoux
2002
2003
2004
2005

Figure 1:

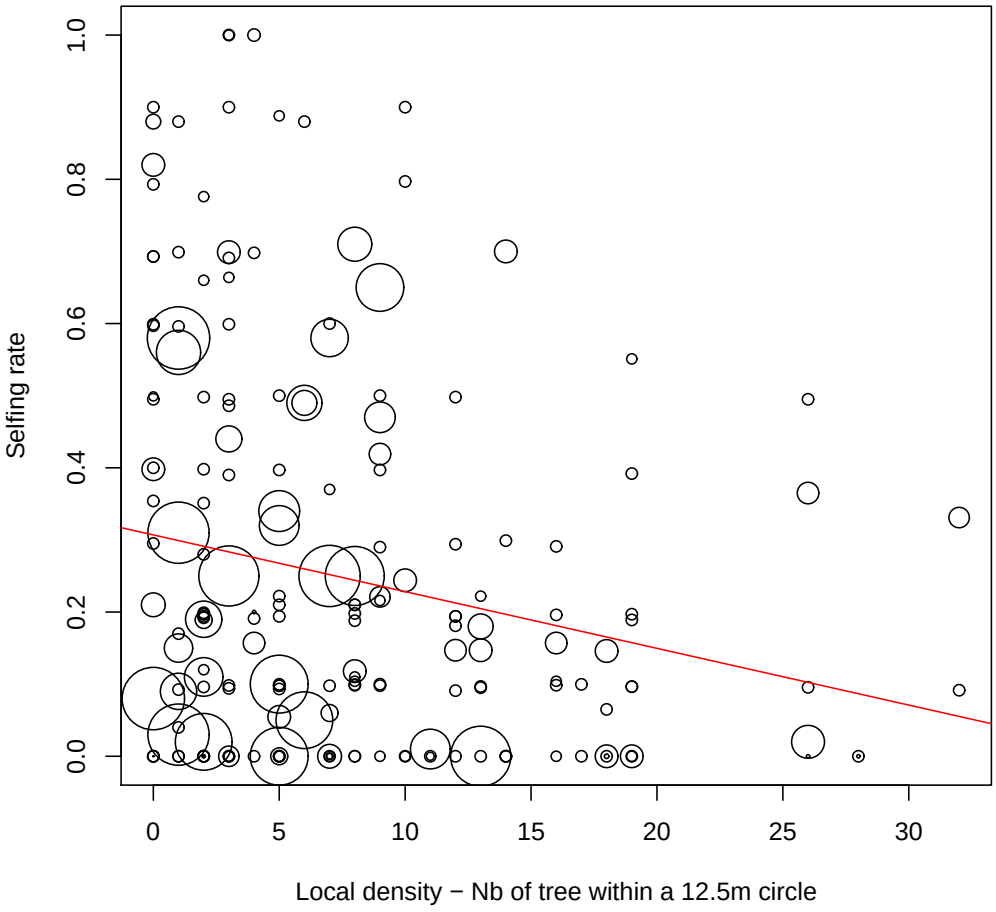

Selfing rate
0.0
0.2
0.4
0.6
0.8
1.0
0
5
10
15
20
25
30
Local density − Nb of tree within a 12.5m circle

Figure 2:

Chapitre 4

Influence de la variabilité inter-individuelle du fardeau génétique sur la détection de la dépression de consanguinité

1 Synthèse

Dans le chapitre précédent nous avons testé la présence de la dépression de consanguinité en effectuant une régression linéaire de la proportion de graines vides par le taux d'autofécondation. N'ayant pas pu mettre en évidence cette dépression, nous avons alors émis l'hypothèse que la méthode pouvait être inadaptée. En effet, si la dépression de consanguinité existe au stade graine et provoque l'avortement des embryons, alors le taux d'autofécondation estimé sur les graines pleines est sans doute sous-estimé par rapport au taux d'autofécondation réel. Nous avons précédemment montré que le régime de reproduction était régi par une loi d'action de masse. Nous avons donc utilisé un proxy du taux d'autofécondation, la production de pollen et la densité locale en donneurs de pollen. Cependant même en utilisant cette méthode nous n'avons pas réussi à détecter d'effet de la dépression de consanguinité au stade graine. Pourtant ce phénomène est relativement fréquent chez les conifères au stade graine et explique souvent la production de graines vides (Sorensen

1969, Charlesworth et Charlesworth 1987, Kormutak et Lindgren 1996, Williams et Savolainen 1996). Afin d'expliquer ce résultat nous avons émis 2 hypothèses :

- Il n'y a pas de dépression de consanguinité.
- Nous n'avons pas réussi à détecter cette dépression de consanguinité.

Cette absence de détection de la dépression de consanguinité pourrait s'expliquer par une variabilité inter-individuelle importante du fardeau génétique.

Afin de tester cette hypothèse et d'un effet de la variabilité du fardeau génétique sur la détection de la dépression de consanguinité, nous avons effectué des simulations. Nous avons simulé des populations définies par un taux d'autofécondation moyen, un fardeau génétique moyen (représenté par un nombre moyen d'équivalents létaux), ainsi que par une variabilité inter-individuelle de ces paramètres décrite par leurs coefficients de variation. A partir de ces valeurs nous avons simulé un nombre N d'individus, pour lesquels nous avons également simulé n descendances, présentant des valeurs sélectives différentes. Nous avons ensuite effectué la même régression linéaire que dans le chapitre 3, c'est-à-dire la régression de la valeur sélective des descendants par le taux d'autofécondation. Nous avons fait varier les 4 caractéristiques de la population et avons également testé plusieurs plans d'échantillonnage qui diffèrent par les nombres et les effectifs des descendances simulés.

La puissance du test de la dépression de consanguinité par régression linéaire est élevé : Sur l'ensemble des simulations nous détectons effectivement de la dépression de consanguinité dans environ 91% des cas (au risque 5%). La puissance de ce test augmente lorsque le fardeau génétique moyen de la population ou le coefficient de variation du taux d'autofécondation sont élevés (figure 2 et tableau 1 de l'article). De plus le coefficient de variation du fardeau génétique est négativement corrélé avec la puissance de ce test (figure 2 et tableau 1). Ainsi la variabilité du fardeau génétique individuel au sein d'une population diminue notre capacité à détecter la dépression de consanguinité. La stratégie d'échantillonnage peut aussi influencer la puissance du test. A effectif constant il est préférable d'augmenter le nombre de descendances échantillonnées plutôt que le nombre d'individus dans chacune de ces descendances.

Ces résultats semblent indiquer qu'il n'y avait effectivement pas d'effet de la dépression de consanguinité au stade graine au sein des populations étudiées dans le chapitre 3 (taux autofécondation moyen $= 0,25$ et coefficient de variation de ce taux $= 100\%$). Dans ce cas la probabilité de détection estimée avec nos simulations varie entre 98 et 100%, avec un écart-type très faible variant entre 10^{-4} et 0.

2 Variability of individual genetic load : consequences for the detection of inbreeding depression through a regression approach

Cette étude a été initiée lors du stage INIP (INitiation à l'Ingénierie de Projet) de Priscille Huot de Longchamp (AgroParisTech) que j'ai encadré au cours de l'année 2008 (10 semaines). Cet article sera prochainement soumis à Genetica.

Variability of individual genetic load: consequences for the detection of inbreeding depression through a regression appoach

Gwendal Restoux[1,2,3], Priscille Huot de Longchamp[1,4],

Etienne Klein[2,5] & Bruno Fady[1,6]

Address:

1 UR 629 Ecologie des Forêts Mediterranéennes, INRA, F-84000 Avignon, France

2 UR 546 Biostatistique et Processus Spatiaux, INRA, F-84000 Avignon, France

3 gwendal.restoux@avignon.inra.fr, corresponding author

4 060102@agroparistech.fr

5 etienne.klein@avignon.inra.fr

6 bruno.fady@avignon.inra.fr

Abstract

Inbreeding depression is a key factor affecting the demography of fragmented populations. Thus measures of inbreeding depression are necessary for conservation biology. In this study on mixed mating species we propose to regress the fitnesses of multiple progenies by the selfing rates of their mothers to infer inbreeding depression at the population level. Using simulations, we investigated how the genetic characteristics of the population, and the individual variability of genetic load in particular, can affect the detection of inbreeding depression when estimated from this method. This method had a large statistical power with about 91% of inbreeding depression detected as significant (at a 5% level). However the variability of individual selfing rate and the mean genetic load of the population had a strong and positive effect on the detection power, while the variability of individual genetic load had a negative effect. However, the sampling design can decrease the effect of the population genetic characteristics, thus assuming a constant sampling effort, we recomand to increase the number of investigated progenies and to decrease the number of individuals within each progeny. Finally we discussed the importance to consider individual variability of genetic load from an evolutionary point of view.

Keywords: Mixed mating system; lethal equivalent; inbreeding depression; genetic load

Introduction

Many plant populations are threatened and suffer from fragmentation leading to size and density decrease, which can affect their survival and persistence. One of the most largely admited cause to expect a relationship between population size and survival, is inbreeding depression (Hedrick (2004) for a review). It results from matings among related individuals and in extreme cases from selfing. In populations with a non-negligible genetic load, higher levels of inbreeding lead to a decrease of fitness through an increase of the number of deleterious homozygotes within progenies. Inbreeding depression is well known and had been studied since the early age of evolution when Darwin noticed some differences in terms of performance and fertility between selfed and outcrossed progeny over many taxa. Inbreeding depression, δ, is expressed as the decrease of fitness of a completely selfed progeny relatively to a completely outcrossed one as follows,

$$\delta = 1 - \frac{\omega_s}{\omega_o}, \tag{1}$$

with ω_s and ω_o the fitnesses of selfed and outcrossed progenies respectively.

From an evolutionary perspective, inbreeding depression is often invoked to explain the evolution of mating system. Indeed, when inbreeding depression occurs, the mating system of an individual can largely influence its fitness since it modifies the proportion of selfed offspring (Jain 1976). Thus, inbreeding depression can influence mating system of population by selectioning for outcrossing (Charlesworth and Charlesworth 1987, Goodwillie et al. 2005, Restoux et al. 2008). However, Lande and Schemske (1985), Schemske and Lande (1985) stated that inbreeding depression is not constant over time, and that genetic load can be purged and then lead to two stable equilibria : strict outcrossing if $\delta > 0.5$ or strict selfing if $\delta < 0.5$ with δ as the initial inbreeding depression. These studies

highlight the tight relationship between mating system evolution and inbreeding depression (Charlesworth and Charlesworth 1987) and emphasize the interest of inbreeding depression study.

Inbreeding depression can be estimated at different level, from individuals to family or populations. To estimate inbreeding depression at the individual level, controled pollination experiments are needed. They enable to estimate genetic load of individuals through the estimation of the number of lethal equivalents (Sorensen 1969, Lindgren 1975, Williams and Savolainen 1996). Lethal equivalent is defined as the number of lethal genes, carried at the heterozygous state within an individual, that would produce the same decrease of fitness in a selfed progeny as the one observed following a selfing experiment. This measure of genetic load does not ignore that genetic load is more often caused by numerous deleterious mutations with weak effects than by few ones with strong effects (Charlesworth and Charlesworth 1987, Husband and Schemske 1996). However it enables to compare genetic load of different individuals even if it implies different genes. This measure is mostly used within breeding programs where few individuals or genotypes (i.e. frequent use of clones) are considered, and thus controled pollination can be conducted (Williams and Savolainen 1996). However, individual estimates of genetic load can be unadapted when studies deal with large natural populations, because (i) it can be too costly to analyse all individuals, and (ii), relative positions of the individuals can strongly affect the estimations of the genetic load and of the inbreeding depression at the population level. When estimated at the population or family level, inbreeding depression is classically measured from different individuals by relating an estimate of their fitness (or a component of it) to their level of inbreeding. Thus good estimates of fitness, ideally over the whole life cycle (Husband and Schemske 1996) and inbreeding are needed. Inbreeding can be estimated by the inbreeding

coefficient, F, which is obtained through either i) correlation of gametes identity estimated from markers (Weir and Cockerham 1984, Robertson and Hill 1984), or ii) the parentage coefficient obtained from pedigree, mainly in the cases of breeding programs. In those cases, the empirical way to describe inbreeding depression is to use linear regression as follow,

$$Y = a + b.F + \epsilon \tag{2}$$

with Y the phenotypic measure, a the intercept, b the regression coefficient, F the inbreeding coefficient and ϵ the residual error term (Williams and Savolainen 1996). Another way to measure inbreeding depression is to use the selfing rates of the parent generation instead of the inbreeding coefficients since these two measures are related (even if it implicitely assumes that the inbreeding coefficient of the parents is constant over all individuals) and the fitnesses of the resulting offsprings. Pollination experiments are a particular case of this method, for which one compares the fitnesses of a strictly selfed progeny to the one of a strictly outcrossed progeny. Selfing rate regression method is mainly used to estimate early inbreeding depression (seed production level) in forest trees (Burczyk et al. 1991, Kormutak and Lindgren 1996, Wang 2003, Restoux et al.).

In conifers, inbreeding depression is assumed to be widespread (Williams and Savolainen 1996) since they are mostly outcrossers and thus purge of deleterious recessive mutations is low (Charlesworth and Charlesworth 1987, Husband and Schemske 1996). Inbreeding depression is then often early expressed through the formation of empty seeds (Charlesworth and Charlesworth 1987, Husband and Schemske 1996, Kärkkäinen et al. 1996) following an early abortion of embryos since no evidence of self-incompatibility mechanisms was found in this order (Sorensen 1969, Williams 2008). Seed production is also affected by inbreed-

ing depression in angiosperms leading to reduced seed-set (Husband and Schemske 1996, Aizen and Harder 2007) but rarely to the production of empty seeds (Wang 2003). Moreover, angiosperms are proner to self-incompatibility and thus it is difficult to distinguish an early abortion due to inbreeding depression from an abortion resulting from cytoplasmic or genetic incompatibilities.

However in some cases, one cannot detect any effect of inbreeding depression in natural populations through a linear relationship between empty seed proportion and the selfing rate, even if it was supposed to occur (Burczyk et al. 1991, Restoux et al.). In this latter study, we did not found any effect of inbreeding depression at the seed stage although occurs at the seedling stage (Restoux et al. 2008). Inbreeding was expected to be high since the studied populations of silver fir, *Abies alba* Miller, results from a recolonization process initiated two generations ago, from a limited number of relictual individuals. Thus inbreeding depression was expected to be strong since a total genetic purge was not expected over a so short period. To test for inbreeding depression we computed a simple linear regression of the individual proportions of empty seeds (i.e. a measure of fitness at the zygote stage) by the individual selfing rates. This regression was not significant at the 5% level (Restoux et al.). This absence of detected inbreeding depression at the seed stage can result from (**i**) the absence of inbreeding depression, **ii**) an unadapted sampling design providing a low statistical power to the regression test or (**iii**) a high variability of individual genetic load that could lead to a non-detected inbreeding depression. This latter hypothesis was also formulated by Burczyk et al. (1991), to explain a similar result for European larch, *larix decidua*.

Here we used simulations to test for the power of detection of inbreeding depression

through the regression of the progeny fitness by the mother selfing rate. In particular we investigated the consequences of genetic load heterogeneity among individuals and other genetic features of the populations (i.e. selfing rates heterogeneity). We used a simple model in which fitness is only affected by the number of lethal equivalents. Using this assumption, we then simulated numerous combinations of selfing rates and lethal equivalent genes carried by an individual to produce progenies with variable fitnesses associated to variable selfing rates. Then we tested for inbreeding depression detection using the regression approach. Finally we also investigated the effects of the sampling design on the ability to detect inbreeding depression.

Material & method

Simulations

The genetic load is measured through the number of lethal equivalent genes, which is defined as the number of heterozygous recessive lethal genes at linkage equilibrium in the parent generation necessary to obtain the same decrease of fitness within a selfed progeny as the one observed. A lethal equivalent gene is assumed to be rare enough in the population to ensure that an outcrossing mating cannot result in an homozygous state of this gene. We defined the ith population with four parameters, the mean selfing rate, S_i, the mean genetic load,L_i, the coefficient of variation of the selfing rate among individuals within the population, CVS_i, and the coefficient of variation of the genetic load among individuals within the population, CVL_i. S_i took values from 0.05 to 0.75 by step of 0.05, and L_i is defined as the population mean number of lethal equivalent and took values from 0 to 15

by step of 1. The coefficient of variation is defined as the ratio of the standard deviation by the mean and is a useful statistic to measure the variations of different phenomenons occuring at different scales since it is unitless. CVS_i values ranged from 0.05 to 0.55 and CVL_i values ranged from 0.05 to 1, both by steps of 0.05.

We sampled $N = (25,\ 50,\ 75\, or\, 100)$ individuals for each population. For each individual, j, of a population i we first drew an individual selfing rate,s_{ij}, ranging from 0 to 1, in a beta distribution $\beta(a_i, b_i)$. The beta distribution is defined with two shape parameters, a_i and b_i related to the mean S_i and the coefficient of variation CVS_i following the relations,

$$E(s_{ij}) = S_i = \frac{a_i}{a_i + b_i}, \tag{3}$$

$$Var(s_{ij}) = (S_i.CVS_i)^2 = \frac{a_i.b_i}{(a_i + b_i)^2.(a_i + b_i + 1)} \quad . \tag{4}$$

This leads to

$$a_i = \frac{1 - S_i(1 - CVS_i^2)}{CVS_i} \tag{5}$$

$$b_i = \frac{(S_i - 1)\,.\,(S_i.\,(1 + CVS_i^2) - 1)}{S_i.CVS_i^2} \quad . \tag{6}$$

Since a_i and b_i are positive for a beta distribution, the range of values for CVS_i is from 0.05 to 0.55 and the range of values of S_i is from 0.05 to 0.75.

For each individual j of the population i we also drew a number of lethal equivalent l_{ij}, as a positive real number, in a gamma distribution $\Gamma(a'_i, b'_i)$, where a'_i is the shape parameter and b'_i is the scale parameter. The parameters a'_i and b'_i are related to the mean and to the coefficient of variation L_i and CVL_i following the equalities,

$$E(u_{ij}) = L_i = a'_i.b'_i \tag{7}$$

$$V(n_{ij}) = (L_i.CVL_i)^2 = a'_i.b'^2_i \quad . \tag{8}$$

This leads to

$$a'_i = \frac{1}{CVL_i^2} \tag{9}$$

$$b'_i = L_i.CVL_i \quad . \tag{10}$$

We computed 100 populations of 25, 50, 75 and 100 individuals for each set of parameters, (S_i, L_i, CVS_i, CVL_i).

We used the classical Mendelian probabilities to infer the expected proportion of the progeny which is homozygous for lethal genes. Thus for l loci the probability of being homozygous for any lethal genes for at least one locus is $1 - \left(\frac{3}{4}\right)^l$. The expected relative fitness (i.e. if the progeny size was infinite) of the progeny of the jth individual of the ith population, ω_{ij}, is thus given by,

$$\omega_{ij} = 1 - \left(s_{ij}.\left(1 - \left(\frac{3}{4}\right)^{l_{ij}}\right)\right) \quad , \tag{11}$$

which ranges from 0 to 1 (figure 1). We used simulations based on this expected fitness to investigate the effects of population features on the ability to detect inbreeding depression, with a constant number of progenies investigated, N, set to 50.

The realized fitness, ω'_{ij}, corresponding to the proportion of the progeny that survives, was drawn in a binomial distribution,

$$\omega'_{ij} \sim \frac{\mathcal{B}(n, \omega_{ij})}{n} \quad , \tag{12}$$

where n is the size of each sampled progeny and ω_{ij} the expected fitness of this progeny. This fitness was used to investigate the effects of the sampling design on the ability to detect inbreeding depresion. In this analysis we simulated 100 repetitions for all the com-

binations of $N = 25$, 50, 75 or 100 progenies, each of size $n = 25$, 50, 75 or 100 individuals.

Detection of inbreeding depression

Test for inbreeding depression

To detect inbreeding depression we computed a linear regression of the expected fitnesses of progenies, ω_{ij}, by the selfing rates, s_{ij}, using the following model,

$$\omega_{ij} = a + b.s_{ij} + \epsilon_{ij} \quad , \tag{13}$$

with a the constant parameter, b the regression parameter of the fitness by selfing rate and ϵ_{ij} is the error term. In this case we only investigated the effects of population characteristics and we fixed the number of progenies investigated to $N = 50$. The same regression was computed for the simulations dealing with the realized fitness, ω'_{ij}, to investigate the effects of sampling design. All the combinations of N and n were then investigated. For all these regressions, we tested for the significance of the detected inbreeding depression using a classical Fisher test, based on a $\mathcal{F}_{1,N-1}$ statistic with 1 and $N - 1$ degrees of freedom. We considered the inbreeding depression as significant for p-values below 0.05.

Effects of population genetic characteristics

Using these tests, we computed the proportion, P_i, of significant regressions over the 100 repetitions for each set of parameters, (S_i, CVS_i, L_i, CVL_i). P_i is the proportion of populations for which we detected inbreeding depression at the 5% level. This quantity represents the statistical power of the inbreeding detection test. Finally, we computed a logistic regression to explain the proportion of simulations for which inbreeding depression was

detected by S_i, L_i, CVS_i and CVL_i and all the interaction terms up to the second order (not detailed in the formula),

$$\begin{cases} P_i \sim \dfrac{\mathcal{B}(100, \pi_i)}{100} \\ ln\left(\dfrac{\pi_i}{1-\pi_i}\right) = \beta_0 + \beta_1.L_i + \beta_2.S_i + \beta_3.CVL_i + \beta_4.CVS_i + ..., \end{cases} \tag{14}$$

with β_x the regression coefficients. We assumed that the distribution of the observations followed a binomial model distribution and used a logit link function. The significance of these regressions was tested using a Type III test to avoid confusion of effects.

Effects of sampling design

We compared the behavior of inbreeding depression detection under various scenarios implying different number of sampled mothers (N) and different progeny sizes (n). We computed a logistic model in which we considered the effects of N and n and their interaction on the proportion of significant regression (equation 13) as follow,

$$\begin{cases} P_{ij} \sim \dfrac{\mathcal{B}(100, \pi_{ij})}{100} \\ ln\left(\dfrac{\pi_{ij}}{1-\pi_{ijk}}\right) = \alpha_{N_i} + \beta_{n_j} + \gamma_{(N\times n)_{ij}}, \end{cases} \tag{15}$$

where N_i is the number of progenies sampled, n_j is the size of each progeny sampled, P_{ij} is the statistical power of the inbreeding depression detection test. We assumed a binomial distribution of the proportion of significant regression, P_{ij} and used a logit link function. N and n were considered as fixed effects.

All simulations and analyses were computed using R statistical langage (R Development Core Team 2007). Simulations and analysis source codes are available upon request.

Results

Effects of population characteristics

The proportion of detected inbreeding depression ranged from 1% to 100% over the population parameters investigated considering the expected fitness and $N = 50$ progenies. The average over all these simulations was about 92%. The power of the detection of inbreeding depression was significantly influenced by all the parameters of the simulations, S,L, CVS and CVL (not shown). But significance level are meaningless on simulated data. The regression model defined by the system of Eq. 14 also showed significant effects of all interactions up to the second order. The coefficient of variation of the individual selfing rates, CVS, the mean genetic load of the population, L and their interaction,$CVS \times L$, appeared to be the most influent variables on the power of the detection of inbreeding depression (figure 2, table 1). The interaction between the mean genetic load of the population and its coefficient of variation, $L \times CVL$, the interaction between the coefficients of variation of individual selfing rates and genetic loads, $CVS \times CVL$ and the coefficient of variation of the individual genetic loads, CVL, also had a strong but weaker influence (figure 2 and 3, table 1). The mean selfing rate of the population, S, and all interaction terms implying this parameter had largely weaker effects on the capacity to detect inbreeding depression than all the other effects previously cited. Increasing the variance of selfing rates (CVS) and the mean number of lethal equivalents (L) both resulted in a strong increase of the power the test. Increasing the variance of the lethal equivalents (CVL) led to a clear decrease of the performance of the test. The positive effect of the coefficient of variation of selfing (CVS) on the statistical power of the inbreeding depression test is stronger when the mean genetic load of the population (L) is low (figure 3 and table 1). The negative effect of the coefficient of variation of the genetic load (CVL) is stronger when the mean genetic

load (L) or the coefficient of variation of selfing rate (CVS) are large (figure 3 and table 1).

Effects of sampling design

The overall statistical power of inbreeding depression detection test (i.e. over all the parameter sets) was largely influenced by the number of progenies sampled and by their sizes. The average power of the detection of inbreeding detection ranged from 67% when we sampled 25 progenies of 25 individuals (i.e. seeds) to 92% when we sampled 100 progenies of 100 seeds (table 2). The general linear model (equation 15) showed positive and significant effects of N and n on the ability to detect inbreeding depression and no effect of the interaction term, $N \times n$ (table 3). The effect of N on the statistical power of the test for inbreeding depression is stronger than the effect of n (tables 3 and 2).

Discussion

Effects of population characteristics

The overall statistical power of the test for detection of inbreeding depression is high (90.9%) when assuming an infinite number of individuals sampled within each progeny. Thus this test for inbreeding depression based on regression of progenies fitnesses by parental selfing rates seems to be appropriate. However in some cases, inbreeding depression was hardly detected or at least detection was very variable depending on the population characteristics (figures 2 and 3).

The coefficient of variation of the individual selfing rates within a population was the most influent variable on inbreeding depression detection (table 1). Indeed if the selfing rate does not vary much, the statistical power to estimate the regression coefficient correctly is low, leading to frequent non-significant regressions. As a consequence, highly variable individual selfing rates permit to better infer inbreeding depression independently of the mean selfing rates at the population level (table 1). Actually, the better design to increase the variance of the selfing rate is to have two groups of progenies, a completely selfed one and a completely outcrossed one. This is exactly what is done during pollination experiments (Kärkkäinen et al. 1996, **?**).

Another very influent variable was the parental mean number of lethal equivalents, representing the mean genetic load of the population. This result was not suprising since genetic load is by definition the basis of inbreeding depression. Thus a higher genetic load led to a larger regression coefficient in the test (equation 13), since for the ith population the regression coefficient, β_1, must be the mean over all the j individuals of $\left(\left(\frac{3}{4}\right)^{l_{ij}} - 1\right)$. This expression tends to $\left(\left(\frac{3}{4}\right)^{L_i} - 1\right)$ if the number of sampled individuals is sufficient. Thus the ability to detect inbreeding depression depends on the mean genetic load of the population since a lower regression coefficient is statistically harder to detect than a larger one.

The mean genetic load of the population (L) acted synergistically with the variability of the individual selfing rates (CVS), leading to a higher detection of the inbreeding depression when both were large, and to lower detection when both were low (table 1). Indeed, in our model the individual selfing rate (s) acts linearily on the fitness of the progeny whereas the individual genetic load (l) acts exponentially on it (equation 11 and figure 1). Thus when the mean genetic load (L) was low, highly variable individual selfing

rates (CVS) resulted in a fewly variable progenies fitnesses, leading to a low significance of the regression of the fitness by the selfing rate (equation 13). Inversely a high mean genetic load and highly variable individual selfing rates led to a high statistical power of the detection of inbreeding depresion.

The main goal of this study was to investigate the effect of inter-individual variability of genetic load on inbreeding depression. It was represented here by the coefficient of variation of the number of lethal equivalents, CVL. This variable had a significant and negative effect on the statistical power of the inbreeding depression test (figure 2 and table 1). A high variability of the individual genetic loads resulted in a low detection of inbreeding depression (as low as 77.1% for $CVL = 1$) induced by an inaccurate estimation of the regression coefficient of fitness by selfing rate since the relationship between fitness and selfing rate was highly variable among progenies (cf the paragraph dealing with the mean genetic load). This effect of inter-individual variability of genetic load was stronger when the mean genetic load (L) or when the variability of selfing rates (CVS) were low (figure 3 and table 1). Indeed, low values of the mean genetic load and high values of its variability among individuals are associated to a bimodal distribution of values of genetic load with half the individuals without any genetic load and half with a low genetic load. Thus it was difficult to detect inbreeeding depression through the regression coefficient of fitness by the selfing rate since the residuals were too large, leading to non significant regression coefficients. We observed a stronger interaction effect between the inter-individual variability of selfing rate and the inter-individual variability of genetic load on the power of detection of inbreeding depression, with a low detection ability when genetic load was highly variable and selfing rate was fewly variable (table 1 and figure 3). In this latter case, fitness was highly variable according to variable genetic loads but the regression coefficient of progenies fitnesses by the parental selfing rates was hardly estimated since the explicative variable was

not much variable, leading to large residual variance and thus to non significant regressions.

Effects of sampling design

The efficiency of inbreeding depression analysis using the regression of fitnesses by selfing rates largely depends on the sampling design. Indeed, the previous commented results were only based on the expected fitnesses of progenies, thus assuming an infinite size of these progenies. However when their sizes, n, varied, the ability to detect inbreeding depression was highly modified, with a better detection when progeny sizes were large, as expected. Also when the number of progenies investigated, N, was low the power of detection of inbreding depression decreased. The overall statistical power of detection of inbreeding depression ranged from about 67% when $N = n = 25$ to about 92% when $N = n = 100$ (table 2). On the one hand, large number of sampled progenies resulted in a better estimation of the mean genetic load of the population and thus of β_1 (cf previous part), while a small N led to a smaller statistical power of the regresion test because it was based on less observations. On the other hand, the size of the sampled progenies also modified the ability to estimate the genetic load but at the individual level (i.e. of the mother). The estimation of individual genetic load can be over-estimated or under-estimated by chance according the size of progenies investigated, and thus if the total number of sampled progenies is large enough, the mean genetic load of the population can be well averaged leading to a good a detection of inbreeding depression. This hypothesis was confirmed by the deviance analysis computed on the proportion of cases in which inbreeding depression was detected (table 3). Thus assuming a constant sampling effort one can better detect inbreeding depression at the population level with a large number of sampled progenies than with a large

progeny size. For example with a sampling effort of 2500 fitness measures, the proportion of detected inbreeding depression was about 86% when $N = 100$ and $Sz = 25$ while this proportion was only 78.33 % when $N = 25$ and $Sz = 100$ (table 2).

Limitations & perspectives

The detection of inbreeding depression through the linear regression of progeny fitness by the selfing rate of the mother appears to be a powerfull approach over a large range of conditions. However some genetic charateristics of populations can modify the statistical power of this test, in particular the selfing rate must be variable enough to investigate a large range of observed selfing rates. Furthermore when the mean genetic load of the population is high or when its variability over individuals is low, the regression test was proner to detect inbreeding depression.

However sampling design can modify the strength of these two latter effects, indeed investigation of a large number of progenies greatly accurates the detection of inbreeding depression since it permits to better estimate the mean genetic load of the population. The sample size of the progenies also plays a role on the ability to detect inbreeding depression since it increases the precision of the progeny fitness estimation, but this parameter has a smaller effect than the number of progenies investigated. Thus assuming a constant sampling effort it is preferable to sample a large number of progenies than a large number of individuals within progenies.

These results are general even if we only considered lethal genes and no polyembryony,

a widespread phenomenon in conifers (Lindgren 1975, Sorensen 1982). But these features would only change the distribution of fitness and would not affect our conclusions. Our conclusions can also be applied to other regression approach to estimate the inbreeding depression, in particular to the regression of fitness by inbreeding coefficient (equation 2, Williams and Savolainen (1996)). In this latter case much attention must be paid to the variability of inbreeding coefficient and to the variability of genetic load, thus the same recommendations about sampling design can be applied for this kind of analysis.

We should stress that regressing the fitness by the selfing rate can be biased since the estimated selfing rate and the really experienced one at the fertilization stage (i.e. primary selfing rate) can be very different if inbreeding depression occurs. Thus the latter in the life cycle it is estimated and the more different they are (Charlesworth and Charlesworth 1987). Charlesworth and Charlesworth (1987) and Maki (1993) proposed two estimators of primary selfing rates but both were more or less explicitely based on the estimation of inbreeding depression itself, thus they can be unadapted for inbreeding depression estimation. However in their study, Husband and Schemske (1996) showed a significant correlation between primary selfing rate and estimated selfing rate for many species, except when they only dealt with gymnosperms probably because of a too small number of studies.

In our study, fitness was only affected by inbreeding depression but in many cases fitness of individual or even the strength of inbreeding depression can be modified by the environment (Hedrick et al. 1999, Bijlsma et al. 2000). The extent of inbreeding depression is larger when individuals experience stressfull conditions such as drought (Cheptou et al. 2000*a*) or competition (Cheptou et al. 2000*b*), thus fitness of individuals is not well described by a binomial distribution. As a consequence, if measures of fitness are

done in various contrasted natural environments, one must account for the environment as co-variable within in the linear model to distinguish a higher genetic load from a higher expression of inbreeding depression under bad conditions.

In the present study we found that the variability of individual genetic loads can affect the detection of inbreeding depression of populations, and thus the detection of the genetic purge. This conclusion is in accordance with the study of Byers and Waller (1999), who found that purging of genetic load was an inconsistent force within populations resulting from different factors such as mating history or population sizes. They suggested to bypass such artefacts by investigating mating history at the population and family levels, and thus revealed in what extent genetic load can have been purged. Finally we can ask the question of inbreeding evolution at the population level when considering on the one hand individual variability of genetic load and restricted gene flow on the other hand which can lead to biparental inbreeding. Thus individual variability of genetic load must be considered as a very influent parameter from an evolutionnary perspective since inbreeding depression is a largely admited force influencing the evolution of sex (Charlesworth and Charlesworth 1987).

Acknowledgments

This study is part of Gwendal Restoux's Ph.D. project, co-funded by The Provence Alpes Côtes d'Azur Regional Government and INRA.

References

Aizen, M. A., and L. D. Harder. 2007. Expanding the limits of the pollen-limitation concept : effects of pollen quantity and quality. Ecology **88**:271–281.

Bijlsma, R., J. Bundgaard, and A. C. Boerema. 2000. Does inbreeding depression affect the extinction risk of small populations ?: predictions from *Drosophila*. Journal of Evolutionary Biology **13**:502–514.

Burczyk, J., G. Kosiński, and A. Lewandowski. 1991. Mating pattern and empty seed formation in relation to crown level *Larix decidua* Mill. clones. Silva fennica **25**:201–205.

Byers, D. L., and D. M. Waller. 1999. Do plant populations purge their genetic load? Effects of population size and mating history on inbreeding depression. Annual Review of Ecology and Systematics **30**:479–513.

Charlesworth, D., and B. Charlesworth. 1987. Inbreeding depression and its evolutionary consequences. Annual Review of Ecology and Systematics **18**:237–268.

Cheptou, P.-O., A. Berger, A. Blanchard, C. Collin, and J. Escarre. 2000*a*. The effect of drought stress on inbreeding depression in four populations of the Mediterranean outcrossing plant *Crepis sancta* (Asteraceae). Heredity **85**:294–302.

Cheptou, P.-O., E. Imbert, J. Lepart, and J. Escarre. 2000*b*. Effects of compettition on lifetime estimates of inbreeding depression in the outcrossing plant *Crepis sancta* (Asteraceae). Journal of Evolutionary Biology **13**:522–531.

Goodwillie, C., S. Kalisz, and C. G. Eckert. 2005. The evolutionary enigma of mixed mating systems in plants : Occurence, theoretical explanations, and empirical evidence. Annual Review of Ecology, Evolution and Systematics **36**:47–79.

Hedrick, P. W. 2004. Recent developments in conservation genetics. Forest Ecology and Management **197**:3–19.

Hedrick, P. W., O. Savolainen, and K. Kärkkäinen. 1999. Factors influencing the extent of inbreeding depression : an example from Scots pine. Heredity **82**:441–450.

Husband, B. C., and D. W. Schemske. 1996. Evolution of the magnitude and timing of inbreeding depression in plants. Evolution **50**:54–70.

Jain, S. K. 1976. The evolution of inbreeding in plants. Annual review of Ecology and Systematics **7**:469–495.

Kärkkäinen, K., V. Koski, and O. Savolainen. 1996. Geographical variation in the inbreeding depression of Scots pine. Evolution **50**:111–119.

Kormutak, A., and D. Lindgren. 1996. Mating system and empty seeds in silver fir (*Abies alba* Mill.). Forest Genetics **3**:231–235.

Lande, R., and D. W. Schemske. 1985. The evolution of self-fertilization and inbreeding depression in plants. I.Genetic models. Evolution **39**:24–40.

Lindgren, D. 1975. The relationship between self-fertilization, empty seeds and seeds originating from selfing as a consequence of polyembryony. Studia Forestalia Suecica **126**:1–24.

Maki, M. 1993. Outcrossing and fecundity advantage of females in gynodioecious *Chionographis japonica* var. *kurohimensis* (*Liliaceae*). American Journal of Botany **80**:629–634.

R Development Core Team, 2007. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. URL `http://www.R-project.org`.

Restoux, G., C. Bonnet, C. Pichot, G. G. Vendramin, E. K. Klein, and B. Fady. ???? Variability of female reproductive success in the conifer tree *Abies alba*: pollen quantity, pollen quality and climate. submitted .

Restoux, G., D. E. Silva, F. Sagnard, F. Torre, E. K. Klein, and B. Fady. 2008. Life at the margin: the mating system of Mediterranean conifers. Web Ecology **8**:94–102.

Robertson, A., and W. G. Hill. 1984. Deviations from Hardy-Weinberg proportions: Sampling variances and use in estimation of inbreeding coefficients. Genetics **107**:703–718.

Schemske, D. W., and R. Lande. 1985. The evolution of self-fertilization and inbreeding depression in plants. II. Empirical observations. Evolution **39**:41–52.

Sorensen, F. 1969. Embryonic genetic load in coastal Douglas-fir, *Pseudotsuga menziesii* var. Menziesii. The American Naturalist **103**:389–398.

Sorensen, F. C. 1982. The roles of polyembryony and embryo viability in the genetic system of conifers. Evolution **36**:725–733.

Wang, K. S. 2003. Relationship between empty seed and genetic factors in European beech (*Fagus sylvatica* L.). Silva Fennica **37**:419–428.

Weir, B., and C. Cockerham. 1984. Estimating F-Statistics for the analysis of population structure. Evolution **38**:6.

Williams, C. G. 2008. Selfed embryo death in *Pinus taeda*: a phenotypic profile. New Phytologist **178**:210–222.

Williams, C. G., and O. Savolainen. 1996. Inbreeding depression in conifers : Implications for breeding strategy. Forest Science **42**:102–117.

Tables & figures

Effects	Regression coefficient	Deviance	F value
S	-0.35	680.84	66.27
L	0.32	854.53	12708.88
CVS	27.60	997.40	23107.42
CVL	-2.86	745.37	4763.47
$S \times L$	0.037	681.17	90.66
$S \times CVS$	6.58	690.85	795.21
$S \times CVL$	-0.71	681.43	109.39
$L \times CVS$	2.06	1025.95	25185.94
$L \times CVL$	-0.36	819.77	10178.91
$CVS \times CVL$	-16.73	768.10	6417.96

Table 1: Regression coefficients of the multiple logistic regression (equation 14) and resulting deviance analysis of the effects of the different population characteristics on the statistical power of inbreeding depression detection test. This analysis was based on the expected fitness of 50 progenies.

		n			
		25	50	75	100
N	25	67.04	73.66	76.61	78.33
	50	78.51	83.57	85.69	86.90
	75	83.52	87.70	89.40	90.36
	100	86.47	90.05	91.52	92.33

Table 2: Proportions of simulations for which inbreeding depression was detected, as a function of the number of progenies sampled (N) and of their sizes (n).

Effects	D.F.	Deviance	F	p
N	3	410134	14458.54	$< 10^{-4}$
n	3	394127	3590.59	$< 10^{-4}$
Nxn	9	388846	1.67	0.09

Table 3: Deviance analysis presenting the effects of the number of investigated progenies (N) and of the size of each of the progeny (n) on the ability to detect inbreeding depression using regression methods.

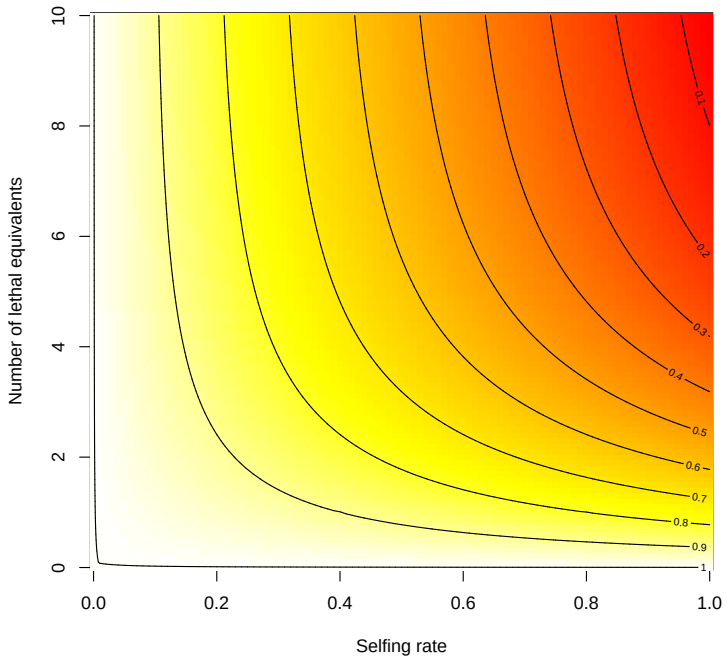


Figure 1: Expected fitness of the progeny of an individual as a function of its selfing rate and of its number of lethal equivalents (see equation 11).

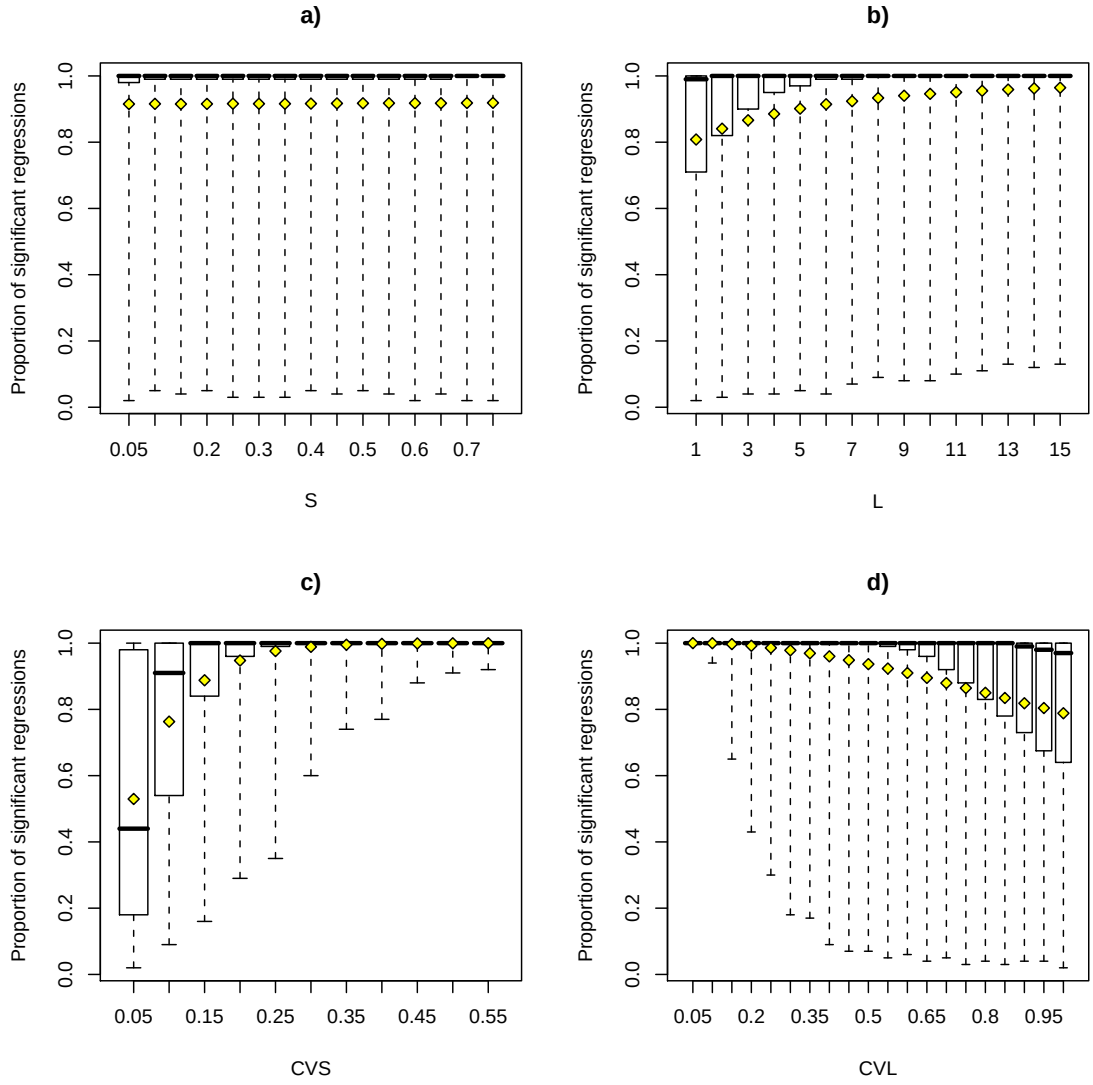


Figure 2: Boxplots representing the statistical power of the inbreeding depression test (i.e. the proportion of cases for which inbreeding depression was detected over the 100 repetitions for each set of parameter) based on the expected fitness of 50 progenies as a function of a) the mean selfing rate of the population, S, b) the mean number of lethal equivalents per individual, L, c) the coefficient of variation of the individual selfing rate over the population, CVS and d) the coefficient of variation of the individual number of lethal equivalents over the population, CVL. The boxes represent the 2nd and the 3th quartile of the distribution separated by the median. The whiskers represent the range of the distribution and the diamonds represent the mean proportion of significant regressions.

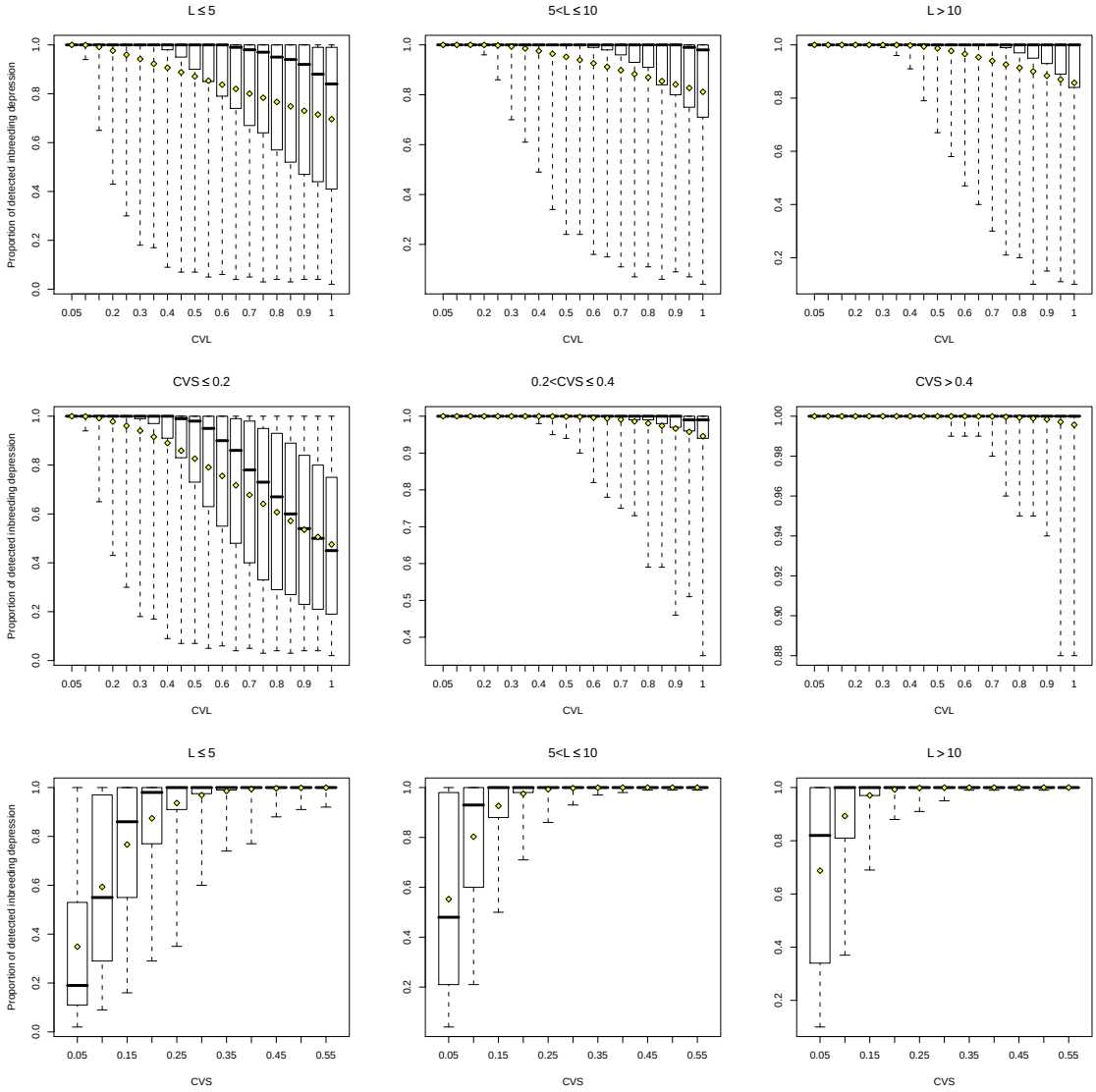


Figure 3: Boxplot representing the statistical power of the inbreeding depression test for the expected fitness of 50 progenies, as a function of the inter-individual genetic load variability, CVL, for different values of the mean genetic load, L, (top row) or for different values of the inter-individual selfing rate variability, CVS, (intermediate row). The bottom row represents the effect of the inter-individual selfing rate variability, CVS, as a function of the mean genetic load, L.The boxes represents the 2nd and the 3th quartile separated by the median, and the whiskers represent the range of the distribution. The diamonds represent the mean statistical power of the inbreeding detection test.

Chapitre 5

Variabilité spatio-temporelle du régime de reproduction et des flux de gènes

1 Synthèse

Dans ce chapitre nous avons caractérisé la structuration génétique spatiale du nuage pollinique de la population de sapin pectiné de la montagne de Lure. Nous avons émis l'hypothèse que chez les espèces pérennes et les arbres en particulier, cette structure était variable dans l'espace mais aussi dans le temps, et que ces variations pouvaient avoir un impact important sur la diversité génétique et la consanguinité de notre population. Une approche spatio-temporelle devrait donc être considérée lors des études visant la conservation de populations d'arbres. Afin de tester cette hypothèse nous avons développé deux approches différentes, que nous avons adaptées pour notre jeu de données (marqueurs haplotypiques à hérédité paternelle) :

- Une méthode permettant l'estimation de la structuration génétique globale, adaptée de l'approche TwoGener (Austerlitz et Smouse 2001; 2002). Nous l'avons modifiée afin de prendre en compte les variations inter-annuelles de composition du nuage pollinique lors des estimations (d'après Irwin et al. (2003)).
- Une méthode d'estimation des paramètres du régime de reproduction (taux

d'autofécondation et corrélation de paternité), dérivée de la méthode MLTR (Ritland 1989; 2002). Nous l'avons modifiée afin de pouvoir estimer les corrélations de paternité à l'échelle individuelle.

De plus, nous avons intégré dans ces deux méthodes la gestion des données manquantes. Nous avons également tenté d'expliquer ces variations par les densités locales et les productions polliniques individuelles (dont le détail des mesures est présenté dans le premier chapitre).

Afin d'éviter les redondances inutiles les résultats ne seront pas présentés dans cette synthèse. Cependant, le lecteur ne souhaitant pas lire de manière approfondie cette partie pourra se reporter à la conclusion de ce chapitre (page 212 et suivantes) ainsi qu'au bilan de la conclusion générale de cette thèse (page 217 et suivantes) où figurent les principaux résultats de cette étude.

2 Introduction

2.1 L'importance des flux de gènes et son lien avec le succès reproducteur

Le réchauffement climatique prévu risque de modifier drastiquement et rapidement les conditions environnementales (IPCC 2007). L'impact de ce changement sera d'autant plus fort sur nos populations qu'elles sont situées en limite de l'aire de répartition de l'espèce (Kawecki 2008) et en régions montagneuses (Parmesan 2006). Les arbres devront s'adapter rapidement au nouveau climat car leur temps de génération est long en comparaison à la vitesse à laquelle ce changement d'environnement va s'effectuer (Petit et Hampe 2006). De plus, ce changement va augmenter le niveau de stress subi par les individus, et donc potentiellement augmenter l'effet de la dépression de consanguinité (Cheptou et al. 2000*a*, Restoux et al. 2008, voir l'introduction et le chapitre 2). Il est donc nécessaire de conserver la diversité génétique qui reflète directement la capacité d'une population à s'adapter à un environnement changeant et à s'y maintenir (Nunney et Campbell 1993, Reed et Frankham 2003). En effet, elle agit à deux niveaux :

- L'hétérozygotie est un bon indicateur de la consanguinité que l'on observe au sein d'une population de taille finie par simple dérive génétique (Spielman et al. 2004).
- La diversité génétique permet aux populations de s'adapter à de nouveaux environnements (Ellstrand et Elam 1993).

La diversité génétique au sein d'une population est largement conditionnée par les flux de gènes (Ellstrand 1992*a*). En effet Wright (1943; 1946) a montré qu'une dispersion limitée induisait un isolement par la distance, et que ce dernier allait diminuer un paramètre clef dans le domaine de la conservation, l'effectif efficace de la population,N_e (Hedrick 2004). Chez les plantes les flux de gènes peuvent s'effectuer par la dispersion des graines et du pollen (Oddou-Muratorio et al. 2001, Chen et al. 2008). La dispersion pollinique en particulier revêt une importance primordiale dans le maintien de la diversité génétique au sein des populations d'arbres

(Ellstrand 1992*b*, Smouse et Sork 2004, Petit et al. 2005).

Les flux de pollen influencent fortement la part relative de pollen importé, et par conséquent le taux d'autofécondation dès lors que le régime de reproduction est gouverné par une loi d'action de masse (Holsinger 1991, voir aussi le chapitre d'introduction). Dans cette étude nous ne définirons pas ce dernier comme étant seulement le taux d'autofécondation mais également par la corrélation de paternité au sein d'une descendance (i.e. la probabilité pour deux graines d'avoir le même père), qui reflète directement la diversité génétique du nuage pollinique échantillonné (Hardy et al. 2004*a*, Robledo-Arnuncio et al. 2004*a*). On déduit de cette corrélation le nombre de pères efficaces, N_{ep}, qui est défini comme le nombre de donneurs de pollen à fécondité égale permettant d'obtenir la même structure génétique que celle réellement observée (Smouse et al. 2001, Austerlitz et Smouse 2001). Dès lors que le nombre de pères efficaces est faible, cela implique un faible nombre de contributeurs ou un déséquilibre fort de contribution des individus à la reproduction.

2.2 Estimation du régime de reproduction et des flux de gènes

2.2.1 Flux de gènes "historiques"

Les flux de gènes, qui influencent fortement le taux d'autofécondation et la corrélation de paternité sont classiquement estimés indirectement en utilisant les indices de structuration génétique de type F-statistiques. Notamment en comparant le F_{st}, qui reflète la divergence génétique entre populations ou individus, aux distances géographiques qui les séparent (Rousset 1997; 2000). Dans le cadre de l'isolement par la distance défini par Wright (1943; 1946), on obtient alors l'augmentation de la différenciation entre populations avec la distance qui les sépare. Cette augmentation est linéaire si l'on considère le logarithme de la distance, on peut alors extraire de la pente :

- Le produit $N_e.\sigma^2$, où σ^2 représente la variance axiale de la dispersion, dans le cas de différences entre populations (Rousset 1997), ou

– le produit $d_e.\sigma^2$, avec d_e représentant la densité efficace, dans le cas de différences entre individus dans une population continue (Rousset 2000).

Dans le deux cas il est impossible d'estimer le paramètre σ^2 indépendamment de N_e ou de d_e. De plus, afin de distinguer la part des flux de gènes imputable à la dispersion du pollen de celle imputable à la dispersion des graines, il est nécessaire de baser ces analyses sur des marqueurs hérités différemment, c'est à dire par voie paternelle ou par voie maternelle, en plus de marqueurs hérités de manière biparentale (Oddou-Muratorio et al. 2001, Petit et al. 2005, Chen et al. 2008).

Les estimations des flux de gènes basées sur les F-statistiques donnent une bonne image de l'évolution passée d'une population et de la "moyenne" des variations démographiques et des flux migratoires historiques (Austerlitz et Smouse 2001). De la même façon il est possible d'estimer les paramètres du régime de reproduction "efficace", notamment le taux d'autofécondation nécessaire à l'obtention de la consanguinité observée *via* le F_{is} au sein d'une population (David et al. 2007). Cependant ces modèles d'estimation sont basés sur des hypothèses d'équilibre migration-dérive et/ou d'absence de sélection au sein des populations. Si les hypothèses d'équilibre peuvent être admises sur des longues périodes de temps puisqu'elle vont intégrer de nombreux épisodes de reproduction ; elles le sont difficilement lors de l'étude des flux de gènes actuels (i.e. "en temps réel"), c'est à dire à l'échelle d'un ou quelques épisodes de reproduction (Sork et al. 1999, Austerlitz et Smouse 2001). Ces hypothèses sont encore plus inadaptées à l'étude d'un seul compartiment du cycle de vie, comme c'est le cas dans cette étude qui ne s'intéresse qu'à la production des graines et à la variabilité de leur viabilité *via* leur diversité génétique notamment (dépression de consanguinité et potentiel évolutif). Ainsi même si les méthodes classiques d'estimation des flux de gènes et de la structuration génétique semblent refléter une histoire démographique plus récente que ce qu'il était précédemment admis (Leblois et al. 2004), elles s'avèrent peu adaptées pour la compréhension de l'évolution actuelle et future des populations notamment en réponse à des changements récents et/ou rapides de leur environnement. Il est donc nécessaire dans ce cas de décrire et d'estimer les paramètres de structuration génétique, de flux de gènes ou du régime de reproduction de manière instantanée.

2.2.2 Les flux de gènes instantanés

Estimation du régime de reproduction L'estimation du régime de reproduction instantané s'effectue en étudiant des descendances (i.e. *progeny array*), et des graines en particulier. Le taux d'autofécondation peut être estimé en comparant le génotype transmis paternellement au descendant à celui de la mère (Ritland et Jain 1981, Ritland 2002). Cette méthode fait l'hypothèse d'un nuage de pollen commun, dont la composition génétique est identique pour tous les individus étudiés, les fréquences alléliques (ou haplotypiques en fonction des marqueurs utilisés) au sein du nuage sont donc estimées conjointement aux taux d'autofécondation. Une autre méthode consiste à utiliser des analyses de paternité qui vont permettre de déterminer l'identité des pères au sein d'une descendance (Ellstrand 1984, Hardy et al. 2004*b*, Robledo-Arnuncio et Gil 2005). Le taux d'autofécondation est alors déduit de la proportion de graines issues d'autofécondation sur le nombre total de graines analysées. Ces deux méthodes permettent également d'estimer la corrélation de paternité (i.e. la probabilité que deux graines partagent le même père). Il est possible en modifiant les méthodes de Ritland et Jain (1981), tout en faisant la même hypothèse concernant le nuage pollinique, d'estimer simultanément aux taux d'autofécondation la corrélation de paternité en comparant les génotypes des descendants pris deux à deux et d'estimer par maximum de vraisemblance la probabilité qu'ils partagent le même père (Ritland 1989; 2002). Les analyses de paternité permettent également de calculer le rapport du nombre de couples de graines partageant le même père sur le nombre total de couples, ce qui correspond à la corrélation de paternité (Hardy et al. 2004*a*, Robledo-Arnuncio et al. 2004*a*). Cependant les analyses de paternité nécessitent un échantillonnage exhaustif de la zone étudiée et notamment de tous les pères potentiels, ce qui n'est pas toujours possible lorsque la population est de grande taille et continue. Finalement les modèles d'appariements, permettent d'estimer le taux d'autofécondation, le taux de migration (i.e. contribution de pollen venant de l'extérieur de la zone étudiée) ainsi que la fécondité relative d'un couple d'individus (Adams et Birkes 1991, Burczyk et al. 1996). Cependant ces modèles font une hypothèse concernant la dispersion du pollen ce qui n'est pas le cas des

deux autres méthodes, et seront décrites un peu plus tard dans ce texte.

Enfin, il est possible d'estimer la corrélation de paternité de manière indirecte par le biais d'estimateurs de la structuration génétique du nuage pollinique ou de la parenté. Dans le premier cas cela s'effectue grâce à une analyse de la variance moléculaire (AMOVA, Excoffier et al. (1992)), où l'on estime la part de variance inter-mères et la part de variance intra-mères, c'est la méthode "TwoGener" (Smouse et al. 2001). On extrait de cette analyse le Φ_{ft}, qui est la corrélation intra-mère des génotypes des contributeurs polliniques. A partir de cet indice il est possible d'estimer la corrélation de paternité au sein d'une mère en faisant l'hypothèse d'isolement par la distance due à une dispersion limitée du pollen (voir la section "matériel et méthode" pour plus de détails). La seconde méthode indirecte pour estimer la corrélation de paternité consiste à utiliser un indice de parenté calculé par couples de graines (Loiselle et al. 1995, Hardy et al. 2004*a*). En utilisant les différentes moyennes des coefficients de parenté par paires au sein d'une descendance et entre les descendances de deux mères, on obtient une approximation des corrélations de paternité au sein d'une descendance ou entre couple de descendances.

Estimation de la dispersion du pollen Comme pour le régime de reproduction, l'estimation de la fonction de dispersion du pollen peut s'effectuer directement par analyse de paternité ou indirectement en se basant sur des indices génétiques synthétiques (Smouse et Sork 2004, Burczyk et Koralewski 2005). Les méthodes directes permettent d'estimer de manière précise la fonction de dispersion mais nécessitent d'échantillonner et génotyper exhaustivement tous les pères potentiels au sein d'une zone définie au préalable. Une analyse de paternité catégorique (i.e. assignation d'un unique père par descendant) des graines prélevées sur les arbres (donc de mère connue) permettra de connaître la position des pères et donc la distribution de la distance parcourue par le pollen (Ellstrand 1984, Streiff et al. 1999, Hardy et al. 2004*b*).

D'autres approches directes sont basées sur la vraisemblance et permettent alors

d'estimer la paternité de manière fractionnelle (Devlin et al. 1988, Smouse et Meagher 1994). En effet, le génotype de chaque descendant possède une probabilité d'avoir été obtenu sachant le père i ($P(Descendant|Pere_i)$), et ce pour chacun des pères recensés. Il est alors possible d'en déduire pour chaque père potentiel, une probabilité d'être effectivement le vrai père du descendant. Une approche similaire permet d'estimer conjointement le taux d'autofécondation, le taux d'immigration (i.e. la proportion de pollen provenant de l'extérieur de la zone d'étude) et la fécondité relative d'un couple père - mère, il s'agit des modèles d'appariement (Adams et Birkes 1991, Smouse et al. 1999). Cette fécondité peut intégrer de nombreux paramètres comme la fertilité du père, le décalage phénologique entre le père et la mère ou la distance qui les sépare (Adams et Birkes 1991). Dans ce dernier cas il est possible de faire l'hypothèse d'un noyau de dispersion (i.e. une famille de fonction de distribution des probabilités de dispersion en fonction de la distance) dont les paramètres sont estimés par maximisation de la vraisemblance (Burczyk et al. 2002; 2004, Oddou-Muratorio et al. 2005, Klein et al. 2008).

Une des méthodes d'estimation indirecte de la fonction de dispersion du pollen ponctuelle (i.e. pour un épisode de reproduction) les plus utilisées est "TwoGener" (Austerlitz et Smouse 2001, Austerlitz et al. 2004). Elle dérive du modèle d'isolement par la distance (Wright 1943; 1946). Il est alors possible d'exprimer la probabilité de copaternité au sein d'une descendance ou par couples de descendances en fonction de la distance qui sépare ces dernières en faisant l'hypothèse que le pollen se disperse suivant une fonction de dispersion choisie et que les pères sont répartis aléatoirement avec une densité d_e. En exprimant les indices de structuration génétique par couples de descendances ($\Phi_{ft_{pw}}$) en fonction des probabilités de copaternités, il est possible d'obtenir une expression théorique des $\Phi_{ft_{pw}}$ en fonction de la distance. La fonction de dispersion est alors estimée en minimisant l'écart entre la distribution des $\Phi_{ft_{pw}}$ estimés et les $\Phi_{ft_{pw}}$ théoriques. La méthode "TwoGener" est décrite de manière plus précise dans le "matériel et méthode" qui suit cette introduction.

La seconde méthode indirecte d'estimation en temps réel de la fonction de dispersion du pollen est "KinDist" (Robledo-Arnuncio et al. 2006). Elle permet une

estimation similaire, mais en se basant non plus sur les $\Phi_{ft_{pw}}$, mais sur des corrélations de paternité par couples de descendances renormalisées (Ψ). Leurs calculs suivent la méthode décrite par Loiselle et al. (1995), Hardy et al. (2004a). KinDist permet d'estimer la fonction de dispersion du pollen sans faire d'hypothèse sur la densité efficace en contributeurs polliniques, contrairement à TwoGener.

Les méthodes indirectes ont l'avantage d'être beaucoup moins coûteuses que les méthodes directes en terme de quantité de données, mais ceci se fait aux dépends de la précision et de la flexibilité. En effet, ces méthodes utilisent un paramètre synthétique (Φ_{ft} ou Ψ) pour représenter la différenciation génétique des nuages et ainsi estimer la fonction de dispersion, alors que les méthodes directes utilisent l'information fournie par chacune des graines échantillonnées pour estimer les paramètres du régime de reproduction et de la fonction de dispersion. Cependant l'échantillonnage peut couvrir une plus large zone et donc être plus représentatif que ne le serait un voisinage dans le cas d'une population continue de grande taille, ce qui est le cas dans cette étude. Nous avons donc choisi d'adopter une approche indirecte afin d'estimer la corrélation de paternité et la courbe de dispersion du pollen, en utilisant la méthode "TwoGener". Pour plus de précisions quant aux différences entre les approches directes et indirectes, je conseille au lecteur les articles de synthèses de Smouse et Sork (2004) et de Burczyk et Koralewski (2005).

2.3 Variabilité spatio-temporelle du régime de reproduction et des flux de gènes

2.3.1 Variabilité spatiale

En effet, si un diagnostic doit être effectué, des solutions doivent aussi être proposées. Or, de nombreux facteurs, variables d'une année à l'autre, peuvent influencer les flux de gènes dus au pollen (i.e. régime de reproduction et dispersion) (Ellstrand 1992a). La densité en conspécifiques florifères semble fortement modifier le régime de reproduction : une faible densité ou un faible effectif de population résulte en général en une augmentation de la corrélation de paternité et du taux

d'autofécondation (Robledo-Arnuncio et al. 2004*a*, par exemple). Une faible densité efficace va également augmenter la distance moyenne de dispersion efficace du pollen (Robledo-Arnuncio et al. 2004*b*, Smouse et Sork 2004, Bittencourt et Sebbenn 2007; 2008). La variation de fertilité mâle des individus est un phénomène répandu chez les arbres (Kang et al. 2003) qui peut modifier la participation relative des individus à la reproduction. Cela peut alors modifier la composition génétique du nuage pollinique échantillonné par les mères, ce qui modifie l'estimation des paramètres du régime de reproduction et de dispersion pollinique (Oddou-Muratorio et al. 2005, Kamalakannan et al. 2007, Klein et al. 2008). La configuration spatiale, et notamment l'agrégation, semble également modifier la structuration génétique spatiale du nuage pollinique et son estimation (Meagher et Vassiliadis 2003, Robledo-Arnuncio et Austerlitz 2006, Fukue et al. 2007). Le décalage phénologique est aussi un facteur déterminant du régime de reproduction car il va diminuer le nombre de contributeur potentiel au nuage pollinique (Hardy et al. 2004*a*, Robledo-Arnuncio et al. 2004*a*, de Lucas et al. 2008). Ainsi une population semblant présenter une densité globale suffisante peut en réalité présenter une forte dérive et une structuration génétique marquée ce qui aura pour effet de diminuer ses capacités adaptatives et d'augmenter la consanguinité. Il est donc primordial de mesurer la structuration génétique spatiale des nuages polliniques et aussi de comprendre ce qui influence cette structuration.

2.3.2 Variabilité temporelle

Pour les espèces à longue durée de vie comme les arbres, la description de la variabilité dans l'espace des flux de gènes et du régime de reproduction une année donnée n'est pas suffisante dans un contexte de conservation. Il est essentiel de considérer l'échelle temporelle dans nos analyses car le régime de reproduction et la dispersion du pollen peuvent également varier dans le temps, au cours de la saison de pollinisation ou bien entre les années (Ellstrand 1992*a*). En effet, le décalage phénologique, et notamment le décalage des périodes de réceptivité des fleurs femelles va modifier la diversité génétique échantillonnée par les individus au cours d'un épisode de reproduction (Hardy et al. 2004*a*). De plus dans le cas des espèces entomophiles, le

comportement et/ou la composition (espèces, densité) des pollinisateurs peut être variable au cours de la saison de reproduction, ce qui va entraîner des différences de pollinisation (Hirao et al. 2006). Finalement les conditions de densité, de fertilité ou de pollinisation (anémophile ou entomophile) peuvent également changer d'une année à l'autre ce qui peut modifier le régime de reproduction et la dispersion efficace du pollen (Kenta et al. 2004, Oddou-Muratorio et al. 2005). Ainsi dans le cas des espèces pérennes, il est nécessaire de comprendre comment évolue dans le temps la structuration génétique du nuage pollinique au sein de la population, afin d'avoir une bonne représentation du déséquilibre de contribution des individus à la reproduction. De plus, cette variabilité inter-annuelle si elle existe peut interagir avec la variabilité spatiale, il est donc également important de caractériser cette interaction.

2.4 Objectifs de cette étude

Dans le second et le troisième chapitre nous avons vu que la densité mesurée à l'échelle du peuplement avait un effet sur la valeur du taux d'autofécondation. De plus nous avons vu que la dépression de consanguinité était présente au stade semis (i.e. faible taux de germination dans les parties du peuplement les moins denses), même si celle-ci semblait inefficace à un stade plus précoce, c'est à dire au moment de la formation de l'embryon. Cependant, nous avons également vu dans le second chapitre que les individus les plus isolés produisaient des semis qui semblaient mieux survivre sur le long terme et notamment en conditions stressantes, que ceux des individus issus de zones où la densité est plus forte. Une explication possible est celle d'une meilleure purge de leur fardeau génétique. Cela pourrait également signifier que pour ces individus, une partie non négligeable des descendants issus de l'allofécondation est peu consanguine, et donc que les nuages polliniques des mères pourraient avoir échantillonné une large diversité génétique. De plus, la proportion de graines vides, le taux de germination ou le taux de survie semblaient très variables entre les individus. Il est donc nécessaire de comprendre l'effet de la densité locale ou de la fertilité des individus sur les paramètres du régime de reproduction.

Dans le troisième chapitre nous avons également observé que la production de graines viables pouvait souffrir de limitation pollinique, et que celle-ci agissait à une échelle locale, ce qui semble assez étonnant chez un conifère. De plus nous avons prouvé que le régime de reproduction répondait à une loi d'action de masse. La composition du nuage pollinique échantillonné est alors le résultat de la dilution de tous les pollen reçus (i.e. allopollen et autopollen). Dans ce cas, il est donc intéressant de comprendre à quelle échelle s'effectue la dispersion efficace (i.e. conduisant à une fécondation) afin de comprendre comment s'effectue le "mélange" du pollen et à quelle échelle spatiale.
Enfin, notre étude s'intéressant à une espèce pérenne, il est nécessaire de s'assurer que les patrons observés sont stables dans le temps afin de pouvoir tirer des conclusions solides sur l'effet des facteurs du milieux et phénotypiques (i.e. densité locale et fertilité individuelle) sur la diversité génétique et la survie de nos populations (voir la partie sur les arbres de l'introduction de cette thèse). De plus, la densité locale et la variabilité inter-annuelle peuvent interagir, conduisant par exemple à une plus grande variabilité inter-annuelle de la diversité génétique échantillonnée chez les individus les plus isolés. En effet, ces derniers ne recevant que peu de pollen, ils pourraient être très sensibles à une variation de fertilité des quelques individus qui les fécondent. Il faut donc tester la présence d'un effet d'interaction entre les variabilités spatiales et temporelles et le quantifier le cas échéant.

Dans ce chapitre nous posons donc les questions suivantes :

- Comment se structure spatialement la diversité génétique du nuage pollinique au sein de notre population ?
- A quelle distance s'effectue la dispersion du pollen ?
- Comment varie le régime de reproduction à l'échelle individuelle ?
- Comment les facteurs du milieu tels que la densité ou la fertilité modifient-ils la structure génétique spatiale du nuage pollinique et le régime de reproduction ?
- Comment varie dans le temps le régime de reproduction et la dispersion du pollen ?
- Existe-t-il une interaction entre la variabilité spatiale et temporelle du régime de reproduction et de la dispersion pollinique ?

Afin de répondre à ces questions nous avons développé de nouvelles méthodes d'estimation :

- Une méthode d'estimation de la structure génétique spatiale et de la fonction de dispersion du pollen. Cette approche repose en partie (i.e. pour l'aspect temporel) sur l'approche développée par Irwin et al. (2003) elle même adaptée de la méthode TwoGener globale (Smouse et al. 2001, Austerlitz et Smouse 2001). Nous l'avons étendue à une approche *pairwise* avec des effectifs déséquilibrés. Cette méthode permet de traîter les données manquantes et les variations temporelles de composition génétique du nuage pollinique. De plus il désormais possible d'estimer la fonction de dispersion à partir de marqueurs haplotypiques à hérédité paternelle.
- Une méthode d'estimation des paramètres du régime de reproduction à l'échelle individuelle. Cette méthode est adaptée de la méthode MLTR développée par (Ritland 1989; 2002), mais elle permet désormais d'utiliser des marqueurs haplotypiques hérités paternellement et de traiter les données manquantes. De plus, l'estimation de la corrélation de paternité à l'échelle individuelle est désormais possible.

3 Matériel & Méthode

3.1 Matériel biologique

Nous avons conduit cette étude sur le site de la montagne de Lure, où la quantité de données disponibles était plus importante que pour le site du Mont Ventoux et les effectifs assez homogènes entre les différentes années. Nous disposons de 50 descendances de sapin pectiné (*Abies alba*, Miller), dont la taille chaque année est comprise entre 2 et 21 graines avec une moyenne de 11 graines par descendance et par an. Pour chacun des arbres mères échantillonnés nous disposons de la densité locale, mesurée comme étant le nombre d'individus producteurs de pollen dans un rayon de 12,5m, ainsi que de la fertilité mâle individuelle chaque année entre 2002 et 2005. Nous disposons des génotypes des mères et des descendants obtenus grâce à 4 marqueurs microsatellites chloroplastiques (Vendramin et al. 1996, Liepelt et al. 2001), qui ont la particularité d'être hérités paternellement chez les conifères (Vendramin et Ziegenhagen 1997). Des précisions complémentaires concernant les mesures effectuées et les protocoles utilisés sont présentés dans le premier et le troisième chapitres.

3.2 Méthode

Nous avons utilisé deux approches pour estimer et modéliser la structuration génétique du nuage pollinique :

- La première approche repose sur la méthode TwoGener (Smouse et al. 2001, Austerlitz et Smouse 2001, Austerlitz et al. 2004). Dans ce cas la structuration génétique du nuage pollinique est estimée de manière globale ou au travers de couples de mères. Cette structuration peut être reliée à un modèle de dispersion.
- La seconde approche repose sur l'estimation des probabilités de copaternité au sein des nuages polliniques, conjointement aux taux d'autofécondation. Ces probabilités de copaternités sont des indicateurs de la diversité génétique des nuages polliniques.

Dans ce dernier cas, l'estimation étant faite au niveau individuel nous avons pu tester leur lien avec la densité locale en conspécifiques ou la fertilité mâle individuelle. Cependant ces deux méthodes n'étant ni adaptées au type de marqueurs dont nous disposions, ni à une approche temporelle, nous les avons donc modifié. De plus la méthode TwoGener ne pouvant gérer l'autofécondation (Burczyk et Koralewski 2005) nous l'avons également modifié pour les données issues d'espèces à régime mixte de reproduction. Toutes les modifications que nous avons appliqué sont détaillées dans la suite de cette partie. Enfin les approches TwoGener et par le régime de reproduction étant complémentaires, leurs résultats seront comparés.

3.2.1 Caractérisation de la dispersion du pollen par l'approche TwoGener

L'approche TwoGener (Smouse et al. 2001, Austerlitz et Smouse 2001; 2002) a été développée afin d'estimer une fonction de dispersion lors d'un événement unique de dispersion à partir de marqueurs diploïdes. Dans cette méthode la fonction de dispersion est estimée à partir de la relation observée entre la différenciation des nuages polliniques des mères, Φ_{ft} et la distance physique qui les sépare. Nous avons modifié cette approche pour l'adapter à nos données issues de marqueurs haploïdes à hérédité paternelle, présentant des données manquantes et récoltées sur plusieurs années. Pour prendre en compte le caractère haploïde de nos marqueurs nous avons recalculé le lien entre le Φ_{ft} et la fonction de dispersion d'après la méthode TwoGener (Austerlitz et Smouse 2001). La variabilité temporelle de la dispersion a été introduite dans le modèle en suivant la décomposition de la variance par AMOVA proposée par Irwin et al. (2003). Cette approche ne décrivant que le calcul du Φ_{ft} global dans le cas d'échantillons équilibrés, nous l'avons donc généralisée aux calculs des Φ_{ft} par couple de mères (Φ_{ft-pw}) introduits par Austerlitz et Smouse (2002) et aux échantillonnages non équilibrés. La gestion des données manquantes au sein des génotypes sera effectuée de manière similaire à l'approche proposée par Smouse et al. (2001) dans le cas d'haplotypes ambigus pour des données diploïdes (i.e. indétermination de l'haplotype paternel). De plus pour comparer les estimations issues du modèle TwoGener, aux estimations des probabilités de copaternité individuelles

(voir 3.2.2 page 165), nous avons effectué l'estimation de la fonction de dispersion de l'allopollen. Pour cela, nous ne prenons pas en compte les graines potentiellement issues de l'autofécondation comme proposé par Carpentier et al. (in prep).

Principe de la méthode TwoGener La méthode TwoGener estime la fonction de dispersion du pollen en exprimant la mesure de différenciation des nuages polliniques des mères, Φ_{ft}, en fonction des probabilités de copaternité au sein de ces nuages. Dans cette méthode, les probabilités de copaternité, (i.e. les probabilités pour deux grains de pollen d'être issus du même père), sont décrites théoriquement en fonction de la densité efficace du peuplement et de la fonction de dispersion (Austerlitz et Smouse 2001). Ces probabilités sont utilisées pour le calcul de Φ_{ft} théoriques qui ne dépendent ainsi que de la densité efficace et de la fonction de dispersion. D'autre part, des Φ_{ft} observés sont estimés à partir des haplotypes paternels des graines échantillonnées (dans la méthode actuelle de marqueurs diploïdes, déduits des génotypes des graines et de leur mères, dans notre cas directement issus de marqueurs haplotypiques à hérédité paternelle). Cette estimation est issue d'une AMOVA (Excoffier et al. 1992) où les nuages polliniques de chaque mère sont considérés comme des groupes et les haplotypes des graines comme des répétitions au sein de chaque groupe (Smouse et al. 2001). En effectuant cette AMOVA, sur l'ensemble des graines échantillonnées, nous pouvons ainsi estimer un Φ_{ft} global, alors qu'en l'appliquant seulement aux graines de deux mères, nous obtenons alors une estimation par couples de mères, les $Phi_{ft_{pw}}$ (*pairwise* Φ_{ft}). Finalement, après avoir choisi une famille de fonctions de dispersion (normale, exponentielle, exponentielle puissance,...), l'estimation de la densité efficace et des paramètres de la fonction de dispersion est obtenue en minimisant la somme des carrés des écarts entre les Φ_{ft} observés et les Φ_{ft} théoriques. Cette méthode peut s'utiliser avec les Φ_{ft} globaux ou les Φ_{ft-pw}. Dans le premier cas, un seul paramètre peut être estimé, le plus souvent la distance moyenne de dispersion. En revanche en utilisant l'ensemble des Φ_{ft-pw}, l'estimation simultanée de plusieurs paramètres est possible (i.e. estimer la densité efficace du peuplement ainsi que les paramètres des fonctions de dispersion : échelle et de forme) (Austerlitz et Smouse 2002, Austerlitz et al. 2004).

Lien entre le Φ_{ft} et la corrélation de paternité : différences entre marqueurs diploïdes et haploïdes La méthode TwoGener repose sur la comparaison entre Φ_{ft} théoriques et observés. Elle a été conçue pour des marqueurs diploïdes. L'estimation des Φ_{ft} observés est similaire pour les marqueurs haplotypiques que pour les marqueurs nucléaires excepté qu'elle se base sur les fréquences haplotypiques et non alléliques puisque les différents loci sont physiquement liés. De plus, il nous a fallu reformuler l'expression des Φ_{ft} théoriques et notamment leur lien avec les probabilités de copaternités.

La proportion de variance expliquée par l'effet mère, le Φ_{ft}, est en fait une corrélation génétique intraclasse, similaire au F_{st}. Soient f_i, f_f, f_t, f_p les probabilités d'identité par état de deux gènes tirés au hasard au sein respectivement d'un individu (i), du nuage pollinique d'une mère (f), du nuage pollinique total des mères considérées (t) ou au sein de la population où ont été échantillonnées les mères (p). On peut aussi définir la diversité au sein de chacun de ces compartiments par : $h_x = 1 - f_x$. Ainsi en utilisant les définitions introduites par Nei (1973) pour les F-statistiques, on peut écrire le Φ_{ft} en fonction des différentes probabilités d'identité :

$$\Phi_{ft} = \frac{f_f - f_t}{1 - f_t} = 1 - \frac{h_f}{h_t} \tag{5.1}$$

Les diversités h_f et h_t peuvent être décomposées en fonction de Q_0, la probabilité de copaternité de deux graines au sein d'une descendance et de Q_t, la probabilité de copaternité de deux graines issues des descendances de deux mères.

Austerlitz et Smouse (2001) traitent le cas de données obtenues à partir de marqueurs diploïdes. Deux gènes tirés au hasard au sein du nuage pollinique d'une mère proviennent du même père avec la probabilité Q_0. Dans ce cas, ils proviennent du même chromosome avec une probabilité 1/2, ils seront alors identiques, ou bien ils proviennent de deux chromosomes homologues avec une probabilité 1/2 et seront différents avec une probabilité h_i. En revanche ces deux gènes proviennent d'un père différent avec une probabilité $1 - Q_0$, ils sont alors différents avec une probabilité h_p. Dans le cas de marqueurs diploïdes, la diversité au sein du nuage pollinique d'une

mère, h_f^{diplo} est donc égale à :

$$h_f^{diplo} = \frac{Q_0.h_i}{2} + (1 - Q_0).h_p \tag{5.2}$$

De la même façon on peut décomposer h_t en fonction de Q_t dans le cas de marqueurs diploïdes de la façon suivante :

$$h_t^{diplo} = \frac{Q_t.h_i}{2} + (1 - Q_t).h_p \tag{5.3}$$

En faisant l'hypothèse que la consanguinité est absente au sein de la population, alors $h_i = h_p$. Dans ce cas, d'après l'équation 5.1 le Φ_{ft}^{diplo} est donc égal à :

$$\Phi_{ft} = \frac{Q_0 - Q_t}{2 - Q_t}, \tag{5.4}$$

Nous avons adapté ces calculs au cas des marqueurs haplotypiques que nous considérons dans cette étude. Comme précédemment lorsque deux gènes sont tirés au hasard chez une mère, ils proviennent du même père avec une probabilité Q_0 et d'un père différent avec une probabilité $1 - Q_0$. Mais dans le cas de marqueurs haploïdes hérités paternellement, si ces deux gènes sont issus du même père, ils seront obligatoirement identiques puisque chaque père ne possède qu'une copie de ce gène. Ainsi nous obtenons h_f^{haplo} telle que :

$$h_f^{haplo} = (1 - Q_0).h_p. \tag{5.5}$$

De même pour l'équation 5.3, le premier terme disparaît et nous obtenons :

$$h_t^{haplo} = (1 - Q_t).h_p. \tag{5.6}$$

Nous pouvons donc remplacer h_f^{haplo} et h_t^{haplo} par leurs expressions dans l'équation 5.1, nous obtenons alors la formule du Φ_{ft}^{haplo} :

$$\Phi_{ft}^{haplo} = \frac{Q_0 - Q_t}{1 - Q_t}. \tag{5.7}$$

Quelque soient les marqueurs utilisés, si la distance moyenne entre les mères $\overline{z}$ est suffisamment grande (de l'ordre de 5 fois la distance moyenne de dispersion d'après Austerlitz et Smouse (2001)) alors Q_t devient négligeable devant Q_0 et 2, on obtient alors :

$$\Phi_{ft}^{diplo} \simeq \frac{Q_0}{2} = \frac{1}{2.N_{ep}} \tag{5.8}$$

$$\Phi_{ft}^{haplo} \simeq Q_0 = \frac{1}{N_{ep}}. \tag{5.9}$$

avec N_{ep}, le nombre efficace de pères, i.e. le nombre de pères équicontributeurs et non apparentés qui dans un régime panmictique créeraient une structuration du nuage pollinique identique à celle observée.

L'approche temporelle du Φ_{ft}. Le but de notre étude est de comprendre comment se structure le nuage pollinique dans l'espace mais également comment cette diversité se structure dans le temps, afin de comprendre si le déséquilibre des contributions à la reproduction est stable d'une année à l'autre, ou au contraire s'il varie. Irwin et al. (2003) ont utilisé un modèle d'AMOVA (Excoffier et al. 1992), dans lequel ils prennent en compte la variabilité spatiale (i.e. inter-mères) ainsi que la variabilité temporelle (i.e. inter-années). Ils ont utilisé le modèle classique de l'AMOVA, qui est construit de manière hiérarchique et dans lequel on peut intégrer 3 niveaux de variabilité : i) entre les groupes, ii) entre les populations au sein des groupes et iii) entre les individus au sein des populations. Ainsi, ces trois niveaux de variabilités sont devenus i) les mères, ii) les années au sein des mères et iii) les grains de pollen au sein des années et au sein des mères. De cette façon il est possible d'estimer la variabilité temporelle du nuage pollinique reçu par les mères. Ils obtiennent alors une estimation de la variation inter-mères (Φ_{ft} globaux), et de la variation génétique inter-annuelle intra-mère du nuage pollinique (Φ_{yf}) exprimées de la façon suivante :

$$\Phi_{ft} = \frac{\widehat{\sigma_f^2}}{\widehat{\sigma_f^2} + \widehat{\sigma_y^2} + \widehat{\sigma_w^2}} \tag{5.10}$$

et

$$\Phi_{yf} = \frac{\widehat{\sigma_y^2}}{\widehat{\sigma_y^2} + \widehat{\sigma_w^2}}, \tag{5.11}$$

avec $\widehat{\sigma_f^2}$, $\widehat{\sigma_y^2}$ et $\widehat{\sigma_w^2}$ étant respectivement les variances estimées entre les mères, entre les années au sein des mères et entre les grains de pollen au sein des années et des mères. Avec cette analyse, Irwin et al. (2003) montrent que lorsque la structuration génétique du nuage pollinique est estimée lors d'un seul épisode de reproduction, la différenciation inter-mères estimée est biaisée et confond alors la variation temporelle et spatiale. Nous avons appelé ce paramètre Φ_{ft}^{nc} (pour non corrigé), il est défini de

la façon suivante :

$$\Phi_{ft}^{nc} = \frac{\widehat{\sigma_f^2} + \widehat{\sigma_y^2}}{\widehat{\sigma_f^2} + \widehat{\sigma_y^2} + \widehat{\sigma_w^2}} \tag{5.12}$$

Ce paramètre correspond à l'espérance du Φ_{ft} mesurée lors d'un seul épisode de reproduction, et qui confond alors la variance inter-mères avec la variance inter-annuelle.

La décomposition de la variance donnée par Irwin et al. (2003) ne montre que le cas simple où les effectifs sont équilibrés pour les différentes mères et pour les différentes années (i.e. 15 mères et 40 descendants par mère pendant 4 années), or nous avons dans notre cas des effectifs déséquilibrés aussi bien entre les années qu'entre les mères. Nous avons donc appliqué les corrections suggérées par Excoffier et al. (1992) pour pouvoir calculer les composantes de la variance à partir des espérances des carrés moyens, tout en prenant compte du déséquilibre d'échantillonnage.
De plus, nous avons étendu l'approche d'Irwin et al. (2003), qui estimait uniquement le Φ_{ft} global, à l'estimation des Φ_{ft} par couple de mères. Ceci nous permet d'obtenir une estimation simultanée de plusieurs paramètres de la dispersion, qui prennent en compte la variabilité temporelle de la dispersion. Nous pouvons ainsi estimer conjointement la densité efficace du peuplement et la fonction de dispersion sans nécessairement fixer son paramètre de forme.

Gestion des données manquantes. Les données manquantes sont un problème récurrent dans les études basées sur les génotypes. Ainsi un jeu de données expérimentales est souvent incomplet et deux solutions sont alors possibles lors de son analyse : i) ne garder que les génotypes complets ou ii) utiliser l'information partielle fournie par les génotypes incomplets. Nous avons choisi cette dernière approche car elle permet d'utiliser au mieux l'information contenue au sein de l'échantillon. Dans le logiciel Arlequin (destiné au calcul d'AMOVA), Excoffier et al. (2005) traitent les données manquantes mais restent très évasifs sur l'approche choisie. Il semblerait qu'ils créent un nouvel haplotype dès qu'un génotype incomplet est rencontré considérant donc l'allèle manquant comme un nouvel allèle (observations personnelles et

déclaration faite par l'auteur dans un forum des utilisateurs de son logiciel). Ainsi, ils négligeraient la probabilité qu'un haplotype incomplet soit identique à un haplotype complet présent dans le jeu de données. Cette approche peut avoir une influence sur la diversité génétique estimée des nuages polliniques. En effet si les haplotypes incomplets sont peu fréquents relativement aux haplotypes complets alors la diversité estimée augmente artificiellement, puisque de nouveaux haplotypes sont crées alors qu'ils peuvent en réalité correspondre à des haplotypes complets déjà échantillonnés. Au contraire si les haplotypes incomplets sont très fréquents relativement aux haplotypes complets, la diversité estimée est artificiellement diminuée puisque tous les haplotypes incomplets pour les mêmes loci sont identiques entres eux, ce qui n'est pas forcément le cas en réalité. Ainsi cette approche introduit un biais dont l'effet n'est pas stable et peu prévisible sur l'estimation de la diversité génétique. De plus il est encore plus imprévisible sur l'estimation des différenciations génétiques résultant de l'AMOVA.

Nous proposons donc de ne pas créer de nouveaux haplotypes mais de distribuer les haplotypes incomplets entre les haplotypes complets compatibles et ce proportionnellement à la fréquence de ces derniers dans le nuage pollinique. Cette approche fait l'hypothèse que notre échantillonnage est suffisamment représentatif du nuage pollinique global réel, i.e. que tous les haplotypes présents ont été observés et que leurs fréquences sont correctement estimées. Cette approche est plus conservative et évite de modifier artificiellement la diversité génétique estimée.

L'approche que nous avons adoptée afin de gérer les données manquantes est similaire à celle adoptée par Smouse et al. (2001) dans le cas des graines pour lesquelles la détermination de l'haplotype paternel est ambiguë (i.e. les graines hétérozygotes pour un locus donné avec le même génotype que leurs mères). Nous illustrerons notre méthode pour le cas où les haplotypes sont obtenus grâce à trois marqueurs. Soit $h_{na} = A_1; B_2; NA_3$ un haplotype incomplet où X_n représente l'allèle X du locus n et NA_n représente une donnée manquante au locus n. Les haplotypes compatibles avec h_{na} sont l'ensemble des haplotypes observés dont les premier et second locus portent respectivement les allèles A_1 et B_2. Connaissant la fréquence des haplotypes

complets du nuage pollinique, nous pouvons donc exprimer la probabilité que h_{na} corresponde à un haplotype compatible donné en fonction de ces fréquences. Cette probabilité sera égale à la fréquence de l'haplotype compatible considéré, divisée par la somme des fréquences de l'ensemble des haplotypes compatibles (cf. tableau 3.2.1).

Nom (h_x)	Haplotype	Freq.	Compatible avec h_{na}	Pr(identique à h_x)
h_1	$A_1; B_2; C_3$	f_1	oui	$\frac{f_1}{f_1+f_2}$
h_2	$A_1; B_2; D_3$	f_2	oui	$\frac{f_2}{f_1+f_2}$
h_3	$B_1; A_2; C_3$	f_3	non	0

TAB. 5.1: Tableau représentant un exemple simple où seulement 3 haplotypes sont présents dans le nuage pollinique. Sont représentés : le génotype, sa fréquence, si celui-ci est compatible avec l'haplotype $h_{na} = A_1; B_2; NA_3$ et enfin la probabilité que h_{na} soit identique à cet haplotype complet.

Pour décrire le calcul de la distance entre deux haplotypes, nous représenterons les haplotypes des graines sous forme de vecteurs. Soit H un vecteur dont chaque élément i représente un haplotype complet observé. Sa taille correspond donc au nombre total d'haplotypes complets observés. Soit Y_{h_x} un vecteur de même taille représentant l'haplotype h_x défini tel que $Y_{h_{x_i}} = Pr(h_{x_i} = H_i)$. Ainsi en reprenant l'exemple du tableau 3.2.1, l'ensemble des haplotypes complets observés est h_1, h_2, h_3, avec l'haplotype $h_1 = A_1; B_2; C_3$. L'haplotype d'une graine sans donnée manquante, correspondant à l'haplotype h_1 sera donc représenté par le vecteur $[1 \quad 0 \quad 0]$. La distance au carré $\delta^2_{h_x h_y}$ entre les haplotypes de deux graines x et y, est définie dans Smouse et al. (2001) comme étant égale à :

$$\delta^2_{h_x h_y} = \frac{1}{2}||Y_{h_x} - Y_{h_y}||^2$$

où $||v||^2$ est la somme de tous les éléments élevés aux carrés de v. Ainsi en l'absence de donnée manquante, $\delta^2_{h_x h_y}$ vaut 0 si les haplotypes sont identiques et 1 s'ils sont différents. Dans le cas où l'haplotype de la graine présente des données manquantes, nous effectuons le même calcul. Ainsi avec l'exemple précédent de h_{na} :

$$Y_{h_{na}} = \begin{bmatrix} \frac{f_1}{f_1+f_2} & \frac{f_2}{f_1+f_2} & 0 \end{bmatrix} \tag{5.13}$$

La distance entre h_1 et $h_n a$ sera donc calculées ainsi :

$$[Y_{h_{na}} - Y_{h_1}] = \begin{bmatrix} \frac{f_1}{f_1+f_2} & \frac{f_2}{f_1+f_2} & 0 \end{bmatrix} - [1 \quad 0 \quad 0] = \begin{bmatrix} \frac{f_2}{f_1+f_2} & \frac{f_2}{f_1+f_2} & 0 \end{bmatrix}$$

$$\delta^2_{h_{na}h_1} = \frac{1}{2}||Y_{h_{na}} - Y_{h_1}||^2 = \left(\frac{f_2}{f_1+f_2}\right)^2. \tag{5.14}$$

C'est alors cette distance entre haplotypes que nous utilisons au sein de l'AMOVA plutôt que celle proposée par Excoffier et al. (1992), afin d'en extraire les Φ_{ft} globaux ou par paires (Φ_{ft-pw}) intégrant ou non l'effet année. Après avoir calculé ces distances entre les haplotypes de toutes les graines, nous utilisons la décomposition des sommes des carrés des distance d'après l'AMOVA (Excoffier et al. 1992) pour estimer les différentes variances intra et inter groupes.

Estimation de la dispersion totale du pollen et de la dispersion issue de l'allopollen En complément de l'estimation classique de la dispersion, effectuée sur l'ensemble des graines échantillonnées, nous avons effectué une estimation de la dispersion de l'allopollen en appliquant la méthode détaillée plus haut sur un jeu de données réduit. Nous avons retiré du jeu de données toutes les graines potentiellement issues de l'autofécondation, c'est à dire toutes celles dont l'haplotype correspond à celui de leur mère. En réalité nous enlevons donc de notre jeu de données les graines réellement issues de l'autofécondation et celles issues de l'allofécondation dont l'haplotype correspond à celui de la mère. Cette méthode conduit à des estimations plus robustes et moins biaisées lorsque l'espèce étudiée pratique l'autofécondation, et a été justifiée théoriquement et par simulations (Carpentier et al. in prep).

Estimation des paramètres de dispersion La méthode TwoGener ainsi modifiée pour prendre en compte les données haplotypiques avec des données manquantes et les variations inter-annuelles de la composition des nuages polliniques nous a permis d'estimer différents paramètres liés à la dispersion que nous allons décrire ici.

Les différents modèles Nous avons utilisé trois modèles différents, en considérant :

- des événements de dispersion annuels indépendants i.e. en effectuant une estimation de tous les paramètres par année,
- l'estimation de la dispersion multi-annuelle sans prise en compte de la variabilité inter-annuelle (i.e. en utilisant les Phi_{ft}^{nc} non corrigés présentés dans l'équation 5.12),
- l'estimation de la dispersion multi-annuelle en prenant en compte la variabilité inter-annuelle du nuage pollinique de chaque mère (i.e. les Φ_{ft} présentés dans l'équation 5.10).

Pour les deux dernières approches une seule fonction de dispersion est estimée.

Les paramètres estimés Nous avons estimé différents paramètres pour chaque modèle en estimant tour à tour la dispersion "totale" puis l'allodispersion (i.e. en excluant l'autopollen). Les paramètres estimés sont :

- **le nombre de pères efficaces**, N_{ep} calculés comme étant égal à l'inverse du Φ_{ft} global d'après l'équation 5.9,
- **les paramètres de la fonction de dispersion et la densité efficace du peuplement** obtenus par minimisation de la somme des carrés des écarts entre les Φ_{ft} par couple théoriques et les Φ_{ft} par couple observés. Afin de modéliser la courbe de dispersion du pollen nous avons choisi des fonctions de la famille des exponentielles puissances, caractérisées par deux paramètres : a le paramètre d'échelle et b le paramètre de forme (Austerlitz et al. 2004). Ces fonctions sont définies de la façon suivante :

$$\gamma^{ep}(x, y; a, b) = \frac{b}{2\pi a^2 \Gamma(2/b)} \exp\left(-\left(\frac{\sqrt{x^2+y^2}}{a}\right)^b\right),$$

  Elles ont pour moyenne $a\frac{\Gamma(\frac{3}{b})}{\Gamma(\frac{2}{b})}$. Nous avons estimé simultanément le paramètre a et la densité efficace d_e, en fixant ou en estimant conjointement le paramètre b. Nous avons fixé b à trois valeurs :

  - $b = 2$, la fonction de dispersion est alors une loi normale,
  - $b = 1$, on obtient alors une fonction exponentielle et,

- $b = 0.5$, la fonction de dispersion est alors à queue lourde (i.e. avec une décroissance plus lente qu'une fonction exponentielle).

Les paramètres des fonctions de dispersion estimées (a, b et d_e) étant difficilement interprétables biologiquement nous les avons convertis en paramètres plus aisément représentables :

- Q_0^γ, la probabilité de copaternité obtenue à partir de la fonction de dispersion,
- $R_{0.95}^\gamma$, le rayon efficace à 95%,c'est dire le rayon du disque centré sur une mère et comprenant 95% des pères,
- $N_{ep0.95}^\gamma$, le nombre de pères efficaces à 95%, c'est dire le nombre de père dans le disque défini précédemment.

En désignant la densité efficace par d_e et le rayon du cercle centré autour d'une mère par r, ces paramètres sont calculés de la manière suivante : D'après (Austerlitz et Smouse 2001),

$$Q_0^\gamma = \frac{1}{d_e} \int \int f^2(x,y)dxdy \tag{5.15}$$

$R_{0.95}^\gamma$ est tel que :

$$\int_0^{R_{0.95}^f} 2\pi r \gamma(r) dr = 0.95 \tag{5.16}$$

enfin

$$N_{ep0.95}^\gamma = d\pi {R_{0.95}^\gamma}^2 \tag{5.17}$$

3.2.2 Estimation du régime de reproduction par les modèles d'appariement

Présentation du modèle utilisé Le modèle précedent ne permet d'estimer que des paramètres commun à toutes les mères. Afin d'estimer et d'expliquer les variations inter-individuelles de composition génétique du nuage pollinique échantillonné et du régime de reproduction, nous avons effectué des estimations par mères. Sur les bases des modèles d'appariement, nous avons développé une nouvelle approche qui estime les probabilités d'autofécondation et de copaternité de chaque mère échantillonnée en se basant sur une vraisemblance par paires de graines (*Pairwise Likelihood*). De plus, nous avons modifié cette approche pour qu'elle puisse s'appliquer à des marqueurs haplotypiques. Comme dans le modèle développé par Ritland et Jain (1981), Ritland (1986) dans MLTR et celui exposé au chapitre 3, nous considérons un nuage pollinique commun et homogène pour toute la population de

mères échantillonnées. Les marqueurs sont haplotypiques à hérédité paternelle et les événements d'autofécondation sont considérés indépendants. Nous avons modélisé les taux d'autofécondation individuels ainsi que les probabilités de copaternité entre graines allofécondées d'une même mère, r_p, (i.e. la probabilité que deux graines allofécondées d'une mère soient issues du même père) et ce pour chacune des mères.

Il est donc possible d'écrire la vraisemblance des haplotypes d'un couple de graines d'une mère donnée en la décomposant en fonction des différents événements ayant pu conduire à l'obtention de ce couple de graines (i.e. pollen provient de la mère, ou d' autres arbres identiques ou non). L'ensemble de ces événements ainsi que leurs probabilités sont présentés dans le tableau 5.2. La vraisemblance, $\mathcal{L}^m_{o,o'}$, de deux graines g^m_o et $g^m_{o'}$ issues de la mère m s'obtient alors en additionnant les probabilités de chacun des événements présentés dans le tableau 5.2.

L'utilisation de la corrélation de paternité de cette façon ne nous permet pas de modéliser la relation qu'il existe entre plus de deux graines d'une même mère (e.g. nous n'avons pas accès à la probabilité pour trois graines allofécondées d'être issues du même père). Ainsi, le choix de ce modèle simple nous permet de limiter à la fois le nombre d'hypothèses faîtes sur la fécondation et le nombre de paramètres à estimer. En contre-partie, nous ne pouvons pas écrire la vraisemblance exacte des génotypes échantillonnés et utiliser les estimateurs classiques du maximum de vraisemblance. En effet la vraisemblance totale du jeux de données est le produit des probabilités toutes les descendances. Or la vraisemblance de chaque descendance est la probabilité du n-uplet composé des n graines que contient la descendance, et dans notre approche nous ne décrivons que les probabilités de chaque couple de graines de la descendance. Ritland (1989) utilise un modèle similaire mais avec une corrélation de paternité identique pour toutes les mères et il a été confronté aux mêmes limites d'un modèle définissant seulement la relation entre deux graines. Ils ont pour cela utilisé une approximation de la vraisemblance basée sur les cumulants d'ordre 1 et 2 (i.e. fonction des vraisemblances d'une ou de deux graines). Nous avons développé une autre approche reposant sur l'utilisation de la vraisemblance des paires de graines (*Pairwise likelihood*). Cette approche est justifiée par des calculs de

convergence (non présentés ici) montrant que lorsque le nombre de graines augmente (et donc le nombre de couples de graines également), les estimateurs du maximum de vraisemblance par paire convergent rapidement vers ceux de la vraisemblance exacte. Finalement la vraisemblance par paire utilisée pour l'estimation est :

$$P\mathcal{L} = \prod_{m \in M} \prod_{\{o,o'\} \in O_m} \mathcal{L}^m_{o,o'}, \tag{5.18}$$

où M est l'ensemble des mères échantillonnées et O_m, l'ensemble des graines échantillonnées de la mère m. Les paramètres estimés sont s_m et r_{p_m}, respectivement les taux d'autofécondation et les corrélations de paternité pour chaque mère m, et f_h, les fréquences dans le nuage pollinique fécondant de chaque haplotype h observé.

Evénements de fécondation possible pour 2 graines	Probabilités
g^m_o et $g^m_{o'}$ sont autofécondées	$\mathbb{1}_{\{h_o=h_m\}}\mathbb{1}_{\{h_{o'}=h_m\}}s^2_m$
g^m_o est autofécondée et $g^m_{o'}$ est allofécondée	$\mathbb{1}_{\{h_o=h_m\}}s_m(1-s_m)f_{h_{o'}}$
g^m_o est allofécondée et $g^m_{o'}$ est autofécondée	$(1-s_m)f_{h_o}\mathbb{1}_{\{h_{o'}=h_m\}}s_m$
g^m_o et $g^m_{o'}$ sont allofécondées issues :	
du même père	$(1-s)^2 r_{p_m}\mathbb{1}_{\{h_o=h_{o'}\}}f_{h_o}$
de pères différents	$(1-s)^2(1-r_{p_m})f_{h_o}f_{h_{o'}}$

TAB. 5.2: **Tableau des probabilités des paternités d'un couple de graines (g^m_o et $g^m_{o'}$) issue de la mère m.** h_o, $h_{o'}$ et h_m représentent les haplotypes paternellement hérités de g^m_o et $g^m_{o'}$ et l'haplotype de la mère respectivement. f_{h_o}, $f_{h_{o'}}$ sont les fréquences haplotypiques des haplotypes h_o et $h_{o'}$ dans le nuage pollinique global de la population. s_m et r_{p_m} sont respectivement la probabilité d'autofécondation et la corrélation de paternité dans la descendance de la mère m. On définit $\mathbb{1}_{\{x\}}$ égal à 1 si la condition $\{x\}$ est vérifiée, égal à 0 sinon.

Estimation des paramètres par l'algorithme Expectation-Maximization

L'algorithme EM est un algorithme itératif de maximisation. Il a été originellement développé pour les données dites "incomplètes" et est très robuste (Dempster et al. 1977). Son utilisation dans le cas de l'estimation de l'autofécondation est très fréquente (Cheliak et al. (1983), Ritland et Jain (1981), voir Ritland (1986) pour une comparaison de cet algorithme à celui de Newton-Raphson avec des modèles

semblables). Cet algorithme a été utilisé et présenté brièvement au chapitre 3 pour l'estimation des taux d'autofécondation en absence de corrélation de paternité. Nous le présentons ici de manière plus détaillée.

L'algorithme de l'EM se décompose en trois étapes : (i) l'initialisation des paramètres à estimer, (ii) le calcul de l'espérance de la vraisemblance d'après les valeurs données aux paramètres connaissant les données observées, (iii) l'estimation de nouvelles valeurs pour les paramètres en maximisant l'espérance de la vraisemblance précédemment calculée. Les étapes (ii) et (iii) sont répétées jusqu'à ce que le gain de vraisemblance obtenu entre deux itérations soit considéré comme négligeable. Sous un certain nombre d'hypothèses de régularité pour le modèle on peut montrer que l'estimation a alors convergé vers les estimateurs du maximum de vraisemblance. L'utilisation de l'EM est pratique lorsque les estimateurs du maximum de vraisemblance peuvent être analytiquement calculés à partir de variables non observées mais dont les espérances peuvent être simplement exprimées en fonction des paramètres du modèle. Dans ce cas l'étape (ii) consiste à calculer l'espérance de ces variables "cachées" sous les paramètres du modèle, et lors de l'étape (iii) de nouveaux paramètres sont estimées à partir des espérances de ce modèle. Dans notre cas, les variables estimées sont les probabilités d'autofécondation et de copaternité de toutes les mères échantillonnées et les fréquences de chaque haplotype potentiellement identifiable avec les marqueurs utilisés au sein du nuage pollinique commun. Les variables cachées sont le nombre de couples de graines participant à chacun des types d'événements de fécondation possibles détaillés dans le tableau 5.2. Les espérances de ces variables cachées sont obtenues en calculant pour chaque couple de graines de chaque mère les probabilités des différents événements de fécondation (tableau 5.2). Les nouveaux estimateurs des paramètres sont alors des ratios entre les espérances des variables cachées, et correspondent aux estimateurs du maximum de vraisemblance (l'estimateur du maximum de vraisemblance de la probabilité d'un événement suivant une loi binomiale est le ratio entre le nombre d'événements réalisés sur le nombre total de tirages).

Gestion des données manquantes Pour prendre en compte l'incertitude engendrée par les données manquantes de certains des haplotypes des graines, nous avons remplacé dans les probabilités des événements de fécondation (tableau 5.2), l'indicatrice de l'égalité de deux haplotypes h_i,h_j ($\mathbb{1}_{\{h_i=h_j\}}$) par sa probabilité. Cette méthode est la même que celle décrite pour la gestion des données manquantes lors des estimations du Φ_{ft} (cf. 3.2.1 à la page 160).

Analyse de la variabilité temporelle. Nous avons utilisé trois modèles différents afin d'évaluer la variabilité temporelle du nuage pollinique global et du régime de reproduction :

- Le modèle 1 : une estimation par année, avec un nuage pollinique différent chaque année, ainsi qu'un s et un r_p pour chaque mère et chaque année.
- Le modèle 2 : une estimation par mère et par année des s et des r_p mais en considérant un nuage pollinique commun pour toutes les années (i.e. les fréquences haplotypiques ne sont estimées qu'une seule fois sur l'ensemble du jeu de données).
- Le modèle 3 : une estimation par mère des s et des r_p commune pour les 4 années, en considérant un seul nuage pollinique

Estimation des intervalles de confiance. Les intervalles de confiance des estimations de s et de r_p ont été réalisés par une procédure de bootstrap qui consiste à tirer avec remise des couples de graines au sein des descendances maternelles puis à réestimer les paramètres. Nous avons effectué 1000 bootstraps par analyse, et ensuite gardé l'intervalle à 95% des valeurs prises par les paramètres lors des différents ré-échantillonnages.

Estimation de la probabilité de copaternité globale Afin de pouvoir comparer les estimations de la corrélation de paternité au sein des descendances avec celles obtenues par l'approche TwoGener, il est nécessaire de modifier le r_p obtenu. En effet cette dernière valeur représente la proportion de pleins frères au

sein de la fraction de la descendance (i.e. la corrélation de paternité) uniquement issue de l'allofécondation, ce qui n'est pas le cas du Q_0 ou du Q_0^γ, obtenus par la méthode TwoGener, qui eux représentent la proportion de pleins frères au sein d'une descendance totale. Il devrait être potentiellement possible de comparer le r_p aux estimations de TwoGener obtenues sur le jeu de données sans les graines potentiellement autofécondées (i.e. sans les graines dont le génotype est le même que celui de leur mère), mais dans ce cas on sur-estime certainement le taux d'autofécondation (i.e. les génotypes descendant-mère peuvent correspondre sans résulter pour autant d'autofécondation). Ainsi à partir des estimations de s et r_p, nous pouvons obtenir la proportion totale de plein frères au sein d'une descendance. On peut la décomposer de la façon suivante :

- Les deux descendants sont issus d'autofécondation avec une probabilité s^2 et partagent donc le même père.
- Les deux descendants sont issus d'allofécondation avec une probabilité $(1-s)^2$ et partage le même père avec une probabilité r_p.

Ainsi la probabilité de copaternité totale au sein d'une descendance ($Q_0^{s-r_p}$) s'exprime de la façon suivante :

$$Q_0^{s-r_p} = s^2 + (1-s)^2 \times r_p. \tag{5.19}$$

Cette valeur est alors comparable aux Q_0 et Q_0^γ, et présente l'avantage d'être estimée au niveau individuel. Ainsi il est possible de comprendre ce qui influence la diversité génétique échantillonnée et la structuration génétique qui en découle afin de l'expliquer par des covariables au niveau individuel, tout en pouvant décomposer l'effet dû à l'autofécondation de celui dû à l'allofécondation (i.e. densité efficace et/ou dispersion limitée).

Analyses statistiques : comparaison des modèles et lien avec les paramètres locaux Dans l'approche par maximum de vraisemblance nous avons intégré l'aspect temporel dans les estimations des taux d'autofécondation, s et des corrélations de paternité, r_p, en estimant ces paramètres selon trois modèles différents. Ces modèles sont emboîtés, du modèle 1 (le plus complet) au modèle 3 (celui avec le moins de

paramètres). Il est donc possible de comparer ces modèles deux à deux en comparant le gain de vraisemblance apporté par l'ajout de paramètres. Nous avons utilisé le test du rapport de vraisemblance classique, en considérant que les rapports des vraisemblances par paires utilisées (*Pairwise likelihood*) étaient des estimations correctes des rapports des vraisemblances exactes. Dans ce cas, afin de tester l'hypothèse H_0 énonçant que les données observées sont issues de $(M-)$ contre l'hypothèse H_1 énonçant qu'elles sont issues de $(M+)$, on utilise la statistique :

$$2.log\left(\frac{L(M+)}{L(M-)}\right) \quad \sim \quad \chi^2_{(p(+)-p(-))}, \qquad (sous \quad H_0) \tag{5.20}$$

avec $L(M-)$ et $L(M+)$ les vraisemblances respectives du modèle le moins complet et du plus complet et $p(+)$ et $p(-)$, les nombres de paramètres respectifs du modèle le plus et le moins complet.

Nous avons ensuite testé comment les estimations des paramètres individuels du régimes de reproduction (s, r_p et $Q_0^{s-r_p}$) dépendent de la densité locale et de la fertilité mâle individuelle. De plus nous avons testé la variabilité temporelle du régime de reproduction, en testant l'effet de l'année sur les valeurs estimées de s, de r_p et de $Q_0^{s-r_p}$. Nous avons utilisé trois modèles d'ANCOVA différents, emboîtés les uns dans les autres et nous avons appliqué une transformation angulaire aux données à expliquer, les paramètres du régime de reproduction (i.e. $arcsin\sqrt{X}$, Dagnélie vol2).

Nous allons décrire ici les trois modèles utilisés pour expliquer le taux d'autofécondation (s). Le premier modèle (S1) considère un effet qualitatif : l'année (A), et deux effets quantitatifs : l'effet de la densité locale (D) et l'effet de la production de pollen (P). Les modèles comprennent également les interactions d'ordre deux pour ces trois effets, nous estimons donc un coefficient de régression de s par la densité locale et la production de pollen par année, ainsi qu'un coefficient de régression du taux d'autofécondation par la variable synthétique $D \times P$. Ce modèle ne comprend

que des effets fixes (à part le terme d'erreur) et s'écrit de la manière suivante :

$$\begin{cases} arcsin\sqrt{s_{ij}} = A_i + \beta_1 D_{ij} + \beta_2 P_{ij} + \beta_{1i} D_{ij} + \beta_{2i} P_{ij} + \beta_3 (D_{ij} \times P_{ij}) + \epsilon_{ij} \\ \epsilon ij \sim \mathcal{N}(0, \sigma_1^2) \\ i = 1, ...4 \quad j = 1, ...Nb.meres, \end{cases} \tag{5.21}$$

avec i l'année, j la mère échantillonnée, β_x les coefficients de régression et ϵ le terme d'erreur. Le second modèle (S2) reprend les effets du précédent en y ajoutant un effet individuel aléatoire (I). Il s'écrit de la façon suivante :

$$\begin{cases} arcsin\sqrt{s_{ij}} = A_i + I_j + \beta_1 D_{ij} + \beta_2 P_{ij} + \beta_{1i} D_{ij} + \beta_{2i} P_{ij} + \beta_3 (D_{ij} \times P_{ij}) + \epsilon_{ij} \\ I_j \sim \mathcal{N}(0, \sigma_1^2), \quad \epsilon ij \sim \mathcal{N}(0, \sigma_2^2), \\ i = 1, ...4, \quad j = 1, ...Nb.meres, \end{cases} \tag{5.22}$$

Enfin le dernier modèle (S3) reprend le modèle S2 mais en considérant un effet aléatoire différent pour chaque année, tel que :

$$\begin{cases} arcsin\sqrt{s_{ij}} = A_i + I_{j|i} + \beta_1 D_{ij} + \beta_2 P_{ij} + \beta_{1i} D_{ij} + \beta_{2i} P_{ij} + \beta_3 (D_{ij} \times P_{ij}) + \epsilon_{ij} \\ I_{j|i} \sim \mathcal{N}(0, \sigma_i^2), \quad \epsilon ij \sim \mathcal{N}(0, \sigma_1^2), \\ i = 1, ...4, \quad j = 1, ...Nb.meres, \end{cases} \tag{5.23}$$

avec la même nomenclature que pour le modèle S1. Nous avons utilisé des modèles similaires pour les corrélations de paternité (r_p et $Q_0^{s-r_p}$), que nous avons appelés R1, R2 et R3 pour r_p et Q1, Q2 et Q3 pour $Q_0^{s-r_p}$ respectivement. Les effets aléatoires ont été testés en comparant les modèles entres eux par un rapport de vraisemblance dont le test est décrit par l'équation 5.20.

4 Résultats

4.1 Estimation de la structure génétique spatiale du nuage de pollen et de la dispersion pollinique

4.1.1 Approche globale

Nous avons appliqué l'approche TwoGener globale pour chaque année entre 2002 et 2005, mais aussi en considérant toutes les années aggrégées. On observe que le nuage pollinique est beaucoup moins différencié entre les différentes mères en 2003 et en 2005, qu'il ne l'est en 2002 et en 2004, avec des valeurs de Φ_{ft}, plus faibles et plus fortes respectivement (tableau 5.3). Ceci aboutit à des nombres de pères efficaces plus faibles en 2002 et 2004 (4,26 et 4,03 respectivement), qu'en 2003 et 2005 (6,97 et 9,45 respectivement). Lorsque toutes les années sont regroupées on observe une différenciation globale plus faible du nuage pollinique échantillonné par les mères avec un $\Phi_{ft} \simeq 12\%$, ce qui résulte en un nombre moyen de pères efficaces de $N_{ep} = 8,26$.

Année	d.d.l. among	d.d.l. within	SSA	SSW	σ_a^2	σ_w^2	Φ_{ft}	N_{ep}
2002	41	380	59,67	133,35	0,11	0,35	**0,23**	**4,26**
2003	19	342	31,25	134,39	0,07	0,39	**0,14**	**6,97**
2004	35	310	50,44	104,97	0,11	0,34	**0,25**	**4,03**
2005	29	259	26,36	108,21	0,05	0,42	**0,11**	**9,45**
Toutes	45	1373	99,20	567,02	0,06	0,41	**0,12**	**8,26**

TAB. 5.3: Tableau des résultats obtenus lors des analyses par la méthode TwoGener globale année par année et également si l'on considère toutes les années mélangées. La dernière colonne représente le nombre de pères efficaces déduits des Φ_{ft} ainsi obtenus.

Nous avons également utilisé la méthode d'AMOVA hiérarchisée afin d'estimer les variabilités inter-mères et inter-annuelles de la composition génétique du nuage pollinique (tableau 5.4). A partir de ce tableau on peut estimer le Φ_{ft}, c'est à dire la proportion de variance expliquée par les mères (equation 5.10), soit $7,83\%$, ce qui résulte en nombre de pères efficaces (N_{ep}) de 12,77. Cette dernière valeur est plus élevée que celles estimées annuellement entre 2002 et 2005 et en mélangeant les an-

nées sans considérer la variabilité inter-annuelle. Il est également possible d'exprimer la variation inter-annuelle du nuage pollinique au sein d'une mère d'après l'équation 5.11, ce qui donne $\Phi_{yf} = 13,9\%$. Ainsi, la variabilité de la composition génétique du nuage de pollen semble bien plus importante entre les années au sein d'une mère qu'entre les mères. Enfin on peut extraire de cette AMOVA le paramètre Φ_{ft}^{nc} qui représente l'espérance du Φ_{ft} mesuré une année donnée et qui ne tiendrait donc pas compte de la variabilité inter-annuelle (Irwin et al. 2003), ainsi d'après l'équation 5.12, $\Phi_{ft}^{nc} = 20,65\%$. Ceci se traduit par nombre de pères efficaces de $N_{ep}^{nc} = 4.84$.

Source de variation	d.d.l.	S.C.E.	Carrés moyens	Variance	% σ_{tot}^2
Inter-mères	45	99,2	2,2	0,04	7,8
Inter-annuelle au sein des mères	82	84,3	1,0	0,06	12,8
Au sein des années	1291	482,7	0,37	0,37	79,4

TAB. 5.4: Tableau des résultats de l'AMOVA hiérarchisée appliquée à l'ensemble du jeu de données, afin d'estimer la structuration génétique globale du nuage pollinique d'un point de vue spatial et temporel.

4.1.2 Approche par couples de mères, Φ_{ft-pw}

Les fonctions estimées annuellement montrent que la dispersion est relativement limitée avec une distance moyenne de dispersion, δ, inférieure à 100m pour les fonctions normale, exponentielle et exponentielle puissance (tableau 5.5). Les fonctions de dispersion ainsi que les densités efficaces estimées sont très variables entre les années. Plus la queue de dispersion est lourde (i.e. plus b est faible) et plus la distance moyenne estimée (δ) est importante (tableau 5.5). Ainsi on observe avec ce modèle de très fortes variations de la distance de dispersion et de la densité efficace. Cette dernière est très différente de la densité réelle mesurée qui est de 186 producteurs de pollen/ha avec des rapports $fracd_{obs}d_e$ compris entre 2,4 et 25 (tableau 5.5). Nous avons écarté certaines estimations des interprétations quand les valeurs de d_e étaient aberrantes, soit trop faibles soit trop fortes. Ainsi, les estimations de l'année 2005 avec la fonction normale et la fonction exponentielle puissance et de l'année 2004 avec la fonction exponentielle puissance, dont les densités efficaces et

les distances moyennes de dispersion étaient aberrantes ne seront pas abordées lors des interprétations (tableau 5.5). L'estimation de la fonction de dispersion selon un noyau exponentiel puissance non contraint s'avère donc difficile, probablement car il y a un paramètre supplémentaire à estimer. Les sommes des écarts au carré sont assez similaires à celles obtenues pour les autres fonctions de dispersion, voir légèrement supérieures dans certains cas, ceci semble indiquer que les estimations n'ont pas convergé vers le maximum absolu, sans doute attirées par un maximum local. Ainsi l'optimisation des paramètres de la fonction de dispersion exponentielle puissance pourrait nécessiter plus de données.

Année	Type fonction	δ	b	d_e arbres/ha	SCE	Q_0^γ	N_{ep}^γ	$R_{0.95}^\gamma$	$N_{0.95}^\gamma$	$N_{0.95}^\gamma/N_{ep}^\gamma$
2002	normale	25,8	2	12,2	1,821	0,155	6,5	50,4	9,7	1,50
2003	normale	15,8	2	60,5	0,373	0,082	12,1	30,9	18,2	1,50
2004	normale	18,4	2	22,8	3,692	0,161	6,2	36,0	9,3	1,50
2005	*normale*	*5,0*	*2*	*821*	*0,336*	*0,061*	*16,4*	*9,8*	*24,6*	*1,50*
non_corrige	normale	22,4	2	12,7	1,393	0,196	5,1	43,8	7,7	1,50
multi_annees	normale	35,6	2	10,9	1,626	0,091	11,0	69,5	16,5	1,50
2002	exp.	31,0	1	10,7	1,379	0,155	6,5	73,5	18,2	2,81
2003	exp.	18,3	1	57,8	0,373	0,082	12,1	43,3	34,1	2,81
2004	exp.	17,4	1	32,7	3,696	0,161	6,2	41,2	17,5	2,81
2005	exp.	42,6	1	14,2	0,337	0,062	16,2	101,0	45,5	2,81
non_corrige	exp.	28,4	1	10,1	1,379	0,196	5,1	67,3	14,3	2,81
multi_annees	exp.	44,8	1	8,72	1,633	0,091	11,0	106,4	31,0	2,81
2002	exp. puiss. $b = 0,5$	53,7	0,5	7,43	1,853	0,155	6,5	161,6	60,9	9,41
2003	exp. puiss. $b = 0,5$	29,7	0,5	45,6	0,373	0,082	12,1	89,3	114,2	9,41
2004	exp. puiss. $b = 0,5$	16,2	0,5	79,0	3,699	0,161	6,2	48,6	58,5	9,41
2005	exp. puiss. $b = 0,5$	77,3	0,5	8,95	0,334	0,062	16,1	232,4	152,0	9,41
non_corrige	exp. puiss. $b = 0,5$	51,7	0,5	6,31	1,368	0,197	5,1	155,4	47,9	9,41
multi_annees	exp. puiss. $b = 0,5$	94,7	0,5	4,06	1,642	0,091	11,0	284,7	103,3	9,41
2002	exp. puiss.	*	*	*	*	*	*	*	*	*
2003	exp. puiss.	25,6	0,58	49,5	0,373	0,082	12,1	73,0	82,92	6,84
2004	*exp. puiss.*	*37,1*	*0,12*	*7850*	*3,700*	*0,161*	*6,2*	*150,3*	*55695,4*	*8957,66*
2005	*exp. puiss.*	*3945,4*	*0,16*	*0,171*	*0,325*	*0,065*	*15,3*	*16204,4*	*14081,5*	*920,99*
non_corrige	exp. puiss.	91,6	0,35	4,09	1,366	0,197	5,1	311,0	124,24	24,45
multi_annees	exp. puiss.	131,9	0,42	2,87	1,642	0,091	10,9	421,7	160,13	14,65

TAB. 5.5: Estimation des paramètres des fonctions de dispersion du pollen pour différents noyaux de dispersion : normal, exponentiel, exponentiel puissance contraint (b=0,5) et exponentiel puissance. Les estimations ont été effectuées à partir des données par année et avec le modèle temporel prenant ou non en compte la variabilité inter-annuelle. En plus des paramètres des fonctions de dispersion (δ, b, d_e) et de leur ajustement (SCE), figurent la densité efficace en contributeurs polliniques, la probabilité de copaternité basée uniquement sur la fonction (Q_0^γ), le nombre de pères efficaces qui en résulte (N_{ep}^γ), le rayon dans lequel se trouvent 95% des pères potentiels ($R_{0.95}^\gamma$), le nombre de pères dans ce rayon ($N_{0.95}^\gamma$) et enfin le rapport entre ces deux dernières valeurs. En italique figurent les estimations aberrantes.

Les estimations intégrant plusieurs années de dispersion présentent des distances de dispersion plus élevées que les estimations annuelles (basées sur les Φ_{ft-pw} ou les Φ^{nc}_{ft-pw}) et ce quelque soient les familles de fonction utilisées. En effet, c'est le début de la courbe de différenciation génétique des nuages polliniques en fonction de la distance qui permet d'estimer la densité et la queue de dispersion, or les figures 5.1 et 5.2, montrent que le plateau est atteint plus rapidement quand les estimations prennent en compte la variabilité temporelle (Φ^{temp}_{ft-pw}) que lorsqu'elles ne la considèrent pas (Φ^{temp}_{ft-pw}).

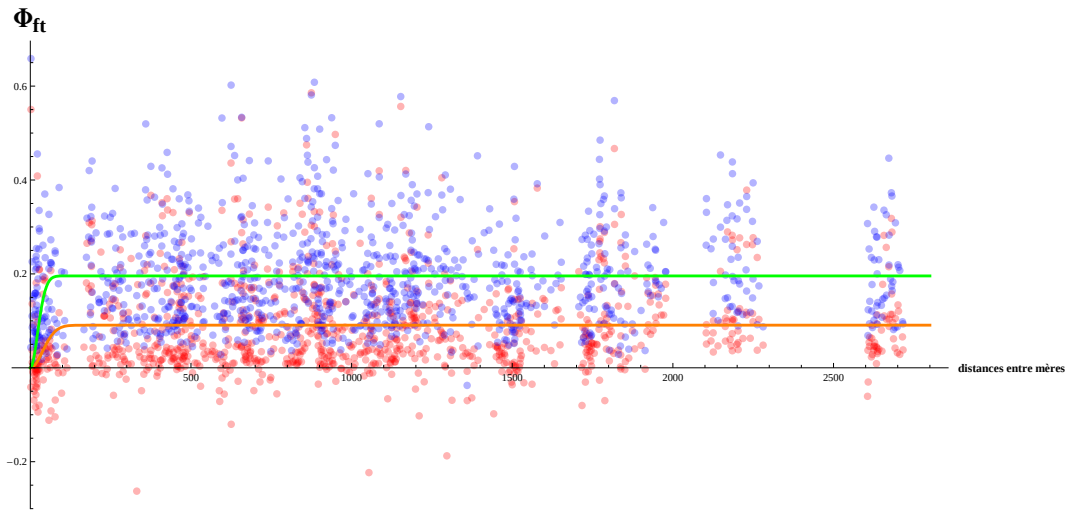


FIG. 5.1: Distribution de la valeur des Φ_{ft-pw} en fonction de la distance. Les points bleus représentent les indices calculés sans la correction (Φ^{nc}_{ft}) alors que les points rouges représentent les indices calculés avec l'estimation de la variabilité temporelle. De même les courbes vertes et oranges représentent les valeurs théoriques des Φ_{ft-pw} d'après le modèle non corrigé et corrigé respectivement pour une fonction de distribution de type X.

La probabilité de copaternité obtenue à partir des fonctions de dispersion estimées (Q^{γ}_{0}) est extrêmement stable quelque soit la famille de fonction utilisée (tableau 5.5). En effet cette estimation est basée sur la valeur du plateau observé lors de la régression des Φ_{ft-pw} par la distance, c'est à dire la valeur de la différenciation génétique du nuage pollinique de deux mères considérées comme indépendantes (voir les équations 5.7, 5.4, 5.9 et 5.8). Or quelque soit la famille de fonctions choisies le plateau est obtenu à la même valeur, en effet toutes les courbes de Φ_{ft} théoriques par

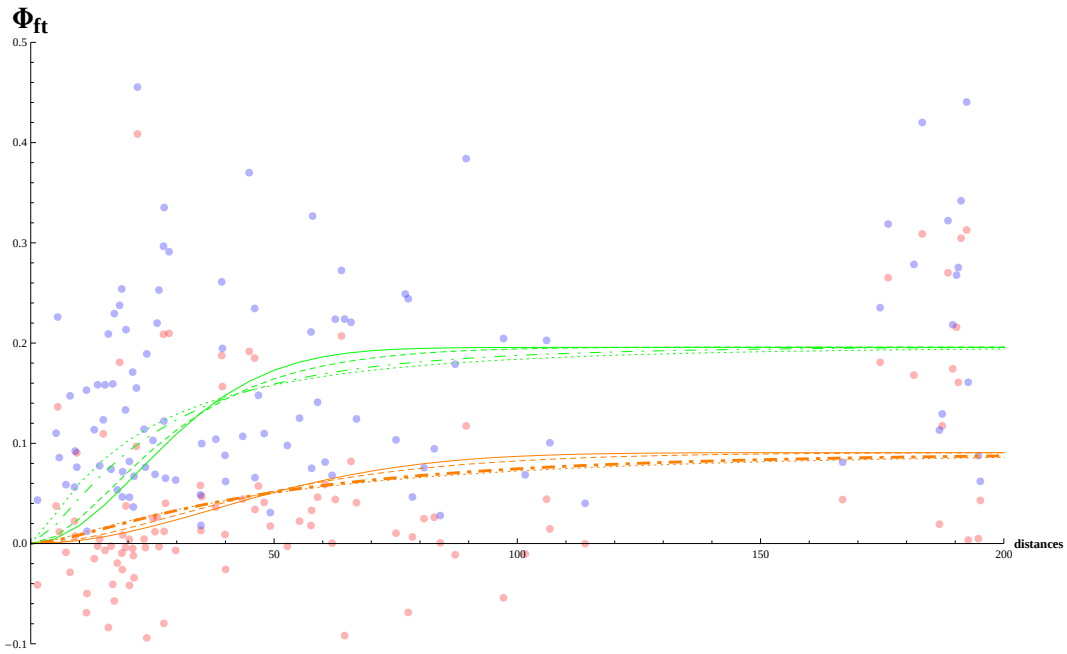


FIG. 5.2: Détail de la distribution de la valeur des Φ_{ft-pw} en fonction de la distance, pour des distances inférieures à 200 m. Les points bleus représentent les indices calculés sans la correction (Φ_{ft}^{nc}) alors que les points rouges représentent les indices calculés avec l'estimation de la variabilité temporelle. De même les courbes vertes et oranges représentent les valeurs théoriques des Φ_{ft-pw} d'après le modèle non corrigé et corrigé respectivement pour différentes lois de distribution. Trait plein = Loi normale ; Tirets = Loi exponentielle ; Tirets et pointillés = Loi exponentielle puissance contrainte avec $b = 0,5$ et Pointillés = Loi exponentielle puissance non contrainte.

paires tendent vers la même asymptote (voir figures 5.1 et 5.2). Ainsi comme pour les estimations des probabilités de copaternité basées sur le Φ_{ft} global, nous obtenons des valeurs de nombres de pères efficaces (i.e. $N_{ep}^{\gamma} = 1/Q_0^{\gamma}$) relativement faibles pour les années 2002 et 2004 (en moyenne 6,5 et 6,2 respectivement) et plus fortes pour les années 2003 et 2005 (en moyenne 12,1 et 16 respectivement). Les valeurs de $N_{0.95}^{\gamma}$ sont plus élevées que les valeurs de N_{ep}^{γ} mais suivent la même tendance que celles observées avec ces dernières à savoir un plus grand nombre de contributeurs pour les années 2003 et 2005 que pour les années 2002 et 2004. La valeur du N_{ep}^{γ} obtenue avec l'estimateur non corrigé (Φ_{ft}^{nc}, équation 5.12) est en moyenne de 5,1 contributeurs efficaces au nuage pollinique et est donc toujours inférieure aux valeurs annuelles. Le nombre de pères efficaces estimé à partir de la méthode TwoGener multi-annuelle est de $N_{ep}^{\gamma-temp} = 11$ contributeurs polliniques en moyenne, ce qui est supérieur à la

moyenne de toutes les années prises individuellement ($N_{ep}^{\gamma} = 10,2$). Cette valeur est inférieure à celle estimée avec l'approche TwoGener multi-annuelle globale qui était de $N_{ep}^{temp} = 12,7$ contre $N_{ep}^{\gamma-temp} = 11$ avec l'approche par couples de descendances. Le rapport $N_{0.95}^{\gamma}/N_{ep}^{\gamma}$ est quant à lui constant et ne semble dépendre que de la queue de dispersion (b).

Année	Type fonction	δ	b	d_e arbres/ha	SCE	Q_0^γ	N_{ep}^γ	$R_{0.95}^\gamma$	$N_{0.95}^\gamma$	$N_{0.95}^\gamma/N_{ep}^\gamma$
2002	normale	27,9	2	14,7	1,240	0,110	9,1	54,4	13,7	1,5
2003	*normale*	*4,3*	*2*	*981*	*0,325*	*0,070*	*14,3*	*8,3*	*21,4*	*1,5*
2004	*normale*	*4,8*	*2*	*392*	*2,341*	*0,138*	*7,3*	*9,4*	*10,9*	*1,5*
2005	normale	28,4	2	23,4	0,307	0,066	15,1	55,5	22,6	1,5
multi_annees	normale	20,7	2	72,7	0,401	0,040	24,8	40,3	37,2	1,5
2002	exp.	37,2	1	10,5	1,235	0,110	9,1	88,2	25,6	2,81
2003	*exp.*	*2,6*	*1*	*3340*	*0,325*	*0,070*	*14,3*	*6,2*	*40,2*	*2,81*
2004	*exp.*	*0,0*	*1*	*2,65e+11*	*2,341*	*0,138*	*7,3*	*0,0*	*20,4*	*2,81*
2005	exp.	32,3	1	23,0	0,307	0,066	15,1	76,7	42,5	2,81
multi_annees	exp.	21,2	1	88,3	0,402	0,040	24,8	50,2	69,9	2,81
2002	exp. puiss. b=0,5	81,5	0,5	4,52	1,227	0,110	9,1	244,9	85,2	9,41
2003	*exp. puiss. b=0,5*	*1,2*	*0,5*	*3,38e+4*	*0,325*	*0,070*	*14,3*	*3,6*	*134,6*	*9,41*
2004	*exp. puiss. b=0,5*	*1,4*	*0,5*	*1,28e+4*	*2,341*	*0,138*	*7,3*	*4,1*	*68,3*	*9,41*
2005	exp. puiss. b=0,5	52,6	0,5	18,1	0,307	0,066	15,1	158,1	142,0	9,41
multi_annees	exp. puiss. b=0,5	25,4	0,5	127	0,402	0,040	24,9	76,5	234,2	9,41

TAB. 5.6: Estimations des paramètres de la fonction de dispersion basées sur le jeu de données ne comprenant pas les graines potentiellement autofécondées. Les paramètres estimés sont les mêmes que dans le tableau 5.5. Les lignes en italique représentent les estimations "aberrantes".

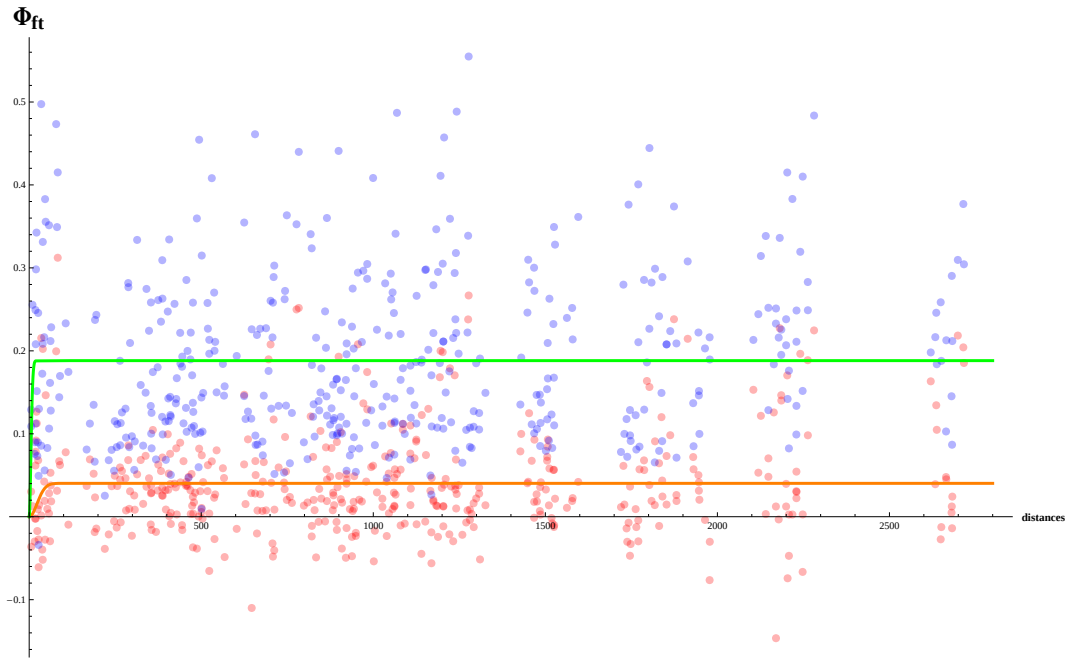


FIG. 5.3: Distribution de la valeur des Φ_{ft-pw} en fonction de la distance sur le jeu de données sans les graines potentiellement autofécondées. Les points bleus représentent les indices calculés sans la correction (Φ_{ft}^{nc}) alors que les points rouges représentent les indices calculés avec l'estimation de la variabilité temporelle. De même les courbes vertes et oranges représentent les valeurs théoriques des Φ_{ft-pw} d'après le modèle non corrigé et corrigé respectivement pour une fonction de distribution de type X.

Lorsque nous enlevons du jeu de données les graines potentiellement autofécondées et que nous estimons la dispersion du pollen, nous obtenons des résultats similaires à ceux obtenus sur le jeu de données complet. Les distances de dispersion estimées annuellement sont également assez faibles et présentent des différences avec celles estimées précédemment plus ou moins marqués : plus fortes en 2005 et plus faibles en 2002 (tableau 5.6). Les densités efficaces présentent les mêmes écarts avec les estimations basées sur le jeu de données complet. Les corrélations de paternité estimées à partir de la fonction de dispersion (Q_0^γ) sont ici encore insensibles à la famille de fonctions de dispersion choisie et sont en général inférieures à celle obtenues lorsque toutes les graines sont utilisées pour l'analyse (tableau 5.6 et figures 5.3, 5.4). En conséquence, les nombres de pères efficaces estimés N_{ep}^γ sont eux plus élevés que précédemment. Les valeurs de $N_{0.95}^\gamma$ sont également plus élevées que celles obtenues sur le jeu de données complet, ce qui semble indiquer un fort taux

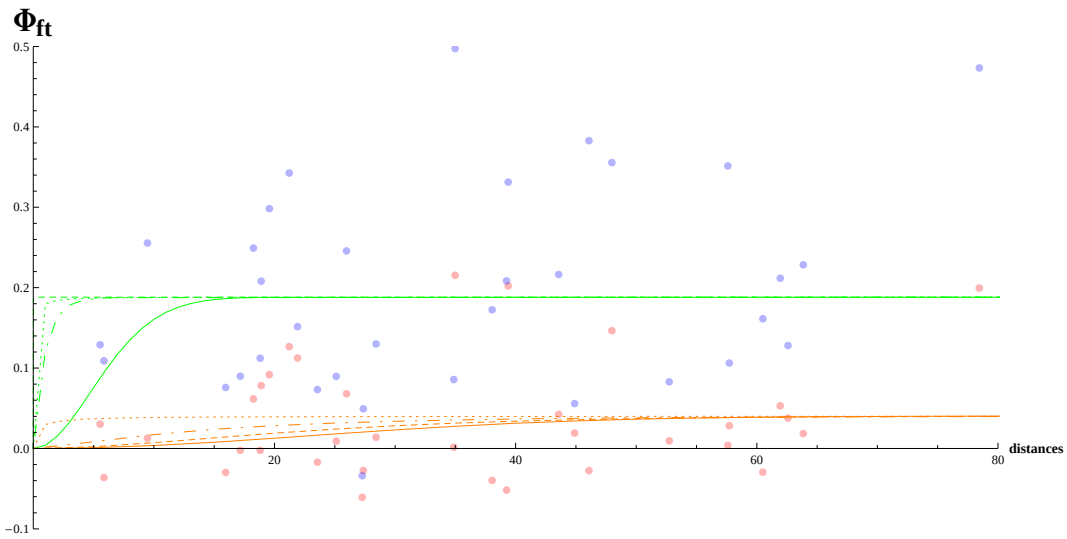


FIG. 5.4: Détail de la distribution de la valeur des Φ_{ft-pw} en fonction de la distance pour le jeu de données sans les graines potentiellement autofécondées, pour des distances inférieures à 80 m. Les points bleus représentent les indices calculés sans la correction (Φ_{ft}^{nc}) alors que les points rouges représentent les indices calculés avec l'estimation de la variabilité temporelle. De même les courbes vertes et oranges représentent les valeurs théoriques des Φ_{ft-pw} d'après le modèle non corrigé et corrigé respectivement pour différentes lois de distribution. Trait plein = Loi normale ; Tirets = Loi exponentielle ; Tirets et pointillés = Loi exponentielle puissance contrainte avec $b = 0,5$ et Pointillés = Loi exponentielle puissance non contrainte.

d'autofécondation. Les années 2002 et 2004 sont celles présentant la plus faible diversité au sein de leur nuage pollinique ($N_{ep}^{\gamma} = 9,1$ et $7,3$), alors que les années 2003 et 2005 montrent une diversité génétique des nuages polliniques plus élevée ($N_{ep}^{\gamma} = 14,3$ et $15,1$ respectivement). Ces tendances sont également observées avec les nombres de pères efficaces au sein de la zone dans laquelle s'effectue 95% de la dispersion ($N_{0.95}^{\gamma}$). Les rapports $N_{0.95}^{f}/N_{ep}^{f}$ restent uniquement dépendants de la queue de dispersion et sont donc identiques à ceux observés sur le jeu de données complet (tableaux 5.6 et 5.5). En revanche, lorsque nous considérons la variabilité temporelle, l'estimation des distances moyennes de dispersion, δ, est relativement stable, ce qui n'était pas le cas lorsque l'autofécondation n'était pas prise en compte lors des analyses (tableaux 5.6 et 5.5). Finalement, les problèmes de convergence que nous avons évoqués lors des estimations des fonctions de dispersion sur le jeu de données complet semblent également plus importants ici car le nombre de graines

sur lequel sont basées les estimations est plus faible (figures 5.3 et 5.4). Ceci est confirmé par la somme des écarts au carré qui reste constante quelque soit la fonction de dispersion ou la densité estimée (tableau 5.6). Ainsi les années 2003 et 2004 présentent des valeurs aberrantes des densités efficaces d_e et des distances moyennes de dispersions δ (tableau 5.6). Pour les mêmes raisons nous ne présentons pas les résultats des estimations basées sur la famille de fonctions des exponentielles puissances non contraintes, ni les résultats obtenus à partir des Φ^{nc}_{ft-pw}.

4.2 Estimation du régime de reproduction

4.2.1 Taux d'autofécondation individuels

Les taux d'autofécondation estimés sont très variables d'un individu à l'autre, allant d'une descendance complètement allofécondée à une descendance complètement autofécondée (figures 5.6 et 5.7). En effet les écart-types σ inter-individuelles des estimations du taux d'autofécondation sont du même ordre de grandeur et même légèrement supérieurs aux moyennes annuelles (tableaux 5.7 et 5.8). Les trois modèles d'estimations renvoient des résultats relativement similaires, avec en moyenne sur les 4 années d'études près d'un quart des descendances individuelles résultant d'événements d'autofécondation : $\overline{s_{mod1}} \simeq 0,22$ pour le modèle 1 et $\overline{s_{mod2}} \simeq \overline{s_{mod3}} \simeq 0,23$ pour les modèles 2 et 3 (tableaux 5.7 et 5.8). Les taux d'autofécondation estimés par le modèle 1 ou le modèle 2 sont variables d'une année à l'autre et présentent une alternance d'années avec de fortes valeurs en 2002 et en 2004, et des estimations bien plus faibles en 2003 et en 2005 (tableaux 5.7 et 5.8 et figure 5.5). L'estimation des taux d'autofécondation individuels grâce à notre méthode gérant les données manquantes et les marqueurs haplotypiques, semble robuste, avec des intervalles de confiance à 95% (i.e. obtenus par bootstrap) relativement faibles (figures 5.6 et 5.7).

4.2.2 Estimation des corrélations de paternité

La variabilité inter-individuelle de la corrélation de paternité au sein du nuage de pollen allofécondant (r_p) est élevée avec valeurs s'échelonnant entre 0 et 1 (figure 5.6 et 5.7). Cependant cette variabilité est légèrement inférieure à celle observée pour les

Année	Nb mères	s	σ_s	r_p	σ_{r_p}	$Q_0^{s-r_p}$	$\sigma_{Q_0^{s-r_p}}$
2002	42	0.27	0.31	0.14	0.13	0.27	0.24
2003	23	0.16	0.21	0.09	0.10	0.15	0.14
2004	39	0.29	0.32	0.16	0.26	0.27	0.27
2005	32	0.11	0.14	0.11	0.13	0.11	0.11

TAB. 5.7: **Modèle 1.**Tableau des probabilités d'autofécondation (s) et de copaternité (r_p et $Q_0^{s-r_p}$) moyennes estimées selon le modèle 1 (i.e. un nuage pollinique global différent pour chaque année). σ correspond aux écart-types de ces estimations sur l'ensemble des mères.

Année	Nb mères	s	σ_s	r_p	σ_{r_p}
2002	42	0.31	0.30	0.16	0.16
2003	23	0.10	0.17	0.14	0.13
2004	39	0.29	0.31	0.21	0.27
2005	32	0.13	0.16	0.12	0.13
Toutes	49	0.23	0.25	0.18	0.16

TAB. 5.8: **Modèle 2 et 3.**Tableau des probabilités d'autofécondation (s) et de copaternité (r_p) moyennes estimées selon le modèle 2 (i.e. un nuage pollinique unique pour toutes les années) pour les quatre premières lignes et selon le modèle 3 pour la dernière ligne. σ correspond aux écart-types de ces estimations sur l'ensemble des mères.

taux d'autofécondation (tableaux 5.7 et 5.8). La corrélation de paternité au sein des nuages d'allopollen (r_p) est assez forte avec en moyenne de 13 à 18% de pleins-frères parmi les descendants issus d'allofécondation selon les modèles utilisés (tableaux 5.7 et 5.8). Les trois modèles d'estimation utilisés montrent une alternance d'années où le r_p est fort (2002 et 2004) et d'autres pour lesquelles cette corrélation est faible (2003 et 2005 ; tableaux 5.7 et 5.8). Les estimations individuelles des corrélations de paternité semblent moins robustes que celles des taux d'autofécondation, avec des intervalles de confiance à 95% plus larges, et qui inclus souvent la valeur 0 (figures 5.6, 5.7).

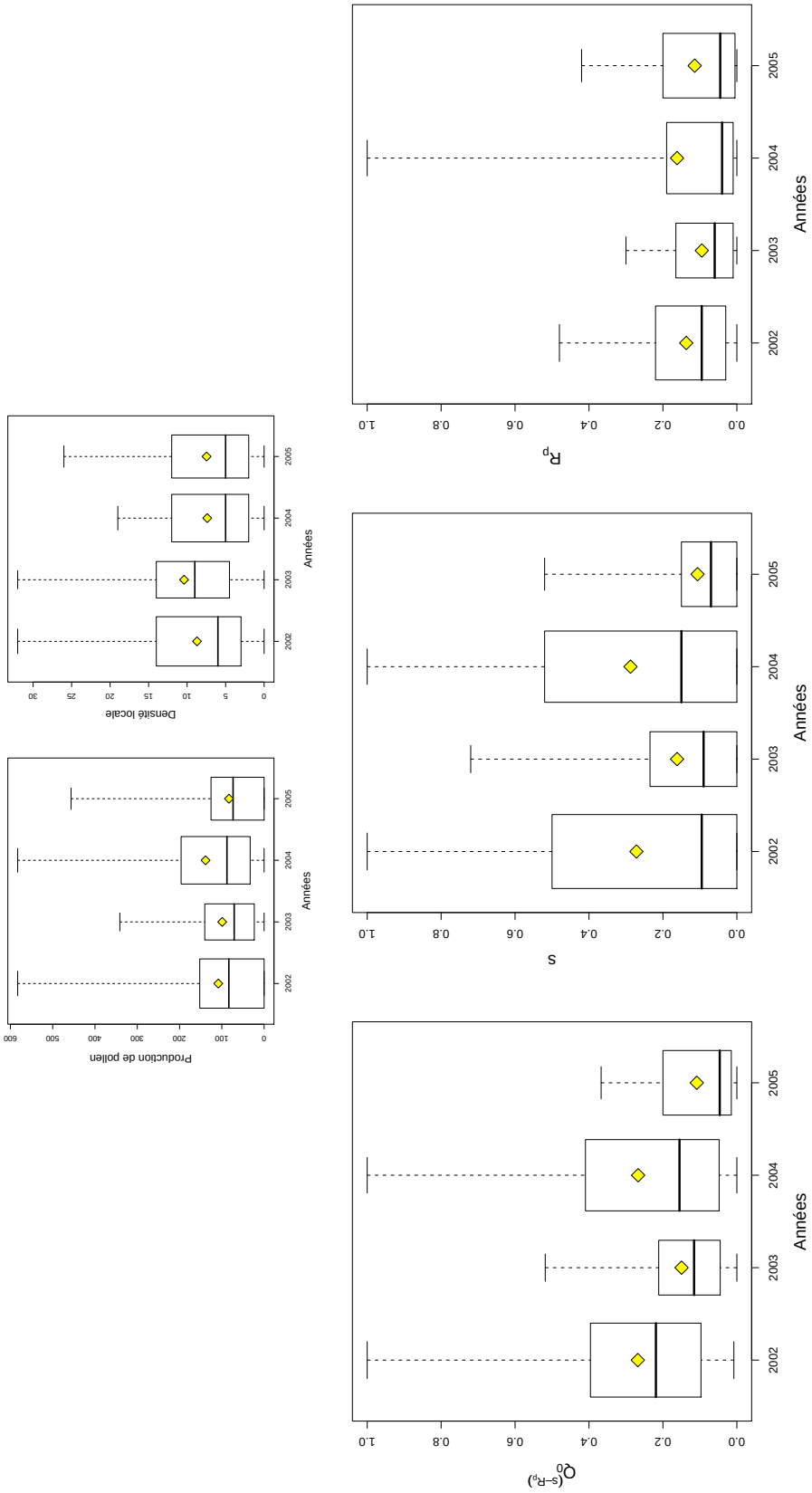


FIG. 5.5: **Modèle 1.** Boxplots des estimations individuelles et par année des productions de pollen (en haut à gauche), de la densité locale des arbres échantillonnés (nombre de voisins dans un rayon de 12,5m ; en haut à droite), de la corrélation de paternité totale (en bas à gauche), du taux d'autofécondation (au centre) et de la corrélation de paternité dans la fraction allofécondée (en bas à droite). Les diamants jaunes représentent les moyennes, les boîtes représentent les seconds et troisièmes quartiles séparés par la médiane, leur largeur est proportionnelle aux effectifs et les moustaches représentent l'intervalle des valeurs observées.

4.2.3 Comparaison des modèles

Nous avons comparé les modèles 1, 2 et 3, qui diffèrent par le nombre de paramètres estimés (voir partie 3.2.2 page 169), par un test du rapport de vraisemblance. La modélisation de paramètres du régime de reproduction individuelles chaque année apporte un gain de vraisemblance significatif par rapport à une modélisation unique pour toutes les années de ces paramètres (modèle 2 vs. modèle 3 : Rapport de vraisemblance $\simeq 3202$; $ddl = 174$; $p < 10^{-4}$). La modélisation de nuages polliniques dont la composition génétique est indépendante entre les années apporte également un gain significatif de vraisemblance par rapport à la modélisation d'un nuage dont la composition est constante pour toutes les années (modèle 1 vs. modèle 2 : Rapport de vraisemblance $\simeq 16717$; $ddl = 86$; $p < 10^{-4}$). Le modèle le plus complet apporte donc toujours un gain significatif de vraisemblance, le modèle 1 est donc le modèle permettant le meilleur ajustement aux données.

4.3 Variabilité spatio-temporelle du régime de reproduction

Nous avons utilisé différents modèles linéaires (S1, S2, S3, R1, R2, R3, Q1, Q2 et Q3) afin d'expliquer les valeurs individuelles des taux d'autofécondation et de corrélations de paternité par les variations inter-annuelles, les productions polliniques et les densités individuelles. Toutes ces analyses sont basées sur les estimations du régime de reproduction obtenues grâce au modèle 1, qui est celui permettant d'obtenir les meilleures estimations (voir partie 4.2.3 page 186).

Effets	d.d.l.	den. d.d.l.	F	p
Année	4	73	15.01	$< 10^{-4}$
Densité locale	1	47	5.32	0.03
Production de pollen	1	73	0.03	0.87
Année x Dens. loc.	3	73	5.89	$1.2.10^{-3}$
Année x Prod. pollen	3	73	1.3	0.28
Dens. loc. x Prod. pollen	1	73	1.36	0.25

TAB. 5.9: **Modèle S2.** Tableau d'ANOVA des effets fixes du modèle S2.

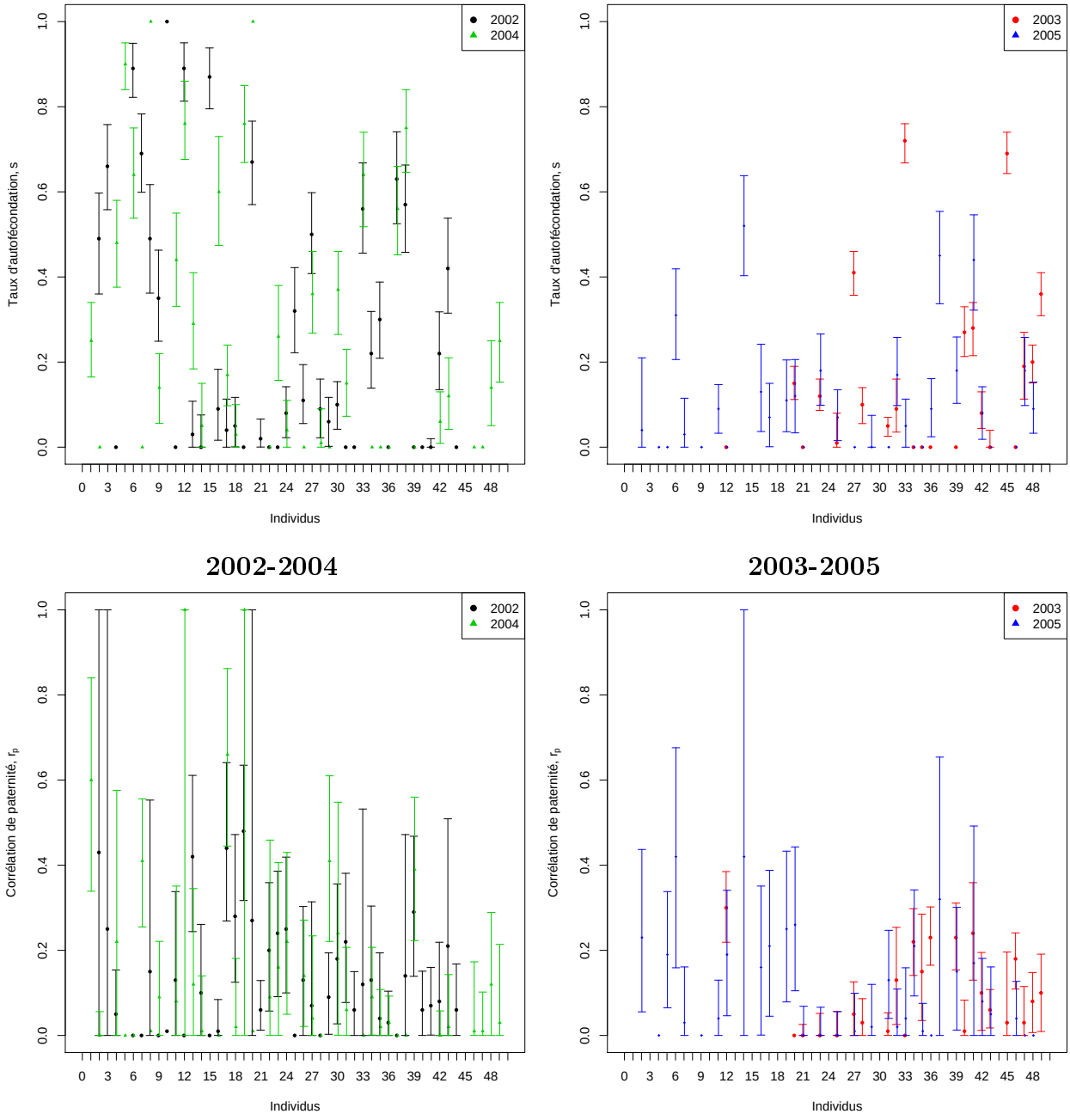


FIG. 5.6: **Modèle 1** Estimation des probabilités d'autofécondation (en haut) et de copaternité (en bas) individuelles et leurs intervalles de confiance à 95% obtenu par bootstrap. Les années 2002 et 2004 sont représentées à gauche et les années 2003 et 2005 à droite.

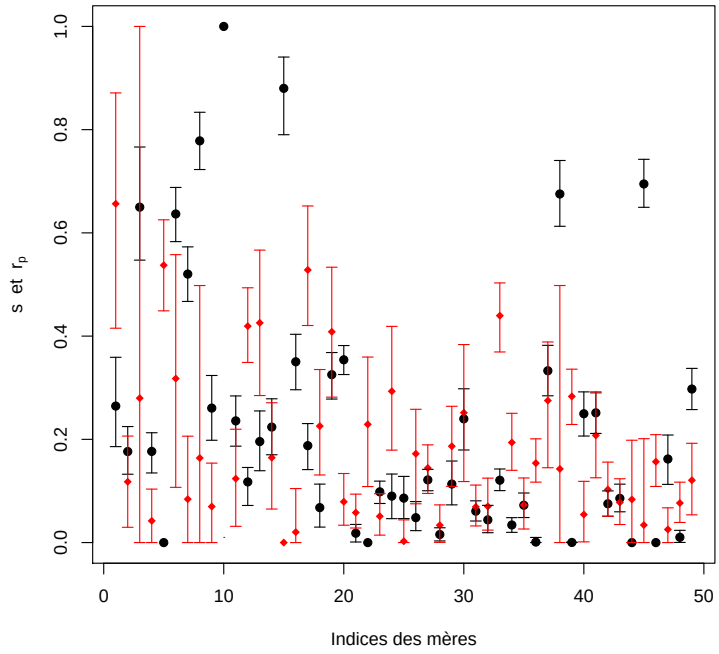


FIG. 5.7: **Modèle 3** Estimation des probabilités d'autofécondation (disques noir) et de copaternité (diamants rouges) et leur intervalles de confiance à 95% obtenu par bootstrap.

4.3.1 Le taux d'autofécondation

L'année a un effet significatif sur les valeurs estimées des taux d'autofécondation (tableau 5.9), avec notamment des effets plus forts des années 2002 et 2004 (valeurs non présentées ici) pour lesquelles nous observons des valeurs moyennes plus fortes que pour les années 2003 et 2005 (tableau 5.7). La densité locale a un effet négatif significatif sur le taux d'autofécondation (tableau 5.9). L'interaction entre la densité locale et l'année a également un effet significatif sur le taux d'autofécondation, avec un effet quasi-nul de la densité en 2003 et en 2005 et un effet négatif en 2002 et en 2004 (valeurs non présentées, voir figures 5.8 et 5.10). En revanche, la production individuelle de pollen ni aucune des interactions impliquant cette variable n'ont d'-effet significatif sur les taux d'autofécondation individuels (tableau 5.9). Cependant la production moyenne de pollen de la population semble avoir un effet positif sur le taux d'autofécondation moyen (figure 5.5). La comparaison de modèles par le test du rapport de vraisemblance a montré que l'effet individuel était significatif (S2 vs.

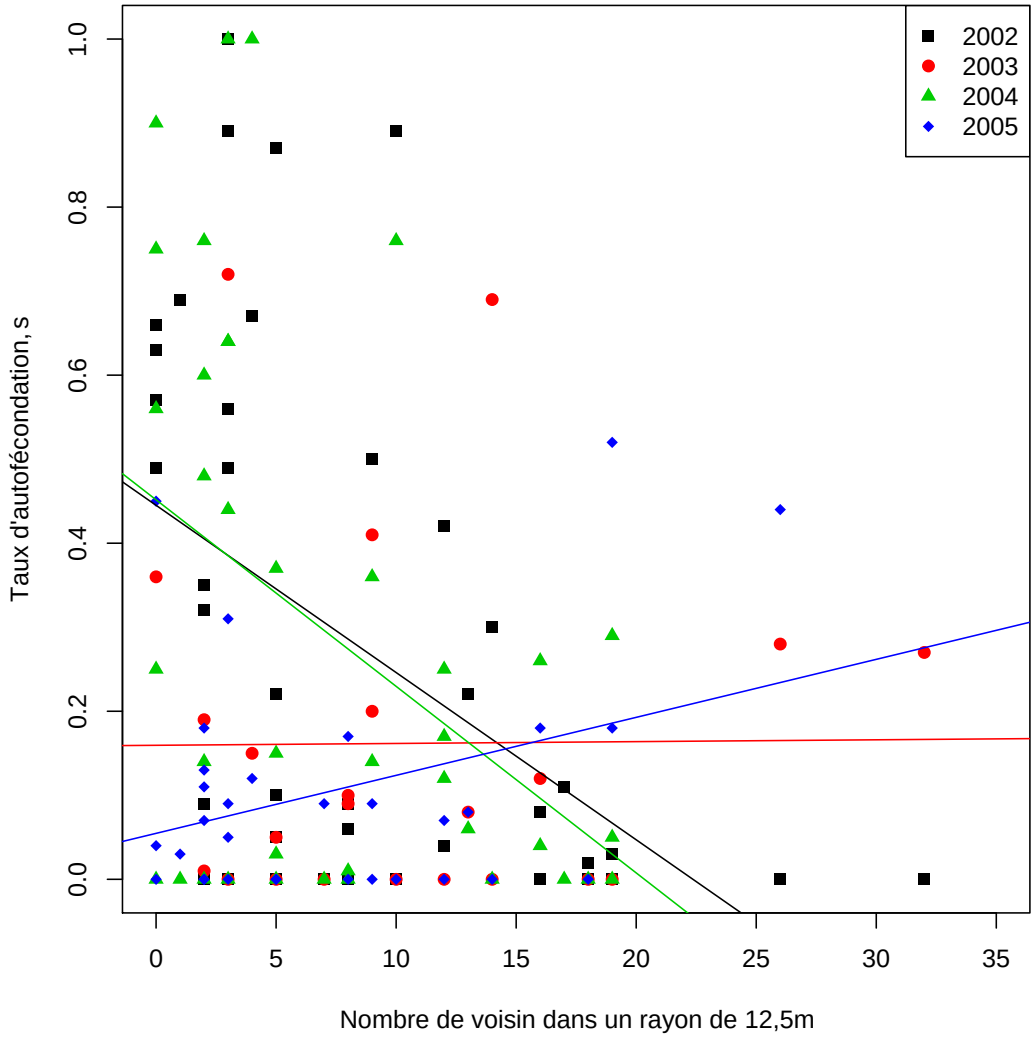


FIG. 5.8: **Modèle 1.** Graphique représentant les taux d'autofécondation individuels (s) en fonction de la densité locale (nombre d'individus producteurs de pollen dans un rayon de 12,5m) pour chaque année. Les droites représentent les régressions linéaires correspondantes par année.

S1 : Rapport de vraisemblance $\simeq 4,25$; $ddl = 1$; $p = 0,04$) et que le modèle le plus parcimonieux était le modèle S2 (S3 vs. S2 : Rapport de vraisemblance $\simeq 16,72$; $ddl = 9$; $p = 0,053$), et donc qu'il n'y avait pas de gain de vraisemblance significatif (au risque 5%) à considérer l'effet individuel conditionnellement à l'année.

4.3.2 La corrélation de paternité

La corrélation de paternité est significativement différente entre les 4 années de l'étude (tableau 5.10), avec des valeurs plus fortes en 2002 et en 2004 qu'en 2003 et 2005 (tableau 5.7). En revanche aucune covariable ni aucune interaction ne semble avoir d'effet significatif sur la probabilité de copaternité (tableau 5.10) et en particulier pas d'effet de la densité locale (figure 5.9 et 5.10). La comparaison des modèles R1, R2 et R3 montre que le modèle R3 est celui présentant la meilleure vraisemblance avec un nombre raisonnable de paramètres (R1 vs. R2 : Rapport de vraisemblance $\simeq 3,9$; $ddl = 1$; $p = 0,048$ et R1 vs. R2 : Rapport de vraisemblance $\simeq 19,6$; $ddl = 9$; $p = 0,02$). Ainsi la variabilité inter-individuelle est significativement différente entre les années. Enfin, il n'y a pas de corrélation entre les r_p et les taux d'autofécondation s individuels ($r = -0,04$, $p = 0,61$).

Effets	d.d.l.	den. d.d.l.	F	p
Année	4	73	12.31	$< 10^{-4}$
Densité locale	1	47	0.41	0.53
Production de pollen	1	73	3.03	0.09
Année x Dens. Loc.	3	73	0.75	0.53
Année x Prod. Pollen	3	73	0.63	0.6
Dens. Loc. X Prod. Pollen	1	73	1.94	0.17

TAB. 5.10: **Modèle R3.** Tableau d'ANOVA des effets fixes du modèle R3, testés par une analyse de type III.

La probabilité de copaternité individuelle intégrant l'allofécondation et l'autofécondation ($Q_0^{s-r_p}$) est également significativement différente entre les années, avec 2002 et 2004 présentant des structurations génétiques plus fortes qu'en 2003 et en 2005 (tableaux 5.11 et 5.7 et figure 5.5). On observe également un effet négatif et significatif de la densité locale et une différence de cet effet entre les années (tableau 5.11 et figure 5.11).

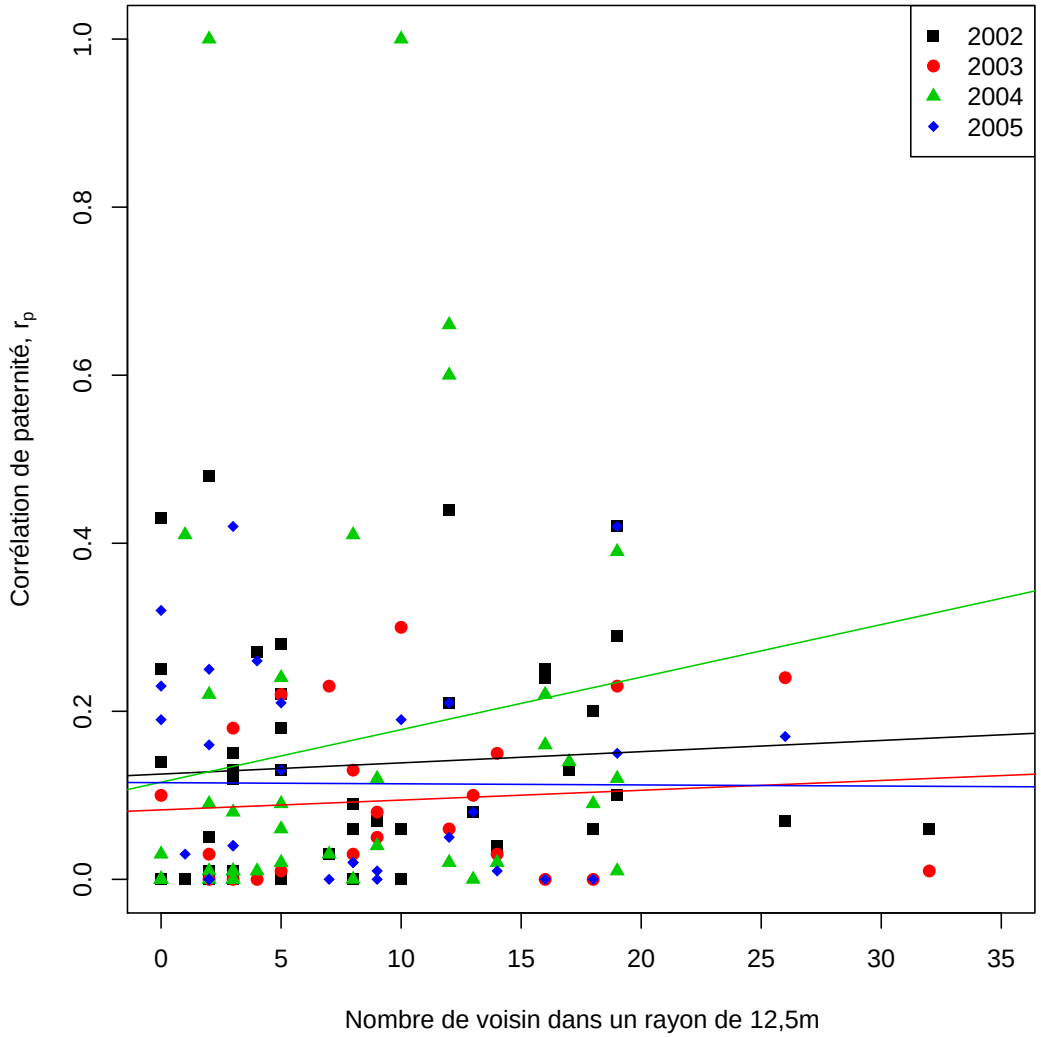
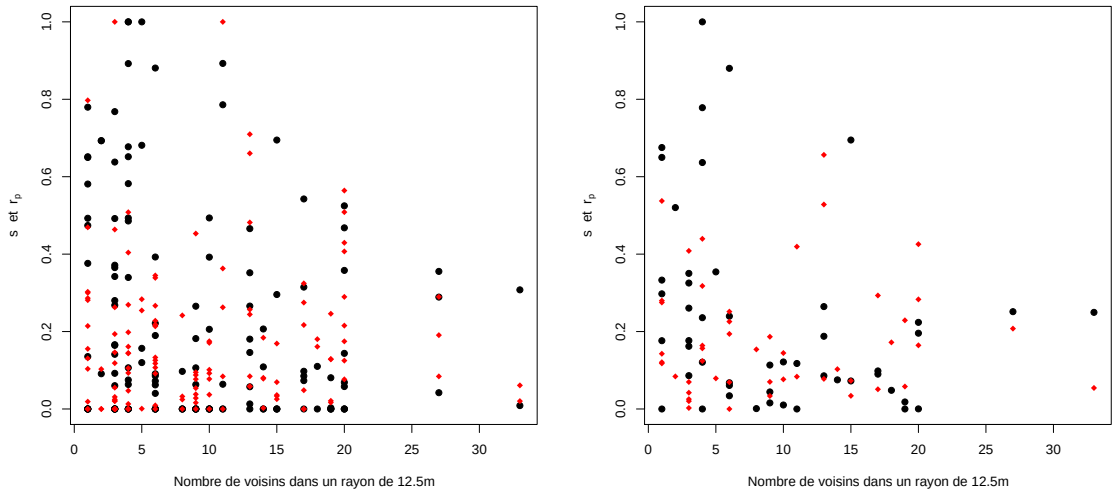


FIG. 5.9: **Modèle 1.** Graphique représentant les corrélations de paternité individuelles estimées sur le nuage pollinique allofécondant (r_p) en fonction de la densité locale (nombre d'individus producteurs de pollen dans un rayon de 12,5m) pour chaque année. Les droites représentent les régressions linéaires correspondantes par année.

Effets	d.d.l.	den. d.d.l.	F	p
Année	4	73	11.6	$< 10^{-4}$
Densité locale	1	47	6.66	0.01
Production de pollen	1	73	0.93	0.34
Année x Dens. Loc.	3	73	2.96	0.04
Année x Prod. Pollen	3	73	3.74	0.01
Dens. Loc. X Prod. Pollen	1	73	2.89	0.09

TAB. 5.11: **Modèle Q3.** Tableau d'ANOVA des effets fixes du modèle Q3, testés par une analyse de type III.



Modèle 2 **Modèle 3**

FIG. 5.10: **Modèle 2 et 3** Estimation des probabilités d'autofécondation (disques noirs) et de copaternité (diamants rouges) en fonction de la densité locale (i.e. nombre de voisins autour de la mère dans un rayon de 12.5m).

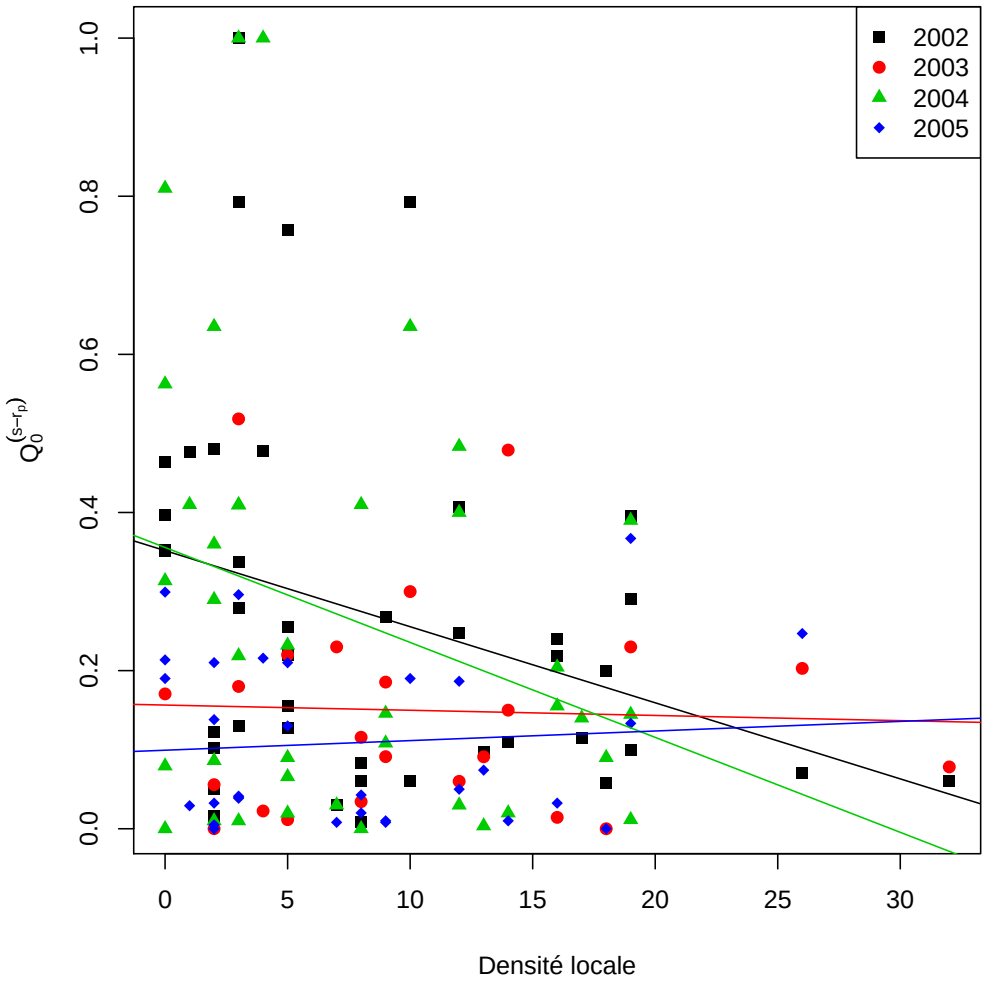


FIG. 5.11: **Modèle 1.** Graphique représentant les corrélations de paternité individuelles estimées sur l'ensemble du nuage pollinique ($Q_0^{s-r_p}$) en fonction de la densité locale (nombre d'individus producteurs de pollen dans un rayon de 12,5m) pour chaque année. Les droites représentent les régressions linéaires correspondantes par année.

5 Discussion

5.1 Structure génétique spatiale du nuage de pollen et dispersion

5.1.1 Structure génétique spatiale du nuage pollinique

Le nuage pollinique fécondant présente une structuration génétique spatiale importante avec des Φ_{ft} globaux compris entre 10,6% et 24,8% suivant les années. Ces valeurs de structuration génétique correspondent à une faible diversité génétique faible au sein des nuages polliniques avec 4 à $9,5$ pères efficaces en moyenne par descendance suivant l'année considérée ($4,9$ pères efficaces en espérance). Il est étonnant d'observer une telle structuration génétique au sein d'une population de conifères présentant une densité aussi élevée (186 sapins producteurs de pollen/ha). En effet, une grande diversité génétique au sein du nuage pollinique est attendue lorsque la densité est forte (Smouse et Sork 2004, pour une revue assez complète). Par exemple Bittencourt et Sebbenn (2007) estiment de 9,5 à 16,2 le nombre de pères efficaces dans une population d'*Araucaria angustifolia* où la densité est de seulement 42,2 arbres/ha. De même, Robledo-Arnuncio et al. (2004*b*) trouvent de 71 à une infinité de pères efficaces par descendance dans une population de pins sylvestres (*Pinus sylvestris*) dont la densité varie en fonction des politiques de gestion forestière de 80 à 315 arbres/ha. En revanche les valeurs que nous avons observées sont de l'ordre de celles observées chez les conifères à des densités très faibles, par exemple Bittencourt et Sebbenn (2008) estiment un nombre moyen de pères efficaces de 6,4 avec une densité en conspécifiques de seulement 3,71 arbres/ha au sein d'une population continue.

L'estimation de la probabilité de copaternité Q_0^γ, permet d'intégrer les estimations de la fonction de dispersion et de la densité efficace dans un seul paramètre de structuration. De plus cette estimation de la probabilité de copaternité est meilleure que celle obtenue avec l'approche globale car elle considère explicitement les distances entre les mères, contrairement à l'approche globale qui considère cette distance constante et égale à la moyenne (Austerlitz et Smouse 2002). Dans notre étude

la structuration génétique spatiale du nuage pollinique estimée par la fonction de dispersion est très robuste et insensible au modèle de dispersion choisi. Elle est moins forte que celle observée avec la méthode globale ou avec le $Q_0^{s-r_p}$. Ceci conduit à des nombres moyens de pères efficaces variant de $6,2$ à $16,4$ ($5,1$ en espérance), ce qui est plus cohérent avec les estimations faites chez d'autres conifères mais demeure faible (Smouse et Sork 2004). Un tel patron de structuration génétique spatiale du nuage pollinique ne peut s'expliquer que par une densité ou une dispersion efficaces limitées.

5.1.2 Dispersion du pollen

En utilisant la méthode décrite par Austerlitz et Smouse (2001), il est possible d'estimer la distance moyenne de dispersion (δ) en fixant au préalable la densité efficace en contributeurs de pollen. En effet en faisant l'hypothèse d'un noyau normal de dispersion on obtient la relation suivante :

$$\delta = \sqrt{\frac{1}{8Q_0 d}}, \tag{5.24}$$

avec d la densité efficace en contributeurs polliniques et Q_0 la probabilité de copaternité estimée avec la méthode TwoGener globale ($\Phi_{ft} \simeq Q_0$, voir équation 5.9). Ainsi en utilisant cette relation avec la densité de 186 contributeurs de pollen par hectare mesurée précédemment, les structurations génétiques spatiales observées correspondent à des distances moyennes de dispersion très faibles ($5,4$, 7, $5,3$ et $8,1$ mètres pour les années 2002, 2003, 2004 et 2005 respectivement et 5,65 m en espérance d'après le Φ_{ft}^{nc} global). Cependant la dispersion du pollen est souvent modélisée par une distribution plus leptokurtique permettant de mieux représenter les événements de dispersion à plus longue distance (i.e. moins fréquents). Ainsi de la même façon que pour la loi normale, Austerlitz et Smouse (2001) expriment la relation entre la distance moyenne de dispersion et la probabilité de copaternité globale pour une fonction de distribution exponentielle, telle que :

$$\delta = \sqrt{\frac{1}{2\pi Q_0 d}}. \tag{5.25}$$

Dans ce cas, en utilisant la même densité que précédemment, les distances moyennes de dispersion estimées sont plus élevées mais demeurent très faibles avec des valeurs

inférieures à la dizaine de mètres ($6,1$, $7,8$, 6 et $8,6$ mètres pour les quatre années d'études respectivement et 6,38 m en espérance d'après le Φ_{ft}^{nc} global). Nous obtenons des distances moyennes de dispersion tout aussi faibles avec les valeurs de la corrélation de paternité totale ($Q_0^{s-r_p}$) obtenues par le modèle d'appariement ($5,07$, $6,8$, $5,07$ et $7,95m$ pour la loi normale ; $5,7$,$6,8$,$5,7$, et $9m$ pour la loi exponentielle pour les années 2002, 2003, 2004 et 2005 respectivement). Ce qui est extrêmement court pour une espèce anémophile et notamment pour un conifère dont le pollen possède deux ballonnets permettant une dispersion à relativement longue distance (Schwendemann et al. 2007). Les distances de dispersion observées chez les conifères sont en général de plusieurs dizaines de mètres (Savolainen et al. 2007, Petit et Hampe 2006, Smouse et Sork 2004). Chez le pin sylvestre Robledo-Arnuncio et al. (2004*b*) estiment une distance moyenne de dispersion comprise entre 17 et 29 mètres suivant la politique de gestion menée (avec un noyau de dispersion normal), ou d'environ 83m dans une petite population de la même espèce (Robledo-Arnuncio et Gil 2005, par analyse de paternité, en enlevant les graines autofécondées et avec une fonction exponentielle puissance : $b = 0,67$). Chez d'autres espèces comme le pin maritime (*Pinus pinaster*) la distance moyenne de dispersion estimée par de Lucas et al. (2008) est plus élevée encore et varie entre 78,4 et 174,4 m (avec des exponentielles puissances très leptokurtiques : $b << 0,5$). Enfin Bittencourt et Sebbenn (2008) estiment une distance moyenne de dispersion d'environ 85m avec une loi normale et d'environ 98m avec une loi exponentielle avec la même approche que celle que nous avons employée (i.e. TwoGener global avec pour densité 1 ou 3 arbres/ha).

La méthode TwoGener par couples de mères (Austerlitz et Smouse 2002, Austerlitz et al. 2004) permet d'estimer la densité efficace conjointement aux paramètres de la fonction de dispersion. Dans ce cas, l'estimation de la dispersion pollinique est plus en accord avec celles obtenues chez d'autres espèces de conifères avec des espérances des distances de dispersion annuelle (i.e. basée sur le Φ_{ft}^{nc}) de plusieurs dizaines de mètres ($22,4$, $28,4$ et $51,7$ mètres pour les lois normale, exponentielle et exponentielle puissance contrainte avec b=0,5). Les densités efficaces qui résultent de ces estimations sont beaucoup plus faibles que les densités mesurées *in natura* ($12,7$, $10,1$ et $6,31$ pères/ha pour les fonctions normale, exponentielle et exponentielle puis-

sance contrainte). Ces deux estimations sont très fortement liées, une densité efficace forte induisant une distance moyenne de dispersion (δ) faible et inversement. C'est pourquoi Robledo-Arnuncio et al. (2006) ont développé une méthode permettant d'estimer la fonction de dispersion indépendamment de la densité efficace (KinDist) qui peut alors être estimée dans un deuxième temps avec la méthode TwoGener. Cette méthode n'est pas encore disponible pour les jeux de données haplotypiques et à hérédité paternelle, elle est donc en cours de modification pour pouvoir être appliquée à notre jeu de données. Cependant il apparaît clairement que les dispersions et densités efficaces sont faibles et que cela doit expliquer en partie la forte structuration génétique spatiale du nuage pollinique que nous observons.

La densité du milieu, et la fermeture de la canopée en particulier, pourraient expliquer la réduction de la dispersion du pollen de manière physique en augmentant la rugosité du milieu, ce qui réduirait le nombre de contributeurs mâles à la descendance d'une femelle (Bittencourt et Sebbenn 2008, Smouse et Sork 2004). Ce phénomène pourrait être d'autant plus vrai chez les conifères que leur caractère résineux pourrait bloquer plus efficacement le pollen en le collant aux branches avec lesquelles il rentre en contact (observations personnelles). Pourtant de Lucas et al. (2008) trouvent pour le pin maritime une corrélation de paternité de seulement $4,9\%$ avec la méthode de TwoGener avec une densité extrêmement forte de 800 à 1000 arbres/ha. Cependant, le sapin pectiné possède un houppier plus étalé en hauteur (forme conique) que celui du pin maritime ; de plus les inflorescences mâles et femelles présentent un étagement spatial au sein de chaque individu, les premières se trouvant à la base et les dernières dans la partie haute du houppier. Cela pourrait accentuer le phénomène de rugosité induit par le feuillage et la résine qu'il contient, et justifier la faible dispersion du pollen estimée.

5.1.3 Impact de l'autofécondation

La corrélation de paternité représente la proportion de pleins frères au sein d'une descendance, et celle-ci peut être le résultat de l'autofécondation ou de l'allofécondation (équation 5.19). Burczyk et Koralewski (2005) montrent que la proportion

relative de génotypes issus d'autofécondation a un effet positif sur la valeur du Φ_{ft} global et donc sur la probabilité de copaternité. Si cela n'a pas d'importance lorsqu'il s'agit d'étudier la consanguinité et la diversité globale du nuage pollinique *via* la corrélation de paternité, cela peut en revanche être un inconvénient lors de l'estimation de la fonction de dispersion. En effet, l'autofécondation n'est pas modélisée par la fonction de dispersion des méthodes indirectes (TwoGener et KinDist), il est donc nécessaire de la modéliser indépendamment sur le pollen allofécondant (Carpentier et al. in prep). Burczyk et Koralewski (2005) proposent une correction du Φ_{ft} qui prend en compte l'autofécondation afin d'estimer la fonction de dispersion avec la méthode TwoGener globale (Austerlitz et Smouse 2001). Cependant l'autofécondation est considérée constante dans la population, ce qui n'est pas notre cas ici. Carpentier et al. (in prep) proposent plusieurs méthodes pour prendre en compte l'autofécondation lors de l'estimation de la fonction de dispersion du pollen par des méthodes indirectes (TwoGener et KinDist). Nous avons choisi d'adopter celle consistant à supprimer du jeu de données toutes les graines dont l'haplotype correspond à celui de leurs mères, c'est à dire toutes les graines potentiellement autofécondées. En effet cette méthode semble robuste (simulations) et permet de prendre en compte l'hétérogénéité du taux d'autofécondation entre les individus (Carpentier et al. in prep).

Les estimations de la corrélation de paternité effectuées sur le pollen allofécondant montrent la même robustesse vis à vis de la famille de fonction de dispersion choisie que celle observée sur le jeu de données complet. En revanche la structuration génétique spatiale du nuage pollinique (Q_0^γ) est plus faible que précédemment. Les différence entre les deux estimations de la corrélation de paternité annuelle (i.e. avec ou sans autofécondation) sont plus importantes en 2002 et en 2004. Ces deux années sont celles pour lesquelles les taux d'autofécondation moyens étaient les plus élevés. Ceci confirme l'effet important du régime de reproduction sur la structuration génétique globale du nuage pollinique. De plus, les estimations moyennes par les modèles d'appariement de la corrélation de paternité, r_p, qui considèrent explicitement l'autofécondation, présentent des résultats similaires à nos estimations basées sur la méthode de Carpentier et al. (in prep). Hardy et al. (2004*a*) mon-

trent par simulation un comportement similaire des estimateurs de la corrélation de paternité Q_0 (Austerlitz et Smouse 2001, Smouse et al. 2001, dans ce cas il s'agit de l'approche globale) et r_p (Ritland 2002; 1989). Cependant, ils montrent que ce dernier estimateur présente toujours des valeurs inférieures aux autres estimateurs. Or dans notre cas nous trouvons un résultat inverse, ceci vient probablement du fait que nous avons retiré du jeu de données des graines qui ne résultaient pas d'un événement d'autofécondation. L'importance du biais induit dépend probablement des fréquences des haplotypes maternels dans le nuage pollinique, celui-ci augmentant avec des haplotypes maternels fréquents. La corrélation de paternité réelle au sein du nuage de pollen allofécondant est donc probablement plus élevée que celle estimée par le Q_0^γ ou le r_p. La structuration génétique du nuage de pollen allofécondant demeure donc élevée.

L'estimation annuelle de la fonction de dispersion s'est avérée difficile lorsque nous avons appliqué la méthode de Carpentier et al. (in prep). En effet, la diminution de la quantité de données disponibles pour l'estimation des paramètres de la fonction de dispersion et de densité locale, a été largement réduite du fait d'un taux d'autofécondation moyen important et d'haplotypes fréquents notamment certaines années. Ceci a conduit à des estimations de la fonction de dispersion peu précises, et tantôt supérieures tantôt inférieures aux estimations faites sur le jeu de données complet. Ce type de méthode nécessite donc un échantillonnage plus important que les méthodes classiques d'estimation de la dispersion du pollen qui ne considèrent pas l'autofécondation. Cependant les distances de dispersion moyennes (δ) ainsi que les densité efficaces (d_e) estimées annuellement semblent faibles ce qui confirment nos précédentes conclusions.

5.1.4 Descripteurs de la fonction de dispersion

L'estimation de la corrélation de paternité s'appuie sur la fonction de dispersion estimée avec l'approche TwoGener par couples. En revanche, sa conversion en nombre de pères efficaces ne fait aucune hypothèse quant au modèle de dispersion choisi, et correspond donc au nombre d'individus contribuant de manière identique

au nuage pollinique de chaque mère (équations 5.8 et 5.9). Ainsi le nombre de pères efficaces ne "spatialise" pas les individus et ne prend donc pas en compte leurs contributions relatives en fonction de la distance à laquelle ils se trouvent. Nous proposons donc d'utiliser un indice permettant de mieux visualiser la diversité génétique des nuages polliniques. Nous avons défini à partir de la fonction de dispersion, une zone circulaire de rayon $R^{\gamma}_{0.95}$ autour de chaque mère comprenant 95% des pères potentiels (équation 5.16). Le rayon $R^{\gamma}_{0.95}$ ne dépend donc que de la fonction de dispersion estimée. A partir de ce rayon et de la densité efficace estimée nous avons calculé le nombre de père potentiels compris dans cette zone, $N^{\gamma}_{0.95}$. Dans ce cas l'estimation n'est pas basée sur la probabilité de copaternité mais uniquement sur la fonction de dispersion et la densité efficace. Le $N^{\gamma}_{0.95}$ représente donc le nombre de pères potentiels mais en tenant compte du modèle de dispersion, donc tous ces individus ne contribuent pas de manière identique au nuage pollinique échantillonné par la mère. Ainsi les paramètres $R^{\gamma}_{0.95}$ et $N^{\gamma}_{0.95}$ permettent de visualiser plus facilement l'aire de pollinisation et le nombre de contributeurs potentiels.

Finalement, la queue de dispersion est un paramètre déterminant de la structuration génétique spatiale, aussi important que la distance moyenne de dispersion (Klein et al. 2006). Cependant la queue de dispersion est un paramètre assez difficilement interprétable biologiquement, nous proposons donc d'utiliser le rapport $\frac{N^{\gamma}_{0.95}}{N^{\gamma}_{ep}}$ comme clef de lecture de la fonction de dispersion en plus de la distance moyenne de dispersion. En effet, ce rapport est constant pour une valeur de b donnée et ce quelque soit la distance moyenne de dispersion estimée ou le jeu de données considéré. De plus, le rapport $\frac{N^{\gamma}_{0.95}}{N^{\gamma}_{ep}}$ augmente quand la queue de dispersion s'alourdit (i.e. quand b diminue). Cela est du à la participation au nuage pollinique d'un plus grand nombre de pères potentiels dont la contribution est de plus en plus faible en fonction de la distance à laquelle ils se trouvent de la mère. En revanche, nous avons vu que la valeur du nombre de pères efficaces est insensible à la fonction de dispersion. Le rapport du nombre de pères potentiels (à 95%) sur le nombre de pères efficaces, représente donc le déséquilibre de contribution des individus au nuage pollinique et représente donc une bonne représentation biologique de la queue de dispersion.

5.2 Variabilité du régime de reproduction

5.2.1 La corrélation de paternité globale, $Q_0^{s-r_p}$

La valeur de la corrélation de paternité totale estimée à l'échelle individuelle ($Q_0^{s-r_p}$) présente une très forte variabilité avec des valeurs comprises entre 0 et 1 (coefficient de variation $c.v.$ = 89; 93; 100 ; 100% pour les années 2002, 2003, 2004 et 2005 respectivement). La diversité génétique estimée des nuages polliniques échantillonnés par les mères peut donc être nulle avec 100% de pleins frères au sein d'une descendance (probablement moins en réalité, cela dépend de l'échantillonnage) ou au contraire présenter une diversité maximale dans ce cas aucune graine ne partage le même père. Cette variabilité est principalement le résultat de la densité mesurée à l'échelle individuelle (i.e. mesurée dans un rayon de 12,5m autour de chaque mère) qui a un effet négatif significatif sur la corrélation de paternité totale. En effet, la population que nous avons étudiée et qui semblait globalement continue et relativement dense, présente en réalité de fortes hétérogénéités à l'échelle individuelle. La densité peut modifier la quantité et la diversité d'allopollen reçue, ce qui peut influencer le taux d'autofécondation (s) par loi d'action de masse (Mimura et Aitken 2007b, Robledo-Arnuncio et al. 2004a) et la corrélation de paternité au sein de la part allofécondée de la descendance (r_p) (Robledo-Arnuncio et al. 2004a). De plus, la production individuelle de pollen a également un effet positif significatif sur la corrélation de paternité totale certaines années. Il est donc nécessaire d'étudier l'effet de chacune des composantes (s et r_p) et de comprendre leur impact sur la corrélation de paternité totale ($Q_0^{s-r_p}$).

5.2.2 Le taux d'autofécondation, s

La population de sapin pectiné de la montagne de Lure présente un niveau moyen d'autofécondation extrêmement élevé (20,8 à 23% en moyenne sur les 4 années suivant le modèle considéré) par rapport à celui normalement observé chez la plupart des conifères. Ces derniers sont généralement des espèces majoritairement allogames, et d'autant plus lorsque la densité en conspécifiques est aussi forte (Adams et Birkes 1991, voir aussi le tableau 1 du chapitre 2). Les taux d'autofécondation individuels

(s) estimés grâce à notre version modifiée de la méthode développée par Ritland (1989; 2002) montrent également une variabilité inter-individuelle très importante, avec notamment certaines descendances complètement consanguines et d'autres pas du tout. La densité individuelle en individus producteurs de pollen a un effet négatif sur ce taux d'autofécondation. Cela est cohérent avec la plupart des observations faites à ce jour (Naito et al. 2008, Fukue et al. 2007, Mimura et Aitken 2007b, Robledo-Arnuncio et al. 2004a, Franceshinelli et Bawa 2000, voir également le tableau 1 du chapitre 2). Ce résultat confirme notre hypothèse selon laquelle il serait le résultat d'une loi d'action de masse, et ne dépendrait que de la quantité d'autopollen émis relativement à la quantité d'allopollen reçu (Harder et Wilson 1998, Holsinger 1991). Une augmentation de la densité entraîne donc une augmentation de la quantité d'allopollen reçue, ce qui a pour effet de "diluer" l'autopollen. Cependant nous n'avons pas détecté d'effet de la production individuelle de pollen qui va, elle aussi, influencer la quantité relative d'autopollen au sein du nuage de pollen fécondant. La difficulté à détecter l'effet de la production de pollen est probablement due à une estimation inadaptée des densités locales dans le cas de la montagne de Lure. En effet dans le cas d'une loi d'action de masse, les taux d'autofécondation estimés à densité locale très faible devraient être en grande partie influencés par la production individuelle de pollen avec un effet positif de cette dernière. En revanche lorsque la densité est suffisamment forte l'effet de la production individuelle de pollen est négligeable en comparaison à la quantité de pollen reçu. Or si la densité est en réalité différente de celle mesurée au sein du cercle de 12,5m (c.a.d. supérieure à la densité nulle estimée) et qu'il est impossible d'accéder à cette variabilité, il est alors difficile d'estimer correctement l'effet de la production individuelle de pollen (Fukue et al. 2007, ces dernier ont estimés la densité dans des cercles de tailles croissantes). Nous observons justement un effet individuel significatif sur le taux d'autofécondation, ce qui souligne le fait que les densités locales où les productions individuelles de pollen ne suffisent pas à expliquer la variabilité inter-individuelle observée (de Lucas et al. 2008, Hardy et al. 2004a, Robledo-Arnuncio et al. 2004a, décalage phénologique par exemple), ou bien que leurs mesures sont inadaptées. Nos résultats sont en accord avec cette dernière hypothèse puisque nous observons une plus grande variabilité de la valeur du taux d'autofécondation lorsque la densité est faible que lorsqu'elle

est forte (figures 5.8 et 5.10). Ainsi les densités élevées semblent correctement estimées ce qui n'est pas le cas des densités faibles. La densité locale permet donc d'estimer correctement la quantité d'allopollen reçu, en revanche une mesure d'isolement pourrait permettre de mettre en évidence l'effet de l'autopollen. Cette échelle doit dépendre de la distance à laquelle se disperse le pollen et de l'échelle à laquelle s'effectue la structuration spatiale des individus (agrégation). La distance au plus proche voisin (conspécifique) est une mesure d'isolement adaptée, permettant de distinguer des individus présentant des conditions de densités locales similaires, cette mesure est d'ailleurs régulièrement utilisée dans les études concernant le régime de reproduction (Eppley et Pannell 2007, Robledo-Arnuncio et al. 2004*a*, Allison 1990).

5.2.3 La corrélation de paternité au sein du nuage de pollen allofécondant, r_p

La structuration génétique spatiale du nuage pollinique observée dans notre population (tableau 5.3) n'est pas uniquement le résultat de l'autofécondation. En effet les corrélations de paternité individuelles estimées conjointement aux taux d'autofécondation, qui permettent donc de se baser uniquement sur l'allopollen (Ritland 1989; 2002) montrent également une forte structuration génétique du nuage pollinique allofécondant. La proportion moyenne de pleins frères (uniquement pour les individus issus d'allofécondation) au sein des descendances varie entre 9 et 16 % suivant les années considérées. Ces valeurs correspondent à celles observées à densité très faible chez le pin sylvestre, *Pinus sylvestris*, par Robledo-Arnuncio et al. (2004*a*) avec $r_p = 0,196$ pour une densité de 2,4 arbres/ha. Dans cette dernière étude la population est très isolée et de taille réduite avec seulement 35 individus, ce qui peut expliquer la faible diversité génétique observée au sein des nuages polliniques de chaque mère. Cependant les valeurs que nous avons observées sont plus élevées que celles observées chez le pin maritime (*Pinus pinaster*) dans des populations où la densité est extrêmement forte (de Lucas et al. 2008, $r_p = 0,018$, pour une densité de 800 à 1000 arbres/ha). Cette différence est probablement la conséquence de la dispersion du pollen bien plus limitée dans notre cas (distance moyenne de dispersion $\delta = (15 - 81m)$ vs. $\delta = (78 - 174m)$). Il existe également une tendance générale

montrant un effet négatif de la densité globale sur la probabilité de copaternité moyenne, $\overline{r_p}$ (figure 5.5). Cette hypothèse est en accord avec les résultats obtenus chez le pin sylvestre (*Pinus sylvestris*) par Robledo-Arnuncio et al. (2004*a*) ou chez l'épicéa de Sitka (*Picea sitchensis*) par Mimura et Aitken (2007*b*) qui montraient des corrélations de paternité plus fortes dans les petites populations isolées que dans les grandes populations.

Localement la corrélation de paternité est très variable entre les individus. Cela suggère une hétérogénéité importante de la quantité et de la diversité du pollen reçu par chaque individu. En effet, dès lors que l'on considère que le pollen se disperse de la même façon pour tous les individus (i.e. même fonction de dispersion) une telle variabilité de la corrélation de paternité individuelle, indique une forte variabilité de la densité en mâles fertiles effectivement perçue par les mères. Cette variation peut être le résultat de différents phénomènes comme les variations de densité (agrégation notamment), les variations de fertilité ou le décalage phénologique, tous agissant sur la densité efficace en contributeurs de pollen (de Lucas et al. 2008, Robledo-Arnuncio et Austerlitz 2006, Hardy et al. 2004*a*, Robledo-Arnuncio et al. 2004*a*). Cette variabilité doit donc exister à une échelle très locale puisque le peuplement présente une densité globale forte et continue. Pourtant la densité locale n'a pas d'effets significatifs sur les valeurs de la corrélation de paternité. Hardy et al. (2004*a*) montrent également chez la centaurée de la clape (*Centaurea corymbosa*) une absence de corrélation entre la probabilité de copaternité et la densité locale (i.e. rayon de 20m autour de la mère). Cependant même si la majorité du pollen reçu provient de quelques contributeurs majoritaires souvent situés à courte distance (Irwin et al. 2003, Koenig et Ashley 2003), la diversité génétique du pollen reçu par les individus situés à longue distance peut être élevée, de sorte que les individus isolés présenteraient une plus grande diversité génétique au sein du nuage de pollen allofécondant (Klein et al. 2006, Robledo-Arnuncio et al. 2004*a*). On peut donc s'attendre à une relation positive entre la probabilité de copaternité et la densité (Robledo-Arnuncio et al. 2004*a*), dès lors que celle-ci représente correctement l'isolement des individus (i.e. dans le sens où ils doivent être à "longue distance"). Dans ce cas, nous pouvons également faire l'hypothèse qu'une mesure de densité n'est pas

appropriée, mais qu'il faut lui préférer une mesure d'isolement, notamment la distance au plus proche voisin. En effet, nous observons également pour la corrélation de paternité une augmentation de la variance à faible densité par rapport à celle entre les individus présentant de fortes densités locales une année donnée (figures 5.9 et 5.10). De plus, nous n'observons pas de corrélation entre le taux d'autofécondation et la corrélation de paternité ce qui pourrait confirmer que l'effet de la densité sur ces deux paramètres du régime de reproduction est différent : La densité locale ayant un effet plus fort sur le taux d'autofécondation (loi d'action de masse, voir Holsinger (1991)) que sur la corrélation de paternité, ce dernier étant plus sensible à l'isolement (mélange de propagules à longue distance, voir Klein et al. (2006)).

Il est également probable que les variations de fertilités et les décalages phénologiques puissent modifier la densité perçue en contributeurs polliniques. Le décalage phénologique agit à deux niveaux, il peut réduire la densité en contributeurs de pollen (Naito et al. 2008, Robledo-Arnuncio et al. 2006), en même temps cela peut également augmenter la corrélation entre les événements de fécondation en engendrant de la compétition pollinique parmi les individus "phénologiquement compatibles" (i.e. partageant au moins partiellement leurs fenêtres d'émission du pollen et de réceptivité des ovules). En effet les individus fécondants les premiers les ovules ou ceux dont la période de fertilité coïncidera le plus avec la période de réceptivité de la mère seront les contributeurs prédominants au nuage pollinique (Hardy et al. 2004*a*). Ceci revient à des variations de fertilité "temporelles" à l'échelle d'un épisode de reproduction. La fertilité absolue des individus peut aussi largement varier entre les différents individus d'une population (Kang et al. 2003), ce qui peut alors également conduire à l'augmentation de l'apparentement au sein des descendances (Kamalakannan et al. 2007). De tels effets sont probables dans notre cas pour expliquer les probabilité de copaternité élevées au sein des descendances. En effet les individus dans les meilleurs conditions (gradients altitudinaux ou en ressources, statut dominant/dominé au sein de la canopée) produisent en général plus de pollen et le dispersent plus tôt (observations personnelles). La prise en compte des variations de fertilité (temporelles et quantitatives) peut donc s'avérer nécessaire lorsqu'elle est importante et permettre d'augmenter la qualité des estimations ou des interprétations (Klein et al. 2008,

Oddou-Muratorio et al. 2005). Elle pourrait s'effectuer en mesurant la fertilité des voisins, *via* le diamètre des individus ce qui constitue un bon proxy (Kang et al. 2003). En conclusion, les variations de fertilité et de densité doivent probablement être considérés en même temps pour pouvoir expliquer précisément la structuration génétique et le régime de reproduction (de Lucas et al. 2008).

5.3 Variabilité temporelle de la structure génétique spatiale du nuage pollinique

La structuration génétique spatiale du nuage pollinique présente une grande variabilité entre les différentes années lorsqu'elle est estimée avec l'approche TwoGener globale. La corrélation de paternité estimée par les modèles d'appariement ($Q_0^{s-r_p}$ et r_p) montrent une évolution similaire. En effet la variation inter-annuelle de la corrélation de paternité est extrêmement élevée avec une variance temporelle représentant près de 13% de la variance totale et un Φ_{yf} de $13,9\%$. En utilisant le même modèle nous obtenons une estimation du Φ_{ft} multi-année corrigé de $7,8\%$. Ceci montre une structuration génétique du nuage pollinique bien plus faible que celle obtenue lors d'un seul épisode de reproduction ($\Phi_{ft}^{nc} = 20,7\%$). Irwin et al. (2003) démontrent également la présence d'une variabilité inter-annuelle chez l'acacia de Constantinople (*Albizia julibrissin*). Cependant dans leur étude les variations inter-annuelles étaient plus modérées que celles que nous observons chez le sapin pectiné ($\Phi_{ft} = 17,4\%$ et $\Phi_{yf} = 4,1\%$ pour Irwin et al. (2003)). Les auteurs expliquent la variabilité temporelle observée en avançant l'hypothèse d'un nuage pollinique dont la composition varie selon les années. En effet, la forte structuration génétique qu'ils observent ainsi que la faible variabilité temporelle, suggèrent que quelques individus ($N_{ep} = 2,87$) contribuent majoritairement au nuage pollinique fécondant tous les ans et que la composition de ce nuage est modifiée par les contributeurs minoritaires qui peuvent être différents d'une année à l'autre.

Les fonctions de dispersion estimées lors du cumul de plusieurs épisodes de reproduction donnent des distances moyennes de dispersion δ plus élevées que lorsqu'elles

sont estimées lors d'un seul épisode de dispersion. Il est assez évident que le cumul de plusieurs épisodes de reproduction va permettre d'augmenter la distance de dispersion du pollen, car cela va intégrer les années de forte ou de faible production de pollen, de forts vents... En revanche, il est difficile de comprendre biologiquement comment ce cumul de plusieurs épisodes de dispersion permet d'obtenir des distances de dispersion supérieures aux distances maximales estimées une année donnée (tableau 5.5). Notre hypothèse est que la fonction de dispersion estimée par la méthode TwoGener une année donnée, sous-destime la distance parcourue par le pollen. En effet, Irwin et al. (2003) montre que l'estimation du Φ_{ft}^{nc} comprend à la fois la variance inter-mères et la variance inter-annuelle, ce qui ne devrait pas poser de problème pour l'estimation de la fonction de dispersion lors d'un seul épisode de reproduction en *pairwise*, puisque l'effet année est le même pour tous (i.e. seule la valeur de Q_0 devrait être affectée). Cependant il est probable que l'effet des variations inter-annuelles soit différent entre les individus de la population en réponse à des conditions micro-environnementales hétérogènes à l'échelle de la population en terme de densité, de pente, d'exposition, de sols... L'effet des variations inter-annuelles devrait donc être relativement similaire pour des individus proches et très différents pour des individus éloignés. En reprenant l'équation du Φ_{ft}^{nc} (équation 5.12) dans une approche *pairwise* et en considérant une variabilité inter-individuelle de l'effet année, on peut écrire :

$$\Phi_{ft-ij}^{nc} = \frac{\widehat{\sigma_f^2} + \alpha_{ij}\widehat{\sigma_y^2}}{\widehat{\sigma_t^2 ot}} \tag{5.26}$$

avec α un coefficient compris entre 0 et 1, positivement corrélé avec la distance séparant les individus i et j et représentant la différence de l'effet année entre ces deux individus. Dans ce cas, l'augmentation de la différenciation entre les nuages polliniques de deux individus en fonction de la distance qui les sépare (voir la figure 5.1) devrait croître plus rapidement que si elle considérait uniquement la variance inter-mère, puisque α_{ij} augmente parallèlement à $\widehat{\sigma_f^2}$; la distance moyenne de dispersion estimée est alors plus courte. De plus, à grande distance α tend vers 1 et notre résultat en *pairwise* converge vers celui de Irwin et al. (2003) avec une approche globale, résultant en des valeurs de N_{ep} similaires ($5,1$ en *pairwise* vs. $4,84$ en global). Il est donc nécessaire de considérer les variations temporelles lors de l'es-

timation de la fonction de dispersion avec la méthode TwoGener, afin d'obtenir des estimations robustes.

Dans notre étude, la variabilité temporelle de la structure génétique spatiale du nuage pollinique est très élevée et nous avons confronté deux hypothèses :

- La composition génétique du nuage pollinique est constante entre les différentes années. Les différences observées sont alors purement aléatoires, et sont le résultat de l'échantillonnage du nuage de pollen par les différentes mères.
- La composition génétique du nuage pollinique varie au cours du temps. Dans ce cas la variabilité temporelle observée est due à une contribution variable des mâles à la reproduction entre les différentes années.

Nous avons utilisé trois modèles d'estimation des paramètres du régime de reproduction présentant des niveaux de complexité différents. Le modèle le plus complexe s'est avéré significativement meilleur que les autres. Or ce modèle est le seul pour lequel les fréquences haplotypiques au sein du nuage de pollen (i.e. commun à tous les individus) sont estimées chaque année. Ce résultat soutient donc la seconde hypothèse selon laquelle les variations inter-annuelles de structurations génétiques du nuages pollinique résultent effectivement d'une composition différente du nuage fécondant.

Oddou-Muratorio et al. (2005) trouvent également des différences de dispersion chez l'alisier torminale (*Sorbus torminalis*) entre les années 1999 et 2000 avec une approche de type directe. Cette tendance est confirmée par les estimations avec la méthode TwoGener sur le même jeu de données qui montrent également une différence entre ces deux années (Austerlitz et al. 2004). Nous proposons donc ici une méthode permettant de prendre en compte les variations temporelles de la structuration génétique spatiale dans l'estimation des paramètres de dispersion et de densité efficaces. Cette méthode est adaptée de celle développée par Irwin et al. (2003), qui se basait sur TwoGener global (Austerlitz et Smouse 2001, Smouse et al. 2001), et ne permettait donc pas l'estimation d'une densité ou d'une distance de dispersion conjointement. En appliquant la décomposition de la variance moléculaire (Excoffier et al. 1992) aux couples de mères tout en distinguant la variance inter-annuelle de

la variance inter-mère, nous avons pu estimer la densité efficace ainsi que la fonction moyenne de dispersion à l'échelle de plusieurs épisodes de reproduction (tableaux 5.5 et 5.6). Nous avons basé nos comparaisons sur les probabilités de copaternité obtenues grâce à ces estimations (Q_0^{γ}). La prise en compte de plusieurs années, montre globalement une diminution de la structuration génétique spatiale du nuage pollinique avec des corrélations de copaternité plus faibles que celles observées lorsque l'on ne prend pas en compte la variance inter-annuelle ($0,091$ vs $0,196$). Parallèlement nous observons un nombre de pères potentiels qui est doublé. Ce résultat est encore plus marqué lorsque les graines potentiellement autofécondées sont retirées du jeu de données, alors la baisse de structuration génétique du nuage de pollen est plus forte ($Q_0^f = 0,040$ vs $0,188$). Notre population montre donc globalement une structuration génétique relativement élevée, mais cependant beaucoup moins marquée que ce qu'elle semblait être lorsque la variabilité inter-annuelle n'était pas prise en compte.

La structuration génétique spatiale semble présenter une alternance bisannuelle (plus de quatre années seraient nécessaire pour le confirmer), avec des années où la corrélation de paternité est élevée (2002-2004) et d'autres où elle est plus faible (2003-2005), et ce quelque soit la mesure de la corrélation de paternité utilisée (tableaux 5.3, 5.5 et 5.7). De plus le taux d'autofécondation (s) varie dans le même sens que la corrélation de paternité totale ($Q_0^{s-r_p}$) dont nous avons vu qu'il était une composante importante (figure 5.5). Cependant la corrélation de paternité mesurée uniquement sur la partie allofécondée (r_p) montre également une cyclicité bisannuelle, avec une alternance de valeurs moyennes fortes et de d'autres plus faibles. Ainsi cette variabilité n'est pas uniquement le résultat d'une variation du taux d'autofécondation. En effet, l'année a un effet significatif sur l'ensemble des paramètres du régime de reproduction estimés à l'échelle individuelle ($Q_0^{s-r_p}$, s et r_p), ce qui signifie que les moyennes annuelles sont effectivement différentes. Cependant il est important de comprendre ce qui induit cette différenciation et comment cela agit sur chacun des compartiments de la corrélation de paternité totale.

Pour expliquer les variations de la structuration génétique spatiale estimée par

les modèles d'appariement nous avons émis deux hypothèses :

- Les individus de la population présentent des fertilité différentes d'une année à l'autre, ce qui modifie la densité efficace.
- Les individus échantillonnés au cours des différentes années présentent des densités locales différentes.

En effet les arbres forestiers présentent souvent une fructification synchrone et intermittente, avec une alternance d'années de fortes ou faibles productions de graines, connue sous le nom de *masting* (Isagi et al. 1997, Herrera et al. 1998, Kelly et Sork 2002). Les modèles les plus récents pour expliquer le *masting* développés par Satake et Iwasa (2002), invoquent la variabilité des conditions environnementales couplée à une production de pollen permettant la synchronisation des individus entre eux. Ainsi la production de pollen peut suivre un cycle temporel d'alternance de fortes et de faibles productions qui pourrait expliquer le patron de structuration génétique observé. D'autre part les individus échantillonnés ne sont pas forcément les mêmes d'une année à l'autre, ainsi la structuration mesurée a pu être modifiée par le choix des individus chaque année, avec notamment une densité locale moyenne qui peut être plus ou moins élevée. La variance inter-individuelle de la corrélation de paternité totale ($Q_0^{s-r_p}$) est différente chaque année et confirme donc notre hypothèse de l'influence de variables à l'échelle très locale. De plus, les effets de la densité locale et de la production individuelle de pollen sur la corrélation de paternité sont significativement différents entre les 4 années de notre étude (tableau 5.11 ; figure 5.11). L'effet négatif de la densité locale est principalement du au taux d'autofécondation. L'effet de l'interaction entre l'année et la densité locale sur le taux d'autofécondation est extrêmement élevé (tableau 5.9) et semble expliquer une grande partie des différences de variance inter-individuelle observées entre les années. En effet nous avons montré qu'il n'était pas nécessaire de modéliser la variance inter-individuelle conditionnellement à l'année, et donc que cette dernière est correctement estimée par les variations de densité entre les années (i.e. par l'échantillonnage). L'année 2003 est principalement responsable de ces effets d'interaction, avec un effet quasiment nul de la densité qui peut s'expliquer par 2 raisons :

- La moyenne des densités locales des individus étudiés est plus élevée que les autres années (figures 5.8 et 5.5). Ceci est principalement du au faible nombre

d'individus présentant des densités locales très faible et à la présence d'individus à très forte densité. Or lorsque la densité atteint un certain seuil, la proportion d'autopollen par rapport au pollen reçu de l'extérieur varie peu et en conséquence le taux d'autofécondation peu variable dès lors qu'il est régi par un modèle d'action de masse. Il est donc difficile dans ce cas d'estimer l'effet de la densité dans ce cas.

– Moins d'individus ont été échantillonnés en 2003, ce qui a pu rendre l'estimation de l'effet de la densité locale moins robuste cette année là.

La corrélation de paternité au sein du nuage de pollen allofécondant varie peu entre les années comparativement à la variation du taux d'autofécondation mais cette variation est cependant significative. De plus nous observons un effet inter-individuel significativement différent entre les années. Comme précédemment nous pouvons émettre l'hypothèse que les mesures individuelles pourraient être inadaptées pour observer des variations de la corrélation de paternité, et qu'une autre mesure de la densité locale (i.e. isolement et fertilité des individus) serait plus appropriée. Cependant le graphique 5.5 indique tout de même une tendance vers un lien positif entre les moyennes annuelles des probabilités de copaternité et les moyennes annuelles de production de pollen. Une tendance similaire est observée chez l'alisier torminale (*Sorbus torminalis*)par Oddou-Muratorio et al. (2005; 2006), avec un nombre moyen de pères efficaces de 8,76 et 9,52, sur la base d'une analyse de paternité pour les années 1999 et 2000, qui diffèrent par le nombre d'individus principalement producteurs de pollen (33 vs 19% d'individus se comportant comme des mâles alors que les autres individus sont hermaphrodites et produisent moins de pollen). Bien que certainement non significatif, ce dernier résultat et notre étude soulignent l'importance des variations inter-annuelles de fertilité chez les arbres et révèlent peut-être un effet insoupçonné du *masting*. En effet si ce dernier est en partie du à la disponibilité en pollen (Satake et Iwasa 2002, Satake 2004), une année de production massive de graines pourrait s'accompagner d'une faible diversité génétique au sein de ces graines, ce qui rendrait caduque l'avantage évolutif qui lui est associé (Kelly et Sork 2002, optimisation de la fécondité) dès lors que la dépression de consanguinité serait présente. Il est possible d'expliquer cela si l'on considère une fonction de dispersion

différente pour les événements de dispersion à longues distances et ceux à courtes distances. En effet, on peut faire l'hypothèse que la contribution relative des principaux donneurs de pollen, souvent situés à courtes distances (Koenig et Ashley 2003), va dépendre de la distance à laquelle ils se trouvent de la mère. En revanche le pollen issu d'individus situés à plus longue distance va se mélanger de façon plus homogène (Klein et al. 2006) et pourra alors être considéré comme un nuage de pollen dont la composition est indépendante de la distance, c'est à dire un nuage de pollen commun à tous les individus, ou au moins à une échelle plus large que celle de l'individu. C'est l'hypothèse qui est adoptée dans les méthodes directes d'estimation des fonctions de dispersion (Burczyk et Koralewski 2005, Oddou-Muratorio et al. 2005, Burczyk et al. 2004; 1996), avec un taux d'immigration constant et une fonction de dispersion estimée au sein d'un voisinage et qui n'est valable que dans cette zone (Shimatani et al. 2007). Ainsi les années avec les plus fortes productions de pollen présenteraient les plus fortes structurations génétiques car cela renforcerait l'impact des contributeurs polliniques majoritaires localement. Cela aurait pour effet d'empêcher l'import de pollen provenant d'individus éloignés et dont la quantité est négligeable devant celle des contributeurs principaux (y compris l'autopollen), d'autant plus que le nombre d'ovules serait limitant ces années là. Inversement, les années de faible production pollinique permettraient une participation plus homogène des individus au nuage pollinique et donc une structuration génétique spatiale du nuage de pollen fécondant plus faible, en diminuant la compétition pollinique pour les ovules (Knight et al. 2005, Ashman et al. 2004, dans le cadre de la limitation pollinique, ce que l'on a démontré dans notre cas dans le chapitre 3). A cela s'ajoute le fait que les années de forte production de pollen s'avèrent dans notre cas être celles présentant les plus fortes variabilités inter-individuelles du régime de reproduction, ce que l'on sait modifier la densité efficace estimée (Oddou-Muratorio et al. 2005).

6 Conclusion

La population de sapins de la montagne de Lure possède une structuration génétique spatiale du nuage de pollen importante, avec des valeurs de corrélation de

paternité totale (i.e. autofécondation et allofécondation) comprises entre 7,8 et 11% suivant la méthode utilisée. Cette structuration est le résultat d'une dispersion annuelle du pollen très limitée de l'ordre de quelques dizaines de mètres en moyenne (δ compris entre 15 et 77m). Les densités efficaces sont également variables avec un rapport $\frac{d}{d_e}$ compris entre $2,5$ et 25, ce qui dépend probablement des différences inter-annuelles de fertilité et de densité. Ces estimations sont assez étonnantes : En général les conifères possèdent de fortes capacités de dispersion et d'autant plus dans une population avec une densité si élevée (186 sapins/ha). Afin d'expliquer la différence observée entre ce résultat et ceux observées chez les pins (de Lucas et al. 2008, Robledo-Arnuncio et al. 2004a), nous avons émis l'hypothèse que la forme plus allongée du houppier (qui plus est résineux) ainsi que la position basale des strobili mâles chez le sapin pectiné pouvaient engendrer une forte rugosité du milieu limitant la dispersion du pollen au sein des peuplements. Cependant si l'on considère les variations inter-annuelles de la structuration génétique du nuage pollinique lors de l'estimation des fonctions de dispersion, les distances observées sont plus élevées (entre 35 et 95 m suivant la fonction de dispersion choisie) mais demeurent toutefois assez faibles pour un conifère.

En moyenne la structuration génétique spatiale du nuage pollinique est moins élevée que celle observée lors d'un seul épisode de reproduction même si elle reste relativement forte comparée à d'autres espèces de conifères (de Lucas et al. 2008, Robledo-Arnuncio et al. 2004a;b, Bittencourt et Sebbenn 2007). On peut donc conclure que le déséquilibre de contributions des individus au nuage pollinique fécondant semble diminuer avec le cumul des épisodes de reproduction. Il serait donc intéressant de considérer une plus longue période pour tester si le déséquilibre de contribution des individus s'efface complètement à l'échelle de la vie d'un arbre, ce qui permettrait ainsi de maintenir une diversité génétique importante au sein de la population.

La production globale de pollen à l'échelle de la population varie largement d'une année à l'autre. De plus, les années de forte production (2002 et 2004) sont aussi les années présentant la plus forte variance inter-individuelle de production de pollen. Cela semble indiquer que les "mauvaises" années de production de pollen,

le sont pour tous les individus, alors que les "bonnes" années semblent profiter différemment à tous les individus, certains produisant énormément de pollen et d'autres pas du tout. Ces variations globales de la quantité de pollen émis influence fortement la structuration génétique spatiale du nuage pollinique et le régime de reproduction moyen de la population. En effet, les années 2002 et 2004 présentent les plus fortes valeurs de la corrélation de paternité moyenne (quelque soit la méthode d'estimation) et du taux d'autofécondation. A l'échelle individuelle ces années de forte production de pollen, sont également celles pour lesquelles les paramètres du régime de reproduction (corrélations de paternité et taux d'autofécondation) sont les plus variables entre les arbres. Cette augmentation de la structuration génétique globale du nuage pollinique les années de fortes production de pollen peut être expliquée par deux raisons :

- Le taux d'autofécondation augmente par loi d'action de masse. En effet, les mesures de production de pollen ont été effectuées sur les mêmes arbres que ceux pour lesquels nous avons estimé les paramètres du régime de reproduction. Les années de fortes production de pollen sont donc aussi celles pour lesquelles le taux d'autofécondation est en moyenne plus élevé, car dans ce cas la quantité d'autopollen produit augmente relativement à l'allopollen reçu.
- La corrélation de paternité augmente en réponse à une participation inégale des individus à la reproduction. En effet, la forte variance inter-individuelle de production de pollen durant les années de forte production va diminuer le nombre de contributeurs au nuage pollinique fécondant en augmentant l'impact de quelques contributeurs majoritaires au détriment des autres individus. En revanche, les années de faible production de pollen présentent une variance inter-individuelle plus faible donc la participation des individus au nuage de pollen fécondant est plus homogène, ce qui entraîne une corrélation de paternité globale plus faible.

Les années de fortes production de pollen sont donc celles pour lesquelles la diversité génétique échantillonnée est en moyenne la plus faible et pour lesquelles la consanguinité des graines est potentiellement la plus forte.

La variabilité inter-individuelle du régime de reproduction est extrêmement mar-

quée dans notre population avec des corrélations de paternité et des taux d'autofécondation compris entre 0 et 1. La composition du nuage pollinique est donc le résultat de variations locales de l'environnement des individus. Nous avons identifié la densité en conspécifiques comme étant l'un de ces facteurs, avec un effet fort sur le taux d'autofécondation et *de facto* sur la probabilité de copaternité totale. Nous suspectons également que cette densité locale pourrait avoir un effet sur la corrélation de paternité au sein du nuage de pollen allofécondant, mais que celle-ci devrait être mesurée différemment (Robledo-Arnuncio et al. 2004*a*). De plus, la diversité génétique des nuages polliniques mesurée par la corrélation de paternité et par le taux d'autofécondation, est significativement différente entre les 4 années de notre étude. Nous avons effectivement montré par une approche de type maximum de vraisemblance que la composition génétique du nuage pollinique était significativement différente d'une année à l'autre et que les différences observées n'étaient pas le résultat d'un tirage aléatoire au sein d'un nuage pollinique constant. Seules les variabilités de production de pollen entre les différents individus ainsi qu'entre les années peuvent expliquer la différence de diversité génétique observée au sein des nuages polliniques fécondant des différents arbres chaque année. Cette hypothèse est en accord avec la baisse de structuration génétique spatiale du nuage pollinique que nous avons observé lors du cumul de plusieurs épisodes de reproduction. Le régime de reproduction du sapin pectiné semble donc présenter une variabilité spatio-temporelle forte, qui lui permettrait de conserver globalement une diversité génétique en réalité plus importante que celle estimée à l'échelle de la population lors d'un unique épisode de reproduction.

En conclusion, la conservation des peuplements forestiers, notamment de leur diversité génétique, doit impérativement intégrer la variabilité temporelle, du fait des traits d'histoire de vie si particuliers des arbres et notamment leur durée de vie. Ce dernier point est particulièrement important puisque la fertilité peut varier avec l'âge ou entre les années (masting), ce qui peut engendrer un déséquilibre des contributions des individus à la reproduction et donc une baisse de diversité génétique à l'échelle de la population. De plus la densité locale a également un effet fort sur le régime de reproduction et la structuration génétique. Ce facteur est d'un intérêt

particulier puisqu'il est possible d'agir sur celui-ci dans le cadre des politiques de gestion (reforestation, éclaircie, exploitation...). Cependant, la qualité génétique des graines seule n'est pas suffisante pour maintenir une population. En effet, la dispersion des graines déterminera la structure génétique spatiale des adultes, elles doit donc être également estimée afin de caractériser complètement les flux de gènes. De plus, il faudra considérer la disponibilité des sites pour la germination et l'installation de nouveaux semis, car c'est également une étape clef du maintien et de la croissance des populations.

Conclusion

Bilan

Au cours de cette thèse nous avons étudié différentes composantes de la reproduction du sapin pectiné (*Abies alba*) en limite sud de son aire de répartition française d'un point de vue écologique et génétique. Tout d'abord une étude bibliographique nous a permis de mettre en évidence à l'échelle des populations, l'effet négatif de la densité sur le taux d'autofécondation chez plusieurs espèces de conifères. Ainsi les populations marginales, qui présentent souvent une faible densité, pourraient présenter des taux d'autofécondation relativement élevés. Nous avons confirmé cette tendance à une échelle plus fine en étudiant la population de sapin pectiné du Mont Ventoux. En effet, les taux d'autofécondation observés dans les parties de la population les moins denses (i.e. en zone d'expansion, notées LD par la suite, pour *Low Density*) sont significativement plus élevés que ceux observés au coeur du peuplement dans des zones plus matures (notées HD par la suite, pour *High Density*). Cette différence du régime de reproduction s'accompagne d'une baisse globale de près de 20% des taux de germination dans les parties colonisatrices de la population (34 vs 52%). Cependant le suivi de semis issus de ces deux types de peuplements sous différentes conditions de stress hydrique conduit à un résultat inattendu : une meilleure survie des individus issus des peuplements les moins denses au bout de 4 années d'expérimentation (81 vs 63%). Cette différence du niveau de survie entre les deux types de peuplement est principalement causée par le traitement de stress hydrique le plus intense, qui avantage les individus issus de LD par rapport à ceux issus de HD. Finalement, les survies des individus issus de LD et de HD sont assez similaires lorsqu'elles intègrent la phase de germination et celle de survie sur

4 ans sous différents traitements de stress hydrique (28 vs 33%). Un tel résultat souligne les conséquences fortes que peuvent avoir les variations de densité sur le devenir des populations. Les populations à faible densité semblent en effet présenter une forte dépression de consanguinité s'exprimant au moment de la germination. Cependant ce phénomène permet d'éliminer une grande partie des gènes délétères portés par ces individus (i.e. différent du fardeau génétique au niveau de la population), ce qui n'est pas le cas des individus issus des parties à plus fortes densités et qui semblent exprimer leur fardeau génétique plus tard notamment en conditions stressantes. Ainsi dans notre cas, celui d'une espèce sciaphile au moins durant les premiers stades de son développement, une faible densité peut induire une consanguinité élevée mais également des conditions suffisamment stressantes pour purger efficacement les gènes délétères au stade semis à cause d'une faible densité (rayonnement plus élevé...). En revanche, il est possible de poser l'hypothèse parallèle que les individus issus de conditions de densité plus fortes vont être moins consanguins mais en revanche purgeront moins efficacement leur fardeau génétique.

Cette étude ne tenait pas compte de la production de graines qui pouvait être très différente entre les individus présents à faible ou forte densité, et qui donnerait un avantage sélectif à ces derniers. En effet, une faible densité en conspécifiques peut augmenter le taux d'autofécondation et diminuer la quantité de pollen disponible, deux phénomènes pouvant conduire à l'avortement des graines. Nous avons donc étudié la production efficace de graines au travers de la proportion de graines vides au sein de descendances sur le Mont Ventoux et la montagne de Lure entre 2002 et 2005. La production efficace de graines est très variable entre les différentes années, avec notamment une très faible production de graines en 2003, année caractérisée par une sécheresse estivale très marquée, ayant entraîné un avortement important des embryons. Nous observons également une différence importante entre les deux sites avec, en moyenne une production plus importante de graines vides sur le site de la montagne de Lure que sur le Mont Ventoux (59 contre 45% en moyenne sur les 4 années). La montagne de Lure pourrait présenter un fardeau génétique plus élevé que le Mont Ventoux, qui a subi une forte réduction de son effectif puis une recolonisation à partir d'un faible nombre de semenciers (Sagnard et al. 2002).

Ceci pourrait avoir augmenté l'efficacité de la purge du fardeau de mutations, en augmentant la consanguinité par le biais d'un taux d'autofécondation élevé (faible densité) mais aussi du faible nombre de contributeurs au nuage pollinique. Finalement, à l'échelle locale la densité semble avoir un effet positif sur la production efficace de graines. Les conifères ont un régime de reproduction principalement allogame qui diminue l'efficacité de la purge des mutations délétères (Charlesworth et Charlesworth 1987, Williams et Savolainen 1996). Nous n'avons cependant pas mis en évidence la présence de dépression de consanguinité au moment de la formation des graines dans nos populations de sapins. Nous avons donc émis l'hypothèse que la baisse de production de graines efficaces pourrait être le résultat d'une limitation pollinique. Ce résultat est surprenant car les populations étudiées (Lure et le Mont Ventoux) présentent des densités globalement élevées lorsqu'elles sont estimées sur l'ensemble du peuplement. La dispersion du pollen semble donc très limitée, bien que le pollen des conifères et des sapins en particulier soit adapté à l'anémophilie, avec deux ballonnets permettant son transport sur de grandes distances, ceci pourrait induire une structure génétique spatiale importante du nuage de pollen.

Nous avons testé par simulations l'absence de détection de la dépression de consanguinité au stade graine que nous avons observé dans nos populations de la montagne de Lure et du Mont Ventoux. Une hypothèse pouvant expliquer cette absence de détection de la dépression de consanguinité est que la variabilité inter-individuelle du fardeau génétique peut induire une faible puissance au test consistant à régresser la valeur sélective (ici la proportion de graines pleines) sur le taux d'autofécondation. Nous avons donc simulé des descendances d'individus issus de populations dont le taux d'autofécondation et le fardeau génétique (représenté par le nombre équivalent de gènes létaux) étaient caractérisés par une moyenne et une variance (un coefficient de variation inter-individuel). Ensuite nous avons appliqué des régressions linéaires entre la valeur sélective moyenne des descendances et les taux d'autofécondation de leurs parents. Ainsi nous avons étudié la proportion d'études dans lesquelles la dépression de consanguinité était détectée, que nous avons tenté d'expliquer par les moyennes et les variances des taux d'autofécondation et des fardeaux génétiques individuels au sein des populations. Le fardeau génétique moyen de la population ainsi

que la variabilité inter-individuelle du taux d'autofécondation ont un effet positif sur la détection de la dépression de consanguinité. La variabilité inter-individuelle du fardeau génétique a, quant à elle, un effet négatif sur la capacité du test à détecter la dépression de consanguinité par une approche de type régression. Cela confirme notre hypothèse selon laquelle laquelle la variabilité inter-individuelle du fardeau génétique pourrait diminuer la détection de la dépression de consanguinité. De plus, nous avons mis en évidence l'importance du plan d'échantillonnage sur la puissance du test de détection, avec un effet positif important du nombre de descendances échantillonnées et un effet plus modéré de leurs tailles respectives. Cependant la gamme de valeurs que nous observons *in natura* devrait nous permettre d'observer la dépression de consanguinité au stade graines si elle était présente (probabilité de détection comprise entre 98 et 100%). Nous confirmons donc l'absence de dépression de consanguinité au stade graine chez le sapin pectiné dans nos deux populations marginales du sud de la France.

Enfin, nous avons étudié la structuration génétique spatiale du nuage pollinique ainsi que sa composition génétique au sein de la montagne de Lure. Nous avons émis l'hypothèse que la dispersion du pollen pouvait être limitée, du fait de l'impact important de la densité locale sur la production efficace de graines. Dans un premier temps, la dispersion efficace du pollen a été estimée par le biais de marqueurs génétiques. Dans un deuxième temps, nous avons étudié la diversité génétique du nuage pollinique au travers de la probabilité de copaternité, que nous avons décomposée par le taux d'autofécondation et la corrélation de copaternité au sein du nuage de pollen allofécondant. Comme pour la production de graines, nous nous sommes intéressés à ces variations à l'échelle individuelle (et spatiale *de facto*), mais également temporelle (inter-annuelle). Nos estimations de la dispersion efficace confirment notre hypothèse d'une dispersion limitée du pollen chaque année (30m en moyenne). Lorsqu'elle est estimée sur plusieurs épisodes de reproduction en considérant la variabilité inter-annuelle de composition du nuage pollinique, cette dispersion devient plus élevée (77m en moyenne) mais reste en deçà des moyennes observées chez les autres conifères. Afin d'expliquer un tel résultat nous supposons que le houppier du sapin pectiné, avec une forme allongée et une position basale des strobili mâles,

peut provoquer une rugosité du milieu importante, limitant le déplacement du pollen notamment lorsque la densité est élevée. Nous observons également une forte variabilité inter-individuelle de la diversité génétique du nuage pollinique échantillonné par les semenciers avec des corrélations de paternité individuelles globales comprises entre 0 et 100% (20% en moyenne). Un tel patron est principalement le résultat du taux d'autofécondation (21% en moyenne), et est donc très sensible à la densité locale (loi d'action de masse, $p = 0,03$). En revanche, la corrélation de paternité au sein du nuage de pollen allofécondant (13% en moyenne) n'est pas influencée par la densité locale telle que nous l'avons mesurée ($p = 0,53$). Cependant la variabilité inter-individuelle de la probabilité de copaternité est significative et cela pourrait indiquer un effet de variables très locales, comme la densité à une échelle intermédiaire, l'isolement ou la fertilité des contributeurs polliniques proches. La variabilité de fertilité des individus peut également avoir un impact sur la variation du régime de reproduction dans le temps. En effet, en utilisant les paramètres du régime de reproduction, nous avons montré que la composition du nuage pollinique était très variable entre les différentes années. Ces différences sont dues à des contributions différentes chaque année et non à des échantillonnages aléatoires au sein d'un même "pool" pollinique constant d'une année à l'autre. Ces différences sont probablement le résultat de variations inter-annuelles de fertilité, modifiant l'impact des contributeurs polliniques majoritaires et des événements de pollinisation à longue distance. Nous observons effectivement une augmentation de la corrélation de paternité moyenne les années de fortes production de pollen (2002 et 2004) qui sont sont aussi celles présentant les plus fortes variabilités inter-individuelles, ce qui est compatible avec un impact fort des contributeurs majoritaires ces années là.

Au cours de cette thèse nous avons vu que le cumul des épisodes de reproduction permettait de diminuer les différences de contributions des individus. Il serait donc intéressant de prolonger cette expérience sur une durée supérieure afin **i**) de mieux estimer la variabilité inter-annuelle des succès reproducteurs (au moins des déséquilibres de contribution par voie mâle et femelle) et **ii**) de vérifier si le déséquilibre de contributions des individus atteint un seuil ou s'il disparaît complètement à partir d'un nombre suffisant d'épisodes de reproduction. De plus, les variations de

densités locales et de fécondités semblent avoir un effet important sur la production de graines viables et le régime de reproduction. Cependant, les mesures que nous avons effectuées dans cette étude pourraient être améliorées en :

- augmentant le rayon au sein duquel la densité locale est mesurée, 12,5m pouvant être une distance trop courte,
- mesurant l'isolement des individus *via* la distance au plus proche voisin afin d'expliquer les variations de corrélation de paternité,
- mesurant les productions de pollen des voisins et leurs phénologies (i.e. décalage phénologique) afin de mieux comprendre l'impact des contributeurs majoritaires sur la composition génétique des nuages polliniques échantillonnés et confirmer que le régime de reproduction est effectivement régi par une loi d'action de masse.

Nos estimations des paramètres de structuration génétique spatiale du nuage pollinique et du régime de reproduction, pourraient également être plus précises en augmentant le nombre d'arbres échantillonnés (en maximisant la gamme de densité locale), mais aussi en augmentant le nombre de graines génotypées. De telles mesures permettraient de confirmer nos conclusions sur l'effet des variabilités spatio-temporelles sur les différentes composantes de la reproduction du sapin pectiné et sur les espèces pérennes en général.

Perspectives

La marginalité implique des pressions de sélections fortes (stress) et différentes de celles rencontrées dans les populations centrales, qui sont souvent localisées là où les conditions permettent d'atteindre l'optimum écologique (Chuine et Beaubien 2001). Le maintien et la croissance des populations en marge de leurs aires de répartition impliquent qu'elles se sont adaptées localement à leur milieu (Kawecki 2008), ce qui est souvent difficile (Savolainen et al. 2007). Or nos populations, situées à la limite sud de l'aire de répartition du sapin pectiné, présentent des caractéristiques propres aux populations marginales telles que de forts taux d'autofécondation et de fortes corrélations de paternité (Mimura et Aitken 2007*b*) et pourtant semblent con-

forter l'hypothèse d'adaptation locale. En effet, la recolonisation initiée il y a deux générations, notamment sur le Mont Ventoux, semble avoir pleinement porté ses fruits et continue d'opérer, ce qui est révélé par une densité en sapin pectiné élevée. Ainsi ces populations de sapins pectinés disjointes du reste de l'aire de répartition (carte 1.6), ont pu échantillonner suffisamment de diversité génétique lors de l'expansion post-glaciaire pour permettre une sélection efficace après la re-contraction de l'aire ayant conduit à l'isolement de nos populations (Liepelt et al. 2009). Afin de confirmer cette hypothèse d'adaptation locale il est nécessaire de comparer nos populations avec celles présentes au centre de l'aire de répartition ou dans des conditions climatiques différentes. Cela peut s'effectuer en réalisant des analyses de gradients adaptatifs au travers de caractères quantitatifs potentiellement sous sélection (la résistance au stress hydrique par exemple dans notre cas) afin d'extraire des indices de différenciation inter-populations de type Q_{st} (Mimura et Aitken 2007*a*). Il serait également intéressant d'effectuer des tests de provenances, en transplantant des individus issus de différentes parties de l'aire de répartition, afin de voir s'il existe un effet de l'interaction **provenance×environnement** sur les caractères adaptatifs, et le cas échéant confirmer ou infirmer notre hypothèse d'adaptation locale (Savolainen et al. 2007).

Les patrons d'adaptation locale sont souvent considérés à l'échelle de l'aire de répartition de l'espèce (Savolainen et al. 2007) ou sont modélisés par des populations discrètes *via* des modèles en îles où seuls les flux de gènes inter-dèmes sont considérés (Lopez et al. 2008, Alleaume-Benharira et al. 2006). Nos populations sont actuellement en phase de recolonisation et présentent donc deux faciès :

- des parties matures, où la densité est forte et les conditions climatiques sont favorables,
- des zones d'expansion vers les versants sud ou des altitudes plus basses, la densité y est faible et les conditions climatiques plus stressantes.

Ces zones d'expansion constituent alors les limites écologiques et donc potentiellement les marges proximales de nos populations isolées, avec notamment une faible disponibilité en eau. Or nous observons des différences de survie entre les individus issus des zones recolonisatrices et ceux issus des peuplements plus matures. Cette

différence est accentuée lorsque le stress hydrique est marqué, ce qui laisse penser que l'adaptation locale peut avoir lieu au sein des populations et pas seulement entre populations. La mise en place de l'adaptation locale ne dépend pas seulement d'une sélection efficace, mais également de flux de gènes modérés afin "d'alimenter" la diversité génétique et d'éviter la mal-adaptation (Savolainen et al. 2007, Alleaume-Benharira et al. 2006). Nos estimations montrant une dispersion limitée du pollen à l'échelle de la population sont compatibles avec notre hypothèse d'adaptation locale à une échelle fine.

Nous avons découvert sur le Mont Ventoux un individu présentant des inflorescences hermaphrodites, ce qui n'a jamais été mentionné dans la littérature auparavant (figure 5.12). Cet individu se trouve dans une zone très isolée (pas de voisin à moins d'un kilomètre), où la quantité de pollen reçue doit être très faible. En l'absence de dépression de consanguinité, l'autofécondation pourrait donc lui procurer une asurance de reproduction. Les sapins possèdent un mécanisme d'évitement de l'autofécondation passif (étagement des inflorescences mâles et femelles). Il serait donc intéressant de suivre les descendants de cet individu afin de comprendre si ce caractère peut ou non être adaptatif, ce qui corroborerait certains modèles théoriques (Maurice et Fleming 1995) et prouverait que l'adaptation locale peut s'effectuer à des échelles très fines

.

Cette adaptation à une échelle très locale, si elle existe, associée à une grande hétérogénéité du milieu pourrait permettre aux populations marginales de mieux résister à des changements rapides de leur environnement et notamment au réchauffement climatique. Ainsi les populations des limites d'aire de répartition, et en particulier celles situées à la limite arrière, peuvent constituer un réservoir de phénotypes important pour le maintien des populations (Hampe et Petit 2005). Il est donc essentiel de les conserver mais cela implique d'estimer correctement les différentes composantes de la valeur sélective des individus et des capacités adaptatives des populations. Dans ce but l'estimation des flux de gènes au sein de la population est essentielle. Bien que notre étude soit limitée à l'étude des flux polliniques, les flux de gènes peuvent également être le résultat de la dispersion des graines (Oddou-

FIG. 5.12: Photographie d'une inflorescence de l'individu hemaphrodite découvert sur le Mont Ventoux en 2007. On distingue à gauche une inflorescence hermaphrodite avec les strobili mâles (couleur rouge) à l'apex de l'inflorescence, et l'initiation du cône femelle (couleur verte) à la base de cette dernière. *Restoux*

Muratorio et al. 2001). C'est pourquoi, il est nécessaire d'estimer les flux de graines pour comprendre l'impact de la structuration génétique du nuage pollinique sur la structuration génétique des individus adultes. En effet, si la dispersion des graines est limitée, la forte corrélation de paternité que nous observons se traduira par des tâches de semis très apparentés ce qui aura pour effet d'augmenter rapidement la consanguinité à chaque épisode de reproduction. A l'inverse, une dispersion importante (i.e. leptokurtique) et hétérogène permettra de conserver une grande diversité génétique, au risque de provoquer de la mal-adaptation. Il doit donc exister des distances de dispersion du pollen et des graines pour lesquelles la diversité est importante et la sélection efficace, soient les conditions nécessaires à l'adaptation locale. Ces distances idéales dépendent principalement de l'hétérogénéité du milieu, qui lorsqu'elle augmente diminue l'impact des flux de graines sur la mal-adaptation (Lopez et al. 2008). L'hétérogénéité du milieu doit donc être caractérisée, et notamment la présence de microsites favorables à l'installation des semis, qui peuvent modifier les conclusions tirées à partir des estimations de diversité génétique, de dispersion et de production de graines (Clark et al. 2007; 2001). Le recrutement des individus au sein de la population s'avère être effectivement un paramètre crucial

du maintien des populations (Wang et Smith 2002). Le succès reproducteur des arbres varie largement entre les individus, notamment en fonction de leur fertilité mais aussi du nombre d'épisodes de reproduction auxquels ils participent au cours de leur vie, et donc de leur survie et de leur croissance. Les études de conservation impliquent donc l'étude de ces caractères mais aussi la compréhension du lien qu'il existe entre ces caractères et la diversité génétique mesurée (Le Corre et Kremer 2003). Cela peut s'effectuer par des mesures d'héritabilité *in situ* qui porteraient par exemple sur la croissance, la tolérance à la sécheresse ou la fertilité.

En conclusion, nos résultats laissent présager que les populations marginales de sapin que nous avons étudiées présentent des patrons d'adaptation locale marqués et à une échelle très fine. Cependant, certains aspects du cycle de vie restent à étudier, en particulier la dispersion des graines et les caractères adaptatifs tels que la résistance à la sécheresse ou la croissance afin de confirmer cette hypothèse. Si tel est le cas, le caractère marginal de nos populations et la grande hétérogénéité du milieu qui lui est associée, conféreraient à ces populations une capacité adaptative suffisante pour résister au réchauffement climatique rapide prévu pour les prochaines décennies (Sagnard et al. 2002, dans cette étude les populations du Mont Ventoux possèdent les meilleures valeurs sélectives pour de nombreux traits). Les politiques de restauration des terrains de montagne initiées au XIXeme siécle s'avèreraient alors être un succès et servirait d'exemple de restauration naturelle "assistée" réussi pour les politiques de gestion future. Celles-ci devront également prêter un intérêt particulier aux populations marginales à l'avenir, car elles pourraient constituer un réservoir de diversité génétique sélectionnée intéressant pour les campagnes de restauration futures (Hampe et Petit 2005). Finalement la densité semble un facteur clef du maintien des populations car elle modifie le succès reproducteur des individus au travers de la limitation pollinique mais également au travers de la diversité génétique et de la consanguinité. Une gestion à visant à homogénéiser la densité au sein des populations est donc à proscrire. Les gestionnaires forestiers devraient plutôt privilégier le maintien de densités hétérogènes qui pourrait constituer une méthode efficace de conservation de la diversité génétique et des populations.

Bibliographie

Adams, W., et D. S. Birkes, 1991. Biochemical markers in the population genetics of forest trees, Chapter estimating mating patterns in forest tree populations, pages 157–172 . SPB Academic Publishing.

Agren, J. 1996. Population size, pollinator limitation, and seed set in the self-incompatible herb *Lythrum salicaria*. Ecology **77** :1779–1790.

Aizen, M. A., et L. D. Harder. 2007. Expanding the limits of the pollen-limitation concept : effects of pollen quantity and quality. Ecology **88** :271–281.

Alleaume-Benharira, M., I. R. Pen, et O. Ronce. 2006. Geographical patterns of adaptation within a species' range : interactions between drift and gene flow. Journal of Evolutionary Biology **19** :203–215.

Allison, T. D. 1990. Pollen production and plant density affect pollination and seed production in *Taxus canadensis*. Ecology **71** :516–522.

Ashman, T.-L., T. M. Knight, J. A. Steets, P. Amarasekare, M. Burd, D. R. Campbell, M. R. Dudash, M. O. Johnston, S. J. Mazer, R. J. Mitchell, M. T. Morgan, et W. G. Wilson. 2004. Pollen limitation of plant reproduction : Ecological and evolutionnary causes and consequences. Ecology **85** :2408–2421.

Austerlitz, F., C. W. Dick, C. Dutech, E. K. Klein, S. Oddou-Muratorio, P. E. Smouse, et V. L. Sork. 2004. Using genetic markers to estimate the pollen dispersal curve. Molecular Ecology **13** :937–954.

Austerlitz, F., S. Mariette, N. Machon, P. H. Gouyon, et B. Godelle. 2000. Effects of Colonization Processes on Genetic Diversity : Differences Between Annual Plants and Tree Species. Genetics **154** :1309–1321.

Austerlitz, F., et P. E. Smouse. 2001. Two-Generation analysis of pollen flow across a landscape. II. Relation betwen Phi-ft, pollen dispersal and interfemale distance. Genetics **157** :851–857.

Austerlitz, F., et P. E. Smouse. 2002. Two-Generation analysis of pollen flow across a landscape. IV. Estimating the dispersal parameter. Genetics **161** :355–363.

Baker, H. G. 1955. Self-Compatibility and Establishment After 'Long-Distance' Dispersal. Evolution **9** :347–349.

Baker, H. G. 1967. Support for Baker's law-as a rule. Evolution **21** :853–856.

Bittencourt, J., et A. Sebbenn. 2007. Patterns of pollen and seed dispersal in a small, fragmented population of the wind-pollinated tree *Araucaria angustifolia* in southern Brazil. Heredity **99** :580–591.

Bittencourt, J. V. M., et A. M. Sebbenn. 2008. Pollen movement within a continuous forest of wind-pollinated *Araucaria angustifolia*, inferred from paternity and TwoGener analysis. Conservation Genetics **9** :855–868.

Blondel, J. 2007. La biodiversité animale sur le mont Ventoux. Forêt Méditerranéenne **28** :359–368.

Brook, B. W., D. W. Tonkyn, J. J. O'Grady, et R. Frankham. 2002. Contribution of inbreeding to extinction risk in threatened species. Conservation Ecology **6**.

Burczyk, J., W. Adams, G. Moran, et A. Griffin. 2002. Complex patterns of mating revealed in a *Eucalyptus regnans* seed orchard using allozyme markers and the neighbourhood model. Molecular Ecology **11** :2379–2391.

Burczyk, J., W. Adams, et J. Y. Shimizu. 1996. Mating patterns and pollen dispersal in a natural knobcone pine (*Pinus attenuata* Lemmon.) stand. Heredity **77** :251–260.

Burczyk, J., et T. E. Koralewski. 2005. Parentage versus two-generation analyses for estimating pollen-mediated gene flow in plant populations. Molecular Ecology **14** :2525–2537.

Burczyk, J., G. Kosiński, et A. Lewandowski. 1991. Mating pattern and empty seed formation in relation to crown level *Larix decidua* Mill. clones. Silva fennica **25** :201–205.

Burczyk, J., A. Lewandowski, et W. Chalupka. 2004. Local pollen dispersal and distant gene flow in Norway spruce (*Picea abies* [L.] Karst.). Forest Ecology and Management **197** :39–48.

Byers, D. L., et D. M. Waller. 1999. Do plant populations purge their genetic load ? Effects of population size and mating history on inbreeding depression. Annual Review of Ecology and Systematics **30** :479–513.

Carcaillet, C., et S. Muller. 2005. Holocene tree-limit and distribution of *Abies alba* in the inner French Alps : Anthropogenic or climatic changes ? Boreas **34** :468–476.

Carpentier, F., J. Chadœuf, et E. K. Klein. in prep. Estimates of backward pollen dispersal curve : why and how considering selfing explicitly ? .

Castric, V., et X. Vekemans. 2004. Plant self-incompatibility in natural populations : a critical assessment of recent theoretical and empirical advances. Molecular Ecology **13** :2873–2889.

Caswell, H. 2001. Matrix population models (second edition). Sinauer associates, Sunderland, MA.

Charlesworth, D., et B. Charlesworth. 1987. Inbreeding depression and its evolutionary consequences. Annual Review of Ecology and Systematics **18** :237–268.

Cheikh Al Bassatneh, M., B. Fady, S. Simon-Tessier, et T. Tatoni. 2007. Biodiversité floristique et gestion sylvicole dans les systèmes forestiers supraméditerranéens et montagnards de la montagne de Lure (sud-est de la France). Ecologia Mediterranea **33** :29–42.

Cheliak, W. M., K. Morgan, C. Strobeck, F. C. H. Yeh, et B. P. Dancik. 1983. Estimation of mating system parameters in plant populations using the EM algorithm. Theoretical and Applied Genetics **65** :157–161.

Chen, X., X. Fan, et X. Hu. 2008. Roles of seed and pollen dispersal in natural regeneration of Castanopsis fargesii (Fagaceae) : Implications for forest management. Forest Ecology and Management **256** :1143–1150.

Cheptou, P.-O. 2004. Allee effect and self-fertilization in hermaphrodites : Reproductive assurance in demographically stable populations. Evolution **58** :2613–2621.

Cheptou, P.-O., A. Berger, A. Blanchard, C. Collin, et J. Escarre. 2000*a*. The effect of drought stress on inbreeding depression in four populations of the Mediterranean outcrossing plant *Crepis sancta* (Asteraceae). Heredity **85** :294–302.

Cheptou, P.-O., E. Imbert, J. Lepart, et J. Escarre. 2000*b*. Effects of competition on lifetime estimates of inbreeding depression in the outcrossing plant *Crepis sancta* (Asteraceae). Journal of Evolutionary Biology **13** :522–531.

Cheptou, P.-O., et D. J. Schoen. 2007. Combining genetics and demographical approaches in evolutionary studies of plant mating systems. Oïkos **116** :271–279.

Chuine, I., et E. G. Beaubien. 2001. Phenology is a major determinant of tree species range. Ecology Letters **4** :500–510.

Clark, C., J. Poulsen, D. Levey, et C. Osenberg. 2007. Are plant populations seed limited ? A critique and meta-analysis of seed additions experiments. The American Naturalist **70** :128–142.

Clark, J. S., M. Lewis, et L. Horvath. 2001. Invasion by extremes : Population spread with variation in dispersal and reproduction. The Amercian Naturalist **157** :537–554.

Clark, J. S., M. Lewis, J. S. McLachlan, et J. HilleRisLambers. 2003. Estimating Population Spread : What Can We Forecast and How Well ? Ecology **84** :1979–1988.

Coulson, T., T. Benton, P. Lundberg, S. Dall, et B. Kendall. 2006. Putting evolutionary biology back in the ecological theatre : a demographic framework mapping genes to communities. Evolutionary Ecology Research **8** :1155–1171.

Darwin, C. R. 1876. The Effects of Cross and Self-Fertilisation in the Vegetable Kingdom. John Murray.

David, P., B. Pujol, F. Viard, V. Castella, et J. Goudet. 2007. Reliable selfing rate estimates from imperfect population genetic data. Molecular Ecology **16** :2474–2487.

de Jong, T. J., J. C. Batenburg, et P. G. Klinkhamer. 2005. Distance-dependent pollen limitation of seed set in some insect-pollinated dioecious plants. Acta Oecologica **28** :331–335.

de Lucas, A., J. Robledo-Arnuncio, E. Hidalgo, et S. González-Martínez. 2008. Mating system and pollen gene flow in Mediterranean maritime pine. Heredity **100** :390–399.

Dempster, A. P., N. M. Laird, et D. B. Rubin. 1977. Maximum likelihood from incomplete data via the EM algorithm (with discussion). Journal of the Royal Statistical Society, B **39** :1–38.

Devlin, B., K. Roeder, et N. Ellstrand. 1988. Fractional paternity assignment : theoretical development and comparison to other methods. Therotical and Applied Genetics **76** :369–380.

Ellstrand, N. C. 1984. Multiple paternity within the fruits of wild radish, *Raphanus sativus*. The American Naturalist **123** :819–828.

Ellstrand, N. C. 1992*a*. Gene flow among seed plant populations. New Forests **6** :241–256.

Ellstrand, N. C. 1992*b*. Gene flow by pollen : Implications for plant conservation genetics. Oïkos **63** :77–86.

Ellstrand, N. C., et D. R. Elam. 1993. Population genetic consequences of small population size : implications for plant conservation. Annual Review of Ecology and Systematics **24** :217–242.

Eppley, S. M., et J. R. Pannell. 2007. Density-dependent self-fertilization and male versus hermaphrodite siring success in an androdioecious plant. Evolution **61** :2349–2359.

Excoffier, L. G. Laval, et S. Schneider. 2005. Arlequin ver. 3.0 : An integrated software package for population genetics data analysis. Evolutionary Bioinformatics Online **1** :47–50.

Excoffier, L., P. E. Smouse, et J. M. Quattro. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes : Application to human mitochondrial DNA restriction data. Genetics **131** :479–491.

Fady, B., I. Forest, I. Hochu, A. Ribiollet, J. de Beaulieu, et P. Pastuszka. 1999. Genetic differenciation in *Abies alba* Mill. populations from Southeastern France. Forest Genetics **6** :129–138.

Fagan, W. F., R. S. Cantrell, et C. Cosner. 1999. How habitat edges change species interactions. The American Naturalist **153** :165–182.

Feurdean, A., et K. J. Willis. 2008. Long-term variability of Abies alba in NW Romania : implications for its conservation management. Diversity and distribution **14** :1004–1017.

Fisher, R. A. 1941. Average excess and average effect of a gene substitution. Annals of Eugenics **11** :53–63.

Franceshinelli, E. V., et K. S. Bawa. 2000. The effect of ecological factors on the mating system of a South American shrub species (*Helicteres brevispira*). Heredity **84** :116–123.

Fukue, Y., T. Kado, S. L. Lee, K. K. S. Ng, N. Muhammad, et Y. Tsumura. 2007. Effects of flowering tree density on the mating system and gene flow in *Shorea leprosula* (Dipterocarpaceae) in peninsular Malaysia. Journal of Plant Research **120** :413–420.

Geng, Q., C. Lian, S. Goto, J. Tao, M. Kimura, M. S. Islam, et T. Hogetsu. 2008. Mating system, pollen and propagule dispersal, and spatial genetic structure in a

high-density population of the mangrove tree *Kandelia candel.* Molecular Ecology **17** :4724–4739.

Gigord, L., C. Lavigne, et J. A. Shykoff. 1998. Partial self-incompatibility and inbreeding depression in a native tree species of La Reunion (Indian Ocean). Oecologica **117** :342–352.

Gilpin, M., et M. Soulé, 1986. Conservation biology : The science of scarcity and diversity, Chapter minimum viable populations : Processes of species extinction, pages 19–34 . Sinauer.

Goodwillie, C., S. Kalisz, et C. G. Eckert. 2005. The evolutionary enigma of mixed mating systems in plants : Occurence, theoretical explanations, and empirical evidence. Annual Review of Ecology, Evolution and Systematics **36** :47–79.

Goodwillie, C., et M. C. Knight. 2006. Inbreeding depression and mixed mating system in *Leptosiphon jepsonii* : A comparison of three populations. Annals of Botany **98** :351–360.

Guende, G. 1978. Sensibilité des milieux et impacts des activités humaines sur le massif du Ventoux. La Terre et la Vie **supplément** :39–66.

Hampe, A., et R. J. Petit. 2005. Conserving biodiversity under climate change : the rear edge matters. Ecology Letters **8** :461–467.

Hanski, I., et O. Ovaskainen. 2003. Metapopulation theory for fragmented landscapes. Theoretical Population Biology **64** :119–127.

Harder, L., et W. Wilson. 1998. A Clarification of Pollen Discounting and Its Joint Effects with Inbreeding Depression on Mating System Evolution. The American Naturalist **152** :684–695.

Hardy, O. J., S. C. González-Martínez, B. Colas, H. Fréville, A. Mignot, et I. Olivieri. 2004*a*. Fine-scale genetic structure and gene dispersal in *Centaurea corymbosa* (Asteraceae). II. Correlated paternity within and among sibships. Genetics **168** :1601–1614.

Hardy, O. J., S. C. Gonzalez-Martínez, H. Fréville, G. Boquien, A. Mignot, B. Colas, et I. Olivieri. 2004*b*. Fine-scale genetic structure and gene dispersal in *Centaurea corymbosa* (Asteraceae) I. Pattern of pollen dispersal. Journal of Evolutionary Biology **17** :795–806.

He, T., et P. E. Smouse. 2002. Paternity analysis in *Ophiogon xylorrhizus* Wang et Tai (*Liliaceae* s.l.) : Selfing assures reproductive success. Journal of Evolutionnary Biology **15** :487–494.

Hedrick, P. W. 2004. Recent developments in conservation genetics. Forest Ecology and Management **197** :3–19.

Hedrick, P. W., O. Savolainen, et K. Kärkkäinen. 1999. Factors influencing the extent of inbreeding depression : an example from Scots pine. Heredity **82** :441–450.

Herrera, C. M., P. Jordano, J. Guitian, et A. Traveset. 1998. Annual variability in seed production by woody plants and the masting concept : Reassessment of principles and relationship to pollination and seed dispersal. The American Naturalist **152** :576–594.

Hirao, A., Y. Kameyama, M. Ohara, Y. Isagi, et G. Kudo. 2006. Seasonal changes in pollinator activity influence pollen dispersal and seed production of the Alpine shrub *Rhododendron aureum* (Ericaceae). Molecular Ecology **15** :1165–1173.

Hirayama, K., K. Ishida, S. Setsuko, et N. Tomaru. 2007. Reduced seed production, inbreeding, and pollen shortage in a small population of a threatened tree, *Magnolia stellata*. Biological Conservation **136** :315–323.

Hirayama, K., K. Ishida, et N. Tomaru. 2005. Effects of pollen shortage and self-pollination on seed production of an endangened tree, *Magnolia stellata*. Annals of Botany **95** :1009–1015.

Holsinger, K. E. 1991. Mass-action models of plant mating systems : The evolutionary stability of mixed mating systems. The American Naturalist **138** :606–622.

Inventaire Forestier National, editor. 2008. La forêt en chiffres et en cartes.

IPCC, 2007. Climate Change 2007 : Synthesis Report. Technical report.

Irwin, A. J., J. L. Hamrick, M. J. W. Godt, et P. E. Smouse. 2003. A multiyear estimate of the effective pollen donor pool for *Albizia julibrissin*. Heredity **90** :187–194.

Isagi, Y., K. Sugimura, A. Sumida, et H. Ito. 1997. How does masting happen and synchronize ? Journal of Theoretical Biology **187** :231–239.

Jensen, A. 1974. Leslie matrix models for fisheries studies. Biometrics **30** :547–551.

Johnson, J. A., R. E. Tingay, M. Culver, F. Hailer, M. L. Clarke, et D. P. Mindell. 2009. Long-term survival despite low genetic diversity in the critically endangered Madagascar fish-eagle. Molecular Ecology **18** :54–63.

Kalisz, S., et D. W. Vogler. 2003. Benefits of autonomous selfing under unpredictable pollinator environments. Ecology **84** :2928–2942.

Kamalakannan, R., M. Varghese, et D. Lindgren. 2007. Fertility variation and its implications on relatedness in seed crops in seedling seed orchards of *Eucalyptus camaldulensis* and *E-tereticornis*. Silvae Genetica **56** :253–259.

Kang, K.-S., A. Bila, A. Harju, et D. Lindgren4. 2003. Estimation of fertility variation in forest tree populations. Forestry **76** :329–344.

Kärkkäinen, K., V. Koski, et O. Savolainen. 1996. Geographical variation in the inbreeding depression of Scots pine. Evolution **50** :111–119.

Kawecki, T. J. 2008. Adaptation to Marginal Habitats. Annual Review of Ecology, Evolution and Systematics **39** :321–342.

Kelly, D., et V. L. Sork. 2002. Mast seeding in perennial plants : Why, how, where ? Annual Review of Ecology and Systematics **33** :427–447.

Kennedy, B. F., et E. Elle. 2008. The reproductive assurance benefit of selfing : importance of flower size and population size. Oecologia **155** :469–477.

Kenta, T., Y. Isagi, M. Nakagawa, M. Yamashita, et T. Nakashizuka. 2004. Variation in pollen dispersal between years with different pollination conditions in a tropical emergent tree. Molecular Ecology **13** :3575–3584.

Klein, E. K., N. Dessassis, et S. Oddou-Muratorio. 2008. Pollen flow in the wild-service tree, *Sorbus torminalis* (L.) Crantz. IV. Whole interindividual variance of male fecundity estimated jointly with the dispersal kernel. Molecular Ecology **17** :3323–3336.

Klein, E. K., C. Lavigne, et P.-H. Gouyon. 2006. Mixing of propagules from discrete sources at long distance : comparing a dispersal tail to an exponential. BMC Ecology **6**.

Knight, T. M., J. A. Steets, J. C. Vamosi, S. J. Mazer, M. Burd, D. R. Campbell, M. R. Dudash, M. O. Johnston, R. J. Mitchell, et T.-L. Ashman. 2005. Pollen limitation of plant reproduction : Pattern and process. Annual Review of Ecology, Evolution and Systematics **36** :467–497.

Koenig, W. D., et M. V. Ashley. 2003. Is pollen limited ? The answer is blowin' in the wind. Trends in Ecology and Evolution **18** :157–159.

Konnert, M., et F. Bergmann. 1995. The geographical distribution of genetic variation of silver fir (*Abies alba, Pinaceae*) in relation to its migration history. Plant Systematics and Evolution **196** :19–30.

Kormutak, A., et D. Lindgren. 1996. Mating system and empty seeds in silver fir (*Abies alba* Mill.). Forest Genetics **3** :231–235.

Krouchi, F., A. Derridj, et F. Lefevre. 2004. Year and tree effect on reproductive organisation of *Cedrus atlantica* in a natural forest. Forest Ecology and Management **197** :181–189.

Lande, R. 1988. Genetics and demography in biological conservation. Science **241** :1455–1460.

Lande, R., S. Engen, B. Saether, F. Filli, E. Matthysen, et H. Weimerskirch. 2002.

Estimating density dependence from population time series using demographic theory and life-history data. The American Naturalist **159** :321–337.

Lande, R., S. Engen, B.-E. Saether, et T. Coulson. 2006. Estimating density dependence from time series of population age structure. The American Naturalist **168** :76–87.

Lande, R., et D. W. Schemske. 1985. The evolution of self-fertilization and inbreeding depression in plants. I.Genetic models. Evolution **39** :24–40.

Lazaro, A., et A. Traveset. 2006. Reproductive success of the endangered shrub *Buxus balearica* Lam. (*Buxaceae*) : pollen limitation, and inbreeding and outbreeding depression. Plant Systematics and Evolution **261** :117–128.

Le Corre, V., et A. Kremer. 2003. Genetic Variability at Neutral Markers, Quantitative Trait Loci and Trait in a Subdivided Population Under Selection. Genetics **164** :1205–1219.

Leblois, R., F. Rousset, et A. Estoup. 2004. Influence of Spatial and Temporal Heterogeneities on the Estimation of Demographic Parameters in a Continuous Population Using Individual Microsatellite Data. Genetics **166** :1081–1092.

Leclerc-Potvin, C., et K. Ritland. 1994. Modes of self-fertilization in Mimulus guttatus (Scrophulariaceae) : A field experiment. American Journal of Botany **81** :199–205.

Leslie, P. 1945. On the use of matrices in certain population mathematics. Biometrika **33** :183–212.

Liepelt, S., R. Cheddadi, J. L. de Beaulieu, B. Fady, D. Gomory, E. Hussendorfer, M. Konnert, T. Litt, R. Longauer, R. Terhurne-Berson, et B. Ziegenhagen. 2009. Postglacial range expansion and its genetic imprints in Abies alba (Mill.) - A synthesis from palaeobotanic and genetic data. Review Of Palaeobotany And Palynology **153** :139–149.

Liepelt, S., V. Kuhlenkamp, M. Anzidei, G. G. Vendramin, et B. Ziegenhagen. 2001.

Pitfalls in determining size homoplasy of microsatellites loci. Molecular Ecology Notes **1** :332–335.

Lindgren, D. 1975. The relationship between self-fertilization, empty seeds and seeds originating from selfing as a consequence of polyembryony. Studia Forestalia Suecica **126** :1–24.

Liu, T.-S. 1971. A monograph of the genus *Abies*. Department of Forestry, College of agriculture, National Taiwan University, Taipei, Taiwan, China.

Lloyd, D. G. 1979. Some reproductive factors affecting the selection of self-Fertilization in plants. The American Naturalist **113** :67–79.

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

Lopez, S., F. Rousset, F. Shaw, R. Shaw, et O. Ronce. 2008. Migration load in plants : role of pollen and seed dispersal in heterogeneous landscapes. Jouranl of Evolutionary Biology **21** :294–309.

Maurice, S., et T. H. Fleming. 1995. The effect of pollen limitation on plant reproductive systems and the maintenance of sexual polymorphisms. Oïkos **74** :55–60.

McCarthy, B., et J. A. Quinn. 1990. Reproductive ecology of *Carya* (*Juglandaceae*) - Phenology, Pollination, and breeding system of 2 Sympatric tree species. American Jouranl of Botany **77** :261–273.

McLachlan, J. S., J. S. Clark, et P. S. Manos. 2005. Molecular indicators of tree migration capacity under rapid climate change. Ecology **86** :2088–2098.

Meagher, T. R., et C. Vassiliadis, 2003. Genes in the environment, Chapter spatial geometry determines gene flow in plant populations, pages 76–90 . Blackwell Publishing, Oxford, UK.

Mimura, M., et S. N. Aitken. 2007*a*. Adaptive gradients and isolation-by-distance with postglacial migration in *Picea sitchensis*. Heredity **99** :224–232.

Mimura, M., et S. N. Aitken. 2007*b*. 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.

Minami, S., et A. Azuma. 2003. Various flying modes of wind-dispersal seeds. Journal of Theoretical Biology **225** :1–14.

Morgante, M., G. Vendramin, et P. Rossi. 1991. Effects of stand density on outcrossing rate in two Norway spruce (*Picea abies*) populations. Canadian Journal of Botany **69** :2704–2708.

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

Naito, Y., M. Kanzaki, H. Iwata, K. Obayashi, S. L. Lee, N. Muhammad, T. Okuda, et Y. Tsumura. 2008. Density-dependent selfing and its effects on seed performance in a tropical canopy tree species, *Shorea acuminata* (Dipterocarpaceae). Forest Ecology and Management **256** :375–383.

Nei, M. 1973. Analysis of gene diversity in subdivided populations. PNAS **70** :3321–3323.

Nunney, L., et K. A. Campbell. 1993. Assessing minimum viable population size : Demography meets population genetics. Trends in Ecology and Evolution **8** :234–239.

Oakley, C. G., et A. A. Winn. 2008. Population-level and family-level inbreeding depression in a cleistogamous perennial. INTERNATIONAL JOURNAL OF PLANT SCIENCES **169** :523–530.

Oddou-Muratorio, S., E. K. Klein, et F. Austerlitz. 2005. Pollen flow in the wild-service tree, *Sorbus torminalis* (L.) Crantz. II. Pollen dispersal and heterogeneity in mating success inferred from parent-offspring analysis. Molecular Ecology **14** :4441–4452.

Oddou-Muratorio, S., E. K. Klein, B. Demesure-Musch, et F. Austerlitz. 2006. Real-time patterns of pollen flow in the wild-service tree, Sorbus torminalis (Rosaceae). III. Mating patterns and the ecological maternal neighborhood. Am. J. Bot. **93** :1650–1659. URL `http ://www.amjbot.org/cgi/content/abstract/93/11/1650`.

Oddou-Muratorio, S., R. Petit, B. Le Guerroue, D. Guesnet, et B. Demesure. 2001. Pollen- versus seed-mediated gene flow in a scattered forest tree species. Evolution **55** :1123–1135.

Okubo, A., et S. Levin. 1989. A theoretical framework for data-analysis of wind-dispersal seeds and pollen. Ecology **70** :329–338.

Ovaskainen, O., et I. Hanski. 2004. From individual behavior to metapopulation dynamics : Unifying the patchy population and classic metapopulation models. The Amercian Naturalist **164** :364–377.

Owens, J. N., T. Takaso, et C. J. Runions. 1998. Pollination in conifers. Trends in Plant Science **3** :479–485.

Papageorgiou, A. C., A. Vidalis, O. Gailing, I. Tsiripidis, S. Hatziskakis, S. Boutsios, S. Galatsidas, et R. Finkeldey. 2008. Genetic variation of beech (Fagus sylvatica L.) in Rodopi (NE Greece). European Journal Of Forest Research **127** :81–88.

Parmesan, C. 2006. Ecological and Evolutionary Responses to Recent Climate Change. Annual Review of Ecology, Evolution and Systematics **37** :637–669.

Petit, R., J. Duminil, S. Fineschi, A. Hampe, D. Salvini, et G. Vendramin. 2005. Comparative organization of chloroplast, mitochondrial and nuclear diversity in plant populations. Molecular Ecology **14** :689–701.

Petit, R. J., et A. Hampe. 2006. Some Evolutionary consequences of being a tree. Annual Review of Ecology, Evolution and Systematics **37** :187–214.

Petit, R. J., E. Pineau, B. Demesure, R. Bacilieri, A. Ducousso, et A. Kremer. 1997. Chloroplast DNA footprints of postglacial recolonization by oaks. Proceedings Of

The National Academy Of Sciences Of The United States Of America **94** :9996–10001.

Pichot, C., C. Bastien, F. Courbet, B. Demesure-Musch, P. Dreyfus, B. Fady, N. Frascaria-Lacoste, S. Gerber, F. Lefevre, M.-E. Morand-Prieur, S. Oddou, E. Tessier Du Cros, et A. Valadon. 2006. Déterminants et conséquences de la qualité génétique des graines et semis lors de la phase initiale de régénération naturelle des peuplements forestiers. Les actes du BRG **6** :277–297.

Porcher, E., et R. Lande. 2005*a*. The evolution of self-fertilization and inbreeding depression under pollen discounting and pollen limitation. Journal of Evolutionary Biology **18** :497–508.

Porcher, E., et R. Lande. 2005*b*. Reproductive compensation in the evolution of plant mating systems. New Phytologist **166** :673–684.

Reed, D., et R. Frankham. 2003. Correlation between Fitness and Genetic Diversity. Conservation Biology **17** :230–237.

Reed, D., E. Lowe, D. Briscoe, et R. Frankham. 2003. Inbreeding and extinction : Effects of rate of inbreeding. Conservation Genetics **4** :405–410.

Restoux, G., D. E. Silva, F. Sagnard, F. Torre, E. K. Klein, et B. Fady. 2008. Life at the margin : the mating system of Mediterranean conifers. Web Ecology **8** :94–102.

Ritland, K. 1986. Joint maximum likelihood estimation of genetic and mating structure using open-pollinated progenies. Biometrics **42** :25–43.

Ritland, K. 1989. Correlated matings in the partial selfer *Mimulus guttatus*. Evolution **43** :848–859.

Ritland, K. 2002. Extensions of models for the estimation of mating systems using n independent loci. Heredity **88** :221–228.

Ritland, K., et S. Jain. 1981. A model for the estimation of outcrossing rate and gene frequencies using n independent loci. Heredity **47** :35–52.

Robledo-Arnuncio, J., et L. Gil. 2005. Patterns of pollen dispersal in a small population of *Pinus sylvestris* L. revealed by total-exclusion paternity analysis. Heredity **94** :13–22.

Robledo-Arnuncio, J.-J., R. Alía, et L. Gil. 2004*a*. Increased selfing and correlated paternity in a small population of predominantly outcrossing conifer, *Pinus sylvestris*. Molecular Ecology **13** :2567–2577.

Robledo-Arnuncio, J. J., et F. Austerlitz. 2006. Pollen dispersal in spatially aggregated populations. The American Naturalist **168** :500–511.

Robledo-Arnuncio, J. J., F. Austerlitz, et P. E. Smouse. 2006. A new method of estimating the pollen dispersal curve independently of effective density. Genetics **173** :1033–1045.

Robledo-Arnuncio, J. J., P. E. smouse, L. Gil, et R. Alía. 2004*b*. Pollen movement under alternative silvicultural practices in native populations of Scots pine (*Pinus sylvestris* L.) in central Spain. Forest Ecology and Management **197** :243–253.

Roff, D. A. 2002. Inbreeding depression : Tests of the overdominance and partial dominance hypotheses. Evolution **56** :768–775.

Rousset, F. 1997. Genetic Differentiation and Estimation of Gene Flow from F-Statistics under Isolation by Distance. Genetics **145** :1219–1228.

Rousset, F. 2000. Genetic differentiation between individuals. Jouranl of Evolutionary Biology **13** :58–62.

Sagnard, F., 2001. Dynamique de recolonisation des pinèdes par la hêtraie-sapinière : Etude de la structure des peuplements forestiers et de l'évolution génétique des populations de sapin pectiné (*Abies alba* Mill.) sur le Mont Ventoux. Ph.D. thesis, Université d'Aix-Marseille III.

Sagnard, F., C. Barberot, et B. Fady. 2002. Structure of Genetic diversity in Abies alba Mill. from southwestern Alps : multivariate analysis of adaptive and non-adaptive traits for conservation in France. Forest Ecology And Management **157** :PII S0378–1127(00)00664–2.

Sagnard, F., C. Pichot, P. Dreyfus, P. Jordano, et B. Fady. 2007. Modelling seed dispersal to predict seedling recruitment : Recolonization dynamics in a plantation forest. Ecological Modelling **203** :464–474.

Satake, A. 2004. Modeling spatial dynamics of episodic and synchronous reproduction by plant populations : the effect of small-scale pollen coupling and large-scale climate. Population Ecology **46** :119–128.

Satake, A., et Y. Iwasa. 2002. The synchronized and intermittent reproduction of forest trees is mediated by the Moran effect, only in association with pollen coupling. Journal of Ecology **90** :830–838.

Savolainen, O., T. Pyhajarvi, et T. Knurr. 2007. Gene flow and local adaptation in trees. Annual Review of Ecology, Evolution and Systematics **38** :595–619.

Schemske, D. W., et R. Lande. 1985. The evolution of self-fertilization and inbreeding depression in plants. II. Empirical observations. Evolution **39** :41–52.

Schossig, G., 2004. *Abies alba* et les sapins méditerranéens : Y-a-t-il un risque de pollution génétique sur la montagne de Lure ? Master's thesis, Université Claude Bernard, Lyon I.

Schwartz, M. W., S. M. Hermann, et P. J. V. Mantgem. 2000. Population persistence in Florida Torreya : Comparing modeled projections of a declining coniferous tree. Conservation Biology **14** :1023–1033.

Schwendemann, A. B., G. Wang, M. L. Mertz, R. T. McWilliams, S. L. Thatcher, et J. A. Osborn. 2007. Aerodynamics of saccate pollen and its implications for wind pollination. American Journal Of Botany **94** :1371–1381.

Shaberg, P. G., D. H. DeHayes, G. J. Hawley, et S. E. Nijensohn. 2008. Anthropogenic alterations of genetic diversity within tree populations : implications for forest ecosystem resilience. Forest Ecology and Management **256** :855–862.

Shimatani, K., M. Kimura, K. Kitamura, Y. Suyama, Y. Isagi, et H. Sugita. 2007. Determining the location of a deceased mother tree and estimating forest regen-

eration variables by use of microsatellites and spatial genetic models. Population Ecology **49** :317–330.

Shimono, A., et I. Washitani. 2002. Factors affecting variation in seed production in the heterostylous herb *Primula modesta*. Plant Species Biology **22** :65–76.

Simon, L., V. Clément, et P. Pech. 2007. Forestry disputes in provincial France during the nineteenth century : the case of the Montagne de Lure. Journal of Historical Geography **33** :335–351.

Smouse, P., et T. Meagher. 1994. Genetic analysis of male reproductive contributions in *Chamaelirium luteum* (L.) Gray (Liliaceae). Genetics **136** :313–322.

Smouse, P. E., R. J. Dyer, R. D. Westfall, et V. L. Sork. 2001. Two-Generation analysis of pollen flow accross a landscape. I. Male gamete heterogeneity among females. Evolution **55** :260–271.

Smouse, P. E., T. R. Meagher, et C. J. Kobak. 1999. Parentage analysis in *Chamaelirium luteum* (L.) Gray (Liliaceae) : why do some males have higher reproductive contributions ? Journal of Evolutionary Biology **12** :1069–1077.

Smouse, P. E., et V. L. Sork. 2004. Measuring pollen flow in forest trees : an exposition of alternative approaches. Forest Ecology and Management **197** :21–38.

Sorensen, F. 1969. Embryonic genetic load in coastal Douglas-fir, *Pseudotsuga menziesii* var. Menziesii. The American Naturalist **103** :389–398.

Sorensen, F. C. 1982. The roles of polyembryony and embryo viability in the genetic system of conifers. Evolution **36** :725–733.

Sorensen, F. C., et J. E. Webber. 1997. On the relationship between pollen capture and seed set in conifers. Canadian Journal of Forest Research **27** :63–68.

Sork, V., J. Nason, D. Campbell, et J. Fernandez. 1999. Landscape approaches to historical and contemporary gene flow in plants. Trends in Ecology and Evolution **14** :219–224.

Soulé, M. E. 1985. What is conservation biology ? Bioscience **35** :727–734.

Spielman, D., B. Brook, et R. Frankham. 2004. Most species are not driven to extinction before genetic factors impact them. PNAS **101** :15261–15264.

Streiff, R., A. Ducousso, C. Lexer, H. Steinkellner, J. Gloessl, et A. Kremer. 1999. Pollen dispersal inferred from paternity analysis in a mixed oak stand of *Quercus robur* L. and *Q. petraea* (Matt.) Liebl. Molecular Ecology **8** :831–841.

Tollefsrud, M. M., R. Kissling, F. Gugerli, O. Johnsen, T. Skroppa, R. Cheddadi, W. O. van Der Knaap, M. Latalowa, R. Terhurne-Berson, T. Litt, T. Geburek, C. Brochmann, et C. Sperisen. 2008. Genetic consequences of glacial survival and postglacial colonization in Norway spruce : combined analysis of mitochondrial DNA and fossil pollen. Molecular Ecology **17** :4134–4150.

Tsitrone, A., S. Duperron, et P. David. 2003*a*. Delayed selfing as an optimal mating strategy in preferentially outcrossing species : Theoretical analysis of the optimal age at first reproduction in relation to mate availability. The American Naturalist **162** :318–331.

Tsitrone, A., P. Jarne, et P. David. 2003*b*. Delayed selfing and resource reallocations in relation to mate availability in the freshwater snail *Physa acuta*. The American Naturalist **162** :474–488.

Vendramin, G., et B. Ziegenhagen. 1997. Characterisation and inheritance of polymorphic plastid microsatellites in *Abies*. Genome **40** :857–864.

Vendramin, G. G., L. Lelli, P. Rossi, et M. Morgante. 1996. A set of primers for the amplification of 20 chloroplast microsatellites in *Pinaceae*. Molecular Ecology **5** :595–598.

Vergeer, P., R. Rengelink, A. Copal, et N. J. Ouborg. 2003. The interacting effects of genetic variation, habitat quality and population size on performance of *Succisa pratensis*. Journal of Ecology **91** :18–26.

von Aderkas, P., G. Rouault, R. Wagner, R. Rohr, et A. Roques. 2005. Seed

parasitism redirects ovule development in Douglas fir. Proceedings Of The Royal Society B-Biological Sciences **272** :1491–1496.

Wang, B. C., et T. B. Smith. 2002. Closing the seed dispersal loop. Trends in Ecology and Evolution **17** :379–385.

Wang, K. S. 2003. Relationship between empty seed and genetic factors in European beech (*Fagus sylvatica* L.). Silva Fennica **37** :419–428.

Williams, C. G., et O. Savolainen. 1996. Inbreeding depression in conifers : Implications for breeding strategy. Forest Science **42** :102–117.

With, K. A. 1997. The theory of conservation biology. Conservation Biology **11** :1436–1440.

Wright, S. 1943. Isolation by distance. Genetics **28** :114–138.

Wright, S. 1946. Isolation by distance under diverse systems of mating. Genetics **31** :39–59.