



HAL
open science

Effets de la dispersion du pollen à longue distance sur les capacités d'adaptation de populations de hêtre commun le long d'un gradient altitudinal

Julie Gauzere

► **To cite this version:**

Julie Gauzere. Effets de la dispersion du pollen à longue distance sur les capacités d'adaptation de populations de hêtre commun le long d'un gradient altitudinal. Sciences du Vivant [q-bio]. Université Montpellier 2 (Sciences et Techniques), 2014. Français. NNT : . tel-02793712

HAL Id: tel-02793712

<https://hal.inrae.fr/tel-02793712>

Submitted on 5 Jun 2020

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.

THÈSE

Pour obtenir le grade de
Docteur

Délivré par le
**Centre international d'études supérieures en
sciences agronomiques**
Montpellier

Préparée au sein de l'école doctorale SIBAGHE
Systèmes Intégrés en Biologie, Agronomie, Géosciences, Hydrosiences et
Environnement

Et de l'unité de recherche :
URFM Écologie des forêts méditerranéennes

Spécialité : **Écologie, Évolution, Ressources Génétiques,
Paléontologie**

Présentée par **Julie GAÜZERE**

**Effets de la dispersion du pollen à longue distance
sur les capacités d'adaptation de populations de
Hêtre commun le long d'un gradient altitudinal**

Soutenue le 31 octobre 2014 devant le jury composé de

M. Arndt HAMPE, DR, INRA Bordeaux

Rapporteur

M. Benoît PUJOL, CR, CNRS Toulouse

Rapporteur

Mme Anne CHARMANTIER, DR, CNRS Montpellier

Examinateur

Mme Emmanuelle PORCHER, CR, Muséum d'Histoire
naturelle Paris

Examinateur

M. Etienne KLEIN, DR, INRA Avignon

Directeur de thèse

Mme Sylvie ODDOU-MURATORIO, CR, INRA Avignon

Co-directeur de thèse

THÈSE

Pour obtenir le grade de
Docteur

Délivré par le
**Centre international d'études supérieures en
sciences agronomiques
Montpellier**

Préparée au sein de l'école doctorale SIBAGHE
Systèmes Intégrés en Biologie, Agronomie, Géosciences, Hydrosiences et
Environnement

Et de l'unité de recherche :
URFM Écologie des forêts méditerranéennes

Spécialité : **Écologie, Évolution, Ressources Génétiques,
Paléontologie**

Présentée par **Julie GAÜZERE**

**Effets de la dispersion du pollen à longue distance
sur les capacités d'adaptation de populations de
Hêtre commun le long d'un gradient altitudinal**

Soutenue le 31 octobre 2014 devant le jury composé de

M. Arndt HAMPE, DR, INRA Bordeaux

Rapporteur

M. Benoît PUJOL, CR, CNRS Toulouse

Rapporteur

Mme Anne CHARMANTIER, DR, CNRS Montpellier

Examinateur

Mme Emmanuelle PORCHER, CR, Muséum d'Histoire
naturelle Paris

Examinateur

M. Etienne KLEIN, DR, INRA Avignon

Directeur de thèse

Mme Sylvie ODDOU-MURATORIO, CR, INRA Avignon

Co-directeur de thèse

REMERCIEMENTS

Difficile d'être exhaustif dans ces remerciements tant les rencontres, échanges et soutiens ont été nombreux durant ces quatre années passées à l'URFM. Je vais tout de même me lancer dans cet exercice périlleux...

Je tiens tout d'abord à remercier les membres de mon jury, Arndt Hampe, Benoît Pujol, Anne Charmantier et Emmanuelle Porcher, d'avoir accepté d'évaluer ma thèse.

Un immense merci à mes directeurs de thèse, Étienne et Sylvie, tout d'abord pour m'avoir accordé votre confiance durant un stage de M2, puis en faisant tout le nécessaire pour que je puisse poursuivre en thèse. Je ne pense pas qu'on puisse rêver meilleur co-encadrement, tant vous avez su être complémentaire sans jamais être en contradiction. Merci de vous être toujours rendu disponibles, malgré, Étienne, tes responsabilités de DU, et, Sylvie, lors de la période où tu encadrais 3 thèses de front ! Merci pour vos grandes qualités pédagogiques (Étienne j'aurai du garder précieusement tous les schémas que tu griffonnais pendant nos réunions) et humaines qui rendaient nos échanges toujours très agréables. Après plusieurs années de collaboration (déjà très fructueuses), je suis très fière d'avoir été la première thésarde que vous ayez co-encadré. J'espère que vous renouvelerez cette expérience afin que d'autres étudiants puissent bénéficier de ce cadre de travail privilégié.

Merci aux membres de mes comités de thèse, Ophélie Ronce, Bénédicte Quilot, Frédéric Mortier, ainsi qu'à Hendrik Davi, François Lefèvre, Oliver Brendel et Laurène Gay, avec qui les discussions et travaux sont allés bien au delà des comités. Merci également à Frédéric Jean, Christian Pichot, Bruno Fady, Éric Rigolot, Denis Allard et Christina Garcia pour nos échanges, discussions et collaborations.

Le bon déroulement de cette thèse a été très largement dû à la qualité des équipes techniques d'Avignon et d'ailleurs. Pour cela, merci à tout le personnel de l'ex- pépinière d'État d'Aix les Milles (merci beaucoup à Patrice et Marie d'avoir fait le déplacement pour ma soutenance) ; merci aux collègues de l'UEFM : Fred (encore), Mehdi, Norbert, Olivier G., Frank, Marianne et Olivier C. ; enfin merci au personnel de la plateforme d'écologie fonctionnelle

de Nancy pour leur accueil et pour m'avoir formé à l'art de la micro-pesée. Un grand merci à Anne et Matthieu pour leur aide dans le travail de mise en plaque et de génotypage (sur matériel frais et sec!).

Merci aux femmes de L'URFM et BioSP qui m'ont facilité la vie pendant ces 4 ans, et là je pense aux gestionnaires de projet, Marie-Claude et Sylvie J., et aux informatiennes Valérie et Malika.

Et comment parler de cette thèse sans parler de la PhD dream team and co. ! Merci infiniment à Aurore et Marion, mes deux superbes collègues de bureau, à Cindouille, Mijou, Guigeeek, Hadrien, Marianne, Pierrot. Je pense que la qualité et le plaisir que j'ai eu à réaliser cette thèse n'auraient pas été les mêmes sans nos discussions, co-formation à \LaTeX , apéros, soirées cinés et jeux, et weekend-excursions dans les coins les plus pommés de France. Mijou, je ne pourrais jamais suffisamment te remercier d'avoir retardé ta fin de thèse pour pouvoir rédiger avec moi ce fameux été 2014 ! Merci aux super post-doctorants et nouvelles doctorantes de l'équipe, Katy et Philippe, Anne et Guia, pour nos discussions plus ou moins professionnelles mais toujours enrichissantes. Merci également aux stagiaires qui sont venus animer nos étés : Jonathan, Marie, François M., Célia.

Merci aux ambianceurs de la salle café et plus généralement des couloirs de l'URFM, Laurent, Thomas, Marion S., Gema, Anne-Sophie, Renée, William, Arnaud, Roland, Florence...

Au delà de cette sphère professionnelle, je tiens à remercier chaleureusement les personnes qui me soutiennent et me supportent depuis de nombreuses années déjà : tout d'abord ma famille (peu de personnes peuvent, par exemple, se vanter d'avoir une maman qui ait relu entièrement sa thèse ou une petite sœur qui ait assisté à l'élaboration du plan de l'oral!); puis mes amis, Lola, Laure, Sophie, Paté, Bapt, Pauline, Mimi, Mamoule, Fayette, Ophé, Luc, Gayell, Gazelle, Kuz. Vous êtes de vrais bols d'oxygène au quotidien ! Enfin, mon remerciement le plus simple et profond ira à François ; merci pour tout.

INTRODUCTION



1.1 CONTEXTE GÉNÉRAL

Changements climatiques et capacités adaptatives des forêts

1.1.1 Les changements climatiques et leurs conséquences

Les observations climatiques sur le dernier siècle permettent d'affirmer que les activités anthropiques ont généré des modifications du climat sans précédent (d'après le GIEC¹ ; IPCC 2013). Ces changements climatiques se sont notamment manifestés par une augmentation globale des températures (+ 0,85 [0,65 ; 1,06] °C au cours de la période 1880-2012 ; Figure 1.1 ; IPCC 2013) et du nombre d'anomalies thermiques : les trente années entre 1983 à 2012 constituent probablement la période la plus chaude qu'ait connue l'hémisphère Nord depuis 1 400 ans. Les activités anthropiques seraient à l'origine d'une augmentation de 40 % de la concentration de dioxyde de carbone (CO₂). Ce CO₂ a été à la fois accumulé dans l'atmosphère terrestre, augmentant l'effet de serre, absorbé par les océans, entraînant une acidification des eaux, et stocké par les écosystèmes terrestres.

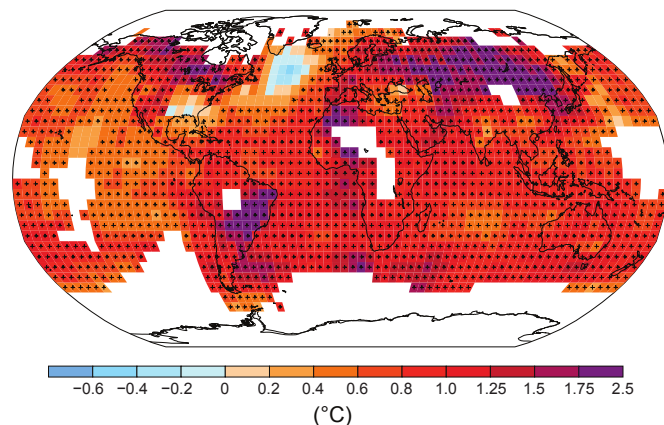


FIGURE 1.1 – Évolution des températures en surface terrestre et océanique observée entre 1901 et 2012. Figure extraite de IPCC 2013.

Tous les scénarios climatiques prédisent que la concentration en CO₂ atmosphérique et les températures globales vont continuer à augmenter dans le siècle à venir (IPCC, 2013). Récemment, Loarie *et al.* (2009) ont estimé que les enveloppes thermiques pourraient se déplacer en moyenne de 42 km sur la surface du globe en un siècle. D'après le GIEC, ces changements climatiques rapides, les perturbations connexes (tempêtes, sécheresses, feux incontrôlés, insectes,

1. Groupe d'experts intergouvernemental sur l'évolution du climat (IPCC en anglais)

etc...) et d'autres facteurs de changement à l'échelle planétaire (changement d'usage des sols, fragmentation de l'habitat naturel, surexploitation des ressources, etc...) devraient générer une modification des aires de répartition des espèces et un risque croissant d'extinction. En se basant sur des modèles de déplacement des enveloppes bioclimatiques des espèces, des études prédisent que le risque d'extinction pourrait atteindre 30 à 40 % d'ici 2050 (Thomas *et al.*, 2004). Dans ce contexte, la question de la vitesse à laquelle les espèces peuvent répondre aux changements environnementaux est plus que jamais au centre des intérêts.

1.1.2 Mécanismes de réponse des espèces aux changements environnementaux

Trois mécanismes peuvent entrer en jeu dans la réponse des espèces aux changements environnementaux : la plasticité phénotypique, la migration et l'évolution génétique (Figure 1.2; Jump & Penuelas 2005). Dans la plupart des espèces, les individus peuvent tout d'abord tolérer la variabilité du climat grâce à leur plasticité phénotypique, i.e. leur capacité à modifier leur phénotype en fonction des conditions environnementales. Néanmoins, au delà d'une certaine gamme de variation environnementale, cette réponse peut ne pas suffire à permettre le maintien des espèces. La survie des espèces à long terme va dépendre de leurs capacités de migration (i.e., le déplacement physique des populations) pour suivre le déplacement des enveloppes bioclimatiques et d'adaptation génétique aux nouvelles conditions pour permettre le maintien des populations dans leur environnement actuel. À petite échelle temporelle la mutation a des effets négligeables et les capacités d'adaptation génétique reposent essentiellement sur la diversité génétique présente au sein des populations et sur le brassage génétique lors des événements de reproduction.

Une manière de comprendre le rôle de ces différents mécanismes dans l'adaptation des espèces modernes au climat futur est d'étudier comment elles se sont maintenues dans le passé. Une première constatation est que la majorité des espèces ayant vécu sur terre sont aujourd'hui éteintes; les réponses sont donc limitées. Le quaternaire est la période la plus récente ayant connu de fortes fluctuations climatiques à cause de la succession de nombreuses périodes glaciaires et interglaciaires. Durant les cycles glaciaires-interglaciaires, la vitesse des changements climatiques a été très variable, avec à la fois des variations climatiques très abruptes (plus de 10 ° C en quelques décennies sur certaines zones géographiques; Petit *et al.*

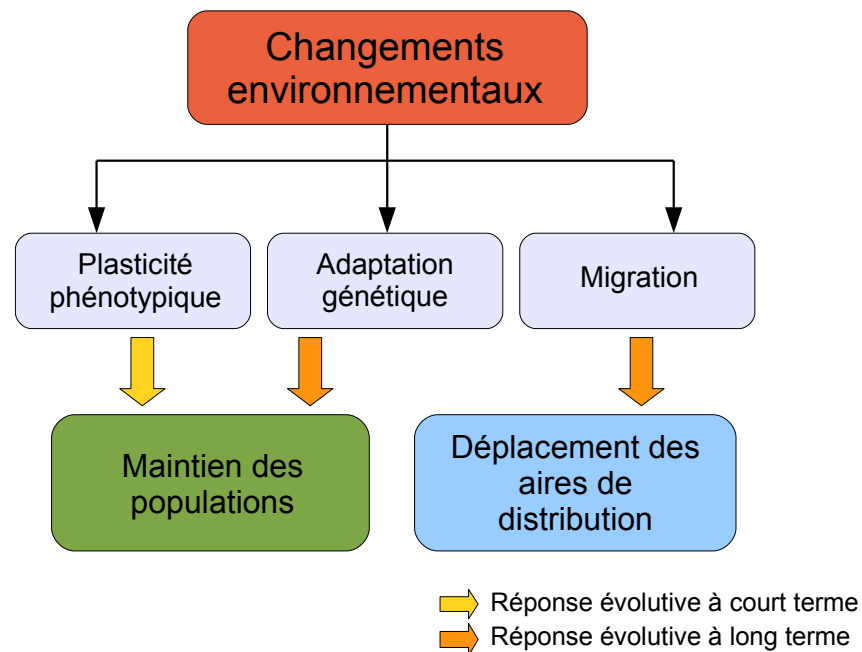


FIGURE 1.2 – Mécanismes impliqués dans la réponse aux changements environnementaux permettant la survie des espèces par le maintien des populations dans leur environnement actuel ou le déplacement des aires de distribution.

2008), et des périodes où le climat a été quasi-constant (variation de 1°C par millénaire ; Davis & Shaw 2001). Les données fossiles montrent que durant cette période les espèces ne s'étant pas éteintes ont connu d'importantes modifications de leurs aires de répartition, en se déplaçant notamment vers et à partir de zones climatiquement favorables appelées "zones refuges". Ces résultats soulignent le rôle majeur de la migration dans la réponse des espèces aux changements environnementaux. Cependant, le processus d'adaptation génétique des espèces à leur nouvel environnement a probablement été largement sous-évalué dans la réponse aux changements climatiques passés (Davis & Shaw, 2001; Hampe & Petit, 2005). Les incertitudes sur le rôle de l'évolution génétique des populations dans le passé, et par extrapolation dans le futur, peuvent notamment s'expliquer par l'a priori hérité de la vision Darwinienne de l'évolution, selon laquelle ce processus ne peut pas permettre une réponse sur des pas de temps courts ("we see nothing of these slow changes in progress, until the hand of time has marked the long lapse of ages" ; Darwin 1859 cité dans Hendry & Kinnison 1999). Pourtant, il existe aujourd'hui des preuves empiriques que l'évolution génétique peut se produire à l'échelle de quelques générations seulement (Berteaux *et al.*, 2004; Kitano *et al.*, 2008). Ce processus d'évolution sur des échelles temporelles courtes est communément appelé "microévolution" (même si ce terme a tout d'abord été défini pour caractériser les changements se produisant

entre populations au sein d'une même espèce), et focalise aujourd'hui l'attention des biologistes de l'évolution (Hendry & Kinnison, 1999).

À l'heure actuelle nous disposons de preuves que les changements climatiques contemporains ont déjà généré des réponses adaptatives reposant sur les trois mécanismes présentés chez certaines espèces (Root *et al.*, 2003; Parmesan, 2006; Chen *et al.*, 2011). Notamment, les aires de distributions de nombreuses espèces terrestres se décalent actuellement vers des latitudes et altitudes plus élevées, avec des vitesses médianes de 16,9 km/décennie et 11,0 m/décennie respectivement (Chen *et al.*, 2011). Dans une localité donnée, de nombreux traits phénologiques, chez les plantes comme les animaux, ont également évolué en lien avec l'augmentation des températures (Figure 1.3 ; Root *et al.* 2003; Berteaux *et al.* 2004). Même si la réponse de ces traits est majoritairement due à une réponse plastique, chez certaines espèces une part significative de cette réponse peut également être expliquée par de l'évolution génétique : chez l'écureuil du Yukon par exemple, l'avancement de 18 jours des dates de parturition a été expliqué à 62 % par une réponse plastique et 13 % par une adaptation génétique (Berteaux *et al.*, 2004).

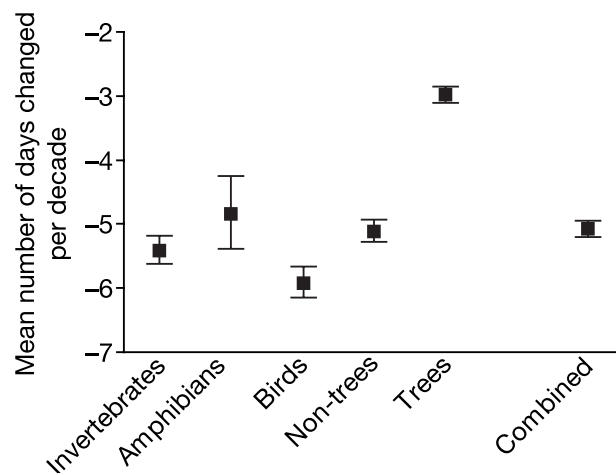


FIGURE 1.3 – Décalages phénologiques moyens (+ erreur standard) en nombre de jours par décennie chez différents taxons. Figure extraite Root *et al.* (2003)

Ces résultats soulignent l'importance de prendre en compte les mécanismes de migration et adaptation au sens large (plasticité phénotypique et adaptation génétique) dans les modèles de prédiction d'aire de répartition future des espèces. Deux grands types d'approches permettent de réaliser ces prédictions : les modèles corrélatifs (souvent appelés "modèles de

niche bioclimatique"²) et les modèles "basés sur les processus"³. Les modèles de niche sont basés sur des modèles statistiques reliant les données de présence-absence ou d'abondance des espèces aux variables climatiques afin de prédire leurs aires de distribution potentielles en climat changeant. Par définition, ces modèles négligent l'effet des processus adaptatifs, même s'ils peuvent être utilisés de manière détournée pour prédire les aires de distribution futures selon des valeurs extrêmes de certains processus adaptatifs. C'est par exemple ce qui peut être réalisé pour la migration, qui peut être considérée soit comme illimitée, soit comme nulle par ces modèles. Les prédictions passent alors de 30-40 % d'extinction dans des scénarios de dispersion nulle à 21-27 % dans les scénarios de dispersion illimitée (Thomas *et al.*, 2004). À l'opposé, les modèles basés sur les processus utilisent une modélisation mécaniste de la réponse des espèces aux variables environnementales ; il existe un grand nombre de modèles, selon les processus pris en compte : e.g., la phénologie chez les arbres (Chuine & Beaubien, 2001), les capacités de migration (Nathan *et al.*, 2011; Corlett & Westcott, 2013). Cependant à l'heure actuelle, il existe peu de modèles capables d'intégrer l'effet de l'adaptation génétique sur l'aire future de distribution d'une espèce donnée (en dehors des modèles théoriques ; voir sous-section 1.2.1.3 ; Duputié *et al.* 2012). De plus, l'efficacité de ces approches et leurs prédictions sur les capacités évolutives des espèces sont fortement limitées par les données et informations disponibles pour les paramétrer.

En outre, les difficultés à prédire les taux d'extinction des espèces et à comprendre le poids que joueront les différents mécanismes de réponse que sont la plasticité, d'adaptation génétique et la migration, sont accentuées par l'interaction complexe qui existe entre ces trois réponses. Notamment, la migration et l'adaptation génétique ne peuvent pas être vues comme de simples alternatives. En effet, l'adaptation génétique a vraisemblablement lieu en même temps que la dispersion pour permettre la colonisation de nouveaux milieux (Davis & Shaw, 2001). À l'inverse, dans une localité donnée, la migration pourrait favoriser la diffusion d'allèles avantageux et permettre aux populations de répondre plus efficacement à la sélection dans un contexte d'environnement changeant. De plus, les capacités de réponse plastique et les traits de dispersion peuvent évoluer sous effet de la sélection (Ghalambor *et al.*, 2007; Charmantier *et al.*, 2008). Ainsi, il semble aujourd'hui crucial de pouvoir comprendre l'interaction entre ces forces évolutives au sein des populations naturelles afin d'être capable faire des inférences

2. "*niche-based models*" en anglais

3. "*process-based models*" en anglais

correctes sur leurs capacités évolutives.

1.1.3 Capacités adaptatives des populations d'arbres et gestion durable des forêts

Dans le monde les forêts couvrent près de quatre milliards d'hectares, ce qui représente 31 % des terres émergées (Figure 1.4; IGN 2013⁴). Cette place prépondérante sur la surface du globe explique que les forêts hébergent la majorité des écosystèmes terrestres et soient les principaux refuges de biodiversité terrestre. En plus de leur rôle écologique, les forêts génèrent de nombreuses activités bioéconomiques, par l'exploitation du bois et leur rôle pour le loisir et la détente (Table 1.1). En Europe par exemple, la filière bois représente près de 300 milliards d'euros et fournit 3 à 5 millions d'emplois (source EFI⁵; www.efi.int).

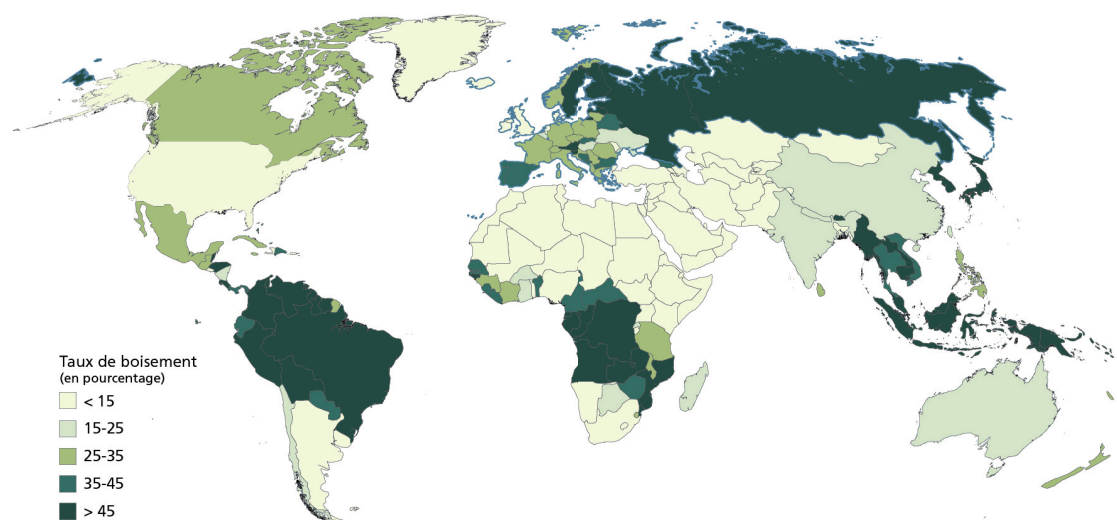


FIGURE 1.4 – Situation des forêts dans le monde et pourcentage de boisement par pays. ©IGN 2013.

Bien que l'augmentation du CO₂ atmosphérique et des températures devraient dans un premier temps augmenter la croissance et la production nette des forêts, à long terme, l'augmentation du risque de sécheresse et autres événements climatiques défavorables (tempêtes, augmentation des feux) devraient avoir un effet négatif sur la production et la survie des forêts (Lindner *et al.*, 2010). En zone Méditerranéenne, où l'on prédit une augmentation des tem-

4. Institut National de l'Information Géographique et Forestière

5. Institut de la Forêt Européenne

Rôles écologiques	Rôles socio-économiques
Hébergement des écosystèmes terrestres, maintien de la biodiversité	Utilisation du bois comme matériau de construction, source d'énergie
Stockage de carbone	Création d'emploi dans des zones rurales
Amélioration du climat, de la qualité de l'air et de l'eau	Lieux de récréation, détente pour le tourisme et les loisirs

TABLE 1.1 – Exemples de bénéfices écologiques et socio-économiques tirés de la forêt (source EFI ; www.efi.int).

pératures annuelles de 3 à 4 ° C et une diminution des précipitations de 20 % dans le siècle à venir (Lindner *et al.*, 2010), le fonctionnement des écosystèmes forestiers devrait être particulièrement affecté par l'augmentation de l'intensité et de la durée des sécheresses, et du risque d'incendie. Dans cette région, considérée comme un point chaud⁶ de biodiversité, 18 % des espèces sont déjà menacées d'extinction à cause de la destruction de leur habitat (FAO, 2013). Dans les zones Méditerranéennes protégées, les effets de la variation du climat devraient être plus forts encore que dans le reste du monde du fait de leurs petites tailles (Loarie *et al.*, 2009).

Du point de vue de leur réponse aux changements climatiques, les arbres présentent beaucoup de caractéristiques paradoxales qui suscitent des incertitudes autour de leurs capacités adaptatives (Petit & Hampe, 2006; Savolainen *et al.*, 2007). En effet, leurs traits d'histoire de vie tels que leur longévité, leur maturité tardive et leur mode de vie sessile, suggèrent que les arbres sont des espèces vulnérables aux changements environnementaux rapides. À l'inverse, la forte plasticité phénotypique, les distances de migrations élevées (la migration par graine pouvant atteindre plusieurs centaines de mètres et la migration par pollen plusieurs kilomètres; Ashley 2010) et les fortes diversités génétiques trouvées au sein des populations d'arbres pourraient faciliter leur réponse évolutive (Hamrick *et al.*, 1992; Petit & Hampe, 2006; Savolainen *et al.*, 2007). Même si de nombreuses études permettent aujourd'hui d'affirmer que la majorité des populations d'arbres ont eu les capacités de s'adapter aux conditions environnementales par le passé (Savolainen *et al.*, 2007; Alberto *et al.*, 2013), les capacités d'adaptation des arbres à des changements environnementaux rapides pourraient être insuffisantes. En effet, des modèles mécanistes ont déjà montré que les distances de migration mesurées à l'heure actuelle chez les arbres ne leur permettraient pas de se déplacer aussi rapidement que leurs

6. *hotspot* en anglais

enveloppes bioclimatiques (Nathan *et al.*, 2011) et que leurs longs temps de génération génèreront une forte maladaptation des populations sur leur aire d'origine (Kuparinen *et al.*, 2010).

Afin de maintenir les services écosystémiques⁷ des forêts, des programmes de gestion des forêts ont vu le jour afin de réduire leur vulnérabilité aux changements climatiques. Les options de gestion adaptative existantes peuvent être classées entre deux extrêmes : celles ayant pour but de favoriser le maintien et l'adaptation locale des espèces, et celles envisageant la substitution des espèces locales par d'autres espèces mieux adaptées aux futures conditions environnementales (un des aspects de la pratique de migration assistée ; Aitken *et al.* 2008). La première approche inclut notamment tous les aspects de gestion sylvicole raisonnée (Lefèvre *et al.*, 2013) ou l'apport de ressources génétiques par flux de gènes assistés entre populations au sein de l'aire de distribution de l'espèce (Aitken *et al.*, 2008). Cependant, il existe des obstacles, limites et coûts à ces programmes de gestion qui rendent ces actions difficiles. Notamment, le manque de connaissance sur le potentiel évolutif des arbres fait que les conséquences et l'efficacité de ces programmes sont difficiles à mesurer. Des études évaluant le potentiel adaptatif des populations d'arbres et les mécanismes favorisant ou défavorisant leur évolution sont donc nécessaires (Hansen *et al.*, 2012).

Le contexte général qui vient d'être présenté est celui dans lequel s'inscrit ma thèse, qui a visé à étudier le potentiel adaptatif d'une espèce d'arbre tempérée (*Fagus sylvatica*) dans un climat montagnard-Méditerranéen. Plus particulièrement, je me suis intéressée à l'interaction qui pouvait exister entre les capacités d'adaptation locale et les flux de gènes et à ses conséquences sur la réponse des populations d'arbres aux changements environnementaux. Dans la section qui suit, je replace ces concepts dans un cadre théorique et empirique général avant de détailler le contexte particulier de mon étude.

7. définit les bénéfices écologiques et socio-économiques que les humains tirent du fonctionnement des écosystèmes

1.2 CADRE THÉORIQUE ET EMPIRIQUE

Adaptation locale : contraintes et facilitations

1.2.1 Adaptation locale, un trade-off entre sélection et flux de gènes

1.2.1.1 Concept d'adaptation locale

La sélection naturelle a pour effet de modifier les fréquences alléliques de gènes codant pour des traits reliés à la valeur sélective des individus, en favorisant les allèles qui apportent un avantage dans un environnement particulier. La valeur sélective⁸ absolue d'un génotype est elle-même définie comme le succès reproducteur des individus portant ce génotype. Elle se mesure par le nombre moyen de descendants viables et reproducteurs laissés par ce génotype à la génération suivante. Dans un environnement hétérogène, la variation spatiale des pressions de sélection naturelle conduit à la différenciation génétique des populations, i.e. à une interaction génotype×environnement pour la valeur sélective. En génétique des populations, l'adaptation locale au sens strict implique l'existence d'interaction génotype×environnement, mais plus encore la notion de pléiotropie antagoniste, i.e. au fait que les allèles des gènes aient des effets opposés sur la valeur sélective des individus dans différents environnements. Un génotype est alors supérieur dans un environnement, au détriment de sa valeur sélective dans d'autres environnements (Figure 1.5 ; Kawecki & Ebert 2004).

Les pré-requis à l'émergence de patrons d'adaptation locale incluent l'existence d'une pression de sélection assez forte, d'une diversité génétique pour les gènes non-neutres et de flux de gènes restreints. En effet, dans de tels systèmes, les fréquences des allèles à effets pléiotropiques sont déterminées par un équilibre MIGRATION-SÉLECTION (Figure 1.5 ; Lenormand 2002). Les flux de gènes, en homogénéisant les fréquences alléliques entre populations localement adaptées, vont représenter la principale force contre-balançant le déséquilibre des fréquences alléliques généré par l'action de la sélection (Lenormand, 2002). De plus, en introduisant des allèles sous-optimaux, ayant été sélectionnés dans d'autres conditions environnementales, la migration va avoir pour effet de diminuer l'adaptation des populations à leur environnement ; on parle alors de fardeau de migration⁹ (Lenormand, 2002).

La diversité génétique locale étant le carburant du changement génétique, à terme, la diver-

8. "*fitness*" en anglais

9. "*migration load*" en anglais

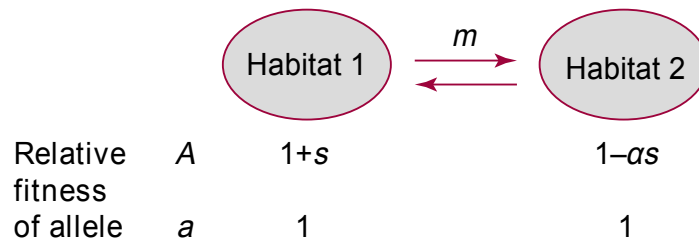


FIGURE 1.5 – Modèle simple de migration-sélection avec deux populations et deux allèles, sans effet de la densité sur la sélection (modèle softselection). La fixation de l’allèle A dans la population 1 et de l’allèle a dans la population 2 va à la fois dépendre de l’effet de A dans chacun des deux environnements (α), de l’intensité des flux de gènes entre les deux populations (m) et de celle de la pression de sélection (s). Tiré de Lenormand (2002)

sité des gènes codant pour les traits fortement reliés à la valeur sélective des individus devrait complètement disparaître par effet de la sélection. Le maintien de forts niveaux de diversité génétique intra-population en présence de différenciation adaptative, comme cela est généralement observé chez les arbres, reste un paradoxe important en biologie évolutive (Le Corre & Kremer, 2012). L’apport de variation génétique à court terme pourrait notamment être expliqué par les flux de gènes entre populations (Hamrick *et al.*, 1992). La fragmentation actuelle des paysages, en réduisant l’apport de diversité génétique par les flux de gènes, pourrait être une des contraintes majeures à l’adaptation génétique des populations aux changements climatiques futurs (Davis & Shaw, 2001).

En théorie, les flux de gènes peuvent également favoriser la diffusion d’allèles avantageux dans l’espace. Dans le cas particulier d’environnements variables à la fois dans l’espace et le temps, les flux de gènes entre populations localement adaptées pourraient favoriser la diffusion d’allèles favorables pré-adaptés aux conditions environnementales futures (Davis & Shaw, 2001; Aitken *et al.*, 2008). Dans le contexte des changements climatiques actuels, cette théorie est très séduisante puisqu’elle suppose que les flux de gènes pourraient contenir des allèles pré-adaptés aux conditions environnementales chaudes et accélérer l’adaptation des populations dans les zones de moyennes à hautes latitudes et altitudes (Kremer *et al.*, 2012). Plus que de permettre le maintien des populations dans un environnement se translatant dans l’espace, ces diffusions de gènes pré-adaptés au nouvel environnement pourraient faciliter la colonisation à l’avant des marges et l’extension des aires de distribution (Davis & Shaw, 2001). Les événements de dispersion à longue distance pourraient tout particulièrement favoriser ces diffusions de gènes avantageux.

D'autres effets bénéfiques des flux de gènes sur l'adaptation sont souvent avancés comme la capacité de contrer les effets de la dépression de consanguinité dans les populations de faible taille (Ingvarsson & Whitlock, 2000). Dans ces populations, les flux de gènes permettent également d'éviter les variations de fréquences alléliques aléatoires et la fixation des allèles défavorables dues à l'effet de la dérive génétique (Alleaume-Benharira *et al.*, 2006). Chez les arbres, le régime de reproduction majoritairement allofécond maintient un fort fardeau génétique du à la présence d'allèles délétères récessifs (i.e. les descendants issus d'autofécondation ou du croisement entre parents apparentés présentent de plus faibles valeurs sélectives). Chez les populations de petites tailles ou présentant des régimes de reproduction mixte, on peut donc s'attendre à ce que les effets de la migration soient particulièrement bénéfiques.

L'étude du concept d'adaptation locale peut sortir de ce formalisme volontairement simpliste de la génétique des populations, décrivant l'effet d'un gène sur la valeur sélective, en utilisant des approches de génétique quantitative, décrivant l'effet d'un ou plusieurs traits quantitatifs sur la valeur sélective.

1.2.1.2 Formalisation de la réponse à la sélection par les modèles de génétique quantitative

La génétique quantitative offre un cadre conceptuel de choix pour comprendre la réponse à la sélection des traits quantitatifs. Un trait quantitatif est un trait présentant une variation continue au sein d'une population (contrairement aux traits discrets), généralement due à un déterminisme complexe, à la fois multigénique et environnemental. La plupart des traits phénotypiques observables sur les individus d'une espèce, et en particulier ceux impliqués dans la réponse adaptative, sont quantitatifs. L'intérêt principal de la génétique quantitative est de permettre la décomposition de la variabilité phénotypique d'un trait mesuré dans une population ou groupe d'individus en composantes génétiques (variance génétique V_G) et environnementales (variance environnementale V_E ; Encadré A), sans avoir à connaître les gènes responsables du déterminisme de ce caractère. La variance génétique, qui peut elle-même être décomposée en composantes additive (V_A) et de dominance (V_D), est en théorie directement reliée au polymorphisme des gènes codant pour le trait (Figure 1.6). La mesure d'héritabilité au sens strict (h^2) d'un trait au sein d'une population permet de quantifier la part du phénotype transmis par les gènes à effets additifs : $h^2 = V_A/V_P$. L'intérêt de l'héritabilité

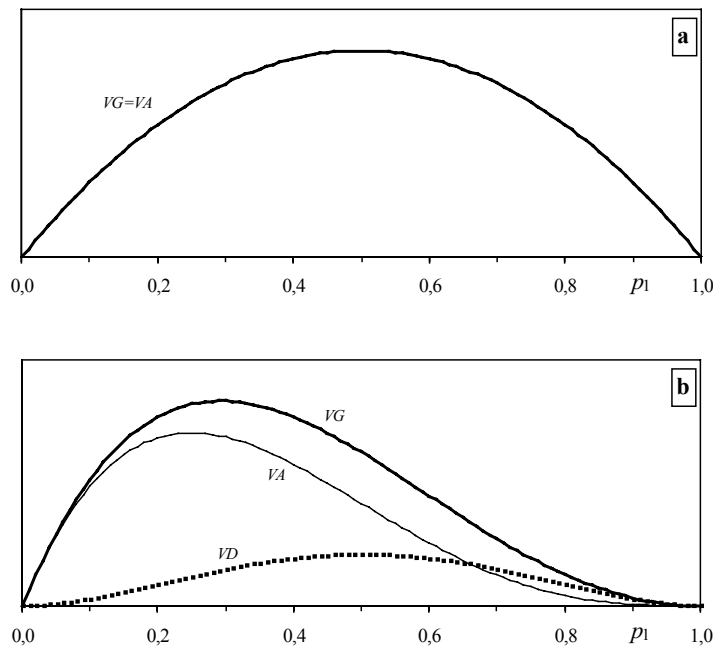


FIGURE 1.6 – Cas a : additivité stricte des allèles des locus, b : dominance complète. En ordonnée on retrouve l'évolution des variances génétiques et en abscisse la fréquence d'un allèle à un locus biallélique. Tiré de Verrier *et al.* (2001).

est notamment de permettre la prédiction de la réponse à la sélection "immédiate" (R) à travers l'équation du sélectionneur, appelée ainsi du fait de son utilisation dans un premier temps par le milieu de l'amélioration animale et variétale :

$$R = h^2 S \quad (1.1)$$

avec S le différentiel de sélection, qui est la différence entre la moyenne de la population des parents avant et après sélection (une mesure de l'intensité de la sélection).

En écologie évolutive cette équation est modifiée pour prédire le taux d'évolution génétique d'un trait soumis à la sélection naturelle sous la forme :

$$R = h^2 V_P \beta \quad \text{ou} \quad R = V_A \beta \quad (1.2)$$

L'intensité de la sélection naturelle sur un trait quantitatif est de manière simple mesurée par la relation linéaire qui existe entre la valeur phénotypique du trait et la valeur sélective, aussi appelée gradient de sélection linéaire (β). Cette équation 1.2 permet notamment d'appliquer l'équation du sélectionneur à un autre type de sélection que celle par troncature classiquement

Encadré A

DÉCOMPOSITION DE LA VARIATION PHÉNOTYPIQUE

Le modèle de génétique quantitative caractérise l'influence du génotype (G) et de l'environnement (E) sur la valeur phénotypique (z) comme :

$$z = G + E$$

Pour un individu diploïde, la valeur génétique à un locus l , $G_{i,j,l}$, est classiquement décomposée comme :

$$G_{i,j,l} = \alpha_{i,l} + \alpha_{j,l} + \delta_{i,j,l}$$

avec $\alpha_{i,l}$ et $\alpha_{j,l}$ les effets moyens de l'allèles au locus l et $\delta_{i,j,l}$ l'effet de leur interaction. La valeur génétique additive, A , est définie par la somme des effets moyens des gènes ($A = \sum_l \alpha_{i,l} + \alpha_{j,l}$) et la valeur de dominance, D , par la somme des effets d'interaction aux loci ($D = \sum_l \delta_{i,j,l}$). Ainsi, la valeur génétique totale peut être écrite comme (Figure 1.7) :

$$G = A + D$$

À partir de ces modèles, on peut partitionner la variance phénotypique d'un trait, V_P , en composantes de variance génétique additive V_A , variance de dominance, V_D , et variance environnementale, V_E :

$$V_P = V_A + V_D + V_E$$

en supposant que les composantes de covariance, $cov(A, D)$, $cov(A, E)$ et $cov(D, E)$ soient nulles.

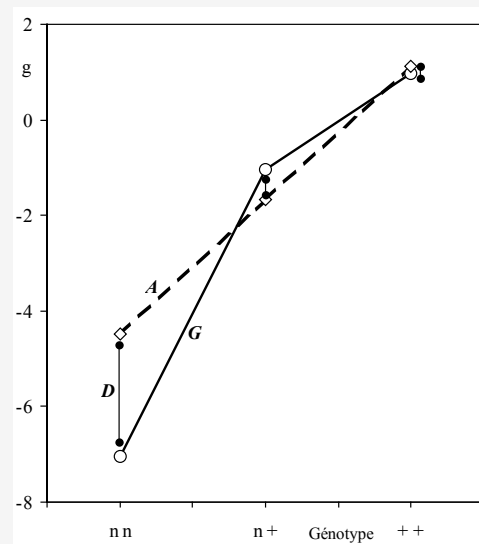


FIGURE 1.7 – Valeurs génétiques additives (A), de dominances (D) et génétiques totales ($G = A + D$) dans le cas du gène du nanisme chez la souris (aux allèles "n" et "+"). L'axe des ordonnées représente l'effet du génotype sur le poids des individus (gramme) par rapport à la moyenne générale de la population. Tiré de Verrier *et al.* (2001).

utilisée en amélioration. La valeur sélective étant généralement affectée par plusieurs traits quantitatifs non-indépendants, un modèle plus correct de changement moyen multivarié de valeurs de traits sur une génération s'écrit (Lande & Arnold, 1983) :

$$\Delta\bar{z} = \mathbf{G} \times \beta \quad (1.3)$$

avec $\Delta\bar{z} = \{\Delta z_1, \Delta z_2, \dots, \Delta z_n\}$ le vecteur des changements moyens de n traits, \mathbf{G} la matrice de variance-covariance additive génétique des traits et $\beta = \{\beta_1, \beta_2, \dots, \beta_n\}$ le vecteur des gradients de sélection directionnels des n traits.

D'un point de vue pratique, les effets génétiques et environnementaux sur la valeur phénotypique peuvent être dissociés par la connaissance des covariances génétiques et environnementales entre apparentés, $cov_G(X, Y)$ et $cov_E(X, Y)$ respectivement (Lynch & Walsh, 1998) :

$$cov(X, Y) = cov_G(X, Y) + cov_E(X, Y) \quad (1.4)$$

$$cov_G(X, Y) = 2\phi_A V_A + \phi_D V_D \quad (1.5)$$

avec $cov(X, Y)$ la covariance phénotypique des individus X et Y , ϕ_A le coefficient de simple parenté et ϕ_D le coefficient de double parenté. Dans les populations naturelles, un des enjeux principaux est d'être capable de distinguer ces covariances génétiques et environnementales entre apparentés car : d'une part, les relations de parentés ne sont pas forcément facilement retraçables, d'autre part, l'environnement naturel est souvent très complexe à caractériser, et, finalement, la répartition spatiale in situ des individus n'est pas randomisée, et des apparentés peuvent se retrouver dans des environnements semblables rendant difficile la distinction entre $cov_G(X, Y)$ et $cov_E(X, Y)$. Pour ces raisons, afin de pouvoir distinguer les ressemblances phénotypiques liées aux gènes ou à l'environnement, beaucoup d'études utilisent des dispositifs dits de "familles", qui consistent à échantillonner des groupes de parenté connue (e.g. familles de plein-frères, de demi-frères), placées dans des conditions environnementales contrôlées (e.g. jardin commun pour les plantes) ou des conditions semi-naturelles (e.g. nids artificiels pour les oiseaux ; Kruuk 2004). Différents modèles statistiques permettent alors de dissocier les causes de similarités entre apparentés pour permettre l'estimation de h^2 et de la matrice \mathbf{G} : la régression parent-enfant, le modèle famille ou encore le modèle animal (Encadré B).

Encadré B

PRINCIPAUX MODÈLES STATISTIQUES EN GÉNÉTIQUE QUANTITATIVE

Les modèles en génétique quantitative ont deux intérêts primordiaux, 1) parvenir à dissocier les effets génétiques et environnementaux sur l'expression du trait Y et 2) estimer la variance génétique additive au sein d'une population, V_A , pré-requis pour savoir comment le trait va répondre à la sélection.

La régression entre apparentés Principalement utilisée pour comparer des parents et leurs descendants, la régression parent-enfant est basée sur un modèle statistique simple reliant la valeur phénotypique du parent X à celle de son descendant, Y :

$$Y_i = bX_i + \sigma_i$$

avec b la pente de la régression et σ un effet aléatoire résiduel avec $\sigma \sim N(0, V_E)$. Comme $b = cov_G(Y_i, X_i)/V(X_i)$ et $cov_G(Y_i, X_i) = 1/2V_A$ ($\phi_A = 1/2$ entre parent et descendant ; equation 1.5), on peut facilement déduire la variance génétique de ce modèle de régression :

$$V_A = 2bV(X_i)$$

Le modèle famille Le modèle famille consiste à analyser la variance inter-famille (V_F) directement reliée à la covariance génétique intra-famille (cov_G) et donc à la variance génétique par l'équation :

$$\begin{aligned} V_F &= cov_G(X_i, X_{i'}) \\ &= 2\phi_A V_A + \phi_D V_D \end{aligned}$$

avec ϕ_A le coefficient de simple parenté et ϕ_D le coefficient de double parenté. Ainsi, pour des familles de plein-frères on peut déduire que $V_F = 1/2V_A + 1/4V_D$ et pour des

familles de demi-frères $V_F = 1/4V_A$ ($\phi_D = 0$). L'estimation de V_F passe généralement par l'utilisation de modèle linéaire mixte où l'effet aléatoire Famille, F , est supposé tiré dans $F \sim N(0, V_F)$:

$$Y_{ji} = \mu + F_j + \sigma_{ji}$$

avec Y_{ji} la valeur phénotypique de l'individu i appartenant à la famille j , μ un effet moyen et σ_{ji} un effet aléatoire résiduel ($\sigma \sim N(0, V_R)$).

Le modèle animal Le modèle animal, récemment appliqué au domaine de la génétique quantitative en biologie évolutive, est maintenant très largement utilisé au sein de cette communauté (Kruuk, 2004). Ce modèle permet d'utiliser les covariances phénotypiques entre toutes les paires d'apparentés au sein d'une population, à l'inverse de la régression parent-enfant ou du modèle famille qui n'utilisent que les covariances de certains type de parentés. Le modèle animal est donc très performant pour extraire toute l'information sur les apparentements disponibles. Sous sa forme la plus simple, le modèle décompose la valeur phénotypique d'un individu i comme :

$$Y_i = \mu + a_i + \sigma_i$$

avec a_i un effet aléatoire représentant la valeur génétique de l'individu i et σ_i un effet aléatoire résiduel. Les a_i sont supposés tirés dans $a \sim N(0, A \times V_A)$ et les σ_i dans $\sigma \sim N(0, I_{de} \times V_R)$, avec A la matrice des apparentements génétiques additifs entre individus et I_{de} la matrice identité. Les coefficients de simple parenté regroupés dans A sont généralement inférées par une information de type pédigrée afin de déduire V_A .

Aujourd'hui cependant, le développement de marqueurs moléculaires apportent une information sur l'apparentement génétique entre individus (données génomiques ; Visscher *et al.* 2006) et de méthodes statistiques permettant de pouvoir prendre en compte l'influence de variabilités environnementales dans les modèles de génétique quantitative (modèle animal ; Kruuk 2004) est très prometteur quand au fait de pouvoir réaliser dans les années à venir des études de génétiques quantitatives en populations naturelles, même en l'absence d'une information préalable sur le pédigrée des individus.

1.2.1.3 Apport des modèles théoriques sur les limitations des aires de distribution

Afin de déterminer quels étaient les principaux facteurs limitant l'adaptation locale, les études théoriques, basées sur des modèles de génétique quantitative évolutive qui intègrent le formalisme de la génétique des populations et de la génétique quantitative, ont notamment cherché à comprendre pourquoi l'adaptation échouait aux limites des aires de distribution des espèces (Bridle & Vines, 2007). Les populations devenant de plus en plus petites et fragmentées aux bordures des aires de distribution, deux explications sont classiquement avancées pour expliquer le manque d'adaptation locale : (1) les effets de dérive génétique limitent la fixation d'allèles bénéfiques et empêchent l'adaptation et l'expansion des populations en marges ; (2) la connexion par les flux de gènes des populations marginales aux populations centrales, bien adaptées à leur environnement, apporte un flux continu d'allèles mal-adaptés qui freine la fixation des allèles localement adaptés par action de la sélection (Bridle & Vines, 2007). La modélisation de la démographie et l'évolution d'un trait quantitatif au sein de populations distribuées dans un paysage à variation continue des conditions environnementales permet de comprendre la formation de marges de distribution dans un environnement spatialement hétérogène. En utilisant ce type de modèle, Kirkpatrick & Barton (1997) ont montré que les flux de gènes asymétriques du centre vers les marges contribuaient à la mal-adaptation et au frein de l'expansion des populations marginales. Ce modèle, qui suppose la variance génétique constante dans l'espace, met en évidence les effets défavorables des flux de gènes sur l'adaptation locale. Barton (2001) a par la suite étendu ce modèle pour inclure les effets de la taille de population et de la migration sur les variances génétiques. Il trouve alors que l'effet bénéfique des flux de gènes sur l'augmentation de la variance génétique au sein des populations marginales l'emporte sur les effets défavorables de la migration, les populations

atteignant leur optimum phénotypique en tout point du gradient environnemental, même si ce dernier est très abrupt. Ainsi, à l'inverse du modèle de Kirkpatrick & Barton (1997), les modèles prenant en compte les effets de la migration sur l'augmentation des variances génétiques soulignent le rôle bénéfique de ce processus sur les capacités d'adaptation. Notamment, l'augmentation de la variance génétique au sein des populations est attendue comme d'autant plus importante que les populations connectées par des flux de gènes subissent des pressions de sélection divergentes (Pease *et al.*, 1989; Barton, 2001).

Les covariances génétiques entre traits quantitatifs sont, au même titre que les flux de gènes entre populations localement adaptées, des forces qui peuvent théoriquement contraindre la réponse adaptative des populations. Dans une étude récente, Duputié *et al.* (2012) ont étendu le modèle théorique univarié d'évolution (Pease *et al.*, 1989; Barton, 2001) afin d'étudier l'évolution de plusieurs traits en réponse aux pressions de sélection locale et aux flux de gènes. Leur conclusion est que les contraintes multitraits et la structure de la sélection (e.g., directionnelle, stabilisante) peuvent fortement diminuer les capacités de persistance dans un environnement changeant. Tout comme les modèles univariés, les modèles multitraits dégagent que des taux intermédiaires de migration favorisent l'adaptation des populations périphériques.

Chez les plantes, les effets bénéfiques et défavorables des flux de gènes pourraient varier selon les contributions de la dispersion par graine et pollen. En effet, un grain de pollen apporte moitié moins d'allèles migrants qu'une graine. La sélection pourrait aussi être moins efficace à supprimer les allèles délétères partiellement masqués par les hybrides et donc le fardeau de migration pourrait être plus important lors de migration par flux de pollen. Or chez beaucoup de plantes, et chez les arbres notamment, on sait que la dispersion à moyenne et longue distance se fait essentiellement par pollen (Ashley, 2010). Récemment, Lopez *et al.* (2008) ont montré que les effets de la dispersion par pollen et graine sur la variance génétique et le fardeau génétique étaient identiques à moins que l'environnement soit très hétérogène. Dans ces derniers cas, le fardeau génétique est supérieur si la dispersion se produit essentiellement par pollen.

Toutes ces études théoriques ont apporté un socle important de connaissances et d'attendus sur les effets de la migration sur l'adaptation locale qui attendent maintenant d'être testés empiriquement (Bridle & Vines, 2007). Dans les deux paragraphes suivants, je présente

brièvement les approches expérimentales possibles pour aborder ces questions.

1.2.1.4 Mesures expérimentales de l'adaptation locale

Comparaison des performances en transplantations croisées

D'après le cadre théorique qui vient d'être présenté, l'adaptation locale se manifeste par le fait que dans chaque localité les individus "locaux" présentent une meilleure valeur sélective que les individus "immigrants". Ainsi, la majorité des études empiriques cherchant à détecter des patrons d'adaptation locale comparent les performances moyennes de populations originaires de différents habitats dans l'ensemble des habitats dont elles sont issues (manipulations dites de "transplantations croisées"; Figure 1.8 a, b; Kawecki & Ebert 2004) . Comme précédemment décrit, tous les types d'interaction population×environnement ne reflètent pas forcément des patrons d'adaptation locale; en effet les patrons où la valeur sélective moyenne de la population est meilleure dans son environnement d'origine que dans un environnement

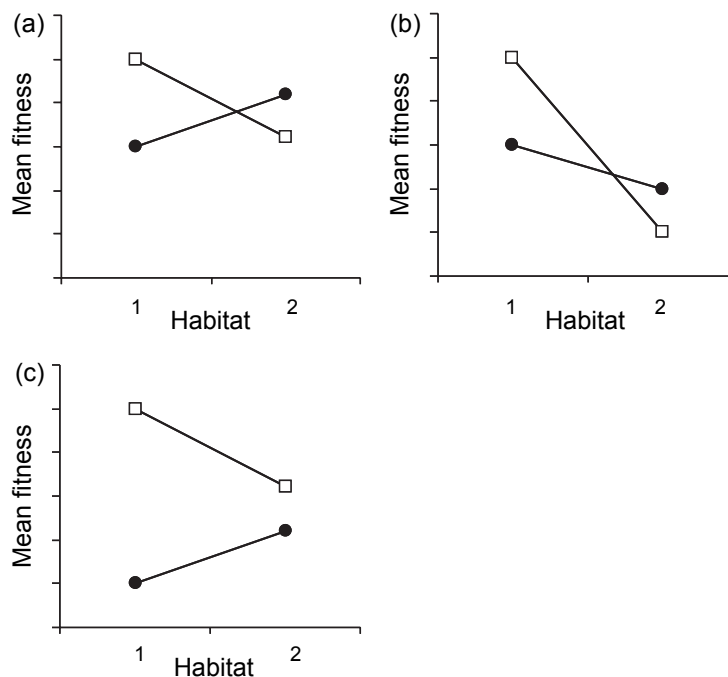


FIGURE 1.8 – Patrons hypothétiques d'interaction population×habitat (ou environnement) pour la valeur sélective. Les carrés blancs représentent la moyenne de valeur sélective de la populations provenant de l'habitat 1 et les cercles noirs la moyenne de valeur sélective pour la population provenant de l'habitat 2. Les patrons représentés par les cas (a) et (b) satisfont le critère "local v.s. étranger" alors que le cas (c) satisfait le cas "'habitat d'origine" v.s. "habitat étranger". Tiré de Kawecki & Ebert (2004).

de transplantation (comparaison "habitat d'origine" versus "habitat étranger"), i.e. une comparaison des performances d'une population entre habitats et non pas des performances des populations au sein d'un habitat, ne constitue pas rigoureusement une preuve d'adaptation locale (Figure 1.8 c ; Kawecki & Ebert 2004). Chez les plantes, les transplantations croisées ont permis de révéler de nombreuses interactions population×environnement le long de gradients environnementaux (71 % des études ; Leimu & Fischer 2008). Cependant, le test rigoureux d'adaptation locale, i.e. le fait que les populations "locales" aient de meilleures performances que les populations "étrangères", représente seulement 45 % des comparaisons entre paires de populations (Leimu & Fischer, 2008). Ces résultats indiquent que l'adaptation locale chez les plantes est finalement moins fréquente que ce que suggérerait une utilisation trop relâchée du concept.

Étude de la différenciation génétique des traits quantitatifs

Du fait de la difficulté de mise en œuvre de l'expérimentation de transplantation croisée pour certaines espèces (e.g., comment et quand mesurer la valeur sélective des espèces pérennes), une autre approche consistant à étudier la différenciation génétique des traits quantitatifs est fréquemment utilisée (Savolainen *et al.*, 2007). Cette approche ne permettant que de tester une interaction population×environnement, elle ne constitue cependant pas une preuve d'adaptation locale d'après la définition de Kawecki & Ebert (2004). Néanmoins, elle permet de dissocier les deux forces pouvant conduire à une perte de variance génétique intra-population et une augmentation de la variance entre populations : la dérive génétique et la sélection naturelle. En environnement commun, la comparaison d'individus issus de populations subissant différentes conditions environnementales permet de pouvoir décomposer la part de variance génétique au sein des populations (V_{GW}) et entre populations (V_{GB}). De cette décomposition de variance, un indice statistique appelé Q_{st} a été créé afin de quantifier la divergence génétique des traits quantitatifs (Lande, 1992; Spitze, 1993) :

$$Q_{st} = \frac{V_{GB}}{V_{GB} + 2V_{GW}} \quad (1.6)$$

Afin de pouvoir dissocier les effets de la sélection et de la dérive sur la différenciation génétique des traits quantitatifs, l'indice de Q_{st} est généralement comparé à son analogue quantifiant la proportion de variance génétique neutre mesurée entre populations, le F_{st} de

Wright (Wright, 1951). Le test de différenciation adaptative des traits repose alors sur un écart à l'attendu neutre de différenciation sous l'action de la dérive génétique seule : $Qst = Fst$ (Lande, 1992). Cet attendu neutre, qui a été démontré sous différents scénarios démographiques (Lande, 1992; Whitlock, 1999), offre un cadre robuste pour aller tester différents attendus sur l'importance relative de la sélection naturelle et de la dérive génétique sur les changements évolutifs.

Pour les traits neutres additifs, Qst et Fst devraient être égaux ; le test $Qst \neq Fst$ indique une action vraisemblable de la sélection naturelle sur la différenciation des traits. Plus précisément, si la divergence aux traits quantitatifs excède celle aux marqueurs neutres, i.e. $Qst > Fst$, alors le patron de différenciation a probablement été causé par l'action de sélection *divergente* ; si à l'inverse la divergence des traits entre populations est inférieure à celle attendue sous action de la dérive génétique seule, le patron a probablement été causé par l'action de sélection *stabilisante*. De nombreuses études en biologie évolutive ont utilisé ce test de Fst - Qst pour rechercher des traces de sélection sur les traits quantitatifs (Leinonen *et al.*, 2008). Des méta-analyses montrent que les patrons de différenciation adaptative des traits sont communs dans la nature (Merila & Crnokrak, 2001; McKay & Latta, 2002; Leinonen *et al.*, 2008; Alberto *et al.*, 2013).

La méthode de comparaison Fst - Qst n'est cependant pas applicable à toutes les études de biologie évolutive, d'une part car elle demande d'être capable de bien séparer les composantes génétiques et plastiques à l'origine de la différenciation des traits (Pujol & Pannell, 2008), et, d'autre part, car elle nécessite le suivi de nombreuses populations afin de pouvoir faire des inférences correctes ($n > 10$; O'Hara & Merila 2005; Whitlock & Guillaume 2009) et il est rare de pouvoir acquérir d'aussi gros jeux de données dans les populations naturelles. Notez que ces limites statistiques ont récemment été contournées par le développement d'une procédure Bayésienne combinant un modèle de divergence génétique neutre (modèle F ; Gaggiotti & Foll 2010) et un modèle de génétique quantitative à l'échelle individuelle (modèle animal) pour détecter des traces de la sélection sur la différenciation entre populations (Ovaskainen *et al.*, 2011).

1.2.1.5 Preuves empiriques de l'impact des flux de gènes

Quelques études corrélatives sont parvenues à mettre en évidence que les niveaux d'adaptation locale et de différenciation des traits fonctionnels augmentaient avec la distance entre populations, et donc, indirectement, dépendaient de l'intensité des échanges de gènes entre ces populations (Hendry & Taylor, 2004; Sambatti & Rice, 2006). Des preuves empiriques corroborent donc l'idée que des flux de gènes trop importants peuvent constituer un frein à l'établissement d'un patron d'adaptation locale. Cependant, un problème majeur de l'approche indirecte utilisée par ces études est qu'elle néglige un biais potentiel dû à la covariation spatiale des pressions de sélection divergentes et de l'intensité flux de gènes : ainsi, les divergences génétiques moins importantes entre populations proches géographiquement pourraient aussi être dues à des pressions de sélection moins hétérogènes entre des environnements proches (Garant *et al.*, 2007). De plus, la proportion d'études montrant un effet des flux de gènes sur les patrons d'adaptation locale par rapport au nombre n'en détectant pas reste faible. Les méta-analyses sur ce sujet soulignent cette tendance, en montrant que la distance entre populations n'est pas le facteur principal expliquant les degrés d'adaptation des populations à leur environnement (Figure 1.9 ; McKay & Latta 2002; Leimu & Fischer 2008). Finalement, les études empiriques ont fourni peu de preuves que la migration freine l'adaptation locale.

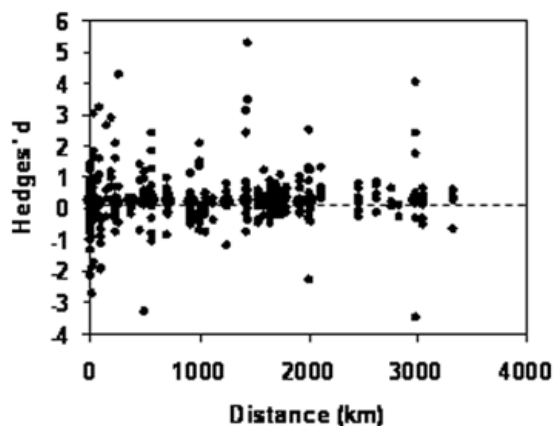


FIGURE 1.9 – Relation entre la distance géographique et l'adaptation locale. Le d de Hedge est un coefficient mesurant la différence entre les performances moyennes des individus locaux v.s. étrangers. Tiré de Leimu & Fischer (2008).

D'autres études empiriques ont également testé les attendus sur les effets favorables des flux de gènes sur les capacités évolutives des populations. Chez les animaux, des manipulations d'évolution expérimentale contrôlant la migration et les pressions de sélection au sein de

populations artificielles ont déjà permis de mettre en évidence les effets bénéfiques de la migration sur le taux d'évolution (e.g. chez la drosophile ; Swindell & Bouzat 2006). Cependant, ces études ne permettent pas de répondre à la question de l'impact réel du flux de gènes sur l'adaptation des populations naturelles. Chez les plantes, le rôle des flux de gènes est généralement étudié grâce à des manipulations dites de "croisements contrôlés". Les croisements entre individus d'une même zone géographique et d'individus séparés par de longues distances sont ensuite comparés aux croisements "naturels" (chez les plantes appelés de "pollinisation libre") pour évaluer le rôle efficace de la migration dans les populations naturelles (Nilsson, 1995). Ces études sont en quantités restreintes du fait des difficultés de mise en œuvre des croisements contrôlés, notamment chez les arbres du fait de la localisation des fleurs et des forts taux d'avortement des fruits. Une étude récente chez l'eucalyptus commun (*Eucalyptus globulus*) a permis de mettre en évidence un avantage sélectif aux individus issus de pollinisation à longue-distance (Costa e Silva *et al.*, 2014). Ainsi, chez les populations présentant des niveaux de consanguinité non-négligeables, les effets bénéfiques de l'hétérosis¹⁰ seraient suffisants pour dépasser le fardeau génétique lié à la migration.

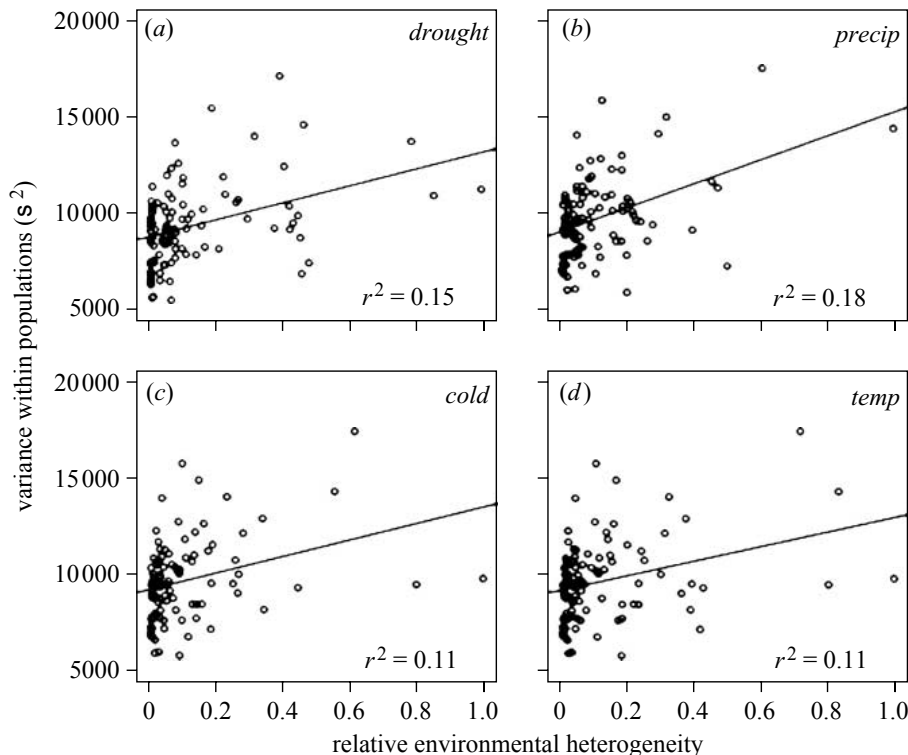


FIGURE 1.10 – Relation entre la variance mesurée au sein des populations et l'hétérogénéité régionale pour 4 variables climatiques. Tiré de Yeaman & Jarvis (2006).

10. hétérosis (ou vigueur hybride) est l'augmentation de la valeur sélective des individus issus de croisements entre lignées étant chacune consanguines

Une étude intéressante chez le pin tordu *Pinus contorta* a permis de tester l'attendu théorique selon lequel plus la différenciation génétique entre populations est importante, plus les flux de gènes ont un effet sur l'augmentation de la variance génétique au sein des populations (Yeaman & Jarvis, 2006). L'étude s'est basée sur un dispositif de jardin commun permettant de mesurer la variance génétique de la croissance pour 142 populations et sur la caractérisation de l'hétérogénéité des conditions environnementales dans la région autour des habitats d'origine. Les auteurs montrent qu'une forte corrélation existe entre l'hétérogénéité régionale et la variance génétique, suggérant que les flux de gènes entre populations subissant des pressions de sélection hétérogènes contribuent à expliquer les niveaux de variance génétique au sein des populations de pin (Figure 1.10 ; Yeaman & Jarvis 2006).

La plupart des études empiriques se sont focalisées sur les conséquences mal-adaptatives des flux de gènes. Ceci peut notamment être expliqué par la plus grande difficulté méthodologique à tester les attendus théoriques des effets bénéfiques des flux de gènes sur l'adaptation. Même si les effets bénéfiques des flux de gènes sont souvent supposés supplanter les effets négatifs, notamment chez les arbres (Kremer *et al.*, 2012), il existe un important manque de preuves empiriques directes de l'effet des flux de gènes en populations naturelles.

1.2.2 Adaptation locale de populations d'arbres aux variations climatiques

Parmi les situations empiriques pouvant générer des patrons d'adaptation locale, je me suis particulièrement focalisée sur la différenciation de traits quantitatifs écophysiologiques le long de gradients altitudinaux chez les arbres.

1.2.2.1 Réponse adaptative des traits fonctionnels aux variations environnementales chez les plantes

Comme nous l'avons vu précédemment, les effets de la sélection et l'adaptation génétique peuvent être révélés par l'étude des traits impliqués dans ces adaptations et une approche de génétique quantitative. Dans le cas de la réponse au climat, le choix de traits adaptatifs pertinents chez les plantes peut s'appuyer sur les connaissances issues de l'écophysiologie, i.e. l'étude de la relation entre la plante et son environnement abiotique. En écophysiologie, le concept de trait fonctionnel définit tout trait phénologique, physiologique ou morphologique,

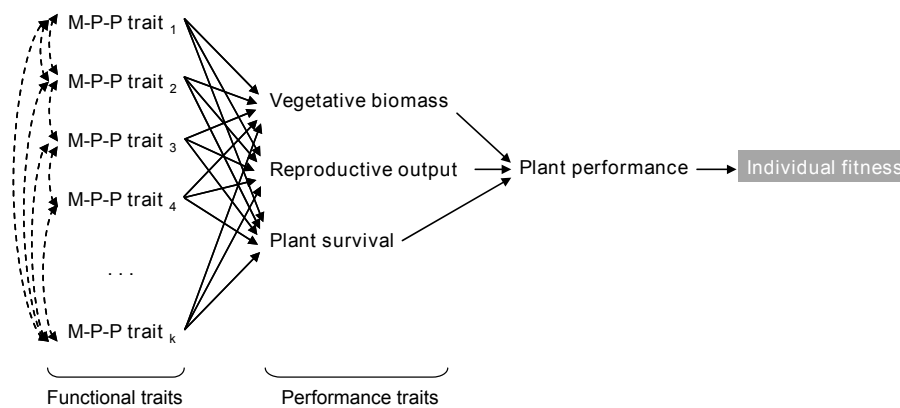


FIGURE 1.11 – Concept de trait fonctionnel chez les plantes. Extrait de Violle *et al.* (2007).

affectant indirectement la valeur sélective des individus par son effet sur des traits de performances, i.e. des traits de survie, croissance et reproduction (Figure 1.11 ; Violle *et al.* 2007). Chez les plantes, les performances photosynthétiques sont impliquées dans la valeur sélective des individus car elles déterminent en partie la quantité de substrat carboné disponible pour la croissance et la reproduction. La température et la concentration en CO_2 sont des facteurs clefs qui gouvernent la productivité des plantes (Gerhart & Ward, 2010). À l'inverse la disponibilité en eau limite le fonctionnement des plantes terrestres, puisque l'absorption du carbone au niveau des stomates se fait au détriment de pertes en eau. En milieu tempérée, la variation des températures est souvent corrélée à la disponibilité en eau : les milieux les plus chauds sont généralement aussi les plus secs et les milieux les plus froids ont des précipitations plus élevées. Ainsi, aux deux extrêmes de l'aire de répartition des espèces de plantes tempérées les enjeux évolutifs vont être de s'adapter à des milieux chauds contraints en eau (plus précisément aux sécheresses estivales) et à des milieux froids (plus précisément aux froids hivernaux).

Au niveau physiologique, les performances photosynthétiques des plantes peuvent être évaluées par des traits tels que l'assimilation nette de CO_2 (A_{max}) qui dépend de la quantité d'azote par unité de masse surfacique foliaire (N_a ; 20 à 30 % de l'azote foliaire chez les plantes étant représenté par la RubisCO, enzyme clef de la photosynthèse), mais aussi de la conductance stomatique (g_w). g_w détermine également la propension qu'a une plante à évacuer de l'eau par la transpiration. Ainsi l'efficacité d'utilisation de l'eau (EUE), qui est le rapport entre A_{max} et g_w , est un trait physiologique particulièrement intéressant en condition de sécheresse, puisqu'il traduit le compromis entre gain carboné par la photosynthèse et perte en eau par la transpiration. Au niveau morphologique, la masse surfacique

foliaire (LMA pour leaf mass area ; $LMA = Poids/Surface\text{foliaire}$) est liée aux performances photosynthétiques (une feuille avec un LMA élevé aura tendance à avoir des taux photosynthétiques plus élevés), mais aussi à la résistance à la dessiccation (Wright *et al.*, 2004). Parmi ces traits fonctionnels, A_{max} , g_w et donc EUE , ne sont pas mesurables sur un nombre suffisant d'individus pour réaliser des analyses de génétique quantitative et étudier leur divergence génétique en populations naturelles (paragraphe 1.2.1.4) ; c'est la raison pour laquelle des substituts de ces traits fonctionnels sont souvent utilisés : par exemple la composition isotopique du carbone foliaire ($\delta^{13}C$) est classiquement utilisée comme un proxy de EUE .

Les traits phénologiques, qui déterminent les événements périodiques du cycle de vie par rapport aux variations saisonnières du climat, sont des déterminants majeurs de l'aire de distribution des espèces d'arbres (Chuine & Beaubien, 2001). Chez les plantes, ces événements sont par exemple la floraison, le débourrement végétatif (ou feuillaison), la fructification et la sénescence foliaire. Les événements de débourrement végétatif et de sénescence foliaire vont en particulier définir la durée de la saison de végétation, période durant laquelle les plantes pourront réaliser la photosynthèse, assimiler du carbone et croître (Howe *et al.*, 2003). Ces traits phénologiques ont donc une importance majeure sur la valeur sélective des plantes.

Bien que beaucoup d'études en écophysiologie aient permis de caractériser la réponse plastique des traits morphologiques, physiologiques et phénologiques aux variations environnementales, la valeur adaptative de ces traits fonctionnels reste encore largement inconnue, notamment chez les arbres (mais voir Ramirez-Valiente *et al.* 2009; Vitasse *et al.* 2009; Bresson *et al.* 2011). Globalement, les études existantes montrent que les traits phénologiques et les traits physiologiques et morphologiques foliaires sont sous sélection dans les populations naturelles de plantes, ce qui confirme leur rôle adaptatif (Arntz & Delph, 2001; Geber & Griffen, 2003). Ces études, généralement basées sur la mesure de gradients de sélection phénotypique ou la mesure de la divergence génétique des traits entre populations subissant des conditions environnementales différentes, ont permis de définir un certain nombre d'attendus sur la réponse adaptative des traits fonctionnels à l'environnement dont les principaux sont listés ci-dessous.

L'adaptation aux milieux froids a été relativement bien étudiée chez les plantes et les arbres, notamment pour les traits phénologiques (e.g., Howe *et al.* 2003 pour une revue chez

les arbres). Généralement, les plantes provenant de zones froides (i.e. zones de hautes latitudes ou altitudes) débourrent plus précocement que les plantes de milieux plus chauds lorsqu'elles sont placées dans un même environnement (Howe *et al.*, 2003; Vitasse *et al.*, 2009). Ce résultat traduit une action de la sélection vers des débourrements plus précoces afin de rallonger la durée de la saison de végétation dans des environnements où celle-ci est fortement contrainte par le froid. Ce débourrement plus précoce pourrait être du à un besoin en froid moins important pour lever le processus de dormance des bourgeons, un plus faible besoin en accumulation de température pour déclencher le débourrement végétatif, ou une combinaison de ces deux explications (Howe *et al.*, 2003). Aux très hautes latitudes et altitudes, l'avancement des dates de débourrement peut cependant être contraint par une augmentation des risques de dégâts du gel pendant la feuillaison. Pour les traits phénologiques, l'adaptation au milieu froid se traduit également par une précocité de l'arrêt de la période de croissance (i.e., de la sénescence foliaire ; Howe *et al.* 2003; Vitasse *et al.* 2009). Au niveau physiologique, le froid limitant les capacités photosynthétiques, l'adaptation à ces milieux passe généralement par une augmentation du *LMA* et du contenu en azote des feuilles (adaptation qualifiée de "*stratégie d'acquisition de ressource*" ; Gonzalo-Turpin & Hazard 2009; Bresson *et al.* 2011).

L'adaptation aux sécheresses estivales est beaucoup moins bien connue chez les arbres que l'adaptation au froid hivernal (mais voir Ramirez-Valiente *et al.* 2009). Chez les plantes il a été montré que l'adaptation aux milieux chauds et secs se traduit fréquemment par l'augmentation du *LMA* et de l'efficacité d'utilisation de l'eau (ou $\delta^{13}C$), afin de maintenir de bonnes capacités photosynthétiques tout en limitant les pertes d'eau (Dudley, 1996; Geber & Griffen, 2003; Donovan *et al.*, 2009). L'adaptation dans ces milieux est en revanche très mal caractérisée pour les traits qui ne sont pas des traits morphologiques ou physiologiques foliaires. Notamment, l'existence d'une réponse adaptative des traits phénologiques à la sécheresse est très mal connue, même si certaines études suggèrent que l'avancement des phénologies reproductives (et donc probablement de la phénologie foliaire également) permettrait de bénéficier d'avantage des conditions printanières favorables précédant les sécheresses estivales (stratégie dite "*d'évitement*" de la sécheresse ; Heschel & Riginos 2005; Franks *et al.* 2007).

Une conclusion à ces études est que l'adaptation résulte d'une interaction complexe entre traits pour maximiser la valeur sélective dans un environnement donné (Arntz & Delph, 2001). Dans des environnements où les ressources (e.g., eau, carbone) sont limitantes, l'allocation

de cette ressource va probablement générer des compromis¹¹ entre les traits fonctionnels : par exemple un individu ne pourra pas à la fois investir du carbone dans le développement de son système aérien et racinaire. En écophysiologie les patrons de covariation entre traits fonctionnels ont longtemps été étudiés dans le but de comprendre les stratégies adaptatives entre espèces (Wright *et al.*, 2004). Cependant, ces patrons ne permettent pas de déduire quelles sont les corrélations et stratégies de traits au sein d'une même espèce entre populations. Ainsi, peu de choses sont encore connues sur leurs stratégies d'adaptation multi-traits au sein des espèces de plantes (Figure 1.12 présente les corrélations hypothétiques entre traits fonctionnels chez les plantes).

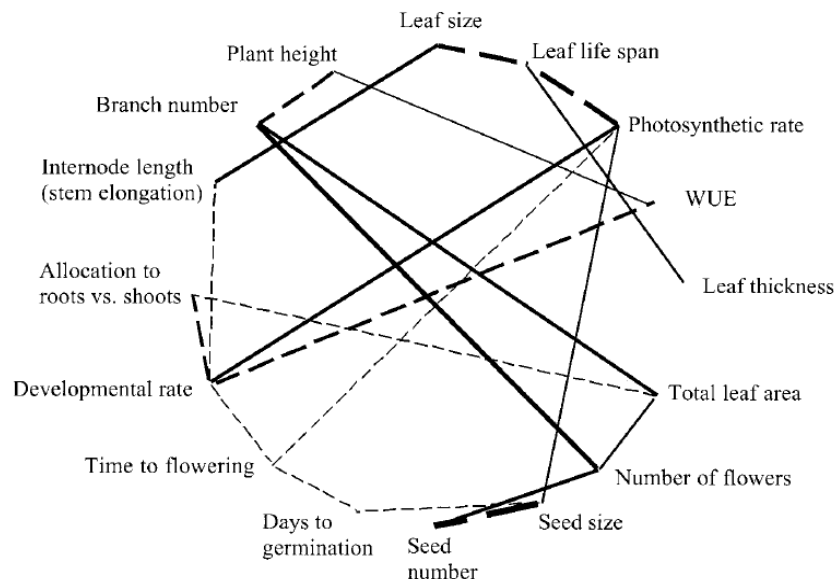


FIGURE 1.12 – Corrélations hypothétiques entre traits fonctionnels chez les plantes. Les traits pleins représentent les corrélations positives et les traits hachurés les corrélations négatives entre traits. L'épaisseur des lignes représente l'importance de ces corrélations. Tiré de Arntz & Delph (2001).

Conjointement à l'effort de mieux comprendre la valeur adaptative des traits fonctionnels des arbres dans les environnements contraignants, il est également nécessaire de mesurer les variances et covariances génétiques de ces traits au sein des populations naturelles afin de comprendre leurs capacités à évoluer génétiquement dans un environnement changeant. En effet, ces corrélations entre traits pourraient fortement affecter les taux d'évolution des traits (comme introduit dans la sous-section 1.2.1.3 ; Blows & Hoffmann 2005; Duputié *et al.* 2012). Les matrices de variance-covariance des traits fonctionnels chez les populations d'arbres ont

11. "trade-off" en anglais

notamment été très peu étudiées ce qui limite notre compréhension de leur réponse adaptative multi-traits aux changements climatiques.

1.2.2.2 Les gradients altitudinaux, modèles pour étudier l'interaction entre flux de gènes et sélection

Chez les espèces longévives, comme les arbres, il est souvent impossible de mesurer l'évolution des traits fonctionnels dans le temps (approche diachronique). Les capacités d'adaptation de ces espèces sont souvent déduites de leur évolution passée en mesurant l'adaptation locale de populations localisées dans différents environnements (Reusch & Wood, 2007). Ces approches dites "synchroniques" substituent donc la variation temporelle de l'environnement par sa variation spatiale¹², avec la logique que la connaissance de la capacité des populations à s'être adaptées dans l'espace nous renseigne sur leurs capacités à s'adapter dans le futur à un changement environnemental équivalent. Les gradients latitudinaux et altitudinaux sont beaucoup utilisés pour mettre en œuvre ces approches synchroniques afin de comprendre la réponse adaptative des espèces aux changements climatiques, puisqu'ils présentent une variation quasi-continue des conditions de température, humidité et précipitation en milieu tempéré (Kozłowski, 1992).

Les études sur les gradients altitudinaux présentent notamment l'intérêt de pouvoir travailler sur des populations subissant de fortes variations de l'environnement sur de courtes distances (Korner, 2007; Keller *et al.*, 2013). Cette particularité est importante lorsque l'on cherche à comprendre la réponse génétique adaptative aux variations environnementales, puisque la variabilité phénotypique mesurée le long de gradients altitudinaux est généralement moins affectée par des effets évolutifs confondants, tels que l'histoire évolutive des populations, que sur les gradients latitudinaux (Keller *et al.*, 2013). De plus, lorsque les variations environnementales sont suffisamment abruptes par rapport aux distances de dispersion, les gradients altitudinaux sont des sites expérimentaux de choix pour étudier l'interaction entre les flux de gènes entre populations et les pressions de sélection naturelles générées par la variation environnementale sur les réponses adaptatives des populations le long du gradient (Figure 1.13; Bridle & Vines 2007).

12. approches de "*space for time*" en anglais

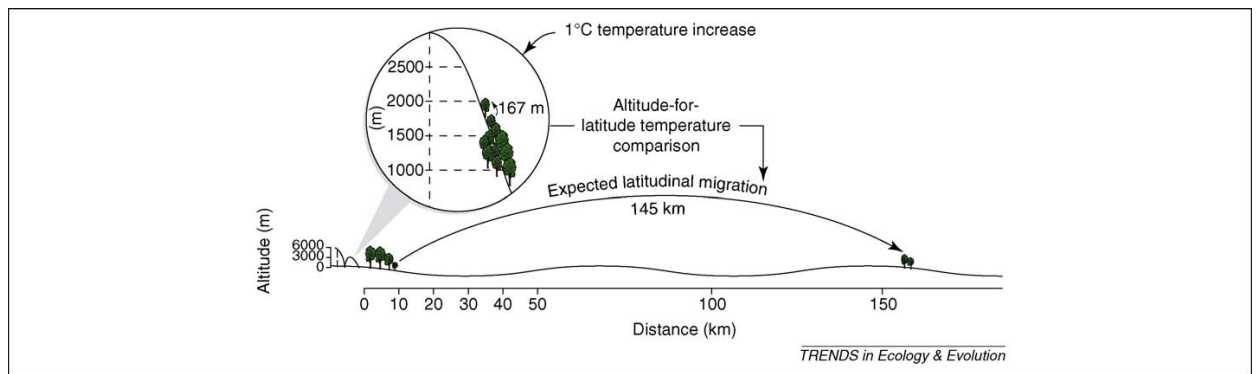


FIGURE 1.13 – Distance séparant deux habitats aux conditions climatiques différent de 1°C en plaine et montagne. En montagne de tels changements de conditions climatiques sont subis à une distance de ~ 167 m alors qu'en plaine cette distance ~ 145 km. Cette figure permet d'illustrer qu'à différence climatique égale, des populations d'arbres en montagne sont plus fortement connectées par des flux de gènes qu'en plaine. Tiré de Jump *et al.* (2009).

1.3 OBJECTIFS ET STRUCTURE DE LA THÈSE

L'objectif de cette thèse a été de mener une étude expérimentale de la différenciation adaptative et la dispersion pollinique dans des populations de hêtre commun (*F. sylvatica*) le long d'un gradient altitudinal afin de comprendre comment les flux de pollen pouvaient contribuer à l'adaptation génétique de ces populations. Parmi les espèces d'arbres forestiers, le hêtre commun est un modèle particulièrement intéressant pour étudier les mécanismes d'adaptation génétique aux variations climatiques puisqu'il s'agit d'une essence forestière majeure en Europe et une espèce connue pour être particulièrement sensible au stress hydrique lors des périodes estivales (Michelot *et al.*, 2012). L'aire de distribution du hêtre s'étend du nord des régions Méditerranéennes au sud de la Scandinavie (Figure 1.14). Alors que dans le nord de l'Europe le hêtre est principalement retrouvé en plaines, en zones Méditerranéennes il se retrouve généralement en altitude. Les effets du changement climatique sur la répartition du hêtre ont déjà été observés, avec en zone Méditerranéenne (nord-est de l'Espagne) un décalage des hêtraies de 70 m vers de plus hautes altitudes (Penuelas & Boada, 2003).

Sur le Mont-Ventoux, un relief culminant à 1911 m d'altitude dans le sud-est de la France,

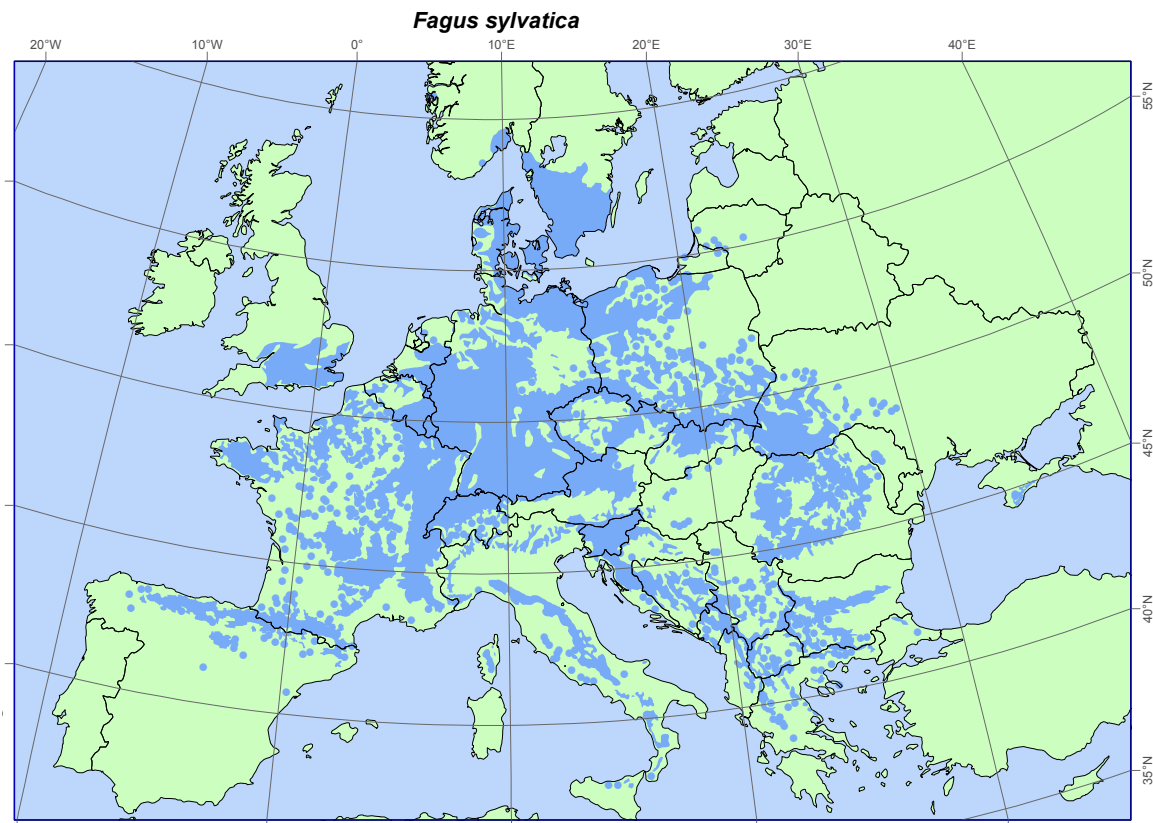


FIGURE 1.14 – Distribution du hêtre commun (*F. sylvatica*). EUFORGEN 2009. www.euforgen.org.

le hêtre se situe à la limite sud occidentale des Préalpes méridionales, dans un bioclimat montagnard-Méditerranéen. Sur le versant nord de ce relief, la hêtraie se répartit de manière quasi-continue de 750 à 1700 m d'altitude. Parmi les populations suivies par l'Unité d'Écologie des Forêts Méditerranéennes (URFM), trois populations de hêtre sur le versant nord du Mont-Ventoux représentent fidèlement la situation que nous souhaitons étudier : la première "N1", de 1,3 ha, se trouve à 1020 m d'altitude, la deuxième "N2", de 2,2 ha, est à 1140 m d'altitude et la troisième "N4", de 0,8 ha, se situe à 1340 m d'altitude (Figure 1.15 et Figure 1.16). Dans chacune des populations des descendance maternelles de 100 graines ont été prélevées sur 20 arbre-mères à l'automne 2009 pour constituer 60 familles issues de pollinisation libre. Ces graines ont été plantées en pépinière d'Aix les Milles au printemps 2010 selon un dispositif classique de génétique quantitative (blocs complets randomisés). A l'été 2010 un stress hydrique a été appliqué sur la moitié du dispositif. L'accroissement des semis a ensuite été mesuré au cours des années 2010 et 2011, la phénologie du débourrement suivie en 2011 et 2012, la sénescence foliaire en 2011 et des caractères relatifs à l'efficacité d'utilisation de l'eau

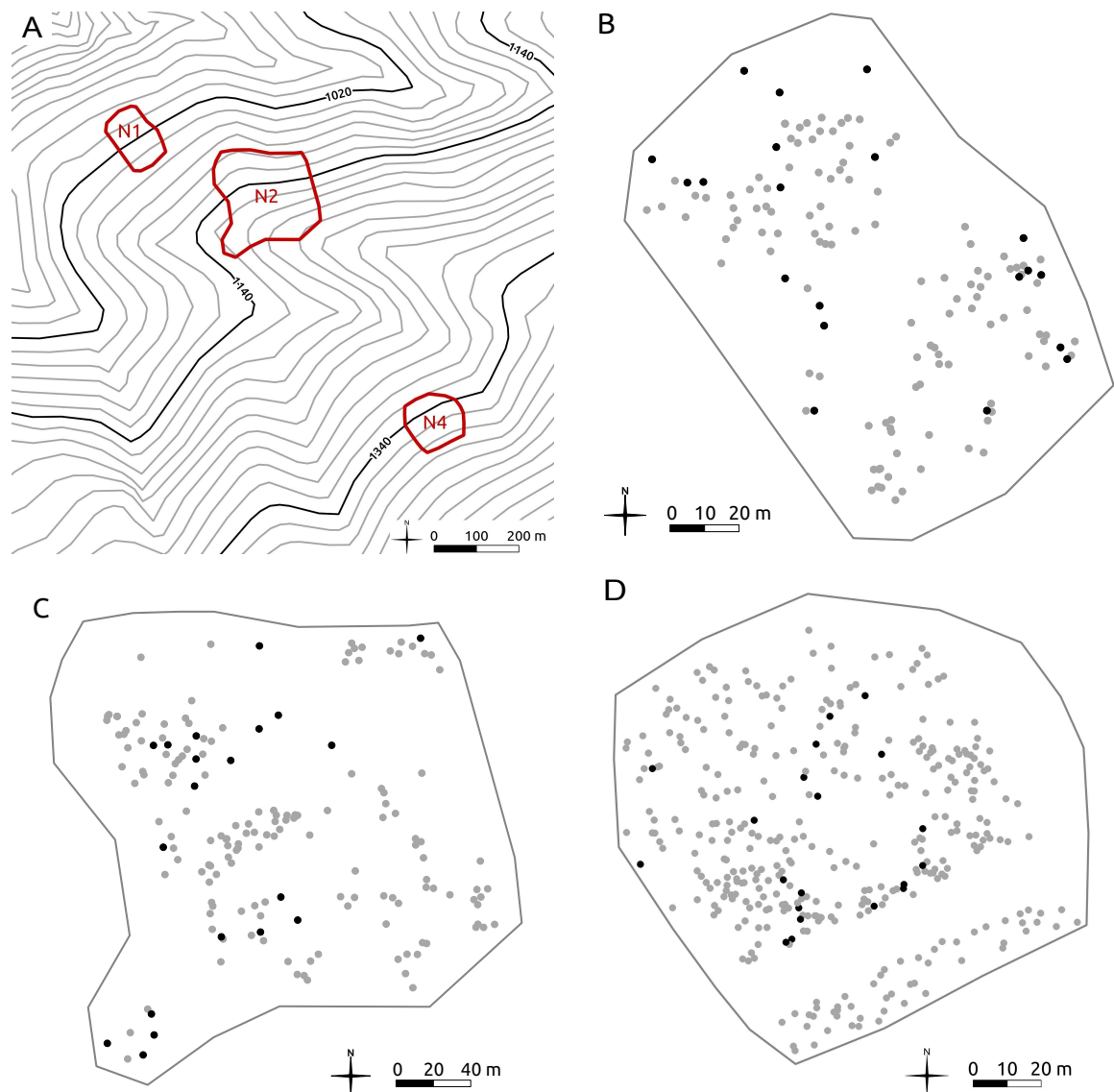


FIGURE 1.15 – Localisation des populations étudiées sur le versant nord du Mont-Ventoux. Au sein de chacune des populations, les arbre-mères suivis pendant cette thèse sont représentés par les points noirs et les hêtres potentiellement reproducteur par les points gris.

($\delta^{13}C$ et masse surfacique foliaire) mesurés après l'été 2011. À la mise en place du dispositif, en moyenne 35 descendants par famille ont été génotypés sur 13 marqueurs microsatellites pour réaliser des assignations de paternité et des estimations d'apparentement, les adultes potentiellement reproductifs dans les populations ayant été exhaustivement génotypés aux mêmes marqueurs. Ces arbres adultes en N1, N2 et N4 ont également été cartographiés (Figure 1.15) et phénotypés sur des caractères liés à la réponse au climat durant la thèse de Aurore Bontemps (Bontemps, 2012).

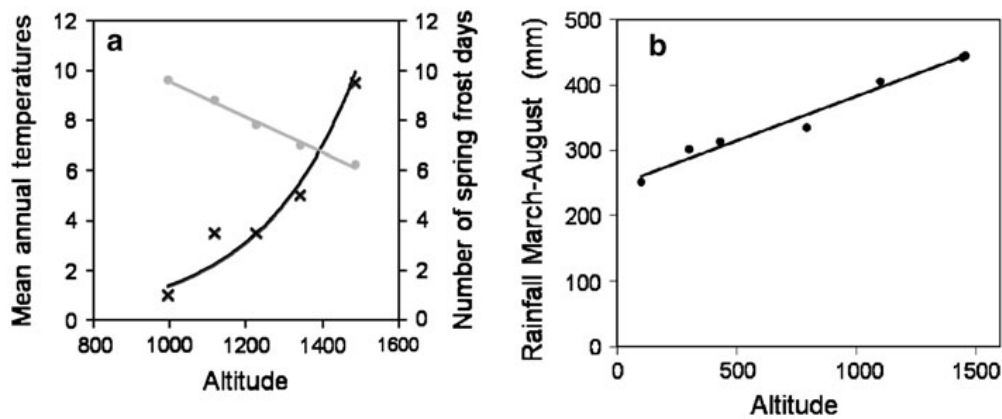


FIGURE 1.16 – a) Effets de l’altitude sur la température annuelle moyenne (points gris) et le nombre de gelées tardives au printemps (croix noires) en 2007 et 2008 sur la face nord du Mont-Ventoux. b) Effets de l’altitude sur les précipitations mesurées en printemps et été sur le Mont-Ventoux pour la période de 2000-2005. Tiré de Cailleret & Davi (2011).

Des expérimentations comparables sur des populations de hêtre réparties le long de gradients altitudinaux ont été déjà réalisées dans le sud-ouest de la France, sur la chaîne montagneuse des Pyrénées (Vitasse *et al.*, 2009; Bresson *et al.*, 2011). Ces études ont notamment permis de caractériser les patrons de variations génétiques pour un certain nombre de traits mesurés dans cette thèse pour le hêtre, comme une variation vers de plus forts *LMA* et contenus en azote foliaire avec l’altitude (Bresson *et al.*, 2011) ou encore une plus forte précocité des phénologies de débourrement végétatif et sénescence foliaire avec l’altitude (Vitasse *et al.*, 2009). Plus globalement, les objectifs de ces études ont été de caractériser la réponse plastique des traits écophysologiques avec les variations environnementales le long de gradients d’altitude et d’étudier la réponse adaptative du hêtre à la variation de température et plus particulièrement au froid hivernal aux hautes altitudes. L’originalité du gradient étudié durant cette thèse est qu’il présente une variation abrupte des conditions environnementales sur une courte distance (~ 1 km séparant N1 et N4) et qu’il est situé en milieu Méditerranéen. Ce gradient est donc à la fois très pertinent pour aller comprendre l’adaptation du hêtre à un gradient de sécheresse et température et étudier l’interaction entre flux de gènes et pressions de sélection, choses qui n’ont jamais été faites chez cette espèce.

Afin d’évaluer empiriquement l’impact des flux de pollen sur les capacités d’adaptation des populations de hêtre commun, nous avons suivi une démarche en 4 étapes, correspondant aux différents chapitres de cette thèse.

Plus précisément, le **Chapitre 2** de cette thèse a cherché à caractériser la dispersion pollinique, la quantité de pollen immigrant interceptée par chacune des populations étudiées et les paramètres individuels du régime de reproduction (nombre de pères efficace, taux d'autofécondation). Sur la base d'un modèle de croisement dans le voisinage (Oddou-Muratorio *et al.*, 2005), je montrerai comment la prise en compte de la variabilité inter-individuelle du régime de reproduction permet de réaliser des comparaisons plus justes de ces paramètres entre populations et permet d'identifier les principaux facteurs écologiques déterminant cette hétérogénéité.

Dans le **Chapitre 3** l'objectif a été de déterminer si les traits fonctionnels potentiellement adaptatifs présentaient une base génétique héritable, et pouvaient donc être transmis par la dispersion pollinique. Une revue bibliographique m'a tout d'abord permis d'identifier les biais classiquement mesurés dans les études de génétique quantitative utilisant des hypothèses sur l'apparement des familles issues de pollinisation libre et récoltées dans les populations naturelles d'arbres. Dans un second temps, l'utilisation d'un modèle de simulation de généalogies m'a permis de comparer différentes méthodes permettant de limiter ces biais méthodologiques en ayant une connaissance incomplète sur l'apparement. Enfin, j'ai appliqué ces méthodes au jeu de données recueilli dans cette thèse afin de mesurer l'héritabilité des traits physiologiques, morphologiques, phénologiques et de croissance mesurés en jardin commun.

L'objectif du **Chapitre 4** a été de tester la divergence adaptative des traits fonctionnels le long du gradient altitudinal du Mont-Ventoux et quantifier le potentiel de réponse de ces traits à la sélection. Ainsi, dans un premier temps j'ai appliqué une approche originale publiée récemment (Ovaskainen *et al.*, 2011) permettant de tester les traces de sélection sur un petit nombre de populations. Dans un second temps, l'analyse des variances génétiques entre populations et l'estimation des matrices de variance-covariance entre traits m'a permis d'évaluer les contraintes potentielles d'évolution multi-traits.

Ces trois premiers chapitres ont permis de tester que le système d'étude "Mont-Ventoux" était pertinent pour étudier l'interaction entre flux de gènes et sélection. Enfin, dans le **Chapitre 5** j'ai testé l'effet de la migration pollinique sur les variances et moyennes génétiques des traits fonctionnels au sein des populations de hêtre le long du gradient altitudinal, en utilisant une approche originale de catégorisation des individus par rap-

port à l'origine de leur père.

Pour **conclure cette thèse**, je synthétiserai les éléments des différents chapitres permettant d'apporter des éléments de réponse sur les capacités adaptatives des populations de hêtre aux changements climatiques et je discuterai de points de perspective et d'ouverture sur ce travail de thèse.

L'**Annexe A** présente un article de revue bibliographique auquel j'ai contribué durant ma thèse, analysant les conséquences évolutives des pratiques de gestion forestière afin de guider ces pratiques dans un contexte de changements climatiques.

DISPERSION POLLINIQUE ET TAUX
D'IMMIGRATION AU SEIN DE
POPULATIONS DE HÊTRE COMMUN



Synthèse

Les objectifs de ce chapitre ont été d'estimer les composantes du régime de reproduction et de caractériser les flux de gènes par pollen au sein de trois populations de hêtre commun le long d'un gradient altitudinal. Les paramètres du régime de reproduction au sens large¹ peuvent être estimés à l'aide de données génétiques et de deux grands types d'approches appelées indirectes et directes (Sork *et al.*, 1999). Les approches indirectes sont basées sur l'analyse de la structure génétique et sur un modèle reliant cette structure aux processus qui la façonnent sous un certain nombre d'hypothèses simplificatrices (notamment l'état d'équilibre). Les méthodes indirectes estiment généralement les flux de gènes historiques entre populations (cumulés sur un grand nombre de générations), à l'exception de l'approche TWOGENER, qui est une approche indirecte contemporaine (Smouse *et al.*, 2001). Les approches directes sont quant à elles basées sur des analyses de parentés pour reconstituer a posteriori les événements de dispersion du pollen et des graines. Elles requièrent donc le génotypage d'un échantillon de graines récoltées sur des arbres récepteurs de pollen (ou arbre-mères) et de la totalité des pères potentiels dans une zone donnée, ainsi que la localisation spatiale de ces adultes reproducteurs. Les méthodes directes, bien qu'elles soient plus lourdes à mettre en œuvre expérimentalement, permettent une caractérisation plus précise des dynamiques de flux de gènes actuels que les approches indirectes (Sork *et al.*, 1999). Parmi les approches directes, le modèle de croisement spatialement explicite (SEMM²; Oddou-Muratorio *et al.* 2005) permet notamment de pouvoir estimer conjointement le noyau de dispersion pollinique (la densité de probabilité qu'un grain de pollen émis au point (0,0) aille fertiliser un ovule en position (x,y)), les composantes du régime de reproduction (taux d'autofécondation, s , et taux d'immigration, m) et les variations inter-individuelles de fertilité mâle liées à des covariables mesurées sur les arbres adultes (e.g., hauteur, diamètre, production de graine ou voisinage de l'arbre source).

Un verrou des approches dérivant du SEMM est que le noyau de dispersion du pollen, et les paramètres du régime de reproduction sont supposés identiques pour tous les individus de la population. Dans ce chapitre je me suis particulièrement intéressée à l'estimation des composantes du régime de reproduction à l'échelle individuelle (m_i et s_i) et aux facteurs écologiques dans le voisinage des arbre-mères déterminant ces variations inter-individuelles. Pour cela, le modèle SEMM a été complexifié, à partir de l'implémentation Bayésienne de ce

1. inclus la reproduction et dispersion
2. pour "*spatially explicit mating model*"

modèle (MEMM³ ; Klein *et al.* 2008).

Sur le gradient altitudinal étudié les populations aux basses altitudes débourrent et fleurissent plus précocement que les populations aux hautes altitudes ($\pm 1,8$ jour de décalage par 100 m ; Davi *et al.* 2011). De plus, chez le hêtre les fleurs femelles sont réceptives ~ 5 jours avant les fleurs mâles (phénomène appelé protogynie ; Merzeau *et al.* 1994). Ici, j'ai donc tout particulièrement cherché à voir quel était l'effet du décalage phénologique avec le voisinage sur la réception de pollen immigrant. J'ai également testé l'effet d'autres variables écologiques, telles que la densité, le nombre de copies clonales et le diamètre des arbre-mères.

Cette étude montre que plus de la moitié du pollen intercepté par les arbre-mères provient de l'extérieur des populations ($m = 56$ %). Au sein des populations, des distances de dispersion par pollen relativement restreintes sont estimées, avec en moyenne $\delta = 46$ m. La queue du noyau de dispersion indique cependant que les événements de dispersion à longue distance ne sont pas négligeables ($b = 0,41$). À l'échelle individuelle, les arbre-mères présentant des phénologie précoces par rapport à leur voisinage interceptent plus de pollen immigrant que les arbres synchrones ou tardifs au sein de la population d'altitude intermédiaire (N2). De plus, les taux moyens d'immigration varient significativement entre populations, avec des taux plus importants mesurés à haute altitude.

Ces résultats permettent d'avancer l'hypothèse que l'effet conjoint du décalage phénologique entre populations dans le temps et de la protogynie pourraient être à l'origine de flux de pollen directionnels du bas vers le haut du gradient, expliquant ainsi les plus forts taux d'immigration observés aux plus hautes altitudes. Cette asymétrie des flux de pollen le long du gradient pourrait être particulièrement avantageuse car elle rendrait plus efficace la diffusion d'allèles favorables pré-adaptés aux conditions environnementales sèches et chaudes vers les hautes altitudes. Cependant, cette approche ne permet pas d'aller distinguer des individus issus de reproduction avec du pollen local ou migrant de moyenne ou longue distance. Dans le **chapitre 5** de cette thèse, je reviendrai sur comment des méthodes d'assignation de paternité catégoriques peuvent permettre d'apporter cette information en complément à l'approche qui a été utilisée ici.

3. pour "*mixed effect mating model*"

Ce travail a donné lieu à la publication d'un article scientifique en 2013 dans le journal *Molecular Ecology*.

Ecological determinants of mating system within and between three *Fagus sylvatica* populations along an elevational gradient

JULIE GAUZERE,*† ETIENNE K. KLEIN*† and SYLVIE ODDOU-MURATORIO*

*INRA, UR629 Ecologie des Forêts Méditerranéennes (URFM), F-84914 Avignon, France, †INRA, UR546 Biostatistique et Processus Spatiaux (BioSP), F-84914 Avignon, France

Abstract

Studies addressing the variation of mating system between plant populations rarely account for the variability of these parameters between individuals within populations, although this variability is often non-negligible. Here, we propose a new direct method based on paternity analyses (Mixed Effect Mating Model) to estimate individual migration (m_i) and selfing rates (s_i) together with the pollen dispersal kernel. Using this method and the KINDIST approach, we investigated the variation of mating system parameters within and between three populations of *Fagus sylvatica* along an elevational gradient. Among the mother trees, s_i varied from 0% to 48%, m_i varied from 12% to 86% and the effective number of pollen donors (Nep_i) varied from 2 to 364. The mating patterns differed along the gradient, the top population showing higher m and lower s , and a trend to higher Nep than the bottom populations. The phenological lag shaped long-distance pollen flow both within population (by increasing m_i at mother-tree level) and between populations (by increasing m at high elevation). Rather than the mate density, the canopy density was detected as a major mating system determinant within population; it acted as a barrier to pollen flow, decreasing the proportion of long-distance pollen flow and increasing s_i . Overall, the effects of ecological factors on mating system were not the same within vs. between populations across the gradient, and these factors also differed from those traditionally found to shape variation at range-wide scale, highlighting the interest of multiscale approaches.

Keywords: bud burst phenology, competitor density, ecological maternal neighbourhood, effective number of pollen donors, mating system, migration rate, multiscale analysis, selfing, spatially explicit mating model

Received 22 March 2013; revision received 21 June 2013; accepted 24 June 2013

Introduction

Mating system critically determines the levels of genetic diversity within and between populations (Loveless & Hamrick 1984; Vekemans & Hardy 2004). Most plant species are hermaphroditic; that is, they reproduce through both self-fertilization (selfing) and mating with other individuals (outcrossing). In this case, mating system is typically characterized by (i) the selfing rate, s ; (ii) the level of correlated paternity among seeds

produced by a single maternal seed-source, r_{gg} (Hardy *et al.* 2004); and (iii) the amount of pollen originating outside of the population, m (Burczyk *et al.* 2002; de-Lucas *et al.* 2008). These mating parameters are linked to the pollen dispersal kernel. For instance, in species with fat-tailed dispersal kernel, that is, important long-distance dispersal abilities (e.g. in wind-pollinated tree species), high m values and reduced r_{gg} values are expected (Klein *et al.* 2006). Also peaked dispersal kernels at short distance lead to L-shaped contributions of pollen donors to the fertilization of a given mother tree, with few major donors and numerous minor contributors (Kang *et al.* 2003; Oddou-Muratorio *et al.* 2005). The

Correspondence: Sylvie Oddou-Muratorio, Fax: +33 (0)4 32 72 29 02; E-mail: oddou@avignon.inra.fr

combination of L-shaped male fecundities, local pollen dispersal and selfing events can result in a decreased effective population size, increased spatial genetic structure and erosion of genetic diversity through increased genetic drift (Vekemans & Hardy 2004). Therefore, the simultaneous investigation of the mating system parameters and the pollen dispersal kernel is expected to precisely characterize the mating system and its evolutionary consequences.

Numerous studies have examined the variations of mating system between populations within species, particularly to reveal the mechanisms that affect the limitation of species distribution areas (Michalski & Durka 2007; Mimura & Aitken 2007) or to understand the impact of habitat fragmentation due to human action on mating parameters (Robledo-Arnuncio *et al.* 2004; Fuchs & Hamrick 2011). Estimations performed at the population scale have generally revealed high variability in outcrossing rates, levels of correlated paternity and pollen flow both between species and between populations within species (El-Kassaby & Jaquish 1996; Piotti *et al.* 2012; Silva *et al.* 2011). This variability primarily reflects the effect of population density (Restoux *et al.* 2009), in addition to the effect of others ecological factors acting at the population scale (e.g. the population size). Experimental studies have shown increased values of s and r_{gg} in populations at range peripheries (Mimura & Aitken 2007) and in populations with a low density (El-Kassaby & Jaquish 1996) or a small size (Robledo-Arnuncio *et al.* 2004).

Additionally, within natural populations, all mother plants experiment different abiotic and biotic environment, which may affect the way they receive pollen. Analysing variation in plant mating system at the mother-plant level is therefore highly relevant (Robledo-Arnuncio *et al.* 2004; Garcia *et al.* 2005; Oddou-Muratorio *et al.* 2006). Accordingly, recent studies have found a large variability in outcrossing rates and r_{gg} at the mother-plant level (de-Lucas *et al.* 2008; Tamaki *et al.* 2009; Silva *et al.* 2011). This variability has been often interpreted as the result of the variability of local ecological factors, in the frame of the mass-action model (Holsinger 1991). This model assumes that the selfing probability is determined by the dilution of the mother tree's own pollen (selfed pollen) on its stigmas by all of the exported pollen (outcrossed pollen). The amount of outcrossed pollen is determined by (i) the amount of pollen exported by reproductive individuals, that is, male fecundity; (ii) the density of reproductive males; and (iii) the probability that the exported pollen will successfully reach a mother-tree's stigmas, which depends on the overlap in flowering and the distance between mates, given a dispersal kernel. The mass-action model can be applied at the population level

(Holsinger 1991) or the individual level. In the latter case, ecological factors within the vicinity of the pollen recipient (typically the number and size of its neighbours) are expected to explain the individual variation in the mating system parameters; these factors were referred to as the 'ecological maternal neighbourhood' by Garcia *et al.* (2005). While a number of studies have focused on individual variability in selfing rates (Franceschinelli & Bawa 2000; Ward *et al.* 2005; Tamaki *et al.* 2009), only a few have investigated the three mating system components, s , m and r_{gg} , at the individual scale and their determinants within populations (Garcia *et al.* 2005; Oddou-Muratorio *et al.* 2006; Silva *et al.* 2011), especially among anemophilous species (but see de-Lucas *et al.* 2008). Moreover, the consequences of the high variability of individual mating system parameters within populations on their comparison between populations have received little attention.

In this study, we investigate the variation of s , m and r_{gg} within and between three populations of *Fagus sylvatica* (the European beech), and we attempt to connect the ecological determinants acting at the individual and population levels. *Fagus sylvatica* is a predominantly outcrossed species, with wind-dispersed pollen. The three studied populations are located in the southern range of the species distribution, across an elevational gradient spanning over ~ 1 km on Mont Ventoux. Along this short gradient, the populations at lower altitudes complete bud burst and flower earlier than the populations at higher altitudes because bud burst is triggered by the heat sum ($a \pm 1.8$ day lag per 100 m and -0.8° per 100 m in altitude; Davi *et al.* 2011). Moreover, in *F. sylvatica*, the female flowers are receptive ~ 5 days before the anthers release their pollen (a phenomenon known as protogyny; Merzeau *et al.* 1994). Accordingly, the populations at medium and high elevations should receive more immigrant pollen than populations at low elevations (hypothesis H1). Furthermore, high temperature and low air humidity have been reported to reduce pollen germination and viability (e.g. Beck-Pay 2012); this may particularly affect mating patterns at low elevation (hypothesis H2). Finally, confronting our results to those of previous mating system studies is expected to shed light on the ecological determinants acting at very local scale (this study) vs. range-wide scale (e.g. Mimura & Aitken 2007).

Mating system parameters can be estimated from genetic markers using three main methods, which essentially differ regarding how they analyse the maternal pollen pool structure. Historically, the multilocus mating system method, implemented in MLTR (Ritland 2002), was the first to estimate mating system parameters (individual s and r_{gg}), using the genotypes of mother trees and their offspring as inputs. The two-generation

models (TWOGENER, Smouse *et al.* 2001), and later the KINDIST approach (Robledo-Arnuncio *et al.* 2007), were then proposed to account for spatial information in analyses of mating system. The correlated paternities, within and between maternal pollen pools, are estimated from genotype data and used together with the location information to estimate pollen dispersal kernel parameters. Finally, Spatially Explicit Mating Models (SEMMs, Burczyk *et al.* 2002; Oddou-Muratorio *et al.* 2005) are paternity-based methods that use the genotypes and spatial positions of mother trees and their offspring as inputs, in addition to those of potential fathers. SEMMs estimate the average values of m , s and dispersal kernel parameters at the population level, and take into consideration that each offspring collected on a mother tree can originate either from an adult tree outside of the sampled population or from a sampled adult via outcrossing or selfing (Burczyk *et al.* 2002). Overall, the MLTR method precisely describes mating components, but conceals spatial and dispersal information. The TWOGENER method accounts for spatial information but only considers r_{gg} and not selfing and immigration. SEMM precisely defines the mating and dispersal components but assumes that the mating parameters are uniform across all of the mother trees in a population. To overcome these methodological limitations, we first developed a new model (MEMMi) that accounts for spatial and dispersal information to estimate dispersal parameters (the mean dispersal distance, δ and the shape parameter, b) at the population level and migration (m_i) and selfing rates (s_i) at the mother-tree level.

Second, we used the MEMMi model to estimate m and s , within the three studied populations of *F. sylvatica*. Additionally, the level of correlated paternity within each maternal pollen cloud (r_{ggi}) was estimated using the KINDIST method (Robledo-Arnuncio *et al.* 2007). Third, we investigated the effect of ecological factors acting within the mother-tree neighbourhood on the variability of m_i , s_i and r_{ggi} . We considered jointly the effects of phenological synchronization with neighbours, local density and diameter of the mother tree, which are expected to be major determinants of individual mating system. Finally, we compared the mating system parameters among the three populations.

Methods

Study species

The European beech, *F. sylvatica*, is a monoecious, predominantly outcrossed species, with selfing rates ranging from 1.2% to 13% (Wang 2003). These low selfing rates probably result from protogyny, and possibly, from a self-incompatibility system (described in the Fagaceae family, Kouno and Mukouda 1985, Yacine &

Bouras 1997). *Fagus sylvatica* is anemophilous, displaying mean pollen dispersal distances ranging from 80 to 184 m (Piotti *et al.* 2012), though in some cases, pollen grains have been reported to disperse up to 1000 km (Belmonte *et al.* 2008). Flowering occurs in April and May, shortly after bud burst (Milleron *et al.* 2012).

Study site and field sampling

We studied three populations on the north face of Mont Ventoux, designated N1 (1.3 ha, 1,020 m a.s.l.), N2 (2.2 ha, 1,140 m a.s.l.) and N4 (0.8 ha, 1,340 m a.s.l.; Fig. 1, Table 1). On this face, the *F. sylvatica* populations range almost continuously from 750 to 1700 m in elevation. All of the potentially reproductive individuals within these populations were mapped (576 in total) and sampled for genetic analyses. *Fagus sylvatica* sometimes produces stump shoots resulting in multiple stems; in such cases, only the largest stem was sampled. In population N2, we verified that the observed multiple stems were clones (data not shown). In 2009, at each site, 20 trees were chosen as mother trees to collect open-pollinated seeds directly from the canopy (Table 2). From these seeds, 1379 offspring successfully germinated (~24 seeds per mother tree; Table 2).

Microsatellite genotyping

The genotypes of the mother trees, offspring and the putative father trees were scored at thirteen microsatellite loci: FS1-15, FS3-04 (Pastorelli *et al.* 2003); sfc0007-2, sfc-1143, sfc-0161 (Asuka *et al.* 2004); mfc7 (Vornam *et al.* 2004); Csolfagus_19, Csolfagus_6, Csolfagus_29, Csolfagus_31 (Lefèvre *et al.* 2012); Csolfagus_7, Csolfagus_25 and Fi05 (G.G. Vendramin, personal communication, Tables S1 and S2, Supporting information). DNA extraction, PCR amplifications and genotype scoring with a MegaBACE 1000 sequencer were performed using the conditions described by Lander *et al.* (2011). Offspring showing genotypes that were inconsistent with the mother-tree genotypes were used to assess the null allele frequencies and genotyping error rates.

Mating variables and dispersal kernel estimates

Selfing and migration rates were jointly estimated with the pollen dispersal kernel in a Bayesian framework using a modified version of the Mixed Effect Mating Model (MEMM) (Klein *et al.* 2008). Originally, the MEMM was based on the SEMM proposed by Oddou-Muratorio *et al.* (2005), where each seed, i , collected on a mother tree, j , can be sired either i) by an adult tree outside the study area (migration), with a probability of m ; ii) by the mother tree (selfing) with a probability of

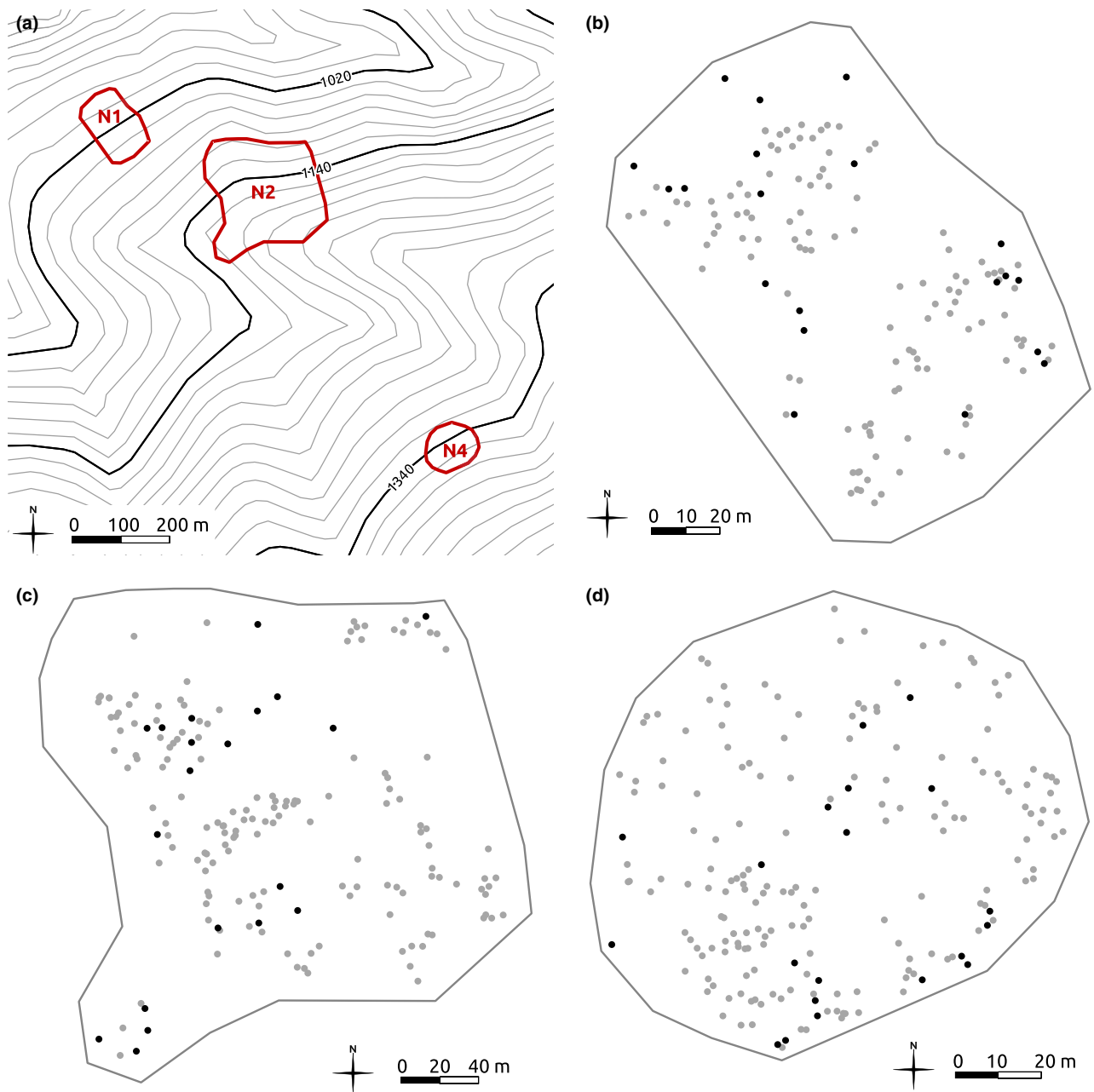


Fig. 1 Location of the three experimental sites on Mont Ventoux (a), and mapping of the *F. sylvatica* trees in each population (b: N1; c: N2; d: N4). The mother trees are indicated with black points, and all of the other adult trees are indicated with grey points.

s ; or iii) by another adult tree within the study site with a probability of $1-m-s$. The likelihood of observing an offspring with a genotype g_o is:

$$L(g_o|F, \sigma^2, a, b, s, m) = sT_1(g_o|g_j, g_j) + mT_2(g_o|g_j, \text{BAF}) + (1-s-m)\sum_k \pi_{jk} T_3(g_o|g_j, g_k) \quad (1)$$

where BAF is the background allelic frequencies outside of the population, and T represents the Mendelian transition probabilities for self-fertilization of mother tree j with a genotype g_j (T_1); for cross-fertilization of female

j with a male outside of the study area (T_2); and for mating between the female j and male k , with a genotype g_k , within the neighbourhood (T_3). Note that this fractional model accounts for cryptic gene flow, that is, that the actual pollen donor can be located outside the population with a non-null likelihood even if a compatible father exists inside the plot.

In the outcrossing case, the probability π_{jk} of each male k , to contribute to the pollen pool of female j is modelled as:

Table 1 Geographical, phenological and phenotypic characteristics of the studied populations

	N1	N2	N4
Altitude (metres)	1020	1140	1340
Dimension (hectare)	1.30	2.20	0.80
Mean temperature	10.09	9.25	7.43
Mean humidity rate	71.7	72.2	73.0
Mean diameter (cm)	25.47	27.37	23.14
Mean conspecific density (/ha)	32.26	20.45	59.33
Mean competitor density (/ha)	230.20	130.16	320.91
Mean number of clonal copies by mother tree	2.55	6.81	1.82
Absolute phenological lag	0.79	1.02	0.66
Basal area (m ² /ha)	13.4	8.5	14.3

Table 2 Details of the sample sizes related to the phenotypic and genetic data obtained for each population

	N1	N2	N4	Total
Number of reproductive adults genotyped	137	170	194	501
Number of seeds genotyped	461	443	474	1378
Number of mother trees phenotyped*	20	16	17	53
Number of reproductive adults phenotyped (including mothers)*	137	136	119	392

*i.e. coordinates, diameter and phenology assessed.

$$\pi_{jk} = \frac{F_k p(d_{jk})}{\sum_{l: \text{father}} F_l p(d_{jl})} \quad (2)$$

where F_k and F_l are the fecundities of father k and l , respectively; d_{jk} and d_{jl} are the distances between mother j and father k and l , respectively; and p is the dispersal kernel.

The pollen dispersal kernel was modelled using the classical exponential power family curve (Klein *et al.* 2006):

$$p(a, b; x, y) = \frac{b}{2\pi a^2 \Gamma(2/b)} \exp\left(-\left(\frac{r}{a}\right)^b\right) \quad (3)$$

with r the distance between the pollen source and the receiver ($r = \sqrt{x^2 + y^2}$); Γ the classical gamma function; a the scale parameter linked to the mean dispersal distance:

$$\delta = \frac{a\Gamma(3/b)}{\Gamma(2/b)} \quad (4)$$

and b the shape parameter ($b > 1$: 'thin-tailed' kernel; $b < 1$ 'fat-tailed' kernel).

Compared with the initial SEMM of Oddou-Muratario *et al.* (2005), which was used to estimate the parameters of the neighbourhood model (m and s , in eqn 1, and δ and b in eqn 3), the Bayesian approach proposed by Klein *et al.* (2008) has the additional advantage of accounting for and estimating the total variance (σ^2) in male fecundity (F_k in eqn 2). For this purpose, the Bayesian framework considers four successive layers (Fig. S1, Supporting information): (i) the *prior* distributions of the parameters; (ii) the parameters a , b , s and m , and the variance of the fecundity σ^2 ; (iii) the latent variables (the individual fecundities F_k); and (iv) the observations (the genotypes of the offspring and the adults). These layers are articulated by eqns (1) and (2). The Bayesian analysis employed a Monte Carlo Markov chain (MCMC), based on the Metropolis–Hasting algorithm.

In this study, we extended this approach to estimate the individual values for each mother tree of the mating system parameters m_i and s_i , instead of the average m and s values at the population level (Fig. S1, Supporting information). For this purpose, we modified the C++ MEMM program (the new code, MEMMi, is available upon request from JG).

We applied both the existing MEMM and the new MEMMi in each of the three populations using a MCMC of 50 000 steps and a burn-in of 5000 steps. We employed a Gamma distribution for the latent variables F_k and uniform prior distributions on the intervals [0, 1], [0, 1], [0, 10 000], [0, 10] and [0, 1000] for parameters s_i , m_i , δ , b and σ^2 , respectively. We derived point estimates for the parameters using the medians of the posterior distributions.

Finally, we employed the genotypes of the sampled outcrossed seeds and mother trees to obtain the effective number of fathers pollinating each mother, Nep_i , with POLDISP software (Robledo-Arnuncio *et al.* 2007). This approach estimates the probability of paternal identity within sibship, also referred to as the paternity correlation (r_{gg}). The level of paternity correlation is directly linked to Nep as: $Nep = 1 / r_{gg}$. This estimate characterizes the effective number of males actually participating in the pollination of a mother tree, whereas the σ^2 estimate describes the variance of the quantity of pollen emitted by the trees. Note that MEMMi can predict the Nep_i of the mother trees from the predicted π_{jk} . These predictions are likely upwardly biased as they largely depend on the model assumptions, potentially missing some mating system determinants in the ecological maternal neighbourhood, and assuming an infinite number of pollen donors outside the populations. We thus preferred relying only on the data by using the POLDISP estimates of r_{gg} .

Ecological variables

We characterized a set of seven phenotypic, phenological and demographic variables for 53 mother trees (Tables 1 and 2).

Mother-tree diameter. Tree diameter is related to tree vigour and is generally well correlated with fecundity (Oddou-Muratorio *et al.* 2005). The diameter of each mother tree was measured at a height of ~1.30 m. If a mother tree displayed multiple stems, the diameter of the principal stem was measured, and the number of clonal stems was counted. The number of clonal copies varied between 0 and 15 per mother tree.

Local density. Using the spatial coordinates, the conspecific local density was estimated based on the number of reproductive *F. sylvatica* neighbours found in disks with a radius of 10, 20, 50, 75 or 100 m around each mother tree. The competitor density was estimated similarly considering all of the trees within the disks (including *F. sylvatica*, *Pinus sylvestris* and *Abies alba*). Both the conspecific and total competitor density differed across the populations: N4 was the most dense population (conspecific density = 59.33 trees/ha) and N2 the least dense (conspecific density = 20.45 trees/ha; Table 1).

Local basal area. This variable combines vigour and density information. The basal area was computed as the sum of surfaces at 1.30 m of all of the reproductive *F. sylvatica* neighbours found in disks with a radius of 10, 20, 50, 75 or 100 m around the mother trees.

Phenological lag. In *F. sylvatica*, the flowering phenology is hard to follow because the succession of the flowering stages is rapid, and the reproductive organs are small. The bud burst phenology, which is easier to observe, was employed to quantify the flowering lag. In spring 2009, the bud burst was recorded on 146 adult trees in population N1, 133 adult trees in N2 and 209 adult trees in N4 (Table 1). The bud burst phenology was characterized using the five stages described by Davi *et al.* (2011): (i) dormant buds, (ii) swelling buds, (iii) broken bud scales, (iv) emerging leaves and (v) spread out leaves. The phenological stage of the upper/lower part of each adult tree was noted on 15 different dates in population N1 (between 3/23 and 5/04/2009), on a single date in population N2 (4/21/2009) and on 13 different dates in population N4 (between 3/24 and 5/14/2009). For populations N1 and N4, a single phenological note, n_p , was computed for each tree as the sum of the phenological stages (1–5) observed over all of the dates; n_p can vary from 0 to 75. For population N2, n_p was simply the phenological stage observed on the unique observation date.

We then derived two measures of the phenological lag between a given mother tree and its neighbours. First, we measured the mean phenological lag, PL_{mean} , which describes whether a mother tree flowers early ($PL_{\text{mean}} > 0$) or late ($PL_{\text{mean}} < 0$) compared to its neighbours. Second, we computed the sum of the absolute phenological lags, $|PL|$, which measures how much a given mother tree is asynchronous from its neighbours. PL_{mean} and $|PL|$ were estimated within a radius of 50, 75 or 100 m (in the 10 or 20 m radius, we missed phenological information for 11 mother trees, so we did not use these distance classes).

Effect of ecological variables on mating patterns

We employed analysis of covariance (ANCOVA) methods to investigate the effects of ecological factors on the mating parameters m_i , s_i and Nep_i within the populations. We used linear models (lm) for m_i and $\log(Nep_i)$ and a generalized linear model (glm) with a quasi-Poisson distribution for s_i to account for overdispersed data. We considered the population of origin as a class variable and the mother-tree diameter, the number of clonal copies, the phenological lag compared with neighbours, the basal area and the neighbourhood density as quantitative variables. To correctly test the population effect in models, we centred the diameter, number of clonal copies, basal area and density variables based on their mean value along the gradient. For the phenological variables, we chose to not interpret the population differences because of the different measures of phenology that were used, so we centred and renormalized PL_{mean} and $|PL|$ within each population. The effect of the phenological variable was treated as linear for $|PL|$ and as linear or quadratic for PL_{mean} , allowing differentiating the effect on mating system of early- and late-flowering behaviour ($PL_{\text{mean}} > 0$ and $PL_{\text{mean}} < 0$, respectively). Moreover, when a phenological variable was present in the model, we always added the interaction effect between the phenology and the population in the model to account for the fact that the phenological measures differed between populations.

For each mating parameter (Y), we used a two-step procedure to conduct a hypothesis testing approach. In a first step, we chose the most relevant neighbourhood size (i.e. 10, 20, 50, 75 or 100 m) for each spatialized explanatory variable (i.e. phenological lag, basal area and densities) and then the most relevant 'density' variables (i.e. among the conspecific density, the total density and the basal area) and 'phenology' variable (i.e. among the PL_{mean} and $|PL|$). This procedure aimed at avoiding that highly correlated explanatory variables (PCA not shown) hide the significant effects of each other. To assist variable selection, we used the Akaike information crite-

rion (AIC) for lm and the quasi-AIC (accounting for the overdispersion of the data) for the glm. Second, we tested our biological hypothesis fitting the following lm and glm complete models. For m and Nep :

$$m \text{ or } \log[Nep] = \text{Intercept} + \text{Population} + \text{Diameter} \\ + \text{Number of clonal copies} + \text{Densities} \\ + \text{Phenology} + \text{Phenology/Population} + \epsilon$$

And for s :

$$s \sim \text{quasi-Poisson}(\lambda)$$

$$\lambda = \text{Intercept} + \text{Population} + \text{Diameter} \\ + \text{Number of clonal copies} + \text{Densities} + \text{Phenology} \\ + \text{Phenology/Population}$$

We used R functions 'lm' and 'glm' for all these analyses using a type III sum of squares.

Results

Pollen dispersal and mating system estimates based on MEMMs

The average pollen dispersal distance, δ , ranged from 35 m (population N4) to 63 m (population N2) (Table 3). The shape parameter, b , of the pollen dispersal kernel was lower than 1 in all of the populations ($b_{N1} = 0.52$, $b_{N2} = 0.27$, $b_{N4} = 0.43$), indicating non-negligible long-distance dispersal events. Population N2 tended to exhibit a lower b ($b_{N2} = 0.27$) and a higher δ ($\delta_{N2} = 63$ m) than the other two populations (Table 3), but the pollen dis-

Table 3 Median MEMM, MEMMi and POLDISP estimates of the dispersal and mating parameters by population, and mean population values based on individual estimates; m : migration rate, s : selfing rate, Nep : effective number of fathers, δ : mean dispersal distance, b : shape parameter for the dispersal kernel, σ^2 : variance of male fecundities

	N1	N2	N4
<i>MEMM</i>			
m	0.48 [0.42; 0.55] <i>a</i>	0.55 [0.50; 0.59] <i>a</i>	0.65 [0.61; 0.70] <i>b</i>
s	0.05 [0.03; 0.08] <i>cd</i>	0.06 [0.04; 0.09] <i>c</i>	0.02 [0.01; 0.04] <i>d</i>
<i>MEMMi</i>			
δ	39 [31; 60] <i>e</i>	63 [37; 294] <i>e</i>	35 [22; 232] <i>e</i>
b	0.52 [0.38; 0.73] <i>f</i>	0.27 [0.11; 0.51] <i>f</i>	0.43 [0.14; 0.8] <i>f</i>
σ^2	1 [1; 1.8] <i>g</i>	1 [1; 3.5] <i>gh</i>	3.8 [2.4; 6] <i>h</i>
Mean (m_i)	0.43	0.53	0.64
Mean (s_i)	0.03	0.06	0.2
<i>POLDISP</i>			
Median (Nep_i)	12.3	15.1	29.2

Differences among populations are represented by letters (populations sharing letters are identical according to the credible intervals).

persal parameters did not significantly differ between the populations (the credibility intervals, CI, overlapped).

When estimated as a single average at the population level with the standard MEMM, the pollen migration rates were $m_{N1} = 48\%$, $m_{N2} = 55\%$ and $m_{N4} = 65\%$ and were significantly different between populations N1 and N4 and between populations N2 and N4 (according to the CI). Similarly, the average s estimates at the population level ranged from $s_{N4} = 2.1\%$ to $s_{N2} = 6.4\%$ and were significantly different between populations N2 and N4 (according to the CI).

MEMMi yielded migration rates (m_i) at the mother-tree level ranging from 12% to 86% with a strong variability within populations (Fig. 2a, Table S3, Supporting information). The migration rates differed significantly (ANOVA P -value = 0.0007) between populations N1 and N4 and between N2 and N4 (with $m_{N1} = 42.8\%$, $m_{N2} = 47.9\%$ and $m_{N4} = 64.3\%$; Fig. 2a). The individual estimate of the selfing rate, s_i , was null for most of the families (37 out of 60); for 17 families, it was within the range estimated in the literature (1.2% to 13%); and for 6 families, it exceeded 13%, with a maximum value of $s = 48\%$ being obtained for a mother tree in population

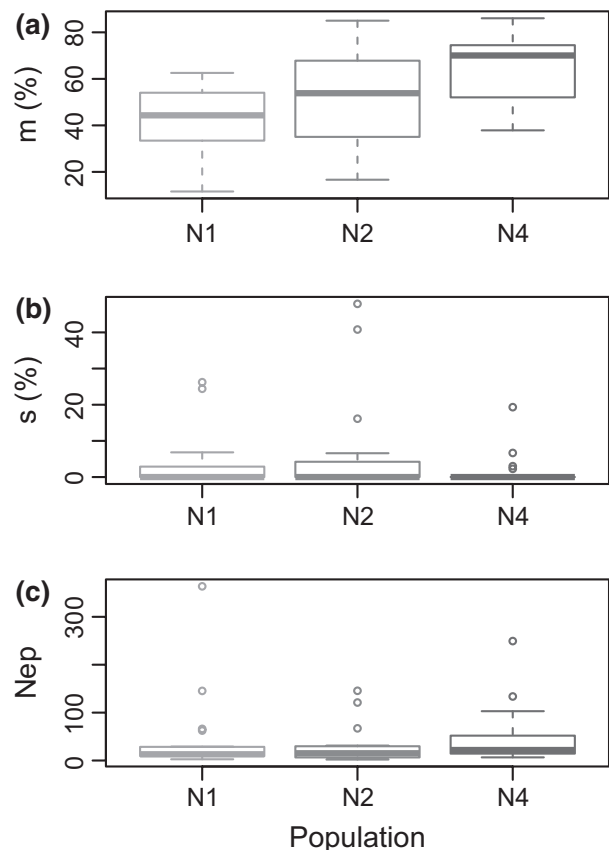


Fig. 2 Statistical distribution of the migration rate (m), selfing rate (s) and the effective number of fathers (Nep) by population.

N2 (Table S3, Supporting information). Moreover, s_i did not significantly differ between the populations (ANOVA P -value = 0.2721), although the mean selfing rate tended to be higher in N2 ($s_{N2} = 6.1\%$) than in the other populations (Table 3). Overall, both the migration and selfing rates showed a strong variability between individuals within each population (Fig. 2a,b).

Diversity within the pollen cloud

At the mother-tree level, the correlated paternities, $r_{gg'}$, obtained with POLDISP ranged from 0.002 to 0.5, and converted into Nep_i between 1.99 and 363.64 equally contributing males (Fig. 2c). Overall, the average Nep did not significantly differ between populations (P -value = 0.1217), although Nep tended to be higher in population N4 ($Nep_{N4} = 19.1$) (Table 3), compared with populations N1 ($Nep_{N1} = 8.8$) and N2 ($Nep_{N2} = 8.5$). However, the variance of male fecundity, σ^2 , was higher in N4 than in N1 and N2 ($\sigma^2_{N1} = 1$ with CI of [1; 1.8], $\sigma^2_{N2} = 1$ [1; 3.5] and $\sigma^2_{N4} = 3.8$ [2.4; 6]; Table 3).

Patterns and lags in bud burst phenology

In populations N1 and N4, for which an exhaustive phenological record is available, the bud burst dynamics (from the first individual at stage 2 to all individuals at stage 5) occurred from 23 March to 4 May 2009 for population N1 and from 24 March to 14 May 2009 for population N4 (Fig. S2, Supporting information). The three populations were then compared on a single key

phenological date, when the mean of the mother-tree phenological stages was between 3 and 4 (the stages where buds are the most sensitive to late frost). On this single key phenological date, the average phenological lags were 4, 9 and 13 days between populations N1 and N2, N2 and N4, and N1 and N4, respectively. The mean absolute within-population phenological lag differed significantly between the populations (P -value = 0.0004; Fig. 3), being highest in population N2 and lowest in the population N4 (Fig. 3).

Impact of ecological variables on m

The population (P -value < 0.0001), the diameter of the mother tree (P -value = 0.0078, slope = 0.7403), the competitor density within a radius of 10 m (P -value = 0.0024, slope = -0.6210; Fig. 4a) and the mean lag of the bud burst phenology between a mother tree and its neighbours within a radius of 100 m (P -value = 0.0068, slope = 5.6459), all significantly explained the variability of m_i in the complete model tested (Table 4). The quadratic effect of the mean phenological lag (P -value = 0.0268, slope = 3.0939) and the interaction effect between the population and the mean phenological lag (P -value = 0.0039) were also significant. Basically, large mother trees located in low-density patches displayed higher m_i values. Moreover, the positive and significant quadratic effect of the phenological lag on m_i indicated that mother trees lagged phenologically compared with their neighbours that received more migrant pollen (Fig. 4b–d). The population/phenology interaction effect indicated a shift of the symmetric axis (x_s) of the quadratic response of m_i to the phenology (Fig. 4b–d). For populations N1 and N4, this value was close to zero ($x_s = -0.28$ and 0.08 , respectively) (Fig. 4b,d), that is, minimum migration rates were found for individuals that were phenologically synchronous, whereas for N2, $x_s = -2.53$, that is, slightly late individuals tended to receive a lower migration rate (Fig. 4c). The number of clonal copies was the only nonsignificant variable. Overall, the complete model tested explained a large part of the variance in m_i ($R^2 = 0.59$).

In contrast to the ANOVA model (Table 3) that was used to test for differences in the average migration rates between populations, the ANCOVA model (Table 4) accounts simultaneously for the local configuration and the population context. The ANCOVA model still showed significant variation of m between the populations (P -value < 0.0001), with population N2 presenting the lowest m ($m_{N2} = 37.4\%$), N1 an intermediate value ($m_{N1} = 39.3\%$) and N4 always displaying the highest value ($m_{N4} = 69.4\%$). Pairwise comparisons showed that m_{N1} and m_{N2} were not significantly different, whereas

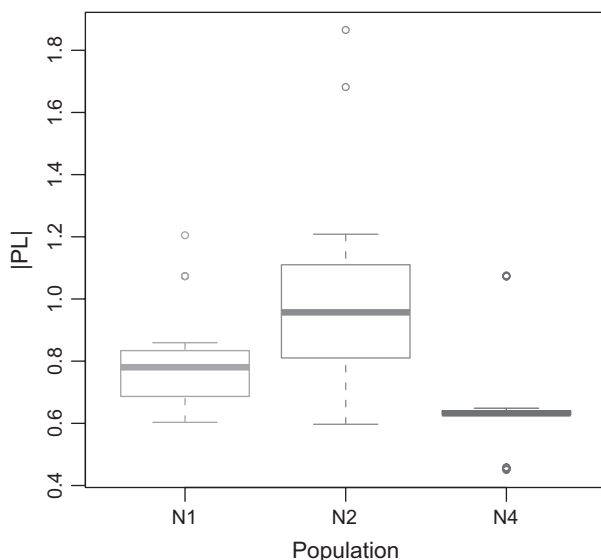


Fig. 3 Statistical distribution of the absolute phenological lags (|PL|) between mother trees and their neighbours within the population, calculated on a key phenological date (17 April for N1, 21 April for N2 and 30 April for N4).

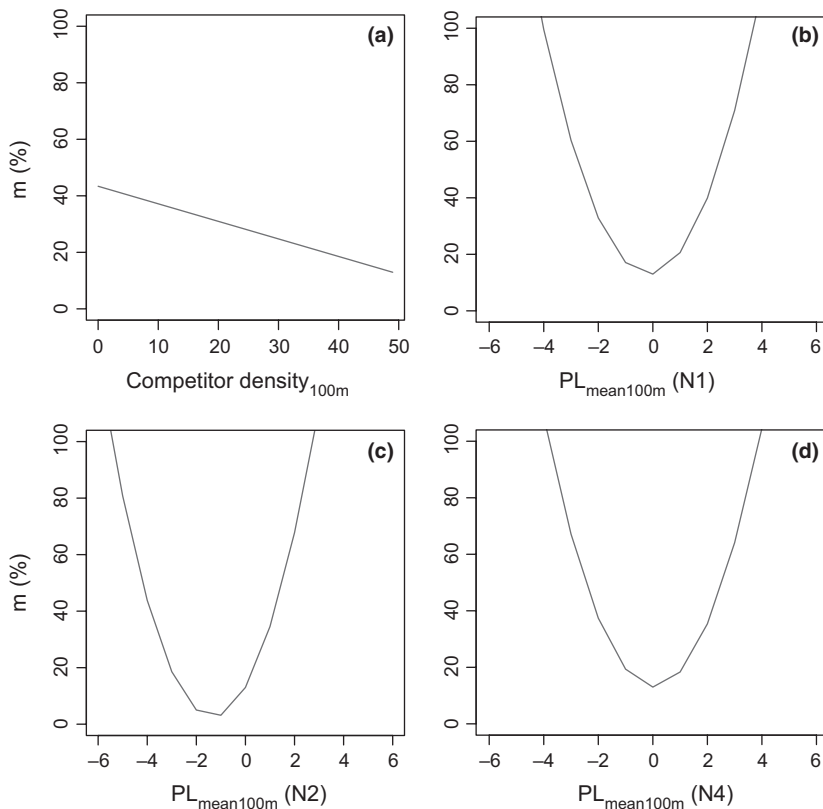


Fig. 4 Effects of ecological variables on migration rate (m): (a) correlation between m and the competitor density; (b), (c) and (d) correlation between m and the mean phenological lag between the mother trees and their neighbours in disks with a radius of 100 m for population N1 (b), population N2 (c) and population N4 (d). Negative values of phenological lag correspond to late-flowering individuals, and positive values to early-flowering individuals.

population N4 displayed a significantly higher m than both other populations.

Impact of ecological variables on Nep

The population (P -value = 0.0722) and the conspecific basal area in a radius of 100 m (P -value = 0.0945, slope = 0.0995) marginally explained the variability of Nep_i (i.e. $0.05 < P$ -value < 0.10 ; Table 4). All the other variables tested, and notably the phenological lag, did not significantly affect Nep_i . Therefore, higher Nep_i values were estimated for mother trees surrounded by large and/or many conspecific pollen donors. The complete model tested explained 32% of the Nep_i variance ($R^2 = 0.32$).

The ANCOVA model tended to accentuate the Nep differences between the bottom populations and the top population, with populations N1 and N2 displaying the lowest mean Nep ($Nep_{N1} = 12.84$ and $Nep_{N2} = 19.29$) and population N4 exhibiting the highest mean Nep ($Nep_{N4} = 32.31$).

Impact of ecological variables on s

All the ecological variables and the interaction factor appeared to have a significant or marginal effect on the s_i variability (Table 4). The mother-tree diameter and the absolute phenological lag had negative main effects

on s_i (slope = -0.0871 and -2.1057 , P -value = 0.0514 and 0.0432, respectively), while the competitor density within a 10-m radius, and the number of clonal copies had positive main effects (slope = 0.0483 and 0.1225, P -value = 0.0795 and 0.0247, respectively). The interaction factor (P -value = 0.0285) showed that the effect of the absolute phenological lag varied between populations. The highest impact of the phenological lag on s_i was found in population N4. Overall, large and isolated mother trees, displaying a low number of clonal copies, exhibited lower s_i .

However, when we removed the mother tree with the highest selfing rate in population N4 (family 47), the phenological variable was no longer significant. Therefore, although we accounted for overdispersion, the link between s_i and the phenology was largely influenced by one outlier family. More generally, no classical distribution satisfactorily fitted the observation of s_i values (many 0 and a large over-dispersion of nonzero values), and P -values here should be considered with care.

Interestingly, the population had a significant impact on s_i (P -value = 0.0200) accounting for the local configuration, which was hidden in the simple model (ANOVA P -value = 0.2721). Nevertheless, the pattern remained the same: population N2 displayed the highest s ($s_{N2} = 8.85$), N1 presented an intermediate value ($s_{N1} = 1.33$) and N4 exhibited the lowest s ($s_{N4} = 0.31$).

Table 4 Covariance analysis table of the complete models tested for each mating system variables (A) m : migration rate, (B) $\log(Nep)$: logarithm of the number of effective fathers and (C) s : selfing rate. PL_{mean} represents the mean phenological lag and $|PL|$ the absolute phenological lag

Variable	d.f.	Main effect	Type III sum of square	F or t value	P
(A) m					
Population	2		7363	23.30	<0.0001
Diameter	1	0.7403	1256	7.79	0.0078
Number of clonal copies	1	-0.2119	26	0.16	NS*
Competitor density _{-10m}	1	-0.6210	1663	10.32	0.0025
$PL_{\text{mean}_{100m}}$	1	5.6459	1303	8.09	0.0067
$(PL_{\text{mean}_{100m}})^2$	1	3.0939	847	5.25	0.0268
$PL_{\text{mean}_{100m}} \times \text{Population}$	2		2032	6.31	0.0039
Error	43		6926		
(B) $\log(Nep)$					
Population	2		6.48	2.79	0.0722
Diameter	1	0.0179	0.80	0.69	NS
Number of clonal copies	1	-0.0667	2.06	1.78	NS
Basal Area _{20 m}	1	0.0995	3.39	2.93	0.0945
Competitor density _{10 m}	1	-0.0281	3.26	2.82	NS
$ PL _{100 m}$	1	0.1986	1.83	1.58	NS
$ PL _{100 m} \times \text{Population}$	2		2.23	0.96	NS
Error	42		48.57		
(C) s					
Population	2			2.23	0.0307
Diameter	1	-0.0871		-2.00	0.0514
Competitor density _{10 m}	1	0.0483		1.79	0.0795
Number of clonal copies	1	0.1225		2.32	0.0247
$ PL _{100 m}$	1	-2.1057		-2.08	0.0432
$ PL _{100 m} \times \text{Population}$	2			2.26	0.0285
Error	46				

*NS' term indicated nonsignificant variables at level 5%.

The mean selfing rates were only significantly different between populations N2 and N4.

Discussion

Plant population biologists have a long-standing interest in mating system, due to its major impact on genetic diversity, the effective population size and, ultimately, the response to selection and the persistence of populations (Loveless & Hamrick 1984). The originality of this study was that it examined the impact of several ecological factors on mating system parameters simultaneously within and between populations, thus providing a multi-scale investigation of mating system variation.

The impact of competitor density on mating system

Our results first revealed the key impact of competitor density on the variation of m_i and s_i within the studied *F. sylvatica* populations. High competitor densities in close neighbourhoods decreased the proportion of migrant pollen (m_i) and tended to increase the probability of self-fertilization (s_i). This could be due to a high canopy

density of nearby competitors, which would block the pollen flow originating from intermediate and long distances. This role of competitors as a barrier to pollen flow is supported by the opposite effects of mother-tree diameter observed on m_i and s_i . Indeed, larger mother trees are likely taller, exhibit a larger crown and display less competition with close neighbours that allow them capturing more pollen from intermediate and long distances. A positive correlation between mother-tree crown size and the outcrossing rate has been reported previously (de-Lucas *et al.* 2008), as has the role of total competitor density as a barrier to gene flow (Garcia *et al.* 2005; Milleron *et al.* 2012), but these effects are considerably less documented than the effect of conspecific density on mating system (e.g. El-Kassaby & Jaquish 1996; Franceschinelli & Bawa 2000; Ward *et al.* 2005). In the studied case, contrary to the classical expectation, we did not observe a decrease of s_i or an increase of Nep_i with an increasing conspecific density, which though showed a wide range of variation (from 2 to 130 conspecific trees in a 50-m radius). This is likely because density is not a limiting factor for reproduction of *F. sylvatica* trees in Mont Ventoux (at least 93 conspecific neighbours were observed within a 200-m

radius). However, the presence of large reproductive trees, expected to have high fecundity, tended to increase the Nep_i of the mother trees in our populations.

Between the populations, the migration rates decreased from the bottom (N1) to the top population (N4) (Fig. 2a). The effects of the competitor density and diameter on the variation in migration rates within population were not consistent overall with the variation in these ecological factors between populations. For instance, population N4 exhibited the highest densities (Table 1), but also presented the highest migration rate ($m_{N4} = 64.3\%$). Therefore, the pollen barrier was not the major factor affecting the mating system at the population level. Furthermore, the mother-tree diameters did not differ between the populations (P -value = 0.219), so this factor did not explain the considerable differences in m between the populations. However, at the population scale, the high distance of pollen dispersal observed for population N2 ($\delta_{N2} = 63$ m and $b_{N2} = 0.27$) may be explained by its low density (Table 1). This would illustrate the negative effect of canopy density on mean dispersal distances, which was previously demonstrated for *F. sylvatica* (Piotti *et al.* 2012; Milleron *et al.* 2012) and other tree species (Robledo-Arnuncio *et al.* 2004; Ward *et al.* 2005; Llorens *et al.* 2012).

The impact of phenology on mating system

Flowering lags are expected to induce temporal patterns of assortative mating, both within and between populations (Gerard *et al.* 2006), and to affect the variations in mating system (Robledo-Arnuncio *et al.* 2004; El-Kassaby & Jaquish 1996; Fuchs *et al.* 2003). However, to our knowledge, no experimental study had previously quantified the impact of the flowering lag on mating system, likely because of the difficulty of observing flowering stages in trees.

Within all of the populations, a phenological lag between a mother tree and its neighbours was found to increase the pollen migration rate, as expected under the mass-action law (due to the decreased amount of local pollen). Additionally, in population N2, late-flowering mother trees were shown to receive less migrant pollen than early-flowering mother trees. The female flowers of *F. sylvatica* are receptive before pollen release (protogyny) (Merzeau *et al.* 1994). Therefore, at a given elevational level, the female flowers of early-flowering trees are receptive when little local pollen is available, and their outcrossed pollen pool should include mostly migrant pollen from populations at lower altitudes. This effect of protogyny was not observed in populations N1 and N4, where the symmetric axes were almost centred on zero. This finding was expected in N1, for which there are no pollen donors below, but not in N4 that is among the last

populations to flower along the gradient. In population N4, the bud burst dynamics were mostly synchronized among individuals (Fig. 3), likely due to the rapid succession of the phenological stages; this could have prevented the development of a clear pattern of variation between m_i and the phenological lag in this population.

Between populations, we observed an increased migration rate from the bottom to the top of the elevational gradient, consistent with the expected effect of protogyny combined with assortative mating due to the elevational variation of temperature (i.e. the hypothesis H1 detailed in the introduction). Protogyny can generally be expected to induce directional pollen flow from the bottom to the top of elevational or latitudinal gradients. It was beyond the scope of this study to fully validate the consequences of protogyny on directional pollen flow, but this could be achieved by surveying the phenology of pollen emission and reception (pollen traps), or by assigning migrant pollen to its population of origin (Davis *et al.* 2004; Nishizawa *et al.* 2005). Interestingly, our multiscale analysis of mating patterns highlighted opposite effects of phenology within and between populations. Indeed, within populations, the phenological lag increased the proportion of immigrant pollen in the pollen cloud of early and synchronized mother trees, while between populations, migration rate was lower at the bottom (the earlier populations) as compared with the top of the gradient (the later population).

We made several hypotheses regarding flowering phenology. First, we assumed that the lags in bud burst phenology between individuals are a good proxy for the lags in reproductive phenology. This relationship is likely to hold true in *F. sylvatica*, where pollen emission has been shown to overlap with the last phases of leaf unfolding (personal observations, Milleron *et al.* 2012). It would be interesting in future studies to analyse this relationship at the individual level. The second assumption was that the comparison of phenology effect between populations is not affected by the ways we measured phenology in the different populations. Within populations, the effects of phenological lag are reliable because phenology was measured similarly for all of the individuals. To deal with the different phenology measures at the population level, we centred and normalized the phenological variable for each population. Unfortunately, this prevented us from testing the absolute phenology effect on mating parameters across the three populations simultaneously.

Advantages of individual-based, multiscale mating system analysis

Using the Bayesian framework proposed by Klein *et al.* (2008), we extended the SEMM to estimate individual

mating system parameters. To our knowledge, this is the first parentage-based modelling approach that provides individual estimates of the m and s parameters. Previous SEMMs assumed independent fecundation events and equal m and s parameters for all of the mother trees within populations. The high within-population variability of the individual mating system estimates observed here showed that this assumption was wrong. Several other studies (de-Lucas *et al.* 2008; Tamaki *et al.* 2009; Garcia *et al.* 2005; Oddou-Muratorio *et al.* 2006; Restoux *et al.* 2009) have suggested that such high interindividual variability in mating system parameters is the rule rather than the exception in trees. In addition, our study highlights some statistical consequences of the departure from the assumption of equal mating system parameters within populations. In particular, not accounting for within-population variation can lead to underestimating the width of the confidence intervals around mating system parameters, thereby overestimating the significance of the differences between populations. For instance, in our study, comparisons based on the initial MEMM concluded that both the migration and the selfing rates differed significantly between the studied populations (Table 3). However, ANOVA based on MEMMi, which accounted for individual variability, showed that only the migration rates, and not the selfing rates, differed among populations.

The MLTR (Ritland 2002) is commonly used to estimate individual selfing rates (Garcia *et al.* 2005; de-Lucas *et al.* 2008; Silva *et al.* 2011), although it does not account for the spatial information. A comparison of the s_i -estimates obtained by MEMMi and MLTR from our data set showed that these two methods lead to divergent estimates at the individual level (e.g. s_i MEMMi = 24% and s_i MLTR = 56% for the mother tree F20; Table S4, Supporting information), but to similar estimates of average selfing rate at population level (Table S4, Supporting information). This divergence is because MEMMi modifies the individual estimates of selfing rate to account for the actual maternal neighbourhood of each mother tree, that is, for the distances of compatible male trees, the individual fecundities of these trees and the pollen dispersal abilities.

Finally, our study shed light on the use of ANCOVA to better connect the ecological determinants of mating system acting at the individual and population levels. The ANCOVA models for the individual values of the migration rates, the effective number of pollinators and selfing rates indicated that they varied between populations, through a marginal or significant effect of the 'population' term on the variation of m_i (P -value < 0001), Nep_i (P -value = 0.0722) and s_i (P -value = 0.031; Table 4). Rather than erasing the population effect,

accounting for the local ecological determinants revealed the effect of the population on Nep_i and s_i , and changed the pattern of variation between populations for m_i . The variation of a mating system variable (e.g. migration rate) as function of an environmental covariate (e.g. competitor density) and of a factor (population) is notoriously difficult to disentangle when the average values of the covariate differ between populations. ANCOVA appears here as a suitable framework to reveal both the effect of populations and the major effect of the covariate that both could be hidden due to the different average covariate values between populations (Fig. 5).

Elevational variation in mating system

As compared to the top population N4, the bottom populations (N1 and N2) presented a lower migration rates and a higher mean s values, and tended to have lower mean Nep . These differences could result from the effect of protogyny combined with assortative mating (as detailed above), and also from the negative effect of climate on pollen viability and germination in bottom populations (hypothesis H2 in the introduction). Directional pollen flow carrying preadapted genes to warmer conditions should thus be particularly favourable for enhancing the adaptation of population N4 to future climatic conditions (Kremer *et al.* 2012). This would be interesting in future studies to test whether populations N1–N4 are locally adapted to the drought gradient, and whether directional pollen flow actually contribute to microevolution of traits involved in response to drought.

Because we investigated only three populations along the elevational gradient, it was beyond the scope of this study to reveal causal variation linked to elevational variations of the climatic and ecological conditions. In addition, the lack of replication prevented us from distinguishing the impact of climate from other ecological factors acting at the population scale, such as the effect of different pollen donor fecundities, the existence of wind-oriented pollen flow among populations or the overlap of phenological windows between populations along the gradient. Still, we excluded the possibility that the population effect reflected the spatial configuration of the populations (i.e. size and density) because the inclusion of variables accounting for the distance to the edge of the population, the distance to the closest neighbour or the conspecific density at a larger scale did not significantly explain variability in the mating system parameters (results not shown).

The mating system variations across the elevational or latitudinal ranges of plant species is currently a very attractive field of research (Michalski & Durka 2007; Mimura & Aitken 2007; Fuchs & Hamrick 2011; Llorens *et al.* 2012). While many range-wide studies have high-

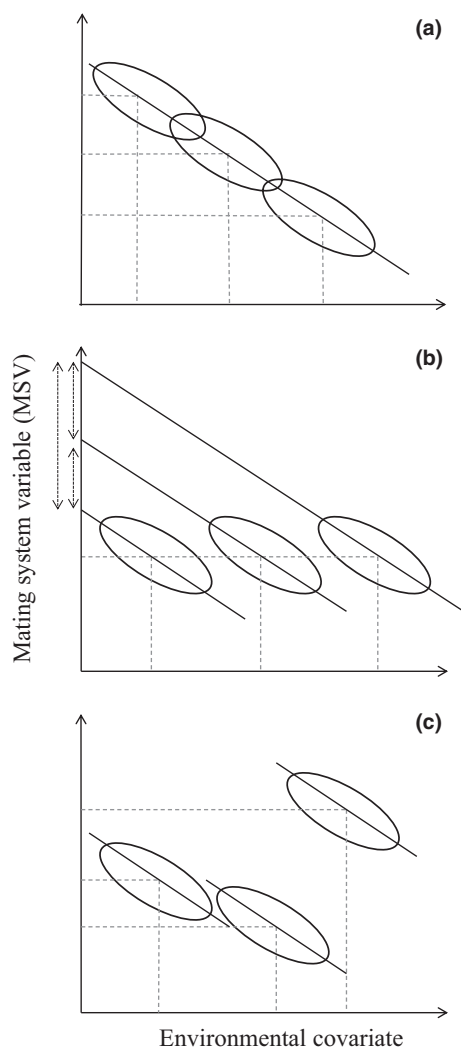


Fig. 5 Schematic representations of the variation of a mating system variable (MSV) in relation to one environmental covariate (x -axis) and a population factor (one oval per population) when the covariate averages differ between populations. The patterns of variation within and between populations may either be consistent (a) or opposite (c). Apart from the simple case where the covariate alone can explain the variation in the MSV (a), both the covariate and the factor effects significantly affect the MSV. When the covariate averages differ between populations in a way where the population effect is counterbalanced by the covariate effect among populations (b), both the factor and the covariate effects tested alone would not be significant. By contrast, the ANCOVA would correctly retrieve the population effects indicated by the dashed arrows. For other type of population effects (c), estimating the effect of covariate alone would lead to an opposite direction of its effect, and population effect test alone (ANOVA) would rank the populations differently compared with ANCOVA. This case corresponds to the relation between density and migration rate in the three studied populations.

lighted the role of reduced mate density and geographical isolation on mating system at range margin, this local study highlighted the role of other ecological factors.

Mate density was not a limiting factor for reproduction of *F. sylvatica* trees in Mont Ventoux, likely because all the studied sites, N1–N4, belong to a single continuous population and that none of this site can be considered as isolated regarding the scale of pollen dispersal. In addition, we showed that the effects of ecological factors were also not the same within vs. between populations across the gradient. These results highlight the interest of multiscale approaches, simultaneously investigating mating system and at least one of its ecological drivers at the individual and population levels. The issue of how to optimize the sampling design (the number of trees per population vs. the number of populations) is still open. However, if the potential ecological drivers of mating system between populations are investigated, mother trees should be sampled over a wide range of variation for these drivers within populations. Finally, the individual variability of the ecological maternal neighbourhood should be accounted for in the analysis. Otherwise individual variability could hide population effects (as for s in this study). Moreover, the interpretation of the impact of ecological factors at the population level could be inverted (Fig. 5).

Acknowledgements

We are grateful to C. Garcia, F. Austerlitz, H. Davi, B. Fady, V. Sork and three anonymous reviewers for comments on the previous versions of this manuscript, and to K. Csillery and C. Lavigne for the discussions on statistical models. We are indebted to A. Bontemps, who collected and compiled the phenological data and adult tree genotypes during her PhD thesis. We thank N. Turion, O. Gilg, F. Rei and F. Jean for field work (mapping, sample and measurements), and M. Correade for assistance with GIS data. We thank A. Roig and M. Lingrand for genotyping and managing the laboratory work and the employees of the nursery of Aix-Les-Milles for raising the seedlings. This study was funded by the EC-supported Network of Excellence Evoltree (GOCE-016322, BEECH initiative), the ERA-Net BiodivERsA LINKTREE project (ANR-08-Biodiversa-006-06), part of the 2009 BiodivERsA call for research proposals and the EMILE project (ANR-09-BLAN-01).

References

- Asuka Y, Tani N, Tsumura Y, Tomaru N (2004) Development and characterization of microsatellite markers for *Fagus crenata* blume. *Molecular Ecology Notes*, **4**, 101–103.
- Beck-Pay SL (2012) The effect of temperature and relative humidity on *Acacia mearnsii* polyad viability and pollen tube development. *South African Journal of Botany*, **83**, 165–171.
- Belmonte J, Alarcón M, Avila A, Scialabba E, Pino D (2008) Long-range transport of beech (*Fagus sylvatica* L.) pollen to Catalonia (north-eastern Spain). *International Journal of Biometeorology*, **52**, 675–687.
- Burczyk J, Adams WT, Moran GF, Griffin AR (2002) Complex patterns of mating revealed in a eucalyptus regnans seed

- orchard using allozyme markers and the neighbourhood model. *Molecular Ecology*, **11**, 2379–2391.
- Davi H, Gillmann M, Ibanez T *et al.* (2011) Diversity of leaf unfolding dynamics among tree species: New insights from a study along an altitudinal gradient. *Agricultural and Forest Meteorology*, **151**, 1504–1513.
- Davis HG, Taylor CM, Lambrinos JG, Strong DR (2004) Pollen limitation causes an Allee effect in a wind-pollinated invasive grass (*Spartina alterniflora*). *Proceedings of the National Academy of Sciences, USA*, **101**, 13804–13807.
- El-Kassaby YA, Jaquish B (1996) Population density and mating pattern in western larch. *Journal of Heredity*, **87**, 438–443.
- Franceschinelli EV, Bawa KS (2000) The effect of ecological factors on the mating system of a South American shrub species (*Helicteres brevispira*). *Heredity*, **84**, 116–123.
- Fuchs EJ, Lobo JA, Quesada M (2003) Effects of forest fragmentation and flowering phenology on the reproductive success and mating patterns of the tropical dry forest tree *Pachira quinata*. *Conservation Biology*, **17**, 149–157.
- Fuchs EJ, Hamrick JL (2011) Mating system and pollen flow between remnant populations of the endangered tropical tree, *Guaiaecum sanctum* (Zygophyllaceae). *Conservation Genetics*, **12**, 175–185.
- García C, Arroyo JM, Godoy JA, Jordano P (2005) Mating patterns, pollen dispersal, and the ecological maternal neighbourhood in a *Prunus mahaleb* L. population. *Molecular Ecology*, **14**, 1821–1830.
- Gerard PR, Klein EK, Austerlitz F, Fernandez-Manjarres JF, Frascaria-Lacoste N (2006) Assortative mating and differential male mating success in an ash hybrid zone population. *BMC Evolutionary Biology*, **6**, 96.
- Hardy OJ, González-Martínez SC, Colas B, Fréville H, Mignot A, Olivieri I (2004) Fine-scale genetic structure and gene dispersal in *Centaurea corymbosa* (Asteraceae). II. Correlated paternity within and among sibships. *Genetics*, **168**, 1601–1614.
- Holsinger KE (1991) Mass-action models of plant mating systems: The evolutionary stability of mixed mating systems. *The American Naturalist*, **138**, 606–622.
- Kang KS, Bila AD, Harju AM, Lindgren D (2003) Estimation of fertility variation in forest tree population. *Forestry*, **76**, 329–344.
- Klein EK, Lavigne C, Gouyon PH (2006) Mixing of propagules from discrete sources at long distance: comparing a dispersal tail to an exponential. *BMC Ecology*, **6**, 3.
- Klein EK, Desassis N, Oddou-Muratorio S (2008) Pollen flow in the wildservice tree, *Sorbus torminalis* (L.) crantz. IV. Whole interindividual variance of male fecundity estimated jointly with the dispersal kernel. *Molecular Ecology*, **17**, 3323–3336.
- Kouno K, Mukouda M (1985) Flowering and seed-setting traits of three broadleaf trees, *Fagus Crenata*, *Cornus controreras* and *Aeculus turbinara*. *Bulletin Tohoku For Tree Breeding Center*, **25**, 74–76.
- Kremer A, Ronce O, Robledo-Arnuncio JJ *et al.* (2012) Long-distance gene flow and adaptation of forest trees to rapid climate change. *Ecology Letters*, **15**, 378–392.
- Lander TA, Oddou-Muratorio S, Prouillet-Leplat H, Klein EK (2011) Reconstruction of a beech population bottleneck using archival demographic information and bayesian analysis of genetic data. *Molecular Ecology*, **20**, 5182–5196.
- Lefèvre S, Wagner S, Petit RJ, de Lafontaine G (2012) Multiplexed microsatellite markers for genetic studies of beech. *Molecular Ecology Resources*, **12**, 484–491.
- Llorens TM, Byrne M, Yates CJ, Nistelberger HM, Coates DJ (2012) Evaluating the influence of different aspects of habitat fragmentation on mating patterns and pollen dispersal in the bird-pollinated *Banksia sphaerocarpa* var. *caesia*. *Molecular Ecology*, **21**, 314–328.
- Loveless MD, Hamrick JL (1984) Ecological determinants of genetic structure in plant populations. *Annual Review of Ecology and Systematics*, **15**, 65–95.
- de-Lucas AI, Robledo-Arnuncio JJ, Hidalgo E, González-Martínez SC (2008) Mating system and pollen gene flow in mediterranean maritime pine. *Heredity*, **100**, 390–399.
- Merzeau D, Comps B, Thiébaud B, Letouzey J (1994) Estimation of *Fagus sylvatica* mating system parameters in natural populations. *Annals of Forest Science*, **51**, 163–173.
- Michalski SG, Durka W (2007) High selfing and high inbreeding depression in peripheral populations of *Juncus atratus*. *Molecular Ecology*, **16**, 4715–4727.
- Milleron M, Lopez de Heredia U, Lorenzo Z *et al.* (2012) Effect of canopy closure on pollen dispersal in a wind-pollinated species (*Fagus sylvatica* L.). *Plant Ecology*, **213**, 1715–1728.
- Mimura M, Aitken SN (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.
- Nishizawa T, Watano Y, Kinoshita E, Kawahara T, Ueda K (2005) Pollen movement in a natural population of *Arisaema serratum* (Araceae), a plant with a pitfall-trap flower pollination system. *American Journal of Botany*, **92**, 1114–1123.
- Oddou-Muratorio S, Klein EK, Austerlitz F (2005) Pollen flow in the wildservice 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, Klein EK, Demesure-Musch B, Austerlitz F (2006) Real-time patterns of pollen flow in the wild-service tree, *Sorbus torminalis* (Rosaceae). III. Mating patterns and the ecological maternal neighborhood. *American Journal of Botany*, **93**, 1650–1659.
- Pastorelli R, Smulders MJM, Vanot Westende WPC *et al.* (2003) Characterization of microsatellite markers in *Fagus sylvatica* L. and *Fagus orientalis* Lipsky. *Molecular Ecology Notes*, **3**, 76–78.
- Piotti A, Leonardi S, Buiteveld J *et al.* (2012) Comparison of pollen gene flow among four European beech (*Fagus sylvatica* L.) populations characterized by different management regimes. *Heredity*, **108**, 322–331.
- Restoux G, Silva DE, Sagnard F, Torre F, Klein EK, Fady B (2009) Life at the margin: the mating system of Mediterranean conifers. *Web Ecology*, **8**, 94–102.
- Ritland K (2002) Extensions of models for the estimation of mating systems using n independent loci. *Heredity*, **88**, 221–228.
- Robledo-Arnuncio JJ, Alia R, Gil L (2004) Increased selfing and correlated paternity in a small population of a predominantly outcrossing conifer, *Pinus sylvestris*. *Molecular Ecology*, **13**, 2567–2577.

- Robledo-Arnuncio JJ, Austerlitz F, Smouse PE (2007) Poldisp: a software package for indirect estimation of contemporary pollen dispersal. *Molecular Ecology Notes*, **7**, 763–766.
- Silva CRS, Albuquerque PSB, Ervedosa FR, Mota JWS, Figueira A, Sebbenn AM (2011) Understanding the genetic diversity, spatial genetic structure and mating system at the hierarchical levels of fruits and individuals of a continuous *Theobroma cacao* population from the Brazilian Amazon. *Heredity*, **106**, 973–985.
- Smouse PE, Dyer RJ, Westfall RD, Sork VL (2001) Two-generation analysis of pollen flow across a landscape. I. Male gamete heterogeneity among females. *Evolution*, **55**, 260–271.
- Tamaki I, Setsuko S, Tomaru N (2009) Estimation of outcrossing rates at hierarchical levels of fruits, individuals, populations and species in *Magnolia stellata*. *Heredity*, **102**, 381–388.
- Vekemans X, Hardy O (2004) New insights from fine-scale spatial genetic structure analyses in plant populations. *Molecular Ecology*, **13**, 921–935.
- Vornam B, Decarli N, Gailing O (2004) Spatial distribution of genetic variation in a natural beech stand (*Fagus sylvatica* L.) based on microsatellite markers. *Conservation Genetics*, **5**, 561–570.
- Wang KS (2003) Mating system in isolated stands of European beech (*Fagus sylvatica* L.). *Forest Genetics*, **10**, 159–164.
- Ward M, Dick CW, Gribel R, Lowe AJ (2005) To self... a review of outcrossing and pollen-mediated gene flow in neotropical trees. *Heredity*, **95**, 246–254.
- Yacine A, Bouras F (1997) Self- and cross-pollination effects on pollen tube growth and seed set in holm oak *Quercus ilex* L (Fagaceae). *Annals of Forest Science*, **54**, 447–462.

J.G. compiled the genetic and ecological data, performed the analyses, wrote the first draft of the text and contributed the the revisions. E.K.K. modified the MEMM code with J.G., and supervised the statistical

analyses. S.O.M. conceived and coordinated the beech experiment. E.K.K. and S.O.M. contributed to the analyses and the revisions of preliminary drafts.

Data accessibility

Sampling locations, microsatellite genotypes and ecological data: Dryad: doi:10.5061/dryad.j1917

MEMMi code: <http://ciam.inra.fr/biosp/node/120>

Supporting information

Additional supporting information may be found in the online version of this article.

Fig. S1 Directed acyclic graph presenting the Bayesian framework used in the new spatially explicit mating model (MEMMi).

Fig. S2 The phenological dynamics of population N1 (red lines) and N4 (blue lines) from 3 March to 11 May 2009.

Table S1 Information for the microsatellite markers used in the study.

Table S2 Number of alleles by locus (k) and the observed (H_{Obs}) and expected polymorphism (H_{Exp}).

Table S3 The selfing rate (s_i), migration rate (m_i) and effective number of pollen donors (Nep_i) estimated per mother tree using MEMMi and KINDIST software.

Table S4 Comparison between the selfing rate (s_i) estimates using MLTR and MEMMi softwares.

Appendix

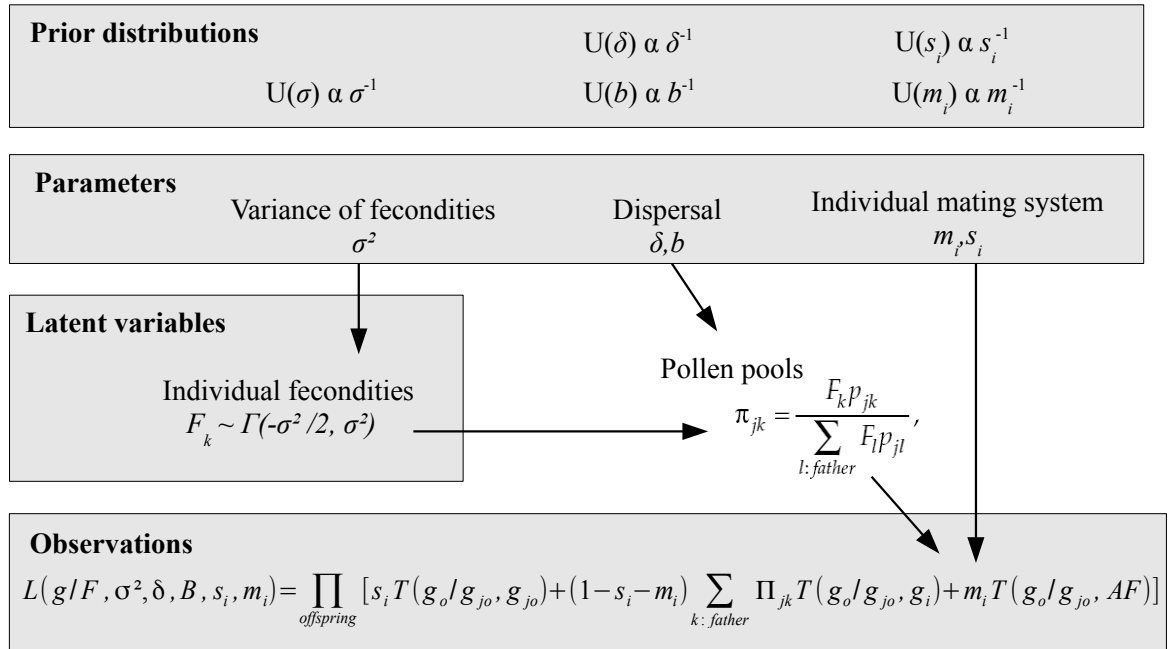


Figure S1 - Directed acyclic graph presenting the Bayesian framework used in the new spatially explicit mating model (MEMMi).

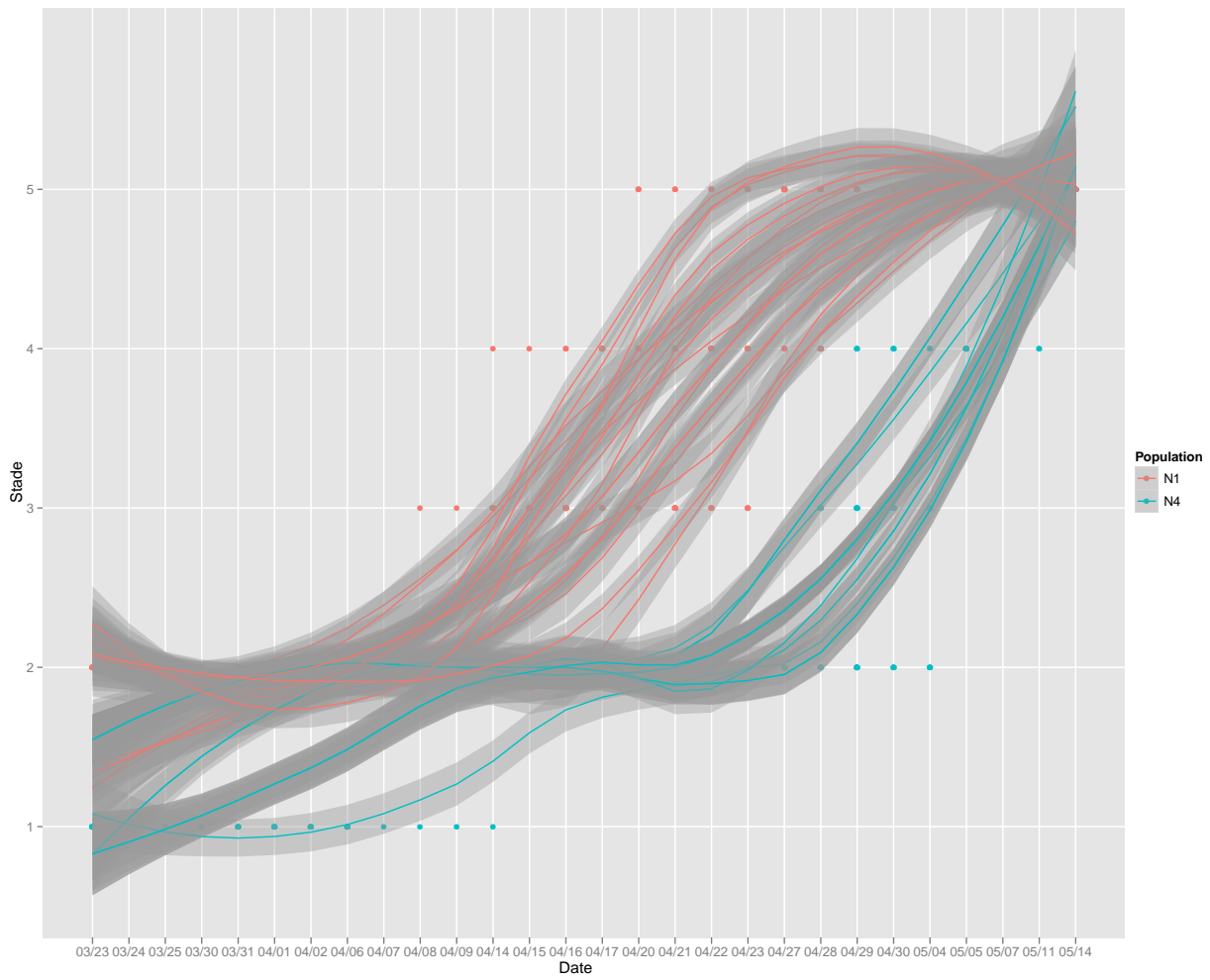


Figure S2 - The phenological dynamics of population N1 (red lines) and N4 (blue lines) from 3 March to 11 May 2009.

ESTIMATION DE L'HÉRITABILITÉ À PARTIR DE FAMILLES ISSUES DE POLLINISATION LIBRE



Synthèse

Ce troisième chapitre de thèse aborde la question de l'optimisation de l'estimation des variances génétiques additives (V_A) et des héritabilités (h^2) des traits quantitatifs à partir de familles issues de pollinisation libre chez les arbres. Plus globalement, les analyses de génétique quantitative chez les populations naturelles de plantes et d'animaux posent la question du degré de précision sur l'information des apparentements génétiques que requièrent ces analyses pour estimer correctement V_A et h^2 (Charmantier & Réale, 2005). Au delà de la quantité idéale d'information, se pose aussi le problème de l'information qu'il est concrètement possible de retracer dans ces systèmes d'étude. Pour beaucoup d'espèces, comme les mammifères, oiseaux ou plantes, on peut généralement facilement reconstituer des fratries maternelles, ce qui a conduit à utiliser des dispositifs dits de "familles" chez ces espèces. En revanche, les autres types de parentés ne peuvent généralement être retracés de manière fiable que par l'utilisation des marqueurs moléculaires. Ainsi, depuis une dizaine d'années maintenant, l'intégration de l'information sur les marqueurs génétiques dans les modèles de génétique quantitative est devenue un véritable enjeu pour améliorer les analyses en populations naturelles (Garant & Kruuk, 2005). De ces études, un consensus fort qui ressort est que les marqueurs moléculaires, lorsqu'ils sont en faible nombre, ne doivent être utilisés que pour reconstruire les relations de parenté par des approches catégoriques (e.g., reconstruction de pédigrées, assignations de paternité), et ne permettent pas d'estimer des coefficients d'apparentement suffisamment correctement (Thomas *et al.*, 2002; Csillery *et al.*, 2006; Gay *et al.*, 2013).

Chez les arbres, l'estimation de l'héritabilité des traits quantitatifs passe souvent par l'analyse de familles issues de pollinisation libre récoltées sur des arbre-mères en populations naturelles et placées en environnement contrôlé dans des dispositifs dits de jardin commun ou ex situ. Dans ces familles, les relations d'apparentement par voie paternelle et les relations multi-générationnelles sont généralement inconnues et négligées par les modèles de génétique quantitative appelés modèles "famille".

Dans la **section 3.1** de ce chapitre, une revue bibliographique en collaboration avec François Lefèvre et Christian Pichot, chercheurs à l'URFM, a été réalisée afin de comprendre quelles étaient les caractéristiques du régime de reproduction des arbres qui pouvaient générer des écarts aux hypothèses simplificatrices faites pour analyser les familles issues de

pollinisation libre. Cette revue a également permis de lister un certain nombre de méthodes permettant de s'affranchir des hypothèses des modèles "famille", souvent basées sur le modèle "animal", modèle défini à l'échelle individuelle permettant de prendre en compte des relations d'apparentements complexes, généralement sous forme de pédigrée. Ce travail a donné lieu à la publication d'un article de revue en 2013 dans le journal *Acta Botanica Gallica*.

Un constat issu de cette revue bibliographique est que peu d'études ont cherché à développer des méthodologies adaptées à l'analyse de dispositifs de grande taille généralement retrouvés chez les arbres (mais voir El-Kassaby *et al.* 2011). Notamment, ces dispositifs ont la spécificité qu'il est très rare de pouvoir génotyper tous les individus suivis et phénotypés en jardin commun.

La section **section 3.2** de ce chapitre présente une approche de simulation ayant pour but de comparer les différentes méthodes d'estimation de l'héritabilité précédemment caractérisées par la revue de la bibliographie en présence de plusieurs écarts aux hypothèses familles (e.g., déséquilibre des contributions paternelles, effets maternels). Une collaboration avec Laurène Gay, chercheuse à l'INRA de Montpellier, m'a permis de complexifier un modèle de simulation de généalogies (Gay *et al.*, 2013) que j'ai ensuite utilisé pour simuler des descendance maternelles issues de l'évolution d'une population adulte en présence d'un déséquilibre des contributions à la reproduction. Dans cette étude, j'ai notamment cherché à déterminer quelles méthodes statistiques permettaient d'utiliser au mieux une information moléculaire incomplète. Le manuscrit présenté dans cette thèse a été soumis au journal *Methods in Ecology and Evolution* en début d'année 2014 et est aujourd'hui en révision pour une resoumission dans un autre journal.

Les résultats du chapitre 2 sur le régime de reproduction des populations de hêtre commun sur le Mont-Ventoux soulignaient une forte variabilité inter-individuelle des composantes du régime de reproduction, et notamment une très forte variabilité des taux d'autofécondation ($0 < s < 48$ %) et du nombre efficace de pères ($2 < N_{ep} < 364$) entre arbre-mères. Ces caractéristiques pourraient fortement affecter les hypothèses du modèle "famille" et conduire à des estimations biaisées de V_A et h^2 .

Dans la section **section 3.3** de ce chapitre, j'ai finalement appliqué les différentes mé-

thodes testées dans la section 3.2 au dispositif de descendance maternelles étudié dans cette thèse, afin de mesurer les écarts aux hypothèses familles générés par le régime de reproduction du hêtre en populations naturelles et d'estimer h^2 pour les différents traits fonctionnels mesurés en conditions contrôlées.

Il ressort principalement de ces résultats que le déséquilibre des contributions paternelles à la reproduction observé au sein des populations étudiées affecte peu les hypothèses du modèle famille. En effet, ce modèle "famille" fournit des estimations de h^2 similaires aux modèles plus sophistiqués prenant en compte l'information sur les apparentements paternels disponibles (modèle animal), sauf dans le cas où les traits présentent des effets maternels (ce qui concerne 3 traits sur les 12 mesurés). Globalement, tous les traits présentent une base génétique significative, ce qui signifie qu'ils peuvent vraisemblablement évoluer sous effet de la sélection et des flux de gènes.

Ce chapitre offre une vision large des différentes analyses réalisables chez des espèces pérennes non-modèles à l'heure actuelle pour estimer de manière fiable le potentiel évolutif des traits quantitatifs. Cependant, avec le développement rapide des données génomiques et la possibilité de marquer de manière dense le génome, de nouvelles questions méthodologiques sont aujourd'hui abordées en génétique quantitative ; l'utilisation des marqueurs moléculaires pour estimer des coefficients d'apparentement sur marqueurs est maintenant envisageable (Visscher *et al.*, 2006) même si pour l'instant il existe peu d'espèces d'arbres pour lesquelles des jeux de marqueurs génomiques (SNPs) conséquents sont disponibles pour un génotypage en routine. Le problème des dispositifs d'arbres reste cependant leur taille et le coût important que constitue le phénotypage et génotypage de l'ensemble des individus récoltés. Le développement des données génomiques chez ces espèces pourrait fortement inciter les études empiriques à optimiser leur échantillonnage en terme de nombre d'individus suivis par famille et nombre de familles suivies par population.

3.1 Biais dans les analyses de génétique quantitative utilisant des familles issues de pollinisation libre provenant de populations naturelles d'arbres



Société botanique de France

Biases in quantitative genetic analyses using open-pollinated progeny tests from natural tree populations

Julie Gauzere^{a,b}, Sylvie Oddou-Muratorio^a, Christian Pichot^a, François Lefèvre^a and Etienne Klein^{b*}

^aINRA, UR629 Ecologie des Forêts Méditerranéennes (URFM), F-84914 Avignon, France; ^bINRA, UR546 Biostatistique et Processus Spatiaux (BioSP), F-84914 Avignon, France

Abstract: In plant quantitative genetic studies conducted *ex situ*, the large number of seeds produced per individual has promoted the use of open-pollinated progeny tests. In subsequent analyses, seeds collected on the same mother-plant are assumed to be half-sibs. The consequences of the departure from half-sib assumption in progeny tests have been investigated since the 1960s using simulation approaches and, more recently, using molecular-based experimental approaches. This review aims to synthesize the results and conclusions of these simulation and empirical studies. We focus on tree species, where controlled crosses are difficult to carry out experimentally and departures from half-sib assumptions occur frequently in natural populations. First, the average level of relatedness expected within maternal progeny for many tree populations is higher than that of half-sibs. This is the consequence of non-random mating resulting from the small number of effective pollen donors per female, unequal male reproductive success and/or selfing. As result, estimates of genetic variance and heritability for quantitative traits may be upward biased. Alternatively, inbreeding depression, dominance effects and the heterogeneity of the male gamete pool among females are often neglected, which may lead to underestimation of the heritability of traits. A correction based on the mean genetic relatedness between offspring and the relatedness between parents is often used to compensate those biases. However, such correction cannot accurately adjust the estimates in situations where variable levels of genetic relatedness among families, dominance effects or inbreeding depression exist within the progeny. An alternative and promising approach is the use of the “animal model” approach, which optimizes the use of molecular data and paternal information to estimate heritability more accurately.

Keywords: additive genetic variance; animal model; common garden; dominance; family model; half-sib family; heritability; inbreeding; selfing; tree

Introduction

Understanding how much of phenotypic variation in natural populations is shaped by genetic or environmental variation and whether the genetic component of phenotypic variation is driven by selection or demography are two major goals of evolutionary quantitative genetics (Kingsolver et al. 2001; Charmantier and Garant 2005). Although quantitative genetics was initially developed in a breeding perspective, its theoretical and experimental tools are now commonly used by evolutionary geneticists to answer their own questions. For example, many studies in wild species have investigated the additive genetic variance (V_A) and the heritability (h^2) of adaptive traits (see Kruuk 2004), the genetic correlations among traits (Etterson and Shaw 2001), and the genetic differentiation

of traits among populations (Ducousso, Guyon and Kremer 1995). The estimation of these quantitative genetic parameters has gained renewed interest to predict the rate of genetic evolution in changing environments (Kingsolver et al. 2001; Conner 2001). The estimation of variance components is based on the knowledge of the level of genetic relatedness (measured by the relatedness coefficient ρ) among the individuals for which phenotypic characters were measured. Theoretical genetic relatedness can be partially or completely known a priori (controlled crosses, maternal families...) and realized genetic relatedness can be estimated a posteriori using molecular markers.

In plants, open-pollinated progeny (OP) designs are widely used to characterize the adaptive potential of populations, because they are easy to perform and V_A

*Corresponding author. Email: etienne.klein@avignon.inra.fr

This study is part of Julie Gauzere PhD thesis, which aims at understanding the role of adaptation and gene flow in the phenotypic variation of traits related to climate response along an altitudinal gradient of a temperate tree species, *Fagus sylvatica*. Sylvie Oddou-Muratorio is a population geneticist studying the impact of pollen and seed dispersal and climate-driven selection on adaptive dynamics of forest trees natural populations under climate change. François Lefèvre's research is targeted at the issues of sustainable management of forest genetic resources and adaptation of the forests to climate change. Christian Pichot is a population geneticist and quantitative geneticist studying the role of forest tree genetic resources on population dynamics. Etienne Klein research is targeted on dispersal of plant propagules (pollen and seeds), and their impact on the genetic composition of populations.

and h^2 can be easily estimated. Offspring within families, collected directly on the mother-plants, are considered to be half-sibs and related to each other with a relatedness coefficient $\rho = 0.25$. In these half-sibs family trials the additive genetic variance is equal to the variance among families (V_F , estimated in a controlled environment) multiplied by the coefficient $\alpha = 1/\rho = 4$ (thus $V_A = 4 \times V_F$).

The analysis of OP designs therefore relies on the main assumption that individuals within families are true half-sibs ($\rho = 0.25$), which in turn generates several other hypotheses on the mating system and the relatedness among mates: (1) the female plants from which the progenies are collected are unrelated, (2) mating occurs with a high number of unrelated males, which are also unrelated to the females, (3) no self-fertilization occurs, and (4) no assortative mating occurs (i.e. preferential mating between males and females). All of these conditions are usually not fulfilled in natural populations, and a first consequence is that the average genetic relatedness among offspring within families (ρ_{mean}) is greater than 0.25 and V_A and h^2 are likely to be overestimated. Additionally, departure from the assumptions listed above will also affect V_F , because the between-family variance component can be inflated by non-additive effects (dominance, epistasis) when families include some proportion of full-sibs. Alternatively, the within-family variance can be inflated by inbreeding depression when crosses occur among relatives or self-fertilization is non-negligible.

These issues are particularly crucial in trees, for which controlled crosses are difficult to manage in situ, and departures from half-sib assumptions frequently occur. As in most plant species, trees generally have mixed-mating systems (i.e. sexual reproduction occurs both through self-fertilisation and outcrossing; Robledo-Arnuncio, Alia and Gil 2004; Ward et al. 2005), and they also generally have a low number of effective pollen donors, unbalanced male reproductive success (Kang et al. 2003; Oddou-Muratorio, Klein and Austerlitz 2005), significant spatial genetic structure (Vekemans and Hardy 2004), distance-limited pollen dispersal (Ashley 2010), and assortative mating induced by flowering synchronization (Soularue and Kremer 2012). Another important source of bias in OP progenies is maternal effects, which increase the phenotypic similarity between maternally related half-sibs, leading to inflated genetic differences among families and, therefore, overestimated V_A and h^2 (Wang, Lechowicz and Potvin 1994). Here, maternal effects will not be treated in detail because they do not depend on mating system or population characteristics, but the possible ways to account for them in quantitative genetic models will be discussed. For tree species, OP designs are both used to investigate natural populations and breeding populations (e.g. in seed orchards).

The issue of the biases in the classical analyses of OP designs has been studied for 50 years, first by

analytical or simulation approaches and then by experimental approaches using highly variable molecular markers (e.g. microsatellites), which have been made available in the last decades for most species and allow patterns of relatedness to be characterized (Loiselle et al. 1995). Here, we propose a review of the main literature dealing with the consequences of departures from half-sib assumptions in tree populations, and investigate the solutions now offered by molecular markers. First, we introduce the two major statistical methods used to analyze OP designs; the family model and the “animal model”. Then, we detail the different departures from half-sib assumptions that could occur in natural populations, their consequences for V_A and h^2 estimates, and existing solutions to correct these biases. Finally, we sum up the principal biases and discuss the use of molecular markers and “animal models” to correct estimates from OP designs.

Statistical approaches for the analysis of OP progenies

Quantitative genetic models decompose the phenotypic variance (V_P) of a given population into genetic (V_G) and environmental (V_E) variance components (assuming no $G \times E$ interaction here):

$$V_P = V_G + V_E$$

The first statistical model developed to analyze OP progenies is the Family model (Falconer and Mackay 1996). It is a random effect linear model which decomposes the value of a trait Y into a mean component (μ), a random Family effect (F_j) containing a part of the genetic effect, and a random residual (e):

$$Y_{ij} = \mu + F_j + e_{ij}$$

In this linear model, the random effect of interest is the family effect and the principal output of the model is the variance among families (V_F). Under controlled conditions, environmental effects are not confounded with family effects, and V_F only results from genetic differences among families.

The variance among families estimated in the linear model is equal to the genetic covariance within the family, also equal to $\rho_{mean} \cdot V_A$:

$$V_F = \rho_{mean} \cdot V_A$$

and

$$h^2 = \frac{V_A}{V_P} = \frac{1/\rho_{mean} V_F}{V_F + V_W} \quad (1)$$

where, V_W is the within-family variance, i.e. the residual variance. In OP designs, the environmental effects and the family structure are controlled for or supposed to be known. As a result, V_A and h^2 estimates can easily be calculated, based on the principles that family variance only includes additive genetic effects (no dominance or maternal effects for example) and that $\rho_{mean} = 0.25$ within families, i.e. they only include half-sibs.

Despite its simplicity and robustness, it is widely recognized that the Family model has limitations; notably it cannot deal with heterogeneous levels of relatedness in experimental designs. The “animal model” approach is increasingly used to overcome these limitations and provide quantitative genetic estimates for cases where the relatedness relationships among individuals are complex and unbalanced (see Kruuk 2004, for a review). Unknown model variables can be predicted using BLUP (Best Linear Unbiased Prediction; Henderson 1975) and variance components can be estimated using maximum likelihood or restricted maximum likelihood procedures. Alternatively, a Bayesian approach can be applied to estimate genetic parameters (Gianola and Fernando 1986; Pujol and Gallaud 2013). This Bayesian framework is particularly suitable for quantitative variables with non-Gaussian distributions or qualitative variables (de Villemereuil, Gimenez and Doligez 2013). The animal model can be written in a simple form as the decomposition of the phenotype y of the individual i :

$$Y_i = \mu + a_i + e_i,$$

where μ is the population mean, a_i is the additive genetic effects of the individual i and e_i is a random residual error. In its simplest writing the animal model is very similar to the Family model; the only difference being that the random effect of interest is the individual additive genetic value and not the family component. The more general form of the animal model is given by a matrix mixed model:

$$y = \beta + u + e \quad (2)$$

where β represents the vector of the fixed effects, e.g. population mean and environmental factors, and u represents the vector of random additive genetic effects and e is a vector of residual errors (Kruuk 2004). The random component, u , of the matrix model has a variance-covariance matrix G , given by $G = A.V_A$, where A is the additive genetic relationship matrix with individual elements $A_{ij} = \rho_{ij}$, and V_A is the additive genetic variance. Resolution methods of the animal model use the inverse of the A matrix (A^{-1}) to predict the breeding values (a_i) of the individuals, or to estimate the V_A . Classically, A^{-1} is directly derived from the pedigree, constructed from observations or paternity/parentage assignation methods. Indeed, skipping the

inversion step from A to A^{-1} has been shown to be more efficient for the analyses of large non-inbred populations (Henderson 1976). More recently, it has been proposed to construct the A matrix directly from pairwise relatedness molecular estimates; this strategy allows accounting for more complex relatedness relationships than the classical offspring-parent, half-sibs or full-sibs (El-Kassaby et al. 2011; Bush et al. 2011), but poses some challenges for the inversion of the A matrix.

Biological processes inducing departures from half-sib assumptions in OP designs and their consequences

Here we considered separately the main characteristics of tree population and mating system leading to departures from the main hypothesis of the classical Family model (Table 1), either by increasing ρ_{mean} (i.e. $\rho_{mean} > 0.25$), creating relatives among families (i.e. the mean genetic relatedness between families, $\rho_b > 0$), or conducting to the expression of dominance effect (i.e. $V_D > 0$) or inbreeding depression in the progenies (i.e. $ID > 0$). Our objective was to quantify the biases they induce on V_A and h^2 estimates.

Self-fertilization

The majority of tree species are monoecious or hermaphrodite, at least partly self-compatible, and the occurrence of self-fertilization events is one of the most frequent departures from half-sib assumptions in natural populations. Indeed, tree species are generally not completely outcrossed or selfed, but show intermediate levels of outcrossing and selfing (Barrett and Harder 1996; Ward et al. 2005), highly variable among species (Restoux et al. 2008). Selfing rates (s) also depend on environmental conditions, and so vary among populations and among individuals within populations (Robledo-Arnuncio, Alia and Gil 2004; El-Kassaby and Jaquish 1996; de-Lucas et al. 2008; Restoux et al. 2008). For example, de-Lucas et al. (2008) observed selfing rates that vary from 0 to about 80% among trees of three different populations of *Pinus radiata*, a predominantly outcrossed species. The literature showed that one of the most important biases on V_A and h^2 estimates results from neglected self-fertilization (Squillace 1974; Askew and El-Kassaby 1994; Borralho 1994; Bush et al. 2011) for reasons detailed below.

Presence of new kinds of relatives in progenies

Progenies collected from natural populations that experiment selfing and outcrossing possibly contain four kinds of relatives within families: half-sibs, self half-sibs, full-sibs and self full-sibs (Squillace 1974). These various relatives have a higher relatedness than half-sibs ($\rho_{full-sib} = 0.5$, $\rho_{self\ half-sib} = 0.408$ and $\rho_{self\ full-sib} = 0.667$; Squillace 1974), which explains why self-fertilization events greatly increase the mean

Table 1. Biological processes leading to departures from half-sib assumptions in natural populations (Yes/No), and their consequences on the genetic variances and the estimation of V_A and h^2 in open-pollinated families (inspired by table 1 in Sorensen and White 1988); with ρ_{mean} the mean genetic relatedness between maternal half-sibs and ρ_b the genetic relatedness between families.

Biological processes	Consequences in the quantitative model				References
	Expression of inbreeding depression	Expression of Dominance effects	$\rho_{mean} > 0.25$	$\rho_b > 0$	
Self-fertilization	Yes	Yes	Yes	No	Squillace 1974 Borrallho 1994 Bush 2011 Askew 1994
Low N_{ep}	No	Yes	Yes	Yes	Squillace 1974 Gaspar 2009 Surlles 1990
Unequal male fecundities	No	Yes	Yes	No	Hansen 2010
Males and females related	Yes	Yes	No	No	Bush 2011 Surlles 1990 Squillace 1974 Fujishima 1972
Females related	No	No	No	Yes	Fujishima 1972 Squillace 1974
Ancestral inbreeding in the parent population	No	No	Yes	No	Bush 2011

genetic relatedness among the offspring. This increase of ρ_{mean} , if neglected, induces an overestimation of V_A and h^2 (Squillace 1974; Surlles et al. 1990). The simulation study of Squillace (1974) showed that h^2 bias becomes non-negligible for traits with high heritability (above 0.20) and medium to high levels of selfing (above an average of 5%; see table 9 in Squillace 1974). For example, for a trait with a “true” $h^2 = 0.5$, a proportion of effective natural selfing of 10% leads to overestimate h^2 by 20% ($h^2_{HS} = 0.6$) when assuming that $\rho_{mean} = 0.25$. Most published experimental studies investigating the effect of selfing rate on V_A and h^2 estimates were generally based on OP progenies from seed orchards (Gaspar et al. 2009; Hansen and Nielsen 2010). They often observed negligible bias in h^2 estimates because of minor departures from half-sib assumptions [$s = 0\%$ and $\rho_{mean} = 0.26$ in Gaspar et al. (2009) in *Pinus pinaster* Ait progenies; $s = 4\%$ and $\rho_{mean} = 0.29$ in Hansen and Nielsen (2010) in *Abies nordmanniana* progenies]. In contrast, the study of Surlles et al. (1990), based on OP progenies from *Robinia pseudoacacia* natural populations, showed that h^2 was overestimated by 38% when assuming $\rho_{mean} = 0.25$, because of the mean selfing rate $s = 21\%$, which resulted in a ρ_{mean} value of 0.34 instead of 0.25. Hence, experimental studies showed contrasted results; OP progenies collected in breeding schemes were less subject to departures from the half-sib assumption and the resulting bias was lower than with progenies from natural populations.

To correct the impact of increased ρ_{mean} due to self-fertilization or other processes on V_A and h^2 , Squillace (1974) proposed to adjust the estimates according to the “real” ρ_{mean} among the offspring (noted r_{00} in the original publication):

$$h^2 = \frac{V_F}{\rho_{mean}(V_w + V_F)}$$

where V_F is the genetic variance among families and V_w is the genetic variance within families. ρ_{mean} can be estimated based on the knowledge of the mating system characteristics of the populations (Squillace 1974, proposed a guideline table which provides ρ_{mean} estimates according to variable mean selfing rates and mean numbers of effective males). Alternatively, ρ_{mean} can be estimated based on molecular data (e.g. Loiselle et al., 1995 or Ritland 1996). This correction has been largely used in experimental or simulation studies, and is denoted as the Squillace correction in the following.

Inbreeding and inbreeding depression in progenies

Self-fertilization generates inbred progenies, i.e. increases the probability that, at a random locus, the alleles within an individual are identical by descent. Biparental inbreeding, i.e. crosses between related individuals, also generates inbreeding in the offspring. The level of inbreeding is generally measured through the heterozygote deficiencies (F_{IS}) or estimated from the multilocus population structure, based for instance on the difference between the single- and multilocus estimates of outcrossing rates (Ritland 2002; David et al. 2007). Inbreeding can have deleterious effects on the offspring phenotype when copies of the same recessive deleterious mutation are found at homozygous state in an individual; a mechanism known as inbreeding depression. Inbreeding depression is expected to be important in tree natural populations (Petit and Hampe 2006). Inbreeding depression is classically defined and measured as the difference between the fitness of an inbred individual and the fitness of an individual from

random mating (Charlesworth and Charlesworth 1987). In experimental studies it is estimated by comparing the trait performances of inbred and non-inbred progenies from controlled crosses (Sorensen and White 1988; Griffin and Cotterill 1988) or by comparing inbred and non-inbred offspring identified by molecular markers within wild open-pollinated progenies, so avoiding confounding between progeny effects and inbreeding depression (Ferriol, Pichot and Lefèvre 2011).

Inbreeding depression increases the within-family variance and so acts to bias downwards the h^2 estimates (Borralho 1994). Using an adjusted average estimate of genetic relatedness in Douglas-fir, Sorensen and White (1988) showed that h^2 was underestimated because of the antagonistic effects between inbreeding depression and ρ_{mean} increase. They warned against the estimation biases that could occur while trying to just correct the h^2 estimates by the ρ_{mean} while not accounting for the existence of inbreeding depression in OP designs. In his simulation study, Borralho (1994) chose to treat the effect of inbreeding depression as a dominance effect (dominance effects were modelled as dependent on the inbreeding coefficient of the offspring; equation 3 in Borralho 1994). His results confirmed that in populations presenting non-negligible selfing rates, inbreeding depression and dominance effects, the Squillace correction (Squillace 1974) did not accurately adjust the h^2 estimates. Overall, inbreeding depression effects in selfed progenies have an underestimation effect that probably (partially) compensates for the overestimation due to the presence of relatives reported by Squillace (1974). However, quantitative investigations of the impact of inbreeding depression are missing, and threshold values involving compensatory effects are unknown.

Heterogeneity of selfing rates and inbreeding depression among families

The heterogeneity of mating systems among families also induces a bias in h^2 and V_A estimates that cannot be corrected using the Squillace approach (Squillace 1974). This was suggested by experimental studies in *Eucalyptus* species comparing OP and controlled-pollinated progenies, where the Squillace correction failed to adjust the bias in estimates based on OP progenies (Griffin and Cotterill 1988; Hodge et al. 1996). This failure was interpreted as resulting from the heterogeneity of selfing rate and inbreeding among families, combined with possible dominance effects. Moreover, Borralho (1994) demonstrated using simulations that the most important uncorrected biases, leading to an overestimation of h^2 , were found in cases where h^2 was low (< 0.25), dominance effects were high, and levels of selfing and inbreeding depressions were variable among families. Interestingly, simulations showed that this heterogeneity only induced uncorrected

bias when traits presented dominance effects (table 2 in Borralho 1974).

Finally, Hodge et al. (1996) confirmed that departure from half-sib assumptions was higher in OP from natural populations than in OP from seed orchards, principally because of higher inbreeding and of more heterogeneous mating systems in natural populations: h^2 estimated from OP progenies from seed orchards were easily adjusted with the Squillace (1974) correction and were comparable to estimates found with controlled crosses in eucalypt species. In contrast, h^2 estimates for OP progenies derived from natural stands were severely inflated, even if high ρ_{mean} among offspring were assumed.

Low number of effective pollen donors

Other processes causing departures from half-sib assumptions are those resulting in a low number of pollen donors contributing to reproduction and unequal representations of the male gamete pool among different females (Squillace 1974; Surles et al. 1990; Gaspar et al. 2009; Hansen and Nielsen 2010; El-Kassaby et al. 2011). In trees, these departures occur frequently in small populations and in populations with unequal male fecundities. Small populations are often observed at the edge of species distributions, in fragmented areas, or simply for species with low density or small range distribution (Tamaki et al. 2009; Fuchs and Hamrick 2011). However, even in large tree populations, the contribution of males to the pollen pool of a given mother-tree is highly uneven, due to highly heterogeneous male fecundities combined with spatial processes (Kang et al. 2003; Oddou-Muratorio, Klein and Austerlitz 2005) and assortative mating. Moreover, the number of effective pollen donors (N_{ep}) pollinating a mother-tree also depends on the mother-tree phenotype (e.g. diameter; Garcia et al. 2005) and on the local environment (e.g. density; El-Kassaby and Jaquish 1996; Restoux et al. 2008). For example, in *Pinus pinaster* populations, de-Lucas et al. (2008) estimated a mean N_{ep} ranging from ~ 15 to 43 fathers between populations, and N_{ep} from ~ 5 to infinite within populations between mother-trees.

Presence of full-sibs within families

Small effective numbers of pollen donors create full-sibs within OP families. The non-negligible proportion of full-sibs can increase ρ_{mean} and lead to overestimating V_A and h^2 (Surles et al. 1990; El-Kassaby et al. 2011), similarly to self-fertilization events. Squillace (1974) estimated that a low effective number of local males ($N_{ep} = 5$) led to the same ρ_{mean} as 10% rate of selfing. This bias can be simply adjusted by the Squillace correction (1974), based on a priori knowledge of ρ_{mean} .

Presence of relatives among families versus heterogeneity of male gamete pool among females

Low N_{ep} generates paternal half-sibs in different families, when mother-trees share the same male contributors, and so generates relatedness among families. This departure from unrelated half-sib family assumption in OP designs is one of the least quantified and documented in the literature. So, in populations with low N_{ep} and homogeneous male contributions to the maternal pollen pools, paternal half-sibs between families could decrease the genetic variance among families, which may balance the overestimation of V_A due to the creation of full-sibs within families. In contrast, the balancing effect of paternal half-sibs should be reduced in populations where the male contributors differ among maternal pollen pools. Hence, the highest biases in V_A and h^2 estimates are expected in populations with low N_{ep} and high heterogeneity of pollen clouds, e.g. in case of heterogeneous male fecundities and restricted pollen dispersal distances. But these predictions would need more thoughtful quantitative investigations from simulations to understand what impact the ratio of full-sibs versus paternal half-sibs has on h^2 estimates.

Expression of dominance effects for full-sibs

For traits with non-negligible dominance genetic controls, neglecting the presence of full-sibs, and hence the dominance effects among these relatives, can inflate V_A due to confounding non-additive effects, and overestimate h^2 . Note that the dominance effects induced by full-sib relatives are opposed to the effects of inbreeding depression occurring in selfed progenies, which can be partially described as dominance effects; the first should lead to overestimate h^2 due to the inflation of between-families variance, whereas the second should bias the estimates downward, due to the inflation of within-family variance. According to Borralho (1994), the presence of variable proportions of full-sibs among families resulting from low N_{ep} and unequal male contributions is expected to induce strong biases in the presence of dominance, difficult to correct with the Family model and the Squillace correction (Squillace 1974).

Assortative mating: a particular case of non-random crosses

A particular case that induces unbalanced crosses between female and male trees is the occurrence of assortative mating. Positive assortative mating occurs when individuals with similar phenotypes mate preferentially. In tree populations, positive assortative mating for phenological traits are likely to occur frequently through preferential mating among individuals showing synchronized flowering (Soularue and Kremer 2012). Assortative mating tends to decrease the N_{ep} per

mother-tree and to increase the heterogeneity of pollen pools across mother-trees. In a seed orchard, Askew and El-Kassaby (1994) used a simulation approach to investigate the effects of reproductive phenology on ρ_{mean} . They concluded that the average difference between uniform phenology and observed phenology for ρ_{mean} was small. Unfortunately, unlike the effects of assortative mating on patterns of pollen flow, the effect of assortative mating on the range of ρ_{mean} and its bias on quantitative genetic estimates has been poorly investigated in natural populations.

Another consequence of assortative mating, rarely mentioned, is the possible reduction of the within-family variance (V_W in Equation 1) and increased variance between families (V_F in Equation 1) for the trait inducing assortative mating (e.g. flowering phenology). Indeed, even if it only negligibly reduces the N_{ep} per mother-tree, assortative mating will lead to overestimating genetic differences among families and the h^2 estimate for this particular trait.

Related or inbred parents

The effect of genetic relatedness between parents on departures from the OP assumptions was less investigated, probably because of the difficulty in estimating it. However, it was recognized as a major source of bias in the analysis of OP progenies from natural populations (Fujishima and Fredeen 1972; Squillace 1974; Surles et al. 1990). Tree populations often show significant spatial genetic structure, i.e. genetic relatedness decreases with geographic distance between individuals. The strength of spatial genetic structure is usually measured by the slope of the regression of genetic relatedness against log of distance (2D habitat), scaled by the level of relatedness at short distance (Sp statistics). Vekemans and Hardy (2004) reanalysed the spatial genetic structure of 23 tree species, and observed a range variation from 0.00567 to 0.05291 for the Sp statistics. Spatial genetic structure is principally created by restricted seed dispersal (Wright 1943) and the mating system characteristics of the population (e.g. selfing rate; Vekemans and Hardy 2004). In natural populations and for self-compatible species, it is likely that levels of relatedness are similar among male and female parents; but for the sake of argument, we treated separately the different kinds of relationships among parents in the following.

Related males and females: creation of inbreeding and inbreeding depression in progenies

In a random mating population, the relatedness among the males and females creates biparental inbreeding and possibly inbreeding depression in the OP progenies, in a similar way to selfing. Selfing is indeed an extreme case of mating between relatives. The effects of inbreeding and inbreeding depression were previously analysed in the selfing section.

Related seed trees: creation of relatives among families

Mother-tree relatedness is another case that can induce strong biases in the analysis of OP designs. It can easily occur if the OP progenies are collected on neighbouring mother-trees. Relatedness among mother-trees does not create inbreeding in progenies, but reduces the genetic variance among families, creating relatedness relationships among all offspring of different families. Squillace (1974) considered a specific case where both the proportion of selfing and genetic relatedness vary among seed-parents, the relatedness among two parents being equal to the mean selfing rate. Results showed that the h^2 overestimation was lower if seed trees were correlated than if they were not (table 9 in Squillace 1974). In this particular case, the genetic relatedness among offspring and among parents had antagonistic effects, the first increasing the V_F and the second decreasing V_F . Furthermore, if mother-trees are related the h^2 biases induced by the half-sib then assumptions were almost negligible, except for the high heritability traits and high proportion of selfing (Squillace 1974).

Related male trees: decrease of N_{ep}

The presence of related male-trees in the population decreases N_{ep} of the population, and so increases ρ_{mean} , reveals dominance effects and possibly creates relatedness among families, with the consequences on V_A and h^2 detailed above.

Ancestral inbreeding: increase of ρ_{mean}

The presence of ancestral inbreeding is frequent in natural populations. For example, in 389 populations of *Fagus sylvatica*, a predominantly outcrossing species, Comps et al. (2001) found that inbreeding, measured by the heterozygote deficit F_{is} , on the basis of the 12 isozyme loci, decreases with time since post-glacial colonization, from 0.06 in southeast Europe to close to 0 in northwest Europe.

In OP designs, ancestral inbreeding in the parent population increases the genetic relatedness among offspring (cf table 3 in Squillace 1974). Indeed, the coefficient of coancestry (θ_{ij}) in an inbred population is classically defined as:

$$\theta_{ij} = \sum_{a,n} (1/2)^n (1 + F_a)$$

where F_a are the inbreeding coefficients of the ancestors a , and n is the number of steps to the common ancestor. As the inbreeding coefficient of natural populations is generally unknown, the F_a are often approximated from the level of effective selfing (s), assuming inbreeding equilibrium (Hedrick and Cockerham 1986):

$$F_{a_{eq}} = \frac{s}{2-s}$$

For inbred individuals, the coefficient of coancestry (θ_{ij}) is linked to the relatedness coefficient as follows: $\rho_{ij} = 2 \theta_{ij}/(1+Fa)$. As ancestral inbreeding only implies an effect on ρ_{mean} it can be corrected with the method of Squillace (1974).

The general effect of relatedness and inbreeding in parent populations

In finite populations with random mating, ignoring the genetic relatedness among parents (among males, females and/or between the males and females) led to either overestimates or underestimates of h^2 (table II in Fujishima and Fredeen 1972).

In a non-random mating population, inbreeding and genetic relatedness among parents increase ρ_{mean} among offspring (Squillace 1974). Moreover, in cases where offspring of different families share related parents, the relatedness among parents decreases the genetic variance among families (Equation 3). So, Squillace (1974) proposed to account for this mean relatedness among parents (noted r_{pp}) in his h^2 correction:

$$V_F = (\rho_{mean} - r_{pp})V_A \quad (3)$$

$$h^2 = \frac{V_F}{(\rho_{mean} - r_{pp})V_w + \rho_{mean}V_F}$$

New methodological developments using the “animal model”

To overcome the assumptions of the Family model and to estimate unbiased V_A and h^2 using OP designs, recent studies proposed using the animal model, which allows accounting for different kinds of relatives within and among families, variable mating systems among families, and/or dominance effects. However, the use of this model depends on the availability of pedigrees or molecular data to estimate relatedness among individuals, at least for the offspring population.

Use of complex pedigrees to estimate the additive effects

As the animal model can incorporate various multigenerational and unbalanced relatedness levels it is particularly suited to account for the different kinds of relatives created by non-random mating systems and relatedness among parents. The simplest use of the animal model in the OP progeny analysis involves providing genetic relationships using a pedigree file with male and female parents of the offspring, classically obtained by observations or paternity assignment and pedigree reconstruction methods using molecular marker data (Hansen and Nielsen 2010; El-Kassaby et al. 2011). It is notoriously more difficult to build a pedigree on the

basis of molecular markers in the presence of biparental inbreeding (Carré et al. 2013).

The combined use of molecular markers and the animal model led El-Kassaby et al. (2011) to propose the new concept of “Breeding without breeding”. This concept is based on the use of offspring with incomplete pedigree information (half-sibs; HS) with a subset of offspring with complete pedigree information (full-sibs; FS), to accurately estimate the heritability of a trait and the breeding values of parents and offspring. A case study on the height of *Larix occidentalis* (Nutt.) seedlings (El-Kassaby et al. 2011), h^2 was overestimated by over 50% when assuming that all offspring were HS. The integration of 10% of paternal information was sufficient to reduce the credibility intervals of the h^2 estimates and to adjust V_A and h^2 values. The optimal use of molecular and phenotypic information is promising, as molecular makers are now available for many non-model species, even though the genotyping of many samples in large quantitative genetic designs is still expensive. However, El-Kassaby et al. (2011) did not compare the HS + FS method with an a posteriori correction of the family variance according to the ρ_{mean} estimated on the subset of genotyped offspring. Except for the estimation of breeding values, it is difficult to evaluate if this method is more appropriate than the use of the family model plus a ρ_{mean} correction to analyse OP designs. The HS + FS method probably provides estimates with the lowest bias, if paternal half-sib relatives are non-negligible and if dominance effects are accounted for in the animal model.

Use of the relatedness relationship matrix to estimate the additive effects

The coefficients of the variance–covariance matrix can be either theoretical values (e.g. 0, 0.25 and 0.5 for unrelated, half-sibs and full sibs, respectively), derived from observational data in a pedigree (A^{-1} is directly derived in this case), or continuously distributed pairwise relatedness coefficients, estimated from molecular markers. The estimation of quantitative genetic parameters using relatedness coefficients is efficient for populations with large variance of relatedness coefficients, e.g. in populations with high selfing and/or inbreeding (Gay, Siol and Ronfort 2013), but not completely adequate for tree populations with large population size and high outcrossing rate. Estimating the variance of relatedness coefficients when average relatedness is low critically depends on the number of molecular markers.

Rather than estimating the whole matrix of pairwise relatedness coefficients, which would require extensive molecular data, Bush et al. (2011) proposed a method to build the A matrix using partial molecular information and mean estimates. This approach follows the rationale of Squillace (1974), but at the family or sub-population level. The approach is illustrated by a case-study on

Eucalyptus cladocalyx, a species with highly variable selfing rates among and within subpopulations (s ranging from 0 to 1). The aim of the authors was to account for variable selfing rates (s) and ancestral inbreeding (F_a) at the family or sub-population levels (see table 1 in Bush et al. 2011) in the V_A and h^2 estimations. However, assuming either constant ($s = 0.36$ and $F_a = 0.25$) or variable selfing rate and inbreeding among families in the A matrix, led to the same h^2 estimations ($h^2 = 0.22$ and 0.21 for diameter, respectively; $h^2 = 0.35$ and 0.32 for height, respectively). This is probably because of the absence of dominance effects: in this case, the simple a posteriori application of a coefficient of relationship of $\rho_{mean} = 0.37$ (estimated according to the mean level of selfing and inbreeding) was sufficient to reach correct h^2 estimates. Only changes in the ranking of breeding values were observed when adjusting the inbreeding parameter at the family level rather than at the population level. Hence, although this new methodology based on the use of the relationship matrix of the animal model seems promising, because it would allow the accounting for several departures from half-sib assumptions at several scales, it needs to be applied on others species and traits presenting dominance effects to appreciate its value compared with the correction proposed by Squillace (1974).

Accounting for non-additive and environmental effects

The animal model allows estimation of the dominance variance, and incorporation of a relationship matrix structuring the dominance effects, as done for the additive effects (Viana et al. 2012). As concluded by Borralho (1994) these dominance effects are very important to consider for avoiding the bias induced by selfing and, more generally, by any mating system varying among families. One of the first experimental studies to account for the paternal information and to test the effects of dominance on h^2 biases with the animal model in an OP design seed orchard showed that dominance effects were negligible for *Abies nordmanniana* (Hansen and Nielsen 2010). But, as mentioned above, seed orchard conditions may not be representative of what happens in natural populations and probably underestimate the bias induced by non-random mating and the expression of dominance effects.

Another advantage of the animal model is the possibility to integrate maternal effects in the model. In the first years of OP progeny tests, these maternal effects are often non-negligible (Wang, Lechowicz and Potvin, 1994; González-Rodríguez, Villar and Navarro-Cerrillo 2011) and inflate V_F , therefore leading to an overestimate of V_A and h^2 . More generally, epigenetic effects have the same impact. Accounting for maternal effects in the Family model requires using a covariate (e.g. seed weights), which may imply important phenotyping effort (González-Rodríguez, Villar and Navarro-Cerrillo 2011). Hence, the animal model is a simple methodological tool

to account for non-additive or environmental effects in quantitative genetic models, simply in adding random effects in Equation (2) and specifying the variance–covariance matrix associated with these effects (Kruuk et al. 2004). Moreover, these relatedness matrices can now be easily constructed based on pedigree (e.g. using the “nadiv” package of R; Wolak 2012), simplifying even further the study of non-additive effects.

Conclusion and perspectives

The perfect analysis of progeny trials probably involves using a complete molecular marker dataset to reconstruct the pedigree of every offspring, or to estimate individual pairwise relatedness coefficients, and to include all this individual information in an animal model. However, some studies point out that caution must be taken against the use of the wrong pedigree information or biased relatedness coefficients, which can lead to greater errors in V_A and h^2 estimates than using incomplete relatedness information (Roughsedge, Brotherstone and Visscher 2001; Kumar and Richardson 2005). In addition, V_A and h^2 estimates in OP progeny tests can be significantly improved using partial molecular marker information and minimum genotyping effort.

Biases in OP progeny tests have been investigated much more frequently for progenies from tree breeding programmes than from natural populations, whereas more important departures from half-sib assumptions are expected in natural than controlled populations (Hodge et al. 1996). This heterogeneity is certainly due to the fact that the OP design was established and principally used by breeders, and one can wonder if this kind of design is really appropriate for progenies collected in natural populations. Other approaches of in situ quantitative genetics may be more appropriate (Gay personal communication; Carré et al. 2013). Yet, to answer this question, more experiments based on natural populations investigating the effects of departure from half-sib assumptions are necessary.

Our review shows that corrections for the majority of the departures from half-sib model assumptions have already been proposed (Table 2). Indeed, the effects on the mean genetic relatedness within (ρ_{mean}) or among (r_{pp}) families can be easily corrected by the method of Squillace (1974). Ancestral inbreeding in the parent population is also either easy to correct via its impact on ρ_{mean} (Squillace 1974) or accounted for in the additive genetic relationship matrix in animal models (Bush et al. 2011). By contrast, dominance effects induced by full-sib relatives cannot be accounted for using the family model and require the use of animal models (Hansen and Nielsen 2010). The only factor for which no correction has been proposed is progenies inbreeding depression, probably because the genetic mechanisms under inbreeding depression are complex and cannot easily be accounted for in quantitative genetic models. However, Borralho (1994) gave an example of how they could be

treated, choosing to define them as dominance effects, depending on the family of origin and the coefficient of inbreeding of the offspring (see equation 3 in Borralho 1994).

In this review, we discuss the mating characteristics that create departures from half-sib assumptions sequentially, but they generally occur simultaneously in natural population. The literature reveals that biases affecting estimates of quantitative genetic parameters vary according to the mating system; therefore it is essential to have a good knowledge of the mating system in the studied populations. Still, as the different departures from half-sib assumptions may have antagonistic or cumulative effects, real biases in analysis of OP families from natural populations are hardly predictable.

In general, the use of the animal model appears very promising to estimate V_A and h^2 because it should allow the consideration of all degrees of relatedness, unequal male contribution, actual level of inbreeding of the studied population, and non-genetic or non-additive random effects (maternal or dominance effects), which are so incidental in plant studies (Kruuk 2004; Wolak 2012; Pujol and Gallaud 2013). However, our review of the literature of open-pollinated maternal families suggests that the “simplest” family model, combined with correction terms based on molecular markers showed satisfactory results even in systems where half-sib assumptions are not valid, provided that these departures only concern the levels of additive genetic relationship between individuals (see the corrections proposed in Table 2). Hence, the major comparative strength of the animal model compared with the family model lies in its capacity to account for non-additive genetic effects in OP progenies analysis and, when these effects are significant, to account for variable levels of inbreeding and genetic relatedness within and between families. Recent studies tended to point out that the genetic variance is mainly additive genetic variance (Hill, Goddard and Vissler 2008), which may advocate for the use of simplest models and appropriate corrections, according to the mean levels of genetic relatedness, to estimate accurate quantitative genetic parameters in half-sib designs.

The conclusions of this review are valid for all perennial and long-generation organisms for which quantitative genetic analysis is performed using family and/or incomplete pedigree information (e.g. paternal relatedness missing), short pedigree information (i.e. no ancestral inbreeding information), or neglecting dominance effects. Even if some studies in animal natural populations have shown that low misassigned paternities or moderately deep pedigree were not issues in the estimation of quantitative genetics parameters (Charmantier and Réale 2005; DiBattista et al. 2009), care should be taken before overlooking the population history, the mating system knowledge, or the test of dominance effects in quantitative genetic analysis.

Table 2. Effects of the departures from half-sib assumptions on the genetic variance within family (V_W), the additive genetic variance (V_A), or the heritability (h^2) estimates. According to the departure different correction are proposed.

Departure	Effect on V_W	Effect on the estimation of V_A	Effect on the estimation of h^2	Correction proposed
$r_w > 0.25$	0	+	+	Squillace (1974) according to the mean genetic relatedness among offspring
$r_b > 0$	0	-	-	Squillace (1974) according to the mean genetic relatedness among parents
Dominance	0	+	+	Use Animal models to account for dominance effects (Hansen and Nielsen 2010)
Inbreeding depression	+	0	-	Define the coefficient of inbreeding of the offspring in the dominance effects (Borralho 1994)
Parents inbred $F_p > 0$	0	+	+	Account for the ancestral inbreeding in the estimation of the mean genetic relatedness and use the correction of Squillace (1974) Or define levels of ancestral inbreeding at the family or sub-population scale in Animal model

The sign “+” corresponds to an overestimation of the parameter, the sign “-” corresponds to an underestimation of the parameter and the sign “0” corresponds to no effect on the parameter.

Author’s contributions

JG, EKK and SOM coordinated the review and JG wrote the text. All co-authors contributed to the revisions of preliminary drafts.

Acknowledgements

We are grateful to Bruno Fady and Katalin Csillery for comments on a previous version of this manuscript, and to the people of the GDR “Quantitative Genetics in Natural Populations” for the discussions about these issues. We also thank B. Pujol and two anonymous reviewers for their comments and suggestions. This study was funded the ERA-Net BiodivERsA TipTree project (ANR-12-EBID-0003), part of the 2012 BiodivERsA call for research proposals.

References

- Ashley, M. V. 2010. “Plant Parentage, Pollination, and Dispersal: How DNA Microsatellites Have Altered the Landscape.” *Critical Reviews in Plant Sciences* 29: 148–161.
- Askew, G., and Y. El-Kassaby. 1994. “Estimation of Relationship Coefficients among Progeny Derived from Wind Pollinated Orchard Seeds.” *Theoretical and Applied Genetics* 88: 267–272.
- Barrett, S. C., and L. D. Harder. 1996. “Ecology and Evolution of Plant Mating.” *Trends in Ecology and Evolution* 11: 73–79.
- Borralho, N. 1994. “Heterogeneous Selfing Rate and Dominance Effects in Estimating Heritabilities from Open-Pollinated Progeny.” *Canadian Journal of Forest Research* 24: 1079–1082.
- Bush, D., D. Kain, C. Matheson, and P. Kanowski. 2011. “Marker-Based Adjustment of the Additive Relationship Matrix for Estimation of Genetic Parameters—an Example Using *Eucalyptus Cladocalyx*.” *Tree Genetics and Genome* 7: 23–35.
- Carré, C., F. Gamboa, B. Pujol and E. Manfredi. 2013. “Genetic Links among individuals: from genealogies to molecular markers” *Acta Botanica Gallica*. doi:10.1080/12538078.2013.809316.
- Charlesworth, D., and B. Charlesworth. 1987. “Inbreeding Depression and Its Evolutionary Consequences.” *Annual Review of Ecology and Systematics* 18: 237–268.
- Charmantier, A., and D. Garant. 2005. “Environmental Quality and Evolutionary Potential: Lessons from Wild Populations.” *Proceedings- Biological Sciences* 272: 1415–1425.
- Charmantier, A., and D. Réale. 2005. “How Do Misassigned Paternities Affect the Estimation of Heritability in the Wild?” *Molecular Ecology* 14: 2839–2850.
- Comps, B., D. Gomory, J. Letouzey, B. Thiebaut, and R. J. Petit. 2001. “Diverging Trends between Heterozygosity and Allelic Richness during Postglacial Colonization in the European Beech.” *Genetics* 157: 389–397.
- Conner, J. K. 2001. “How Strong is Natural Selection?” *Trends Ecology and Evolution* 16: 215–217.
- David, P., B. Pujol, F. Viard, V. Castella, and J. Goudet. 2007. “Reliable Selfing Rate Estimates from Imperfect Population Genetic Data.” *Molecular Ecology* 16: 2474–2487.
- DiBattista, J. D., K. A. Feldheim, D. Garant, S. H. Gruber, and A. P. Hendry. 2009. “Evolutionary Potential of a Large Marine Vertebrate: Quantitative Genetic Parameters in a Wild Population.” *Evolution* 63: 1051–1067.
- Ducouso, A., J. Guyon, and A. Kremer. 1996. “Latitudinal and Altitudinal Variation of Bud Burst in Western Populations of Sessile Oak (*Quercus Petraea* (Matt) Liebl).” *Annals of Forest Science* 53: 775–782.
- El-Kassaby, Y., and B. Jaquish. 1996. “Population Density and Mating Pattern in Western Larch.” *Journal of Heredity* 87: 438–443.
- El-Kassaby, Y. A., E. P. Cappa, C. Liewlaksaneeyanawin, J. Klápště, and M. Lstibůrek. 2011. “Breeding without Breeding: Is a Complete Pedigree Necessary for Efficient Breeding?” *PLoS One* 6: e25737.
- Etterson, J. R., and R. G. Shaw. 2001. “Constraint to Adaptive Evolution in Response to Global Warming.” *Science* 294: 151–154.
- Falconer, D. S., and T. F. C. Mackay. 1996. *Introduction to Quantitative Genetics*. New York: Longman Scientific & Technical edn.
- Ferriol, M., C. Pichot, and F. Lefèvre. 2011. “Variation of Selfing Rate and Inbreeding Depression among Individuals and across Generations within an Admixed *Cedrus* Population.” *Heredity* 106: 146–157.
- Fuchs, E. J., and J. L. Hamrick. 2011. “Spatial Genetic Structure within Size Classes of the Endangered Tropical Tree *Guaiaacum Sanctum* (Zygophyllaceae).” *American Journal of Botany* 97: 1200–1207.

- Fujishima, T., and H. T. Fredeen. 1972. "General Formulae for Estimating Heritability in a Population with Related Parents." *Canadian Journal of Genetics and Cytology* 14: 549–557.
- Garcia, C., J.M. Arroyo, J.A. Godoy, and P. Jordano. 2005. "Mating patterns, pollen dispersal, and the ecological maternal neighbourhood in a *Prunus mahaleb* L. population." *Molecular Ecology* 14: 1821–1830.
- Gaspar, M., A. I. Lucas, R. de Alia, J. Paiva, E. Hidalgo, J. Louzada, H. Almeida, and S. González-Martínez. 2009. "Use of Molecular Markers for Estimating Breeding Parameters: A Case Study in a *Pinus Pinaster* Ait. Progeny Trial." *Tree Genetics and Genome* 5: 609–616.
- Gay, L., M. Siol, and J. Ronfort. 2013. "Pedigree-free estimates of heritability in the wild: promising prospects for selfing populations." *PLoS One* 8: e66983.
- Gianola, D., and R. L. Fernando. 1986. "Bayesian Methods in Animal Breeding Theory." *Journal of Animal Science* 63: 217–244.
- González-Rodríguez, V., R. Villar, and R. M. Navarro-Cerrillo. 2011. "Maternal Influences on Seed Mass Effect and Initial Seedling Growth in Four *Quercus* Species." *Acta Oecologica* 37: 1–9.
- Griffin, A., and P. Cotterill. 1988. "Genetic Variation in Growth of Outcrossed, Selfed and Open-Pollinated Progenies of *Eucalyptus Regnans* and Some Implications for Breeding Strategy." *Silvae Genetica* 37: 124–131.
- Hansen, O., and U. Nielsen. 2010. "Microsatellites Used to Establish Full Pedigree in a Half-Sib Trial and Correlation between Number of Male Strobili and Paternal Success." *Annals of Forest Science* 67: 703.
- Hedrick, P., and C. Cockerham. 1986. "Partial Inbreeding: Equilibrium Heterozygosity and the Heterozygosity Paradox." *Evolution* 40: 856–861.
- Henderson, C. R. 1975. "Best Linear Unbiased Estimation and Prediction under a Selection Model." *Biometrics* 31: 423–447.
- Henderson, C. R. 1976. "A Simple Method for Computing the Inverse of a Numerator Relationship Matrix Used in Prediction of Breeding Values." *Biometrics* 32: 69–83.
- Hill, W. G., M. E. Goddard, and P. M. Vissler. 2008. Data and Theory Point to Mainly Additive Genetic Variance for Complex Traits. *PLoS Genetics* 4, e1000 0008.
- Hodge, G., P. Volker, B. Potts, and J. Owen. 1996. "A Comparison of Genetic Information from Open-Pollinated and Control-Pollinated Progeny Tests in Two *Eucalypt* Species." *Theoretical and Applied Genetics* 92: 53–63.
- Kang, K.-S., A. D. Bila, A. M. Harju, and D. Lindgren. 2003. "Estimation of Fertility Variation in Forest Tree Population." *Forestry* 76 (3): 329–344.
- Kingsolver, J. G., H. E. Hoekstra, J. M. Hoekstra, D. Berrigan, S. N. Vignieri, C. E. Hill, A. Hoang, P. Gibert, and P. Beerli. 2001. "The Strength of Phenotypic Selection in Natural Populations." *The American Naturalist* 157: 245–261.
- Kruuk, L. E. B. 2004. "Estimating Genetic Parameters in Natural Populations Using the Animal Model." *Philosophical Transactions of the Royal Society B: Biological Sciences* 359: 873–890.
- Kumar, S., and T. E. Richardson. 2005. "Inferring Relatedness and Heritability Using Molecular Markers in Radiate Pine." *Molecular Breeding* 15: 55–64.
- Loiselle, B., V. L. Sork, J. D. Nason, and C. Graham. 1995. "Spatial Genetic Structure of a Tropical Understory Shrub, *Psychotria Officinalis* (Rubiaceae)." *American Journal of Botany* 82: 1420–1425.
- de Lucas, A. I., J. J. Robledo-Arnuncio, E. Hidalgo, and S. C. González-Martínez. 2008. "Mating System and Pollen Gene Flow in Mediterranean Maritime Pine." *Heredity* 100: 390–399.
- Oddou-Muratorio, S., E. K. Klein, and F. Austerlitz. 2005. "Pollen Flow in the Wildservice Tree, *Sorbus Torminalis* (L.) Crantz. II. Pollen Dispersal and Heterogeneity in Mating Success Inferred from Parent–Offspring Analysis." *Molecular Ecology* 14: 4441–4452.
- Petit, R., and A. Hampe. 2006. "Some Evolutionary Consequences of Being a Tree." *Annual Review of Ecology, Evolution, and Systematics* 37: 187–214.
- Pujol, B. and J. P. Gallaud. 2013. "A practical guide to quantifying the effect of genes underlying adaptation in a mixed genomics and evolutionary ecology approach." *Acta Botanica Gallica*. 160 (3). doi:10.1080/12538078.2013.799045.
- 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.
- Ritland, K. 1996. "Estimators for Pairwise Relatedness and Individual Inbreeding Coefficients." *Genetical Research* 67: 175–185.
- Ritland, K. 2002. "Extension of Models for the Estimation of Mating Systems Using *N* Independent Loci." *Heredity* 88: 221–228.
- Robledo-Arnuncio, J. J., R. Alia, and L. Gil. 2004. "Increased Selfing and Correlated Paternity in a Small Population of a Predominantly Outcrossing Conifer, *Pinus Sylvestris*." *Molecular Ecology* 13: 2567–2577.
- Roughsedge, T., S. Brotherstone, and P. M. Visscher. 2001. "Bias and Power in the Estimation of a Maternal Family Variance Component in the Presence of Incomplete and Incorrect Pedigree Information." *Journal of Dairy Science* 84: 944–950.
- Sorensen, F., and T. White. 1988. "Effect of Natural Inbreeding on Variance Structure in Tests of Wind-Pollinated Douglas Fir Progenies." *Forest Science* 34: 102–118.
- Soularue, J.-P., and A. Kremer. 2012. "Assortative Mating and Gene Flow Generate Clinal Phenological Variation in Trees." *BMC Evolutionary Biology* 12: 79.
- Squillante, A. 1974. "Average Genetic Correlations among Offspring from Open-Pollinated Forest Trees." *Silvae Genetica* 23: 149–156.
- Surles, S. E., J. Arnold, A. Schnabel, J. L. Hamrick, and B. C. Bongarten. 1990. "Genetic Relatedness in Open-Pollinated Families of Two Leguminous Tree Species, *Robinia Pseudoacacia* L. and *Gleditsia Triacanthos* L." *Theoretical and Applied Genetics* 80: 49–56.
- Tamaki, I., K. Ishida, S. Setsuko, and N. Tomaru. 2009. "Interpopulation Variation in Mating System and Late-Stage Inbreeding Depression in *Magnolia Stellata*." *Molecular Ecology* 18: 2365–2374.
- Vekemans, X., and O. J. Hardy. 2004. "New Insights from Fine-Scale Spatial Genetic Structure Analyses in Plant Populations." *Molecular Ecology* 13: 921–935.
- Viana, J. M. S., R. O. DeLima, V. R. Faria, G. B. Mundim, M. D. Vilela de Ressende, and F. F. eSilva. 2012. "Relevance of Pedigree, Historical Data, Dominance, and Data Unbalance for Selection Efficiency." *Agronomy Journal* 3: 722–728.
- de Villemereuil, P., O. Gimenez, and B. Doligez. 2013. "Comparing Parent-Offspring Regression with Frequentist and Bayesian Animal Models to Estimate Heritability in Wild Populations: A Simulation Study for Gaussian and Binary Traits." *Methods in Ecology and Evolution* 4: 260–275.

- Wang, Z. M., M. J. Lechowicz, and C. Potvin. 1994. "Early Selection of Black Spruce Seedlings and Global Change: Which Genotypes Should We Favor?" *Ecological Society of America* 4: 604–616.
- Ward, M., C. W. Dick, R. Gribel, and A. J. Lowe. 2005. "To Self, or Not to Self ... a Review of Outcrossing and Pollen-Mediated Gene Flow in Neotropical Trees." *Heredity* 95: 246–254.
- Wolak, M. E. 2012. "Nadiv: An R Package to Create Relatedness Matrices for Estimating Non-Additive Genetic Variances in Animal Models." *Methods in Ecology and Evolution* 3: 792–796.
- Wright, S. 1943. "Isolation by Distance." *Genetics* 28: 114–138.

3.2 Amélioration des estimations d'héritabilité en utilisant des informations moléculaires incomplètes dans les tests de descendance : une approche de simulation

Title : Improving heritability estimates in family designs by using incomplete molecular information

Authors : Gauzere J., Oddou-Muratorio S., Gay L., Klein E. K.

Abstract

1. Accurate estimates of heritability (h^2) are necessary to assess adaptive responses of fitness-related traits in changing environments. For plants, h^2 estimates often rely on maternal progenies, assuming that offspring are either half-sibs or unrelated. Actually, plant mating system often departs from half-sib assumptions and incomplete knowledge of the relationships among offspring biases the h^2 estimates.

2. We performed simulations to investigate how molecular markers can improve the analysis of maternal progenies designs. We compared the basic "family method", considering or not adjustments based on average relatedness coefficients (I), and methods based on the animal model (II). The animal model was used with average relatedness information (IIa), or with hybrid relatedness information : mixing one-generation pedigree and family assumptions (IIb), or mixing one-generation pedigree and average relatedness coefficients (IIc). For these methods, we assessed the effect of unequal male reproductive success, inbreeding and maternal effects on h^2 estimates.

3. Our results showed that in presence of unequal male fecundities, if the mothers sample the same pollen pool, neglecting the genetic relatedness between families biased the h^2 estimates derived from the family model. Methods using average marker-based relatedness coefficients were efficient to deal with unequal male reproductive success and ancestral inbreeding in the parental population, but were biased in presence of maternal effects. In the latter case, the use of pairwise genetic information (e.g. pedigrees in animal models) improved the accuracy of the h^2 estimates. We showed that the use of realistic genetic data, i.e. relatedness coefficients estimated from markers or pedigrees with paternity errors, did not penalize the methods, except for the methods using pairwise marker-based relatedness coefficients.

4. A new hybrid approach proposed here, which uses both the pedigree and mean relatedness information, was efficient to deal with any departure from the family assumptions. This approach is particularly interesting to fully use the genetic information available when only a subsample of the offspring are genotyped.

Key words : Quantitative genetics ; Heritability ; Unequal reproductive success ; Maternal Effects ; Inbreeding ; Statistics ; Pedigree-free methods

Introduction

The ability of fitness-related traits to evolve in response to environmental variation depends on the amount of additive genetic variation available and the strength of the selection, as summarized by the breeder equation $R = V_A\beta = V_P h^2\beta$, with R the response to the selection, V_A and V_P the additive and phenotypic variances, h^2 the heritability and β the selection gradient. Despite its simplicity, the robustness of this model and its efficiency in plant and livestock genetic improvement has promoted the estimation of h^2 , V_A and β for potentially adaptive traits in wild populations (Kruuk, 2004). Accurate estimates of these parameters are notably critical when predicting the rate of evolution of a population (Kingsolver *et al.*, 2001; Conner, 2001; Senneke *et al.*, 2004).

Basically, h^2 estimates are derived from the comparison of the phenotypic and genetic resemblances between pairs of relatives, and thus depend on the knowledge of the genetic relatednesses, measured by the relatedness coefficient ρ . Because the accuracy of h^2 estimates increases with the variance in pairwise relatednesses, the sampling of sibling groups has been widely used in quantitative genetic studies (DiBattista *et al.*, 2009; Ramirez-Valiente *et al.*, 2009; Hansen & Nielsen, 2010). In both plant and animal natural populations, the family structure is often made of maternal progenies (Kruuk, 2004; DiBattista *et al.*, 2009). In plant, they are easily sampled by harvesting seeds on open-pollinated mother-plants (Ramirez-Valiente *et al.*, 2009; Hansen & Nielsen, 2010). In animals, offspring that share the same mother are

often easy to observe (e.g. offspring in the same nests, suckling observations for mammal species, offspring gathering around the mother). The most parsimonious assumption is that the offspring within maternal families are half-sibs. The mean genetic relatedness within family (ρ_w) is thus 0.25 and the additive genetic variance can be derived from the variance among families (V_F), as : $V_A = 1/\rho_w \times V_F = 4 \times V_F$, and $h^2 = V_A/V_P$ (Falconer & MacKay, 1996).

Therefore, the estimation of h^2 using V_F in sibling group experiments relies on several assumptions about the mating system and the relatedness among mates. Treating samples within maternal families as "real" half-sib, unrelated to others families, implies that (1) mating events occur with a high number of equally fecund males and without preferential mating, therefore there is no unequal male reproductive success (MRS), (2) mates are unrelated and there is no self-fertilization, and (3) mothers are not inbred. These assumptions are rarely all satisfied in natural populations (Gauzere *et al.*, 2013b; Pemberton *et al.*, 1999) and relationships within and between families are generally complex and unbalanced. Moreover, V_F is assumed to include only V_A component ; yet non-additive variance components, e.g. due to dominance or maternal effects, can only be distinguished from V_A if several classes of relatedness are included (e.g. full-sibs in half-sib designs). In some species, such as monogamous animals, the father identity can be deduced by monitoring during the mating season or observations of parental cares. However, the paternal information derived from observations is not always reliable (Kruuk, 2004), for instance because of extra-pair paternity (Charmantier & Blondel, 2003). In addition, for many species, paternity assignment from direct observation is impossible.

The molecular information offers an alternative to retrieve paternal links within and between maternal progenies. The development of molecular markers, such as microsatellites, has boosted the use of methods to assign paternity (Marshall *et al.*, 1998; Hadfield *et al.*, 2006) and reconstruct pedigrees (Blouin, 2003; Fernandez & Toro, 2006) in non-model species. In return, the possibility to establish pairwise relationships between individuals has stimulated the use of pedigree information in quantitative genetic models, such as the animal model (Kruuk, 2004; Hansen & Nielsen, 2010; DiBattista *et al.*, 2009; Charmantier & Réale, 2005). However, paternity assignment methods require an exhaustive genotyping of the potential parents. Alternatively, sibship reconstruction methods only use the genotypes of the offspring : they

are less costly, but often require an *a priori* information about the categories of relatedness (full-sibs, half-sibs,...) and size of the relatedness groups in the sampled population (Thomas *et al.*, 2000). Finally, wrong sibship reconstruction or parental assignment can generate biases in h^2 estimates (Senneke *et al.*, 2004; Charmantier & Réale, 2005).

To circumvent the problems of the categorical methods, pedigrees in the animal model can be replaced by a matrix of pairwise relatedness coefficients directly estimated from molecular markers (Frentiu *et al.*, 2008; Gay *et al.*, 2013). Different marker-based relatedness or kinship coefficients have been proposed (Loiselle *et al.*, 1995; Ritland, 1996) and an infinite number of markers can theoretically trace the distant relatednesses among individuals. However, this "pedigree-free" approach has not yet reached a consensus because the performance of the different marker-based coefficients varies with the range of relatednesses to estimate, their accuracy are challenged in population with low relatedness variance and the uncertainty in relatedness estimates is not taken into account in the animal model (Gay *et al.*, 2013; Csillery *et al.*, 2006). Moreover, the relatedness matrix constructed from molecular markers is frequently not positive definite, which can generate biases in h^2 estimates due to matrix inversion problems (Frentiu *et al.*, 2008). Consequently, in numerous studies the optimal way to accurately estimate h^2 is still questioned and further studies are needed to compare the different approaches.

In natural tree populations, the large number of reproductive individuals, long-distance pollen dispersal and high migration rates (Ashley, 2010) limit the amount of offspring for which a father can be assigned. Moreover, the long life cycle of tree species often constrains the pedigree reconstruction to one-generation. Thus, the relatedness data sets are generally heterogeneous, with typically two levels of relatedness characterization : (1) offspring with maternal and paternal relatednesses known over one-generation and (2) offspring with only a known mother. Recently, El-Kassaby *et al.* (2011) promoted the use of sparse pedigree data sets, rather than a subset corresponding to the full-connectivity pedigree, to increase the accuracy of the predicted parental breeding values. Even if this method was not termed as "hybrid", this example and the incomplete knowledge of the relatedness relationships in quantitative genetic studies clearly encourage the use of hybrid relatedness information.

Here, we investigate the use of partial relatedness information to improve h^2 estimates in half-sib designs. Over the last 50 years, several methods have been proposed to correct h^2 estimates derived from open-pollinated progenies in trees (Gauzere *et al.*, 2013b). We classified these methods and new approaches in four broad categories according to the model used and the treatment of relatedness information. Type I gathers the methods using the family model and average marker-based relatedness coefficients within and/or between families to adjust the V_F estimates. Type II gathers the methods using the animal model. Type IIa methods use average marker-based relatedness coefficients and type IIb and IIc methods use hybrid relatedness information, including the one-generation pedigree plus half-sib assumptions, and the one-generation pedigree plus average marker-based relatedness coefficients respectively.

We used a simulation approach to compare these methods considering several departures from the half-sib assumptions challenging their performances. First, we addressed the effect of unequal male fecundities and preferential mating on the performance of the type I methods; we expected that methods accounting for genetic relatedness among families should be more suitable. Second, with ancestral inbreeding we expected that type I, IIa and IIc methods, using marker-based relatedness coefficients, should be more efficient than the type IIb method using only one-generation pedigrees. Third, in presence of non-additive effects we expected that methods IIb and IIc using the finest level of genetic information should perform best. Finally, we tested the sensibility of the methods to the use of imperfect genetic data.

Material and method

SIMULATION MODEL

We used the model described in Gay *et al.* (2013) to simulate populations evolving at constant population size for a number N_g of non-overlapping generations. We started from $N_{pop} = 200$ diploid and unrelated individuals at generation 0 with $L_M = 20$ neutral unlinked loci of 10 alleles each and $L_{QTL} = 500$ unlinked Quantitative Trait Loci (QTL) of 5 alleles each. The allele frequencies at the QTL and neutral loci evolved under random drift and without mutation, selection nor migration. The phenotype of an individual ($P_{j,k}$) was defined as :

$$P_{j,k} = \sum_l (a_{j,k,l,i} + a_{j,k,l,i'}) + \epsilon_{j,k}$$

with $a_{j,k,l,i}$ and $a_{j,k,l,i'}$ the allelic effects at the QTL locus l of the alleles i and i' of the offspring k in the family j , randomly drawn from a normal distribution ($N(0, \sigma_a^2)$), and ϵ_k a random effect representing the environmental effect, drawn from a normal distribution ($N(0, \sigma_\epsilon^2)$).

We modified this model to account for unequal male reproductive success (MRS), ancestral inbreeding and maternal effects using different scenarios (Table 3.1). At the last generation, we sampled 15 maternal progenies of 40 phenotyped offspring each. We assumed that only a subset of $N_{th} = 15$ offspring per family have a known one-generation pedigree and genotype at the neutral markers.

Unequal male reproductive success

We modeled individual fecundities for each male k (F_k) using a Dirichlet distribution, with a scale parameter α :

$$(F_1, \dots, F_k, \dots, F_{N_{pop}}) \sim D(\alpha; \frac{1}{N_{pop}}, \dots, \frac{1}{N_{pop}})$$

Low values of α ($\alpha < 1$) generate unequal fecundities, i.e. few highly fecund individuals and many individuals with low fecundities. The effective number of pollen donors (Nep) in the population is given by :

$$Nep = \sum_{k: fathers} F_k^2$$

Thus, Nep decreases for more unequal male fecundities. We considered either equal male fecundities ($\alpha = 5$, scenario 0) or unequal male fecundities ($\alpha = 0.01$ for scenarios *i* and *iii*, $\alpha = 0.1$ for scenarios *ii*).

Besides, unequal MRS can also result from spatial or temporal differentiation of male gamete pools, e.g. due to preferential mating. The mother-trees of the last generation were thus split into N_{PC} groups sharing the same pollen pool. We generated as many distributions of male fecundities as the number of pollen pools (N_{PC}). In populations with unequal male fecundities (low α), the differentiation of pollen pools increases with N_{PC} ($N_{PC} = 1$: a single pollen pool ; $N_{PC} = 15$: pollen pools highly differentiated). In scenarios *i* we investigated $N_{PC} = \{1, 5, 15\}$.

Ancestral inbreeding

For each mating event, selfing can occur with probability s . In scenario *ii* ancestral inbreeding was considered by increasing the number of generations ($N_g = 3$) and allowing selfing with probabilities $s = \{0, 0.2, 0.4\}$. We also considered unequal MRS by using $N_{PC} = 5$ and $\alpha = 0.1$. In all other scenarios we set $s = 0$ to avoid confusion between the effect of selfing and low N_{ep} on the increase of relatedness.

Maternal effects

In scenarios *iii* we included maternal effects on the phenotypes as follows :

$$P_{j,k} = \sum(\alpha_{j,k,l,i} + \alpha_{j,k,l,i'}) + m_j + \epsilon_{j,k}$$

with m_j the maternal effect of the mother j . The maternal effects were normally distributed, $m_j \sim N(0, \sigma_m^2)$, and $\sigma_m^2 = \{0, 0.22, 0.32\}$.

Accuracy of the genetic data

In all simulations, we recorded (1) the multi-generational pedigree to calculate the true relatedness coefficient, $\rho_{k,k'}$, between each pair of individuals (k, k') , (2) the one-generation pedigree, $\rho_{k,k'}^{(1)}$, and (3) the one-generation pedigree with 20 % of paternity error simulated randomly, $\tilde{\rho}_{k,k'}^{(1)}$. The genotypes of the offspring at the neutral markers were used to estimate pairwise relatednesses, $\tilde{\rho}_{k,k'}$, using the approach of Loiselle *et al.* (1995) ($\tilde{\rho}_{k,k'} = 2f_{k,k'}$, with $f_{k,k'}$ the pairwise kinship coefficient of Loiselle). This estimator was used because it does not assume Hardy-Weinberg equilibrium and performs well, even in the presence of rare alleles (Gay *et al.*, 2013).

In the first set of simulations, the performance of the methods were compared under the 4 different scenarios, using the true relatednesses $\rho_{k,k'}$ and $\rho_{k,k'}^{(1)}$ (Table 3.2). Then, we investigated the performance of the methods using imperfect genetic data, i.e. $\tilde{\rho}_{k,k'}$ and $\tilde{\rho}_{k,k'}^{(1)}$, using the scenario *i* with $N_{PC} = 5$. The scenarios were simulated with moderate value of heritability (around 0.33). The simulation program was written in C++ and run in batch using a python script.

ESTIMATION OF HERITABILITY

Based on the simulated data sets, we used 4 categories of methods to estimate h^2 .

Type I : Methods using the classical family model

We used a classical family model, accounting only for the maternal relatedness :

$$P_{j,k} = \mu + Fam_j + \epsilon_{j,k} \quad (3.1)$$

where $P_{j,k}$ is the phenotype of the offspring k from family j , μ is the mean, Fam_j a random family effect with variance V_F and $\epsilon_{j,k}$ a random residual error with variance V_R .

In the "**Fam**" method, h^2 was estimated as :

$$h_{Fam}^2 = \frac{4V_F}{V_F + V_R} \quad (3.2)$$

Based on model (3.2) and Squillace (1974), we applied two corrections on the h^2 calculations.

First, the "**FamAdj**" method considered :

$$h_{adj}^2 = \frac{\frac{1}{\rho_w} V_F}{V_F + V_R} \quad (3.3)$$

where ρ_w is the mean genetic relatedness estimated within families.

Second, the "**FamDoubleAdj**" method corresponded to :

$$h_{doubleadj}^2 = \frac{V_F}{(\rho_w - \rho_b)V_R + \rho_w \cdot V_F} \quad (3.4)$$

where ρ_b is the mean genetic relatedness estimated between families.

The adjustment "FamAdj" was proposed by Squillace (1974) to account for $\rho_w > 0.25$ in sibling designs, due to non negligible selfing or low number of effective pollen donors. The second adjustment, "FamDoubleAdj", is inspired from another correction proposed by Squillace (1974) to account for relatedness among mates or, more generally, for non-null relatedness among families in natural populations. ρ_w and ρ_b were computed by averaging $\rho_{k,k'}$ (or $\tilde{\rho}_{k,k'}$) over all pairs of individuals (k, k') with a known molecular information. These models were fitted in R using the package "lme4".

Type II : Methods using the animal model

We used a set of 3 methods (plus one reference method) based on the animal model without maternal effect :

$$P_{j,k} = \mu + a_{j,k} + \epsilon_{j,k} \quad (3.5)$$

or with maternal effect :

$$P_{j,k} = \mu + a_{j,k} + m_j + \epsilon_{j,k} \quad (3.6)$$

where $P_{j,k}$ is the phenotype of offspring k from family j , μ is the mean, $a_{j,k}$ is the random additive genetic effect for individual k in family j , m_j is the random maternal effect and $\epsilon_{j,k}$ is a random residual error. The additive genetic effects $a_{j,k}$ are normally distributed with mean 0 and a variance-covariance matrix, $G = AV_A$, where A is the additive genetic relationship matrix with individual elements $\rho_{k,k'}$.

The following methods to estimate h^2 differ for the matrix A :

The "**PhiA**" matrix is made of the pairwise relatedness coefficients calculated from the multigenerational pedigree ($\rho_{k,k'}$) for all of the phenotyped individuals. This matrix was used as a reference to evaluate the performance of the other methods since it includes the most accurate knowledge of pairwise relatednesses, generally unknown in experimental studies. Similarly, all the pairwise $\tilde{\rho}_{k,k'}$ were used to test the classical "pedigree-free" approach (Frentiu *et al.*, 2008; Gay *et al.*, 2013), named "**PhiA** $_{\tilde{\rho}_{kk'}}$ " here.

Type IIa - The "**Relatedness**" matrix, inspired from Bush *et al.* (2011), uses the pairwise $\rho_{k,k'}$ (or $\tilde{\rho}_{k,k'}$) recorded for the N_{th} individuals per family with a known molecular information to estimate mean relatedness coefficients at the family level. The A matrix thus contains the $\rho_{w,j}$ estimated for each family j (instead of 0.25) and the $\rho_{b,jj'}$ estimated for each family pair (j, j') (instead of 0 ; Figure 3.1a).

$$\rho_{w,j} = \sum_{k,k' \in family(j)} \frac{\rho_{k,k'}}{n_j(n_j-1)/2}$$

$$\rho_{b,jj'} = \sum_{k \in j} \sum_{k' \in j'} \frac{\rho_{k,k'}}{n_j n_{j'}}$$

Type IIb - The "**HS+FS**" matrix, inspired from El-Kassaby *et al.* (2011), combines the classic half-sib model with the pairwise relatednesses estimated using the one-generation pedigree ($\rho_{k,k'}^{(1)}$ or $\tilde{\rho}_{k,k'}^{(1)}$) for the subset of genotyped individuals, which can be unrelated ($\rho_{k,k'} = 0$), half-sibs (HS, $\rho_{k,k'} = 0.25$) or full-sibs (FS, $\rho_{k,k'} = 0.5$; Figure 3.1b).

Type IIc - The "**Hybrid Relatedness**" matrix mixes the relatedness information contained

in the "HS+FS" and "Relatedness" matrices to optimally use the information contained in the genetic data. As in the "HS+FS" method, the one-generation pedigree provides the $\rho_{k,k'}^{(1)}$ (or $\tilde{\rho}_{k,k'}^{(1)}$) among the genotyped individuals (Figure 3.1c), but the molecular data available from this subset is also used to estimate the mean genetic relatedness coefficients within ($\rho_{w,j}$) and between families ($\rho_{b,jj'}$; Figure 3.1c). For testing the sensitivity to imperfect genetic data, we also replaced the $\rho_{k,k'}^{(1)}$ by the pairwise $\tilde{\rho}_{k,k'}$ coefficients for the N_{th} individuals with a known molecular information, thereby forming a new method called "**Hybrid** $\tilde{\rho}_{k,k'}$ ".

The A matrices for type II methods can be non-positive definite (Frentiu *et al.*, 2008). Rather than forcing the inversion of A in the animal model, we derived \bar{A} , a positive definite matrix, close to A , by modifying the eigenvalues following the method described in the supplementary material. The animal models were fitted using the \bar{A} matrix in AsReml v3.0.5 (Gilmour *et al.*, 2006).

The true h^2 value was estimated as $h^2 = \frac{V_A}{V_P}$, with V_A the variance of the individual allelic effects and V_P the phenotypic variance in the offspring population. The different methods were tested on $n = 50$ replicated simulations per parameter set. We used three performance indices applied to h^2 : (1) the relative bias $b(\hat{\theta})$, defined as $\frac{1}{n} \sum_{i=1}^n \frac{(\hat{\theta}_i - \theta_i)}{\theta_i}$, where $\hat{\theta}$ is the estimator and θ the true parameter value, (2) the mean square error $MSE(\hat{\theta}) = \frac{1}{n} \sum_{i=1}^n (\hat{\theta}_i - \theta_i)^2$ and (3) the 95 % coverage, corresponding to the proportion of times the 95 % confidence intervals contained the true value of the parameter. Confidence intervals were computed only for the methods using the animal model and assuming a normal distribution of the estimates ($h^2 \pm 1.96s.e.$).

Results

In scenario 0, the assumptions of the family model were met ($\rho_w = 0.25$ and $\rho_b = 0.004$). All the methods tested had low negative biases, but not significantly different from zero (Scenario 0, Table 3.3). The methods exploiting average relatedness information (type I and IIa) had a higher MSE than the methods exploiting individual relatedness (type IIb and IIc).

EFFECT OF UNEQUAL MALE REPRODUCTIVE SUCCESS ON h^2 ESTIMATES (SCENARIOS *i*)

In scenarios *i*, ρ_w was close to 0.33 and ρ_b decreased from 0.09 to 0.003 with increasing N_{PC} (Table 3.3). In the family model the MSE and bias increased with the pollen pool differentiation : the family model over-estimated h^2 when $N_{PC} = 15$ ($MSE_{Fam} = 0.027$ and $b(\hat{\theta}_{Fam}) = 28$ % for $\rho_b \approx 0$) but was no more biased when $N_{PC} = 1$ ($MSE_{Fam} = 0.013$ and $b(\hat{\theta}_{Fam}) = -5$ % for $\rho_b = 0.089$). In contrast, the bias of the "FamAdj" method decreased with the differentiation of the pollen pool ($b(\hat{\theta}_{FamAdj}) = 29$ % for $N_{PC} = 1$). The MSE and biases of the "FamDoubleAdj" method were low for all values of N_{PC} ($MSE_{FamDoubleAdj} < 0.017$, $b(\hat{\theta}_{FamDoubleAdj}) < 8$ %).

The type II methods based on pedigree and/or relatedness information performed well regardless of the level of pollen pool differentiation, with negligible bias and low MSE. Regarding the MSE, the "Hybrid" method was always the closest to the "PhiA" method, but it sometimes showed a significant negative bias. More generally, all the methods tended to under-estimate h^2 when ρ_w was high and ρ_b non-null.

EFFECT OF ANCESTRAL INBREEDING ON h^2 ESTIMATES (SCENARIOS *ii*)

Ancestral inbreeding combined with unequal MRS largely increased the mean genetic relatedness within and between families, with $\rho_w > 0.40$ and $\rho_b = 0.016$ (Table 3.4) instead of $\rho_w = 0.25$ and $\rho_b = 0$ under the hypothesis of half-sib families. Inbreeding decreased the performance of the classical family model ($MSE_{Fam} > 0.053$ and $\theta_{Fam} > 49$ % for $\rho_w > 0.40$ and $\rho_b = 0.016$). The adjustment of the family model using ρ_w and/or ρ_b efficiently corrected those biases ($b(\hat{\theta}_{FamAdj}) > -13$ % and $b(\hat{\theta}_{FamDoubleAdj}) < 12$ %), but the "FamDoubleAdj" method presented a higher MSE than the "FamAdj" method ($MSE_{FamAdj} \in [0.012 ; 0.017]$ and $MSE_{FamDoubleAdj} \in [0.021 ; 0.028]$).

Regarding the type II methods, the "Relatedness" and "Hybrid" methods performed well as predicted, but the "FS+HS" method presented increased MSE and bias with increasing inbreeding ($MSE_{FS+HS} > 0.023$). Overall, strong inbreeding tended to decrease the coverage power of the type II methods ($cov \in [0.76 ; 0.82]$ for $s = 0.40$), particularly the "Hybrid" method.

EFFECT OF MATERNAL COMPONENTS ON h^2 ESTIMATES (SCENARIOS *iii*)

For moderate unequal MRS, maternal effects decreased the performance of all methods : MSE were always larger than 0.028 and biases were all significantly different from zero (Table 3.5). As expected with increasing maternal effects, the type I methods were less efficient than the type II methods, in which a random maternal effect was included ($MSE_{Family} \in [0.086 ; 0.238]$ and $MSE_{animal} \in [0.035 ; 0.075]$ for $V_M = 0.08 V_P$). Moreover, family models tended to overestimate h^2 , whereas animal models tended to underestimate h^2 . Among the type II methods, the "Relatedness" method performed worst, with a reduced coverage power and higher MSE in presence of maternal effects ($MSE_{Relatedness} = 0.067$ and $cov_{Relatedness} = 0.56$ for $V_M = 0.04 V_P$).

For the simulated sample design (600 offspring) the inclusion of 37 % of paternal information allowed to correctly estimate h^2 in presence of unequal MRS ($b(\hat{\theta}_{HS+FS}) = -4.94$ %, $b(\hat{\theta}_{Hybrid}) = -12.95$ %; Table 3.1). But this proportion was not sufficient in presence of maternal effects ($b(\hat{\theta}_{FS+HS}) < -25$ %, $b(\hat{\theta}_{Hybrid}) < -21$ %). Figure 3.2 showed the performance of the "Hybrid" method according to variable percentages of paternal information available in presence of maternal effects. Note that even if all offspring had a known pedigree, the h^2 estimate would still be slightly biased ($b(\hat{\theta}_{Hybrid}) = -11.60 \pm 11.16$, with $N_{th} = 40$, equivalent to the "PhiA" method).

EFFECT OF MOLECULAR MARKERS UNCERTAINTIES

Considering the scenario *i* with $N_{PC} = 5$, the use of mean genetic relatedness information based on Loiselle coefficients $\tilde{\rho}_{k,k'}$ instead of the true $\rho_{k,k'}$ did not decrease the performance of methods "FamAdj", "FamDoubleAdj", "Relatedness" or "Hybrid Matrix" (Table 3.6). However, the PhiA $_{\tilde{\rho}_{k,k'}}$ method, i.e. animal model with the full $\tilde{\rho}_{k,k'}$ matrix, performed worse than when the true pairwise relatedness coefficients were used ($b(\hat{\theta}) = -36.49$ %, $cov(\hat{\theta}) = 0.48$). Adding 20 % of pedigree errors did not decrease the performances of the "FS+HS" and "Hybrid" methods. We also compared the "Hybrid" method using the $\tilde{\rho}_{k,k'}^{(1)}$ (pedigree with 20 % error) to the "Hybrid $\tilde{\rho}_{k,k'}$ " using the pairwise $\tilde{\rho}_{k,k'}$ (Loiselle coefficients) for the N_{th} individuals with known molecular information : the latter had worse performances ($b(\hat{\theta}) = -20.33 \pm 7.80$).

Discussion

This study investigated how to improve h^2 estimates using family designs in presence of incomplete molecular information and classical departures from half-sib assumptions. To that purpose, we used a simulation approach allowing to compare estimators derived from different methods according to the true parameter value and thus, we provide accurate advices about the method to choose.

DEVIATIONS FROM EQUAL REPRODUCTIVE SUCCESS

Unequal MRS is classically presented as one of the most common biological process creating departure from half-sib assumptions (Hansen & Nielsen, 2010; El-Kassaby *et al.*, 2011; Pemberton *et al.*, 1999). Unequal MRS results from two main processes : unequal male fecundities and the differentiation of the paternal gamete pools. As expected, type I methods, but not type II methods, were affected by unequal MRS. We showed that unequal male fecundities increases the genetic relatedness within and between maternal families and that the biases generated by these two departures counterbalance each other. Overall, unequal male fecundities generate no bias in the estimates of the family model. Squillace (1974) similarly showed that when the parents of the half-sib families sampled were related (increasing ρ_b) the bias generated by the use of wrong half-sib assumptions was mostly insignificant.

Increasing differentiation of the male gamete pools decreases the genetic relatedness between families in presence of unequal male fecundities. In natural populations, this differentiation is common, due to restricted gamete dispersal, asynchronous reproductive phenologies and selfing (Hardy *et al.*, 2004; de Lucas *et al.*, 2008). These biological processes lead the classical family models to overestimate h^2 .

The departure from the assumption $\rho_w = 0.25$ (or $\rho_w = 0.5$ for full-sib families) has attracted much attention (ElKassaby & Jaquish, 1996; Bush *et al.*, 2011; El-Kassaby *et al.*, 2011). Consequently, studies using classical family models thus often adjust the h^2 estimates according to ρ_w only (equation 3.3 ; Ramirez-Valiente *et al.* 2009; Bush *et al.* 2011). However, in cases where ρ_b is non negligible, the use of single-adjusted h^2 calculations can lead to bias. Thus, our simulations showed that in presence of unequal MRS it is most robust to apply the

"FamDoubleAdj" correction (equation 3.4), accounting for both ρ_w and ρ_b .

AVERAGE MARKER-BASED RELATEDNESSES TO RETRACE MULTIGENERATIONAL PEDIGREE

Several studies showed the interest of multigeneration pedigrees to estimate h^2 in bred or natural populations (Charmantier & Réale, 2005; Atkin *et al.*, 2009). However, when the ability to reconstruct parentage over one or several generations is limited, alternative methods to estimate h^2 must be considered. Here, we tested three "pedigree-free" methods using only average marker-based relatednesses : the "FamAdj", "FamDoubleAdj" and "Relatedness" methods. In presence of unknown levels of unequal MRS and inbreeding, the most parcimonious method is the "Relatedness" method (scenarios *i* and *ii*). The use of average information at the family level ("Relatedness") rather than at the population level ("FamAdj" and "FamDoubleAdj") improves only slightly the estimations. One could have expected a more important quantitative improvement, but Bush *et al.* (2011) similarly showed that correcting the A matrix by mean inbreeding coefficient estimated at the family or population level conducted in roughly the same h^2 estimates.

Moreover, we showed that genotyping few individuals (37 %) to compute marker-based relatednesses ($\tilde{\rho}_{k,k'}$) do not affect much the performance of the methods exploiting average relatedness information. Thus, average marker-based relatednesses correctly retrace the past population history and are efficient substitute to multigenerational pedigrees. The interest of these methods could thus be to re-analyze old common garden designs accounting for non-random mating when parents are no more available. Finally, when "simple" quantitative genetic models are fitted, i.e. only random additive effects are specified, average marker-based relatedness methods perform as well as pedigree-based methods and are much easier to implement.

PEDIGREE OR MARKER-BASED PAIRWISE RELATEDNESSES ?

One main issue in current quantitative genetics is to know if marker-based pairwise relatednesses can replace pedigree information (Gay *et al.*, 2013; Frentiu *et al.*, 2008). Our simulations showed that pairwise pedigree-free methods ("PhiA $\tilde{\rho}_{kk'}$ " and "Hybrid $\tilde{\rho}_{k,k'}$ " methods) had lo-

wer performances than the pedigree based methods with 20 % of wrong pedigree assignments. Yet, 20 % of false paternities can be considered as a high assignment error for parentage analysis using 20 or more polymorphic microsatellites (Harrison *et al.*, 2013). Thus, our results argue in favour of using reconstructed pedigrees rather than pairwise marker-based relatedness coefficients for outcrossed species. Here, we considered 20 neutral molecular markers to mimick experimental studies using microsatellites in non-model species. However, with a much larger number of markers we can expect a better performance of the marker-based pairwise relatedness approach, such as demonstrated in Gay *et al.* (2013) and Csillery *et al.* (2006).

In presence of maternal effects we showed that pairwise relatedness information is necessary to analyze maternal progenies. In informed animal models (equation 3.6) incomplete relatedness matrices led to under-estimate h^2 . Such trend has also been observed by Villemereuil *et al.* (2013) (appendix D). Actually, when relatedness matrices did not contain enough information to dissociate additive and maternal effects, maternal variance is inflated due to confusion. Maternal effects affect many traits and species (Raesaenen & Kruuk, 2007) and can explain a large proportion of the total phenotypic variance (Wilson *et al.*, 2005; Raesaenen & Kruuk, 2007); they are thus non-negligible sources of bias. In populations with unequal MRS, El-Kassaby *et al.* (2011) showed that beyond a given sample size the inclusion of more paternal information did not significantly improve the h^2 estimates (threshold defined at 10 % for their design of 6000 plants). Unfortunately, we did not observe such threshold in presence of maternal effects (Figure 3.2). Therefore, the high amount of pairwise relatednesses necessary to correctly analyze progenies appears as the main limiting factor to study natural populations.

PERFORMANCE OF THE HYBRID METHOD

Finally, our scenarios showed the limits of methods using either one-generation pedigree or average relatednesses information : none appears as a good compromise to deal with every departure from half-sib assumptions. Here, we proposed a new hybrid method that rely on a matrix gathering the one-generation pedigree and the average marker-based coefficients (Figure 3.1). The main drawback of this relatedness matrix is of being non-inversible, due to its particular structure. We proposed a modification of the matrix that solved the problem adequately. Without this approximation the method estimated h^2 with biases ranging from -10 to

-55 % in scenari i (results not shown), whereas our solution lowered it to a maximum of -11 %.

Overall, our simulations showed that the "Hybrid" method is the most robust method to estimate h^2 . Indeed, it performed as well as the best one of its two "parent" methods ("HS+FS" or "Relatedness") in every scenari tested and was not affected by the use of "realistic" genetic data sets ($\tilde{\rho}_{k,k'}^{(1)}$ and average $\tilde{\rho}_{k,k'}$). The negligible effect of paternity errors on the performance of pedigree-based methods confirmed previous findings in birds and sharks (Charmantier & Réale, 2005; DiBattista *et al.*, 2009).

A major strength of the new hybrid method is to fully use the genetic information when only a subset of individuals have been genotyped. Here, we considered a simple case where each individual with a known molecular information could be assigned to a father, but we could easily imagine additional layers of relatedness information in data sets, e.g. a subset of individuals genotyped but non-assigned, from which $\tilde{\rho}_{k,k'}$ coefficients can be estimated. The hybrid approaches are particularly adapted to plant quantitative genetic experiment where complete and multi-generational pedigree are difficult to recover, but we can also imagine its application to animal full-sib families, to efficiently consider extra-pair paternities, ancestral inbreeding or maternal effects.

Although the major methodological developments are focused today on the use of dense genetic markers maps to estimate h^2 using multilocus association models (Yang *et al.*, 2010) and on the optimization of the use of pedigrees and genomic information (Legarra *et al.*, 2009), there are still needs to improve the "standard" animal model to study wild populations of non-bred species, for which genomic data are not yet available. For these species the use of sibling group designs and few highly polymorphic markers will still be the rule for some time and studies like this one, wishing to improve the statistical power of the classical methods, are still welcome.

Acknowledgements

We are grateful to F. Lefèvre, M. Siol and K. Csillery for the discussion and comments on the study and the previous versions of this manuscript. We also thank the people of the GDR "Quantitative Genetics in Natural Populations" for the discussions about these issues.

This study was funded by the ERANet BiodivERsA TipTree project (ANR-12-EBID-0003), part of the 2012 BiodivERsA call for research proposals.

Parameter	Definition	Scenario 0	Scenario i	Scenario ii	Scenario iii
N_g	Number of generation	1	1	3	1
σ_a	Std dev of the allelic effect	0.02	0.02	0.02	0.02
σ_e	Std dev of the environmental effect	0.9	0.9	0.9	{0.84, 0.87}
σ_m	Std dev of the maternal effect	0	0	0	{0.22, 0.32}
s	Proportion of selfing	0	0	{0, 0.2, 0.4}	0
α	Level of unbalanced male fecundities	5	0.01	0.1	0.01
N_{PC}	Number of maternal pollen pools	15	{1, 5, 15}	5	5

TABLE 3.1 – Summary of the parameters for the null scenario and the three scenarios investigating the effect of (i) unequal male reproductive success, (ii) ancestral inbreeding and (iii) maternal effects on h^2 estimates. Scenario 0 : random mating system. Scenario (i) : non-random mating with unbalanced male fecundities and variable levels of pollen pool differentiation. Scenari (ii) : three generations with variable levels of selfing. In scenari (iii) : unequal male reproductive success and three levels of maternal effect variance component V_M .

Method	Type	Relatedness information	
		Perfect	Imperfect
Fam	I	- Family assumptions ($\rho_w = 0.25, \rho_b = 0$)	
FamAdj	I	- average $\rho_{k,k'}$ within all families (ρ_w)	- average $\tilde{\rho}_{k,k'}$ within all families ($\tilde{\rho}_w$)
FamDoubleAdj	I	- average $\rho_{k,k'}$ within and between all families (ρ_w and ρ_b)	- average $\tilde{\rho}_{k,k'}$ within and between all families ($\tilde{\rho}_w$ and $\tilde{\rho}_b$)
Relatedness	IIa	- average $\rho_{k,k'}$ within each family j ($\rho_{w,j}$) and between each pair of families j, j' ($\rho_{b,jj'}$)	- average $\tilde{\rho}_{k,k'}$ within each family j ($\tilde{\rho}_{w,j}$) and between each pair of families j, j' ($\tilde{\rho}_{b,jj'}$)
HS+FS	IIb	- $\rho_{k,k'}^{(1)}$ for genotyped individuals and family assumptions for non-genotyped individuals	- $\tilde{\rho}_{k,k'}^{(1)}$ for genotyped individuals and family assumptions for non-genotyped individuals
Hybrid	IIc	- $\rho_{k,k'}^{(1)}$ for genotyped individuals and $\rho_{w,j}$ and $\rho_{b,jj'}$ for non-genotyped individuals	- $\tilde{\rho}_{k,k'}^{(1)}$ for genotyped individuals and $\tilde{\rho}_{w,j}$ and $\tilde{\rho}_{b,jj'}$ for non-genotyped individuals
Hybrid $\tilde{\rho}_{k,k'}$	IIc		- $\tilde{\rho}_{k,k'}$ for genotyped individuals and $\tilde{\rho}_{w,j}$ and $\tilde{\rho}_{b,jj'}$ for non-genotyped individuals
PhiA/PhiA $\tilde{\rho}_{k,k'}$	Reference	- $\rho_{k,k'}$ for all individuals	- $\tilde{\rho}_{k,k'}$ for all individuals

TABLE 3.2 – Summary of the relatedness information used in the \hat{e} methods tested. Type "I" : methods using the family model ; type "II" : methods using the animal model. For each method two cases are considered : true or imperfect relatedness information. $\rho_{k,k'}$ is the true pairwise relatednesses ; $\tilde{\rho}_{k,k'}$ is the Loiselle estimated pairwise coefficients ; $\rho_{k,k'}^{(1)}$ is the true one-generation pedigree ; $\tilde{\rho}_{k,k'}^{(1)}$ is the one-generation pedigree with 20 % of paternity error

$$\left(\begin{array}{ccccc|ccccc} \mathbf{1} & \mathbf{0.41} & \mathbf{0.41} & 0.41 & 0.41 & \mathbf{0.08} & \mathbf{0.08} & \mathbf{0.08} & 0.08 & 0.08 \\ \mathbf{0.41} & \mathbf{1} & \mathbf{0.41} & 0.41 & 0.41 & \mathbf{0.08} & \mathbf{0.08} & \mathbf{0.08} & 0.08 & 0.08 \\ \mathbf{0.41} & \mathbf{0.41} & \mathbf{1} & 0.41 & 0.41 & \mathbf{0.08} & \mathbf{0.08} & \mathbf{0.08} & 0.08 & 0.08 \\ 0.41 & 0.41 & 0.41 & \mathbf{1} & 0.41 & 0.08 & 0.08 & 0.08 & 0.08 & 0.08 \\ 0.41 & 0.41 & 0.41 & 0.41 & \mathbf{1} & 0.08 & 0.08 & 0.08 & 0.08 & 0.08 \\ \hline \mathbf{0.08} & \mathbf{0.08} & \mathbf{0.08} & 0.08 & 0.08 & \mathbf{1} & \mathbf{0.33} & \mathbf{0.33} & 0.33 & 0.33 \\ \mathbf{0.08} & \mathbf{0.08} & \mathbf{0.08} & 0.08 & 0.08 & \mathbf{0.33} & \mathbf{1} & \mathbf{0.33} & 0.33 & 0.33 \\ \mathbf{0.08} & \mathbf{0.08} & \mathbf{0.08} & 0.08 & 0.08 & \mathbf{0.33} & \mathbf{0.33} & \mathbf{1} & 0.33 & 0.33 \\ 0.08 & 0.08 & 0.08 & 0.08 & 0.08 & 0.33 & 0.33 & 0.33 & \mathbf{1} & 0.33 \\ 0.08 & 0.08 & 0.08 & 0.08 & 0.08 & 0.33 & 0.33 & 0.33 & 0.33 & \mathbf{1} \end{array} \right)$$

(a)

$$\left(\begin{array}{ccccc|ccccc} \mathbf{1} & \mathbf{0.5} & \mathbf{0.5} & 0.25 & 0.25 & \mathbf{0} & \mathbf{0.25} & \mathbf{0} & 0 & 0 \\ \mathbf{0.5} & \mathbf{1} & \mathbf{0.25} & 0.25 & 0.25 & \mathbf{0.25} & \mathbf{0} & \mathbf{0} & 0 & 0 \\ \mathbf{0.5} & \mathbf{0.25} & \mathbf{1} & 0.25 & 0.25 & \mathbf{0} & \mathbf{0} & \mathbf{0.25} & 0 & 0 \\ 0.25 & 0.25 & 0.25 & \mathbf{1} & 0.25 & 0 & 0 & 0 & 0 & 0 \\ 0.25 & 0.25 & 0.25 & 0.25 & \mathbf{1} & 0 & 0 & 0 & 0 & 0 \\ \hline \mathbf{0} & \mathbf{0.25} & \mathbf{0} & 0 & 0 & \mathbf{1} & \mathbf{0.25} & \mathbf{0.5} & 0.25 & 0.25 \\ \mathbf{0.25} & \mathbf{0} & \mathbf{0} & 0 & 0 & \mathbf{0.25} & \mathbf{1} & \mathbf{0.25} & 0.25 & 0.25 \\ \mathbf{0} & \mathbf{0} & \mathbf{0.25} & 0 & 0 & \mathbf{0.5} & \mathbf{0.25} & \mathbf{1} & 0.25 & 0.25 \\ 0 & 0 & 0 & 0 & 0 & 0.25 & 0.25 & 0.25 & \mathbf{1} & 0.25 \\ 0 & 0 & 0 & 0 & 0 & 0.25 & 0.25 & 0.25 & 0.25 & \mathbf{1} \end{array} \right)$$

(b)

$$\left(\begin{array}{ccccc|ccccc} \mathbf{1} & \mathbf{0.5} & \mathbf{0.5} & 0.41 & 0.41 & \mathbf{0} & \mathbf{0.25} & \mathbf{0} & 0.08 & 0.08 \\ \mathbf{0.5} & \mathbf{1} & \mathbf{0.25} & 0.41 & 0.41 & \mathbf{0.25} & \mathbf{0} & \mathbf{0} & 0.08 & 0.08 \\ \mathbf{0.5} & \mathbf{0.25} & \mathbf{1} & 0.41 & 0.41 & \mathbf{0} & \mathbf{0} & \mathbf{0.25} & 0.08 & 0.08 \\ 0.41 & 0.41 & 0.41 & \mathbf{1} & 0.41 & 0.08 & 0.08 & 0.08 & 0.08 & 0.08 \\ 0.41 & 0.41 & 0.41 & 0.41 & \mathbf{1} & 0.08 & 0.08 & 0.08 & 0.08 & 0.08 \\ \hline \mathbf{0} & \mathbf{0.25} & \mathbf{0} & 0.08 & 0.08 & \mathbf{1} & \mathbf{0.25} & \mathbf{0.5} & 0.33 & 0.33 \\ \mathbf{0.25} & \mathbf{0} & \mathbf{0} & 0.08 & 0.08 & \mathbf{0.25} & \mathbf{1} & \mathbf{0.25} & 0.33 & 0.33 \\ \mathbf{0} & \mathbf{0} & \mathbf{0.25} & 0.08 & 0.08 & \mathbf{0.5} & \mathbf{0.25} & \mathbf{1} & 0.33 & 0.33 \\ 0.08 & 0.08 & 0.08 & 0.08 & 0.08 & 0.33 & 0.33 & 0.33 & \mathbf{1} & 0.33 \\ 0.08 & 0.08 & 0.08 & 0.08 & 0.08 & 0.33 & 0.33 & 0.33 & 0.33 & \mathbf{1} \end{array} \right)$$

(c)

FIGURE 3.1 – Representation of the "Relatedness Matrix" (a), "HS+FS Matrix" (b) and "Hybrid Matrix" (c), with $N_{fam} = 2$, $N_{off} = 5$ and $N_{th} = 3$ (genotyped individuals are represented in bold).

Method	Scenario 0		$N_{PC} = 1$		$N_{PC} = 5$		$N_{PC} = 15$	
	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE
Family	-7.26 ± 14.81	0.022	-5.23 ± 11.87	0.013	21.10 ± 15.53	0.029	28.1 ± 13.40	0.027
FamAdj	-7.83 ± 14.70	0.022	-29.52 ± 8.78	0.014	-9.78 ± 11.91	0.016	-4.14 ± 9.89	0.012
FamDoubleAdj	-6.45 ± 14.93	0.022	-7.28 ± 11.30	0.012	-5.82 ± 12.36	0.017	-3.21 ± 9.99	0.012
Relatedness	-6.30 ± 15.02	0.022	-7.32 ± 11.62	0.011	-6.11 ± 10.19	0.012	-5.94 ± 10.81	0.011
HS+FS	-5.91 ± 13.51	0.017	-12.23 ± 10.81	0.010	-4.94 ± 10.02	0.010	9.61 ± 9.76	0.011
Hybrid Matrix	-4.99 ± 13.55	0.017	-11.41 ± 8.09	0.010	-12.85 ± 9.41	0.009	-6.72 ± 7.98	0.008
PhiA	-5.94 ± 10.81	0.011	-8.12 ± 9.63	0.008	-6.03 ± 8.05	0.006	0.31 ± 7.77	0.006
ρ_w	0.25		0.34		0.34		0.33	
ρ_b	0.004		0.089		0.016		0.003	

TABLE 3.3 – Performance (bias and MSE) of the methods under the null model scenario ($N_{PC} = 1$, $\alpha = 5$) and variable levels of pollen pool differentiation ($N_{PC} = 1, 5, 15$) with unequal male fecundities ($\alpha = 0.01$; Scenario i). The mean genetic relatedness within (ρ_w) and among families (ρ_b) is indicated for each parameter set.

Method	$s = 0$			$s = 0.20$			$s = 0.40$		
	Bias	MSE	Cov	Bias	MSE	Cov	Bias	MSE	Cov
Family	7.70 ± 14.32	0.020	-	49.12 ± 16.86	0.053	-	97.84 ± 18.61	0.148	-
FamAdj	-0.97 ± 13.26	0.017	-	-6.69 ± 10.38	0.012	-	-12.55 ± 8.40	0.012	-
FamDoubleAdj	15.49 ± 16.56	0.028	-	11.71 ± 14.15	0.023	-	9.40 ± 12.28	0.021	-
Relatedness	5.53 ± 14.46	0.020	0.90	-3.41 ± 10.26	0.011	0.90	-16.27 ± 7.31	0.011	0.82
HS+FS	1.95 ± 12.44	0.015	0.88	25.30 ± 12.53	0.023	0.92	49.20 ± 12.12	0.046	0.82
Hybrid Matrix	-0.11 ± 12.23	0.015	0.88	-7.15 ± 10.54	0.012	0.92	-16.67 ± 7.28	0.011	0.76
PhiA	1.07 ± 9.50	0.009	0.94	-4.88 ± 6.83	0.005	1	-14.42 ± 6.42	0.008	0.88
ρ_w		0.27			0.40			0.57	
ρ_b		0.016			0.017			0.016	

TABLE 3.4 – Performance (bias, MSE and coverage) of the methods according to ancestral inbreeding ($s = 0.20, 0.40$) in unequal male fertilities scenarios ($\alpha = 0.1, N_{PC} = 5, N_g = 3$; Scenario (ii)).

Method	$V_M = 0$			$V_M = 0.04$			$V_M = 0.08$		
	Bias	MSE	Cov	Bias	MSE	Cov	Bias	MSE	Cov
Family	21.10 ± 15.53	0.029	-	73.56 ± 16.56	0.081	-	128.10 ± 21.30	0.223	-
FamAdj	-9.78 ± 11.91	0.016	-	30.65 ± 12.42	0.028	-	72.43 ± 16.36	0.086	-
FamDoubleAdj	-5.82 ± 12.36	0.017	-	58.91 ± 17.26	0.066	-	123.14 ± 25.91	0.238	-
Relatedness	-6.11 ± 10.19	0.012	0.94	-29.37 ± 22.84	0.067	0.56	-14.16 ± 24.37	0.075	0.58
HS+FS	-4.94 ± 10.02	0.010	0.94	-25.16 ± 17.31	0.038	0.84	-45.50 ± 12.26	0.040	0.76
Hybrid Matrix	-12.85 ± 9.41	0.009	0.92	-21.28 ± 13.11	0.024	0.82	-40.60 ± 11.98	0.035	0.84
PhiA	-6.03 ± 8.05	0.006	1	-11.60 ± 11.16	0.016	0.84	-15.36 ± 8.90	0.012	0.92
ρ_w		0.34			0.34			0.34	
ρ_b		0.016			0.016			0.016	

TABLE 3.5 – Performance (bias, MSE and coverage) of the methods according to two levels of maternal effects ($V_M = 0.04$ V_P , $V_M = 0.08$ V_P) in unequal male fertilities scenarios ($\alpha = 0.01$, $N_{PC} = 5$, $N_g = 1$; Scenario (iii)).

Method	True $\rho_{k,k'}$ coefficients + true pedigree			Loiselle coefficients + true pedigree			Loiselle + 20 % paternal assignment error		
	Bias	MSE	Cov	Bias	MSE	Cov	Bias	MSE	Cov
Family	21.10 ± 15.53	0.029	-	-	-	-	-	-	-
FamAdj	-9.78 ± 11.91	0.016	-	-11.23 ± 11.82	0.016	-	-	-	-
FamDoubleAdj	-5.82 ± 12.36	0.017	-	-3.06 ± 12.69	0.017	-	-	-	-
Relatedness	-6.11 ± 10.19	0.012	0.94	-4.60 ± 11.24	0.014	0.94	-	-	-
HS+FS	-4.94 ± 10.02	0.010	0.94	-	-	-	-3.57 ± 10.34	0.012	0.92
Hybrid Matrix	-12.85 ± 9.41	0.009	0.92	-9.26 ± 9.56	0.010	0.94	-10.19 ± 9.52	0.011	0.92
Hybrid $\tilde{\rho}_{k,k'}$	-	-	-	-20.33 ± 7.80	0.010	0.88	-	-	-
PhIA / PhIA $_{\tilde{\rho}_{k,k'}}$	-6.03 ± 8.05	0.006	1	-36.49 ± 4.10	0.014	0.48	-	-	-
ρ_w	0.34			0.34			0.34		
ρ_b	0.016			0.016			0.016		

TABLE 3.6 – Performance (bias, MSE and coverage) of the methods using realistic genetic data, i.e. Loiselle coefficients for the "FamAdj", "FamDoubleAdj", "Relatedness" and "Hybrid Matrix" methods and pedigree with errors for the "FS+HS" and "Hybrid Matrix" methods. "Hybrid $\tilde{\rho}_{k,k'}$ ": the $\tilde{\rho}_{k,k'}$ coefficients were used for the genotyped individuals instead of the pedigree information $\tilde{\rho}_{k,k'}^{(1)}$. The simulation parameters were $\alpha = 0.01$ and $N_{PC} = 5$ (scenario i).

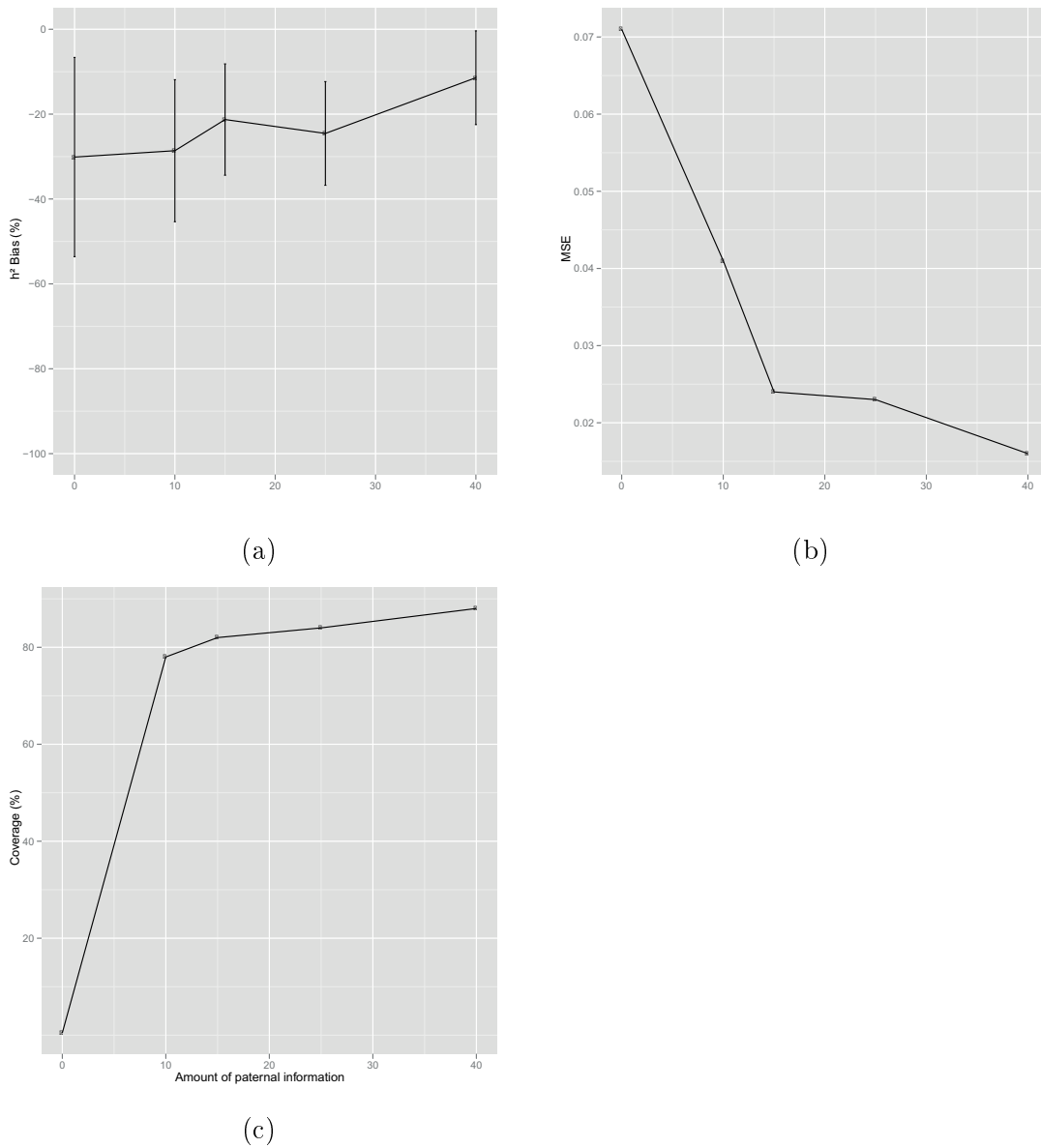


FIGURE 3.2 – Performance of the "Hybrid Matrix" method according to the amount of paternal information available (N_{th}) in scenario *iii* with $V_M = 0.04 V_P$, using the (a) bias, (b) MSE and (c) coverage indices. The case where no paternal information is available correspond to the "Relatedness" method with $N_{th} = 0$.

Appendix A : Algorithm used to substitute a definite-positive relatedness matrix in the animal model

Filling a relatedness matrix A with pairwise relatedness coefficients estimated from molecular markers (relying on Loiselle estimates for instance) frequently leads to a non definite-positive matrix, thus non-inversible (Frentiu *et al.*, 2008). This was particularly frequent for the hybrid matrices investigated in the manuscript (e.g. for the "**Hybrid**" method).

For these problematic cases, AsReml proposes to use the "!NSD" command, which allows the matrix to be negative (semi)definite (Gilmour et al. 2006). But this solution does not provide robust estimates of the variance components : the heritability was largely underestimated (bias= -44.52 %, -19.87 %, -11.06 % for scenario i and $N_{PC} = 1, 5, 15$ respectively). Below we propose a modification of the relatedness matrix that makes it definite-positive.

Consider the symmetric relatedness matrix A of dimension $n \times n$ and the spectral decomposition of A as

$$A = U\Lambda U^T$$

where U is an upper triangular matrix of coefficients representing the eigenvectors and Λ is the diagonal matrix containing the eigenvalues $\lambda_1 \geq \dots \geq \lambda_n$ of A . If A is non definite-positive, the $n - q > 0$ last eigenvalues are negative or null and $\lambda_q > 0$.

We modify the matrix Λ into a diagonal matrix $\bar{\Lambda}$ containing the following values $\bar{\lambda}_1 \geq \dots \geq \bar{\lambda}_n$:

$$\bar{\lambda}_k = \lambda_k, \text{ for } k \leq q$$

$$\bar{\lambda}_k = \lambda_q h^{k-q}, \text{ for } k > q$$

where $h < 1$ controls the decrease of the modified $\bar{\lambda}$'s towards 0. A too small h value can lead to numerical errors. We used $h = 0.9$.

This correction means that the non-positive eigenvalues are replaced by a geometrically decreasing serie of strictly positive eigenvalues.

The modified relatedness matrix \bar{A} is then obtained as

$$\bar{A} = U\bar{\Lambda}U^T$$

This matrix \bar{A} was substituted for A in the animal model analysed with AsReml. All matrix inversions of \bar{A} 's were completed correctly by AsReml (no warning message).

3.3 Déterminismes génétiques, maternels et environnementaux des traits fonctionnels dans des populations naturelles de *Fagus sylvatica*

Title : Genetic, maternal and environmental determinants of growth, morphological, physiological and phenological traits in natural populations of *Fagus sylvatica*

Authors : Gauzere J., Oddou-Muratorio S., Klein E. K.

Introduction

The adaptive potential of natural populations in face of environmental variations will mainly depends on the phenotypic variation of the adaptive trait (V_P), the strenght of the selective pressure due to environmental change (S) and the genetic component of the trait variation (heritability, h^2). The rate of evolution of a given trait (R) can generally be predicted following the simple model : $R = V_P \times S \times h^2$ (Lynch & Walsh, 1998). For long-lived species for which S is hard to measure experimentally, V_P and h^2 are usually considered as useful indicators of the adaptive potential (DiBattista *et al.*, 2009; Scotti *et al.*, 2010). A major challenge for estimating the adaptive potential based on V_P and h^2 is to identify the "good" adaptive traits, that is the traits likely to contribute to adaptation to future climate. In the context of current climate change, interesting candidate traits appear to be the functional traits involved in the carbon assimilation by photosynthesis and in the response to drought because of increased transpiration and water loss, which can conduct to mortality (McDowell *et al.*, 2008). Phenology, by determining the length of the vegetation season, mainly affects the net carbon assimilation and thus the fitness of trees. Leaf ecophysiological traits related to photosynthetic capacity have also important effects on carbon assimilation (Wright *et al.*, 2004). Among these fitness-related traits, phenology is known to be highly affected by temperature changes (Bertin, 2008). On the other hand, in tree species, our current understanding of the response of ecophysiological traits to drought stress is very limited, even though a rich litterature exists on plant responses (Dudley, 1996; Arntz & Delph, 2001; Geber & Griffen,

2003; Donovan *et al.*, 2009). However, this literature is very useful to identify good candidate traits, like the water use efficiency, leaf size and leaf thickness, that will likely affect tree adaptation to future climate.

The estimation of the phenotypic and genetic variances of adaptive traits in natural tree populations are mainly based on open-pollinated families managed in seed orchard trials (Vargashernandez & Adams, 1991; Kumar & Richardson, 2005; Gaspar *et al.*, 2009). Open-pollinated families are classically maternal families and thus, members of the same family are at least related by a coefficient of relatedness $\rho = 0.25$, corresponding to the genetic correlation among half-sibs (Falconer & MacKay, 1996). The great success of these progeny trials is based on their simplicity to set up and the easy calculations linking the genetic variance among families, V_F , to the part of inheritant phenotypic variance, V_A : $V_A = \frac{1}{\rho_w} V_F$, with ρ_w the average genetic relatedness among families. Based on this formula, the heritability of a given trait, h^2 , is generally deduced as : $h^2 = \frac{1/\rho_w \times V_F}{V_F + V_W}$, with V_W the variance within families. These estimates are classically derived from a mixed model considering the identity of the mother as a random effect and called the family model.

Because the true relatedness relationships among offspring is generally unknown, ρ_w is often assumed to be 0.25 and the average genetic relatedness among families, ρ_b , expected to be null. However, these assumptions of true half-sib families are unlikely hold in many natural populations, where mating events do not occur at random (Gauzere *et al.*, 2013b). Moreover, the family model assumes that the variance among families, V_F , only includes additive genetic variance. However, in tree progeny trials, the phenotypes are measured on juveniles, and therefore are expected to be under more non-additive genetic control than adult traits, notably because of maternal effects (Roach & Wiulff, 1987). Classically, finest pairwise relatedness relationships among offspring are used with models defined at the individual level (the so-called animal model) to avoid the use of family assumptions (Kruuk, 2004).

Here, we focus on a major economic tree species in Europe, *Fagus sylvatica* L. (european beech), a temperate species known to be sensitive to drought. On Mont-Ventoux, a mountain in the South-East of France culminating at 1911 m a.s.l, beech populations grow in Mediterranean climate corresponding to the marginal conditions of their ecological distribution range. The aim of this study was to evaluate the heritability of growth, morphological, physiological

and phenological traits for *F. sylvatica* in a Mediterranean ecosystem. To that purpose, we used 60 open-pollinated families from 3 beech populations located in the northern face of the Mont-Ventoux. *F. sylvatica* is a monoecious tree species, predominantly outcrossed and anemophilous (Merzeau *et al.*, 1994; Wang, 2003). However, in the studied sites, Gauzere *et al.* (2013a) have shown a wide variability of the mating system characteristics between trees within population, with selfing rates varying from 0 to 48 % and the number of effective pollen donors varying from 2 to 364 among mother-trees. Thus, we can expect paternal relatednesses to be non-negligible within and between the maternal families collected. More precisely we (i) investigate the departures from the assumptions of true-half-sibs in a progeny trial of beech, (ii) compare different methods to estimate h^2 based the family and animal models previously tested with simulations (Gauzere *et al.* (a) *in prep*), (iii) estimate h^2 for different potentially adaptive traits and (iv) discuss the relevance of the estimates derived from open-pollinated trials.

Material and Method

HALF-SIB TRIAL

The progeny trial used in this study included 60 open-pollinated families of *Fagus sylvatica* from 3 populations, N1, N2 and N4, along an elevational gradient on the north-face of the Mont-Ventoux. In August 2009, in each population 20 highly fertile and randomly distributed trees were chosen as mother-trees to collect on average 344.3 seeds per mother-tree (min= 202, max=733) directly from the canopy. All the seeds were dried to a humidity rate of 8 % and for each family, a sample of 100 seeds was selected to measure the rate of empty and parasited seeds using X-ray photography. The global weight of each sample was also measured. Seeds were then rehydrated and conserved at +4 ° C during 10 weeks to break the dormancy and initiate the germination. In April 2010, a total of 5475 seedlings were successfully germinated (91.25 seedlings per family on average) and were transferred in a common garden at the State nursery of Aix-Les-Milles (43 ° 30'N 5 ° 24'E).

In the common garden, all seedlings were planted in independent pots of 1.2 L with sand substrate and fertilizer, arranged in 50 complete-blocks in which 1.8 seedlings per family on average were randomized. Transplanting was done periodically by blocks, in order to keep the

differences due to germination included in the block effect. Half of the trial (25 blocks, 2783 seedlings) was regularly irrigated while the other half (25 blocks, 2692 seedlings) received a water stress treatment. Note that in 349 pots identified "a posteriori" the fertilizer was missing. The seedlings grew 3 years in these common garden conditions.

PATERNITY ANALYSIS

In total 2088 offspring (on average of 34.8 offspring per family ; Table 3.7) were genotyped at 13 microsatellite markers : FS1-15, FS3-4 (Pastorelli *et al.*, 2003) ; sfc007-2, sfc-1143, sfc-0161 (Asuka *et al.*, 2004) ; mfc7 (Vornam *et al.*, 2004) ; Csofagus-19, Csofagus-6, Csofagus-29, Csofagus-31 (Lefevre *et al.*, 2012) ; Csofagus-7, Csofagus-25 and Fi05 (G.G. Vendramin, personal communication ; see Gauzere *et al.* (2013a) for the genotyping details). The potential reproductive adults of each natural population (690 in total) were exhaustively genotyped at the same markers (Gauzere *et al.*, 2013a). Note that here we used the genotype dataset published in Gauzere *et al.* (2013a) (supplementary material) completed with 710 additional genotyped seedlings and 189 genotyped adults (17 adults added for N1 ; 2 adults added for N2 ; 170 adults added for N4).

This genetic dataset was first used to reconstruct the "one-generation" paternal pedigree for all the genotyped offspring. We used the likelihood-based software CERVUS version 3.0 (Marshall *et al.*, 1998). Using mendelian probabilities of inheritance, this method tests if (H1) a given potential father is the true father against the hypothesis (H2) a random individual in the population is the true father. For each seedling the most-likely male (M1) is compared to the second most-likely (M2) using the criterion Δ , which represents the difference of LOD-score between M1 and M2. The father M1 is assigned if Δ is higher than Δ_c , a critical value evaluated from simulations. For each offspring, paternity was assigned to M1 if (1) more than 6 loci matched between the offspring and its potential father, (2) the Δ of the trio "offspring-mother-father" was significant (i.e. $> \Delta_c$), considering a strict confidence level of 95 % and (3) the potential father belong to the same population than the mother. Δ_c was determined simulating 10000 offspring and considering (i) the global allelic frequencies calculated from the 3 populations, (ii) 0 % of typing error, (iii) 100 % sampling of candidate fathers and (iv) allowing selfing events. Paternity analyses allowed to estimate the realized reproductive success (also called fertility) of each potential reproductive tree within each population.

Fam	Ntot	Ng	Nass	Fam	Ntot	Ng	Nass	Fam	Ntot	Ng	Nass
F01	96	34	15	F21	105	40	12	F41	97	34	14
F02	99	39	33	F22	96	39	7	F42	100	34	11
F03	97	34	16	F23	39	23	11	F43	96	33	19
F04	102	38	20	F24	23	16	11	F44	106	38	18
F05	99	35	17	F25	100	47	22	F45	99	37	18
F06	101	35	21	F26	41	24	7	F46	96	36	16
F07	98	36	17	F27	96	47	19	F47	101	34	24
F08	99	32	13	F28	102	41	26	F48	98	38	16
F09	95	38	19	F29	18	12	10	F49	95	35	8
F10	99	37	33	F30	99	36	21	F50	97	36	19
F11	98	33	16	F31	31	24	13	F51	99	35	10
F12	70	29	17	F32	100	39	8	F52	98	36	15
F13	95	32	16	F33	98	36	10	F53	101	35	15
F14	98	35	24	F34	31	16	11	F54	97	35	11
F15	101	39	23	F35	100	39	13	F55	98	34	13
F16	99	36	21	F36	95	39	19	F56	98	38	20
F17	101	37	14	F37	109	37	12	F57	103	39	10
F18	100	31	22	F38	102	43	18	F58	101	37	20
F19	97	33	19	F39	73	40	7	F59	100	39	22
F20	95	31	21	F40	99	36	30	F60	99	37	17

TABLE 3.7 – Total number of offspring sampled (N_{tot}), genotyped (N_g) and assigned to a father within the population (N_{ass}) per family. Families F01 to F20 were sampled in population N1, families F21 to F40 in population N2, and families F41 to F60 in population N4 (column "Fam").

RELATEDNESS ESTIMATES

The complete genetic dataset (2088 genotyped offspring from the whole trial) was also used to estimate paternal relatedness coefficients within and between maternal families. First, knowing the mother genotypes, the haplotypes corresponding to the paternal contribution were extracted from the diploid genotype of the offspring in double heterozygotes cases (Table 3.8). To deal with ambiguous characterization of the paternal alleles, the paternal haplotypes were described as diploid genotypes : at each locus, if the paternal contribution can be unambiguously deduced, the paternal allele was doubled ; in ambiguous cases the two possible alleles of the father were recorded (Table 3.8 ; following Smouse *et al.* 2001).

Diploid offspring genotype	Diploid maternal genotype	Gametic paternal genotype	Recorded paternal contribution
AA	AA	A	AA
AB	AA	B	BB
AB	AB	A or B	AB
AB	BB	A	AA

TABLE 3.8 – Inference of the paternal contribution to the diploid offspring genotype at a locus knowing the maternal genotype.

These paternal haplotypes were then used to estimate pairwise paternal relatednesses among individuals, $\tilde{f}_{k,k'}$, using the software SPAGEDI (Hardy & Vekemans, 2002) and the kinship coefficient of Loiselle *et al.* (1995). The allelic frequencies of the 690 adults, considered as a single population, were used as the reference allelic frequencies. The genetic relatednesses of pairs of individuals within the same family were then recomputed as : $\tilde{\rho}_{k,k'} = 0.25 + \tilde{f}_{k,k'}$, and for pairs of individuals from different families : $\tilde{\rho}_{k,k'} = \tilde{f}_{k,k'}$. Mean genetic relatedness within and among families ($\tilde{\rho}_w$ and $\tilde{\rho}_b$ respectively) were calculated as :

$$\tilde{\rho}_{w,j} = \sum_{k \neq k' \in \text{family}(j)} \frac{\tilde{\rho}_{k,k'}}{n_j(n_j-1)/2} \quad \tilde{\rho}_w = \sum_{k \in j} \frac{\tilde{\rho}_{w,j}}{n_j}$$

$$\tilde{\rho}_{b,jj'} = \sum_{k \in j} \sum_{k' \in j'} \frac{\tilde{\rho}_{k,k'}}{n_j n_{j'}} \quad \tilde{\rho}_b = \sum_{k \in j} \frac{\tilde{\rho}_{b,jj'}}{n_j}$$

with n_j the sample size of the genotyped offspring of family j .

We analysed the relationship between the average paternal relatedness among families $\tilde{\rho}_b$ and the pairwise distances between mother-trees simply by plotting these two variables (Robledo-Arnuncio *et al.*, 2007). Negative $\tilde{\rho}_b$ were replaced by zero to be used in the quantitative genetic models (following Gay *et al.* 2013).

PHENOTYPIC TRAITS

Growth traits We measured the total growth in height (δH) and diameter (δD) between august 2010 and november 2011. The total biomass production was estimated as : $BioM = h_{2011} \times \pi \times (d_{2011}/2)^2$, where h_{2011} and d_{2011} were respectively the height and diameter measured in November 2011.

Phenological traits The timing of bud burst was monitored weekly from april to may

in springs 2011 and 2012. Six stages were used to follow the bud burst dynamic : 1 : buds are dormant or swelling ; 2 : bud scales are broken ; 3 : at least 15 % of the leaves are emerging ; 4 : at least 50 % of the leaves are emerging ; 5 : leaves are spread out but have not reached their mature sizes (Annexe A). The timing of leaf senescence was monitored weekly from october to november in autumn 2011. Three stages were used to follow senescence dynamics : 1 : leaves have not fallen and are not colored ; 2 : at least 10 % of the leaves are colored or have fallen ; 3 : at least 50 % of the leaves are colored or have fallen (Annexe B). For each survey, phenology was always monitored by the same two groups of observers.

We focused on two critical stages of bud burst and senescence dynamics : for bud burst development, stage 3 corresponds to the most sensitive stage to frost damages ; for leaf senescence, we focused on stage 2 because, below 50 % of leaf coloration, the phenological process can be confound with effects of water stress in summer. We used an linear interpolation to estimate the date of passage from stage 2 to 3, $t_{b2 \rightarrow 3}$, for bud burst phenology and the date of passage from stage 1 to 2, $t_{s1 \rightarrow 2}$, for senescence. In 2011, the duration of the vegetative season was characterised by the lag between the dates of bud burst and leaf senescence as : $VD = t_{s1 \rightarrow 2} - t_{b2 \rightarrow 3}$. Note that the dates of bud burst in spring 2012 presented a larger spread than in 2011 since in winter 2012 many buds had frozen. Damages due to frost were recorded in June 2012 by visually estimating the percentage of empty buds per seedling after bud burst. Only the subset of seedlings with less than 25 % of buds damaged was analysed the second year of monitoring (Table 3.9).

At the end of summer 2011, three light-exposed leaves on the main branch of each seedling were collected to mesure morphological and physiological traits.

Morphological traits We first measured the fresh leaf area (LA) with a planimeter. The leaves were then dried at 60 °C during about 3 days to finally record the leaf dry mass (LM). The leaf mass area was calculated as $LMA = \frac{LM}{LA}$, with LM the dry leaf mass.

Physiological traits Leaf nitrogen and carbon content ($\%N$ and $\%C$ respectively), and carbon isotope composition ($\delta^{13}C$) were measured on a subset of 1574 individuals (1034 in the irrigated and 540 in the water stress treatment), representative of all populations, families and blocks. A sample of 1 mg of dry leaves was analysed using an isotopic ratio spectrometer and $\delta^{13}C$ was calculated as :

Abbreviation	Trait	Category	N_w	Unit
δH	Total height increase	Growth	2403	cm
δD	Total diameter increase	Growth	2403	cm
$BioM$	Total biomass produced	Growth	2406	cm ²
$t_{b2 \rightarrow 3(2011)}$	Bud burst date in 2011	Phenology	2605	Degree days
$t_{b2 \rightarrow 3(2012)}$	Bud burst date in 2012	Phenology	1854	Degree days
$t_{s1 \rightarrow 2}$	Senescence date in 2011	Phenology	2412	Degree days
VD_{2011}	Vegetation season duration in 2011	Phenology	2412	Degree days
LM	Dry leaf mass	Morphology	2518	g
LA	Leaf Area	Morphology	2518	cm ²
LMA	Leaf Mass Area	Morphology	2518	g m ⁻²
$\%N$	Nitrogen content in leafs	Physiology	1038	%
$\%C$	Carbone content in leafs	Physiology	1038	%
$\delta^{13}C$	Water use efficiency	Physiology	1038	‰

TABLE 3.9 – Description of the traits measured, with the number of phenotyped individuals in the irrigated treatment N_w . For $t_{b2 \rightarrow 3(2012)}$, N_w represents the amount of individuals phenotyped with less than 25 % of the buds damaged by the frost.

$$\delta^{13}C = \frac{R_{sa}R_{sd}}{R_{sd}} \times 1000$$

where R_{sa} and R_{sd} are the isotopic ratio $^{13}C/^{12}C$ of the sample and the standard respectively. $\delta^{13}C$ is used as a surrogate of the intrinsic water use efficiency (WUE_i) at the leaf level. WUE_i is defined as the ratio of net CO_2 assimilation rate (A) to stomatal conductance for water vapour (g_w).

For each trait we calculated the global intra-population coefficient of variability, CV , as :

$$CV_t = \frac{\sigma(P_t)}{\mu(P_t)}$$

with $\sigma(P_t)$ and $\mu(P_t)$ the standard deviation and mean of the phenotypic trait t . CV was not calculated for $\delta^{13}C$ because $CV_{\delta^{13}C}$ is directly dependent on the isotopic ratio of the standard used (Brendel, 2014).

HERITABILITY ESTIMATES

Here, we focus on the expression of the genetic and phenotypic variance under irrigated

water condition, which is comparable to classical common garden experiments found in the litterature.

Family model First, we used the classical family model to estimate the genetic variance of each trait P . This model is a mixed model considering the mother identity (Fam_j) as a random effet with variance V_F . For each seedling k of the family j the model writes :

$$P_{p,m,n,j,k} = \mu + B_m + E_n + M_p + Fam_j + \epsilon_{p,m,n,j,k} \quad (3.7)$$

with $\epsilon_{p,m,n,j,k}$ a residual random effect with variance V_R . All the other terms are fixed effects, with μ the mean phenotypic effect, B_m the effect of the blocks, E_n the effect of the

$$\left(\begin{array}{ccccc|ccccc} \mathbf{1} & \mathbf{0.5} & \mathbf{0.5} & 0.25 & 0.25 & \mathbf{0} & \mathbf{0.25} & \mathbf{0} & 0 & 0 \\ \mathbf{0.5} & \mathbf{1} & \mathbf{0.25} & 0.25 & 0.25 & \mathbf{0.25} & \mathbf{0} & \mathbf{0} & 0 & 0 \\ \mathbf{0.5} & \mathbf{0.25} & \mathbf{1} & 0.25 & 0.25 & \mathbf{0} & \mathbf{0} & \mathbf{0.25} & 0 & 0 \\ 0.25 & 0.25 & 0.25 & \mathbf{1} & 0.25 & 0 & 0 & 0 & 0 & 0 \\ 0.25 & 0.25 & 0.25 & 0.25 & \mathbf{1} & 0 & 0 & 0 & 0 & 0 \\ \hline \mathbf{0} & \mathbf{0.25} & \mathbf{0} & 0 & 0 & \mathbf{1} & \mathbf{0.25} & \mathbf{0.5} & 0.25 & 0.25 \\ \mathbf{0.25} & \mathbf{0} & \mathbf{0} & 0 & 0 & \mathbf{0.25} & \mathbf{1} & \mathbf{0.25} & 0.25 & 0.25 \\ \mathbf{0} & \mathbf{0} & \mathbf{0.25} & 0 & 0 & \mathbf{0.5} & \mathbf{0.25} & \mathbf{1} & 0.25 & 0.25 \\ 0 & 0 & 0 & 0 & 0 & 0.25 & 0.25 & 0.25 & \mathbf{1} & 0.25 \\ 0 & 0 & 0 & 0 & 0 & 0.25 & 0.25 & 0.25 & 0.25 & \mathbf{1} \end{array} \right)$$

(a)

$$\left(\begin{array}{ccccc|ccccc} \mathbf{1} & \mathbf{0.5} & \mathbf{0.5} & 0.41 & 0.41 & \mathbf{0} & \mathbf{0.25} & \mathbf{0} & 0.08 & 0.08 \\ \mathbf{0.5} & \mathbf{1} & \mathbf{0.25} & 0.41 & 0.41 & \mathbf{0.25} & \mathbf{0} & \mathbf{0} & 0.08 & 0.08 \\ \mathbf{0.5} & \mathbf{0.25} & \mathbf{1} & 0.41 & 0.41 & \mathbf{0} & \mathbf{0} & \mathbf{0.25} & 0.08 & 0.08 \\ 0.41 & 0.41 & 0.41 & \mathbf{1} & 0.41 & 0.08 & 0.08 & 0.08 & 0.08 & 0.08 \\ 0.41 & 0.41 & 0.41 & 0.41 & \mathbf{1} & 0.08 & 0.08 & 0.08 & 0.08 & 0.08 \\ \hline \mathbf{0} & \mathbf{0.25} & \mathbf{0} & 0.08 & 0.08 & \mathbf{1} & \mathbf{0.25} & \mathbf{0.5} & 0.33 & 0.33 \\ \mathbf{0.25} & \mathbf{0} & \mathbf{0} & 0.08 & 0.08 & \mathbf{0.25} & \mathbf{1} & \mathbf{0.25} & 0.33 & 0.33 \\ \mathbf{0} & \mathbf{0} & \mathbf{0.25} & 0.08 & 0.08 & \mathbf{0.5} & \mathbf{0.25} & \mathbf{1} & 0.33 & 0.33 \\ 0.08 & 0.08 & 0.08 & 0.08 & 0.08 & 0.33 & 0.33 & 0.33 & \mathbf{1} & 0.33 \\ 0.08 & 0.08 & 0.08 & 0.08 & 0.08 & 0.33 & 0.33 & 0.33 & 0.33 & \mathbf{1} \end{array} \right)$$

(b)

FIGURE 3.3 – Representation of the relatedness matrices considering incomplete pedigree information (a) and hybrid relatedness information (b), with $N_{family} = 2$, $N_{offspring} = 5$ and $N_{assigned} = N_{genotyped} = 3$ (represented in bold). In the matrix (b), $\tilde{\rho}_{w,1} = 0.41$, $\tilde{\rho}_{w,2} = 0.33$ and $\tilde{\rho}_{b,12} = 0.08$

presence/absence of the fertilizer and M_p the effects of the observers (only included for the analysis of the phenological traits).

This model was run using the "lme" function of the R software. For each trait, we tested if the random family effect was significant, comparing hierarchical models with and without random family component with the "anova" procedure. The residuals of these family models were also analysed to decide if trait's transformations were necessary. Notably the homoscedasticity was tested analysing the correlation structure between the absolute value of the residuals and the fitted values of the models ; a significant correlation indicating heteroscedasticity of the residual variance.

Family and residual variance components (V_F and V_R) were used to derive classical h^2 estimates, called h_{Fam}^2 , assuming true half-sib families, and corrected h^2 estimates, called $h_{doubleadj}^2$, accounting for average genetic relatedness estimated with the genetic markers $\tilde{\rho}_w$ and $\tilde{\rho}_b$ (described in Gauzere et al. (a) *in prep* ; Squillace 1974) :

$$h_{Fam}^2 = \frac{4V_F}{V_F + V_R} \quad (3.8)$$

$$h_{doubleadj}^2 = \frac{V_F}{(\tilde{\rho}_w - \tilde{\rho}_b)V_R + \tilde{\rho}_w V_F} \quad (3.9)$$

The confidence intervals for h_{Fam}^2 and $h_{doubleadj}^2$ were derived using the delta method, whose general formulation is the following :

$$V(f) \simeq \nabla f \times \Sigma \times \nabla f^T$$

where f is a function of several random variables say a and b . If a , b can be estimated and if Σ is the variance-covariance matrix of a and b , thus $V(f)$ can be deduced using ∇f the vector of derivatives of f according to a and b , Σ and ∇f^T the transpose of ∇f . Here, we applied this method with $f = h^2(V_F, V_R)$, using the variances of (V_F, V_R) estimated by the function "apVar" on a lme object in R.

Animal model Second, we used the animal model to account for the available pairwise relatedness information. This model considered the same fixed effects than the family model

but included a random additive genetic effect $a_{j,k}$ for individual k in family j :

$$\begin{aligned} P_{p,m,n,j,k} &= \mu + B_m + E_n + M_p + a_{j,k} + \epsilon_{p,m,n,j,k} \\ \{a_{j,k}\} &\sim N(0, A \times V_A) \\ \{\epsilon_{p,m,n,j,k}\} &\sim N(0, I_{de} \times V_R) \end{aligned} \quad (3.10)$$

where the variance-covariance matrix of the additive genetic effects, G , is given by $G = A \times V_A$, where A is the genetic relatedness matrix and V_A the additive genetic variance ; the variance-covariance matrix of the residual effects is given by $I_{de} \times V_R$, with I_{de} the identity matrix and V_R the residual variance. The heritability was estimated as :

$$h^2 = \frac{V_A}{V_R + V_A} \quad (3.11)$$

Finally we also used an animal model accounting for a maternal random effect :

$$P_{p,m,n,j,k} = \mu + B_m + E_n + M_p + a_{j,k} + m_j + \epsilon_{p,m,n,j,k} \quad (3.12)$$

with m_j the maternal effect associated to family j . As for the ϵ , all m_j were assumed independent, with variance V_M . Using this model, h^2 and h_{mat}^2 were estimated as :

$$h^2 = \frac{V_A}{V_R + V_M + V_A} \quad (3.13) \quad h_{mat}^2 = \frac{V_M}{V_R + V_M + V_A} \quad (3.14)$$

The influence of maternal effects was tested comparing the hierarchical models (3.10) and (3.12) with a loglikelihood ratio test (Wilson *et al.*, 2010). For traits with significant maternal effects, we tested if the average phenotypic values were correlated to the average weight of the seeds measured at the family level.

Animal models were used considering two different relatedness matrices A : (a) the incomplete pedigree information available for the assigned individuals (method "HSFS" ; El-Kassaby *et al.* 2011 ; Gauzere *et al.* (a) *in prep*), and (b) hybrid relatedness information, i.e. the one-generation pedigree information for the assigned individuals and the average relatedness information within and between each family ($\tilde{\rho}_{w,j}$ and $\tilde{\rho}_{b,jj'}$) for the unassigned and ungenotyped individuals (method "MH" ; Figure 3.3 ; Gauzere *et al.* (a) *in prep*). In order to optimise the use of the available genetic information, $\tilde{\rho}_{w,j}$ and $\tilde{\rho}_{b,jj'}$ were estimated based

on the whole offspring genotyped. We neglected the paternity errors in the pedigree reconstructed from paternity analysis. Indeed, previous simulation studies have shown that these errors faintly affected the h^2 estimates with the animal model (Charmantier & Garant, 2005; DiBattista *et al.*, 2009). Finally, in a simulation study, Gauzere *et al.* (a) (*in prep*) showed that the "MH" method provided a more reliable h^2 estimates than the "HSFS" method if progenies were sampled from populations with related adults; we thus chose this method to estimate the most accurate h^2 values. Analyses of the animal models were performed with the software AsReml v3.0.5 (Gilmour *et al.*, 2006).

Results

MALE REPRODUCTIVE SUCCESS AND VARIATION OF RELATEDNESS

Among the genotyped seedlings of the whole trial, 47.9 % were assigned to a father, 15.8 % were unassigned because they were compatible with several fathers and 36.3 % were unassigned because there was no compatible father in the populations (Table 3.7). More precisely, we get the paternal relatedness information for 658 offspring in the irrigated condition, that is 24 % of the phenotyped individuals in this part of the trial. Male reproductive success was highly unbalanced, with 60 % of the adults with null fertilities, many individuals with low fertilities and few individuals with high fertilities (Figure 3.4).

Average paternity relatednesses estimated from molecular markers showed a weak departure from half-sib assumptions, with $\tilde{\rho}_w = 0.276$ and $\tilde{\rho}_b = 0.001$ (Figure 3.5). Paternal relatednesses tended to increase more $\tilde{\rho}_w$ than $\tilde{\rho}_b$ (deviation of 0.026 and 0.001 respectively). Population N1 tended to have the highest $\tilde{\rho}_w$ and the most variable $\tilde{\rho}_{b,jj'}$ and population N2 the more variable $\tilde{\rho}_w$ (Figure 3.5). At the scale of the Mont-Ventoux, one could have expected maternal pollen pools within each population to be more related than maternal pollen pools from different populations. However, the $\tilde{\rho}_{b,jj'}$ estimated between families from different populations (red bloxplot in the Figure 3.5) were only slightly lower than $\tilde{\rho}_{b,jj'}$ estimated between families from the same population, and even within the same range than the $\tilde{\rho}_{b,jj'}$ estimated within population N1. The correlation between $\tilde{\rho}_{b,jj'}$ and the pairwise distances among mother-trees allow to understand these slight differences of pollen pool structures within and between populations. Indeed, we can see that $\tilde{\rho}_{b,jj'}$ decreased quickly with distance. Above

the threshold of $d_{PP} \sim 50$ m, that corresponds to the average pollen dispersal distance ($\delta = 46$ m on average ; Gauzere *et al.* 2013a), relatednesses among the maternal pollen pools were

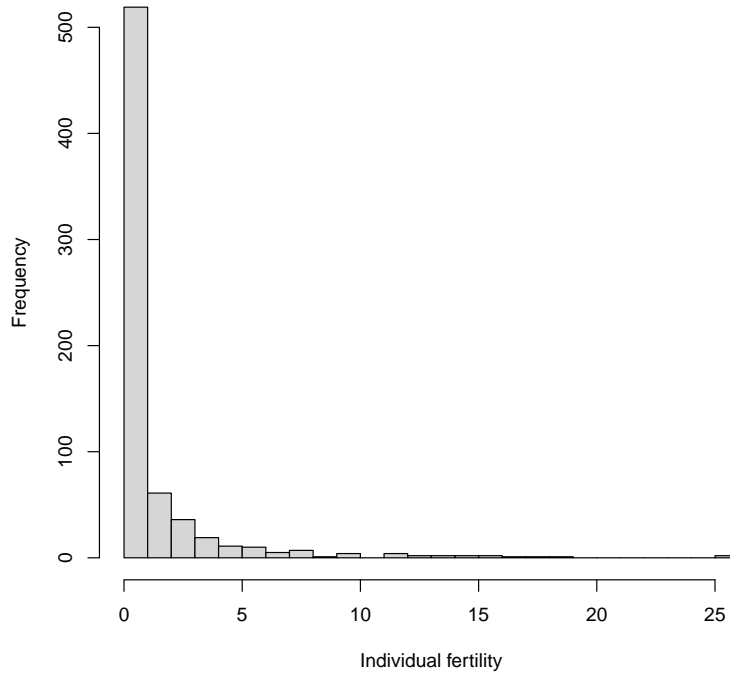


FIGURE 3.4 – Distribution of the individual male fertilities estimated with the paternity analyses.

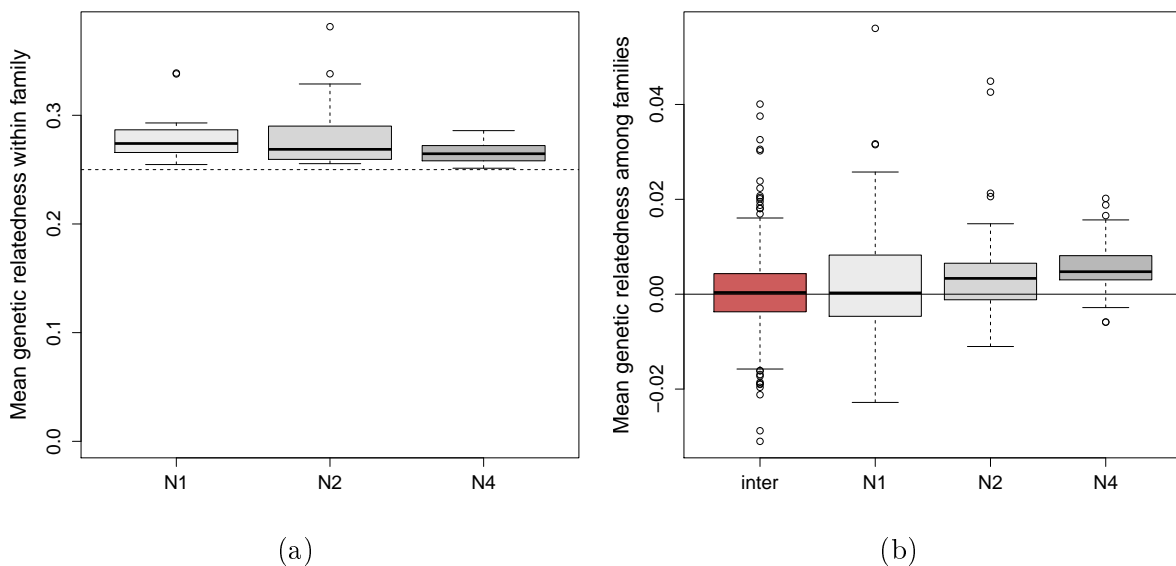


FIGURE 3.5 – (a) Variation of average genetic relatedness within families, $\tilde{\rho}_{w,j}$, within-population. The dashed line represents the relatedness value of 0.25. (b) Variation of average genetic relatedness between families, $\tilde{\rho}_{b,jj'}$, within- and between-populations (grey and red boxplots respectively).

almost no more affected by the distance (Figure 3.6).

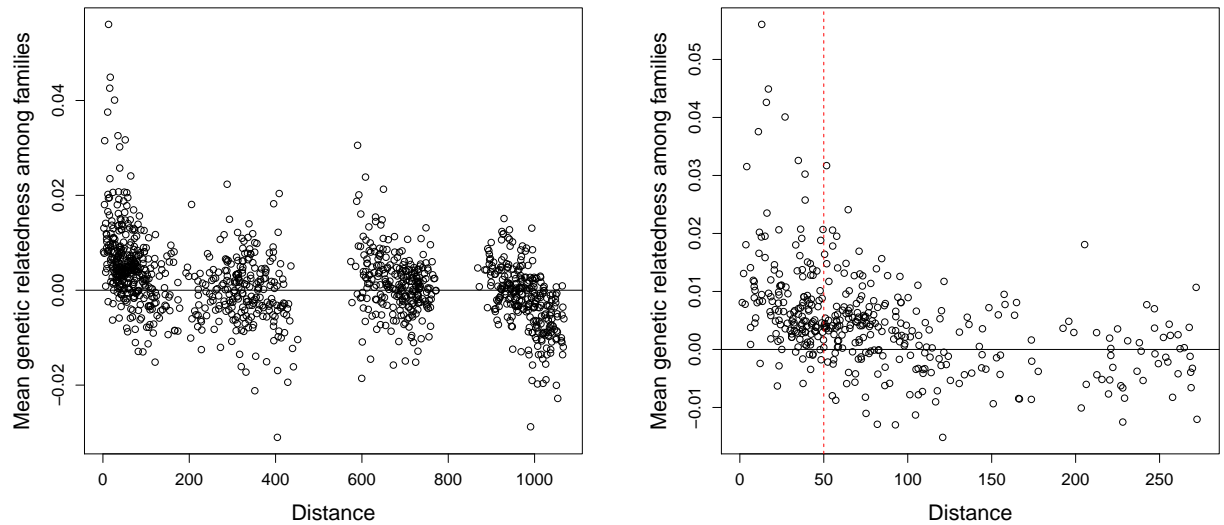


FIGURE 3.6 – Relationship between the mean genetic relatedness between a pair of families and the distance separating the pair of mother-trees. An enlargement on the distances from 0 to 275 m allows to define a threshold around 50 m (indicated by the red dashed line) above which the mean genetic relatedness among families no more depends on the pairwise distances among mother-trees.

HERITABILITY ESTIMATES

The traits measured in the common garden presented a large phenotypic variation ($CV_{mean} = 0.31$; Table 3.10). The growth traits, $BioM$, δH and δD , presented the highest phenotypic variation (more than 42 %). The carbon content and the date of leaf senescence (and thus the vegetation season duration) presented the lowest phenotypic variation (less than 3 %).

Three traits were transformed to improve the fit of the linear mixed models : δH with a logarithm transformation, LM and LA with a square-root transformation. For all traits, the inclusion of a random family component improved the models, indicating that V_A and h^2 were significantly different from zero (results not shown). Significant maternal effects were found for 3 traits : strong maternal effects for 2 growth traits, δD and $BioM$, and weaker maternal effects for $t_{b2 \rightarrow 3(2011)}$. Yet, no maternal effect was detected the second year of bud burst monitoring ($t_{b2 \rightarrow 3(2012)}$). Among these 3 traits, only the mean phenotypic values of $BioM$ were significantly correlated to the average weight of the seeds per families (Figure 3.7). This phenotypic correlation was positive, indicating that families with the highest seed weight also

Trait	mean	minimum	maximum	CV
δH	1.474	0.096	31.083	0.824
δD	1.089	0.084	3.585	0.425
$BioM$	15.785	0.122	148.494	0.832
$t_{b2 \rightarrow 3(2011)}$	251.047	155.099	407.657	0.172
$t_{b2 \rightarrow 3(2012)}$	341.238	154.360	640.004	0.226
$t_{s1 \rightarrow 2}$	3171.505	2997.012	3250.858	0.021
VD_{2011}	2921.650	2651.925	3095.759	0.026
LM	0.184	0.008	0.648	0.449
LA	31.570	3.850	85.860	0.388
LMA	57.463	4.668	150.979	0.166
$\%N$	2.138	1.340	3.040	0.124
$\%C$	46.008	38.610	51.850	0.025
$\delta^{13}C$	-29.105	-33.450	-25.760	-

TABLE 3.10 – Descriptive statistics for the traits analysed, with CV , the coefficient of variation.

Trait	p-val(Mat)
δH_{ln}	0.135
δD	3.771e-05
$BioM$	2.251e-05
$t_{b2 \rightarrow 3(2011)}$	0.023
$t_{b2 \rightarrow 3(2012)}$	1
$t_{s1 \rightarrow 2}$	0.357
VD_{2011}	0.223
$LM_{\sqrt{}}$	1
$LA_{\sqrt{}}$	1
LMA	0.284
$\%N$	0.587
$\%C$	0.544
$\delta^{13}C$	0.764

TABLE 3.11 – Test of maternal effects for each trait. ln and $\sqrt{}$ indicate that a transformation has been applied.

had the highest biomass production at the seedling stage after 2 years of growth.

For traits unaffected or slightly affected by maternal effects, the different methods produced similar h^2 estimates. Even the confidence intervals around the estimates were quite similar between the family models (h_{Fam}^2 and $h_{doubleadj}^2$) and the animal models (h_{HSFS}^2 and h_{MH}^2). The largest differences between the family and animal methods were found for the traits highly affected by maternal effects. The family and hybrid methods presented a difference of 0.69 for $h_{\delta D}^2$, 0.58 for h_{BioM}^2 , and 0.37 for $h_{t_{b2 \rightarrow 3(2011)}}^2$. Globally, the heritabilities estimated with the family method (equation 3.8) tended to be higher and to have larger confidence intervals than the estimates derived from the animal methods. The two methods using the animal model, i.e. the methods using the incomplete pedigree (method HSFS) or hybrid relatedness information (method MH), provided similar h^2 estimates. The heritabilities estimated with the HSFS method tended to be slightly higher than those estimated with MH method.

Focusing on estimates from the hybrid method, trait heritabilities were low to moderate (global mean = 0.22). The log-transformed height increment presented the highest heritabi-

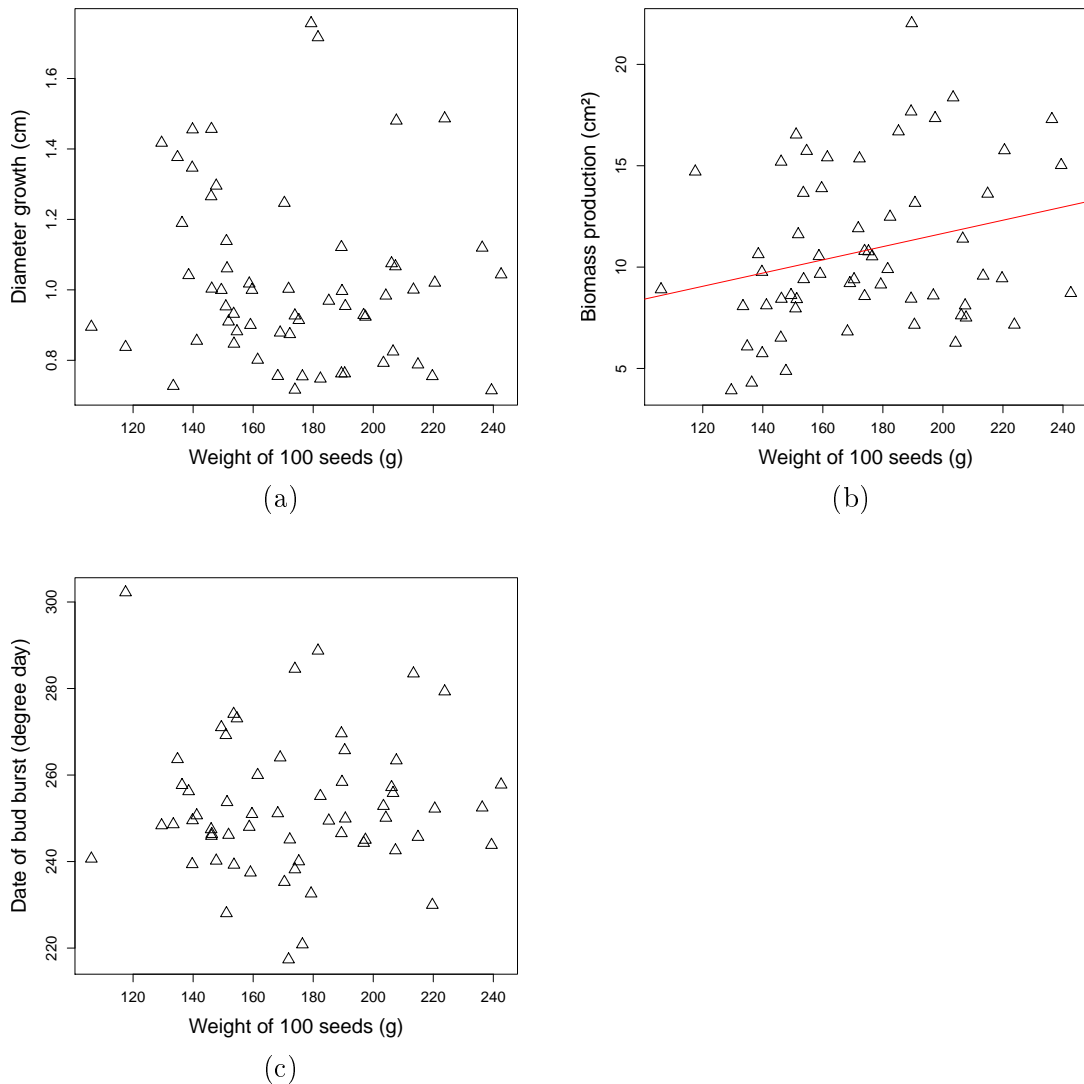


FIGURE 3.7 – Phenotypic correlation between the average seed weight (in g) and the (a) average diameter growth (cm), (b) average biomass production (in cm^2) and (c) average bud burst date at the family level (in degree days). A significant correlation was found only for the biomass production (trend line represented in red) with a p-value = 0.045.

lity with $h_{\delta H}^2 = 0.35$ and the total production of biomass lowest heritability with $h_{BioM}^2 = 0.13$. For the traits affected by maternal effects, the proportion of maternal variance explaining the total phenotypic variance was similar to the additive genetic variance estimated (Figure 3.8), highlighting the non-negligible contribution of the maternal effects in the family effects. For the traits presenting significant maternal effects the h^2 were lower ($mean_{h^2_{mat}} = 0.14$) and the estimates were less precise (higher confidence intervals) than for the other traits (Figures 3.9 and 3.10).

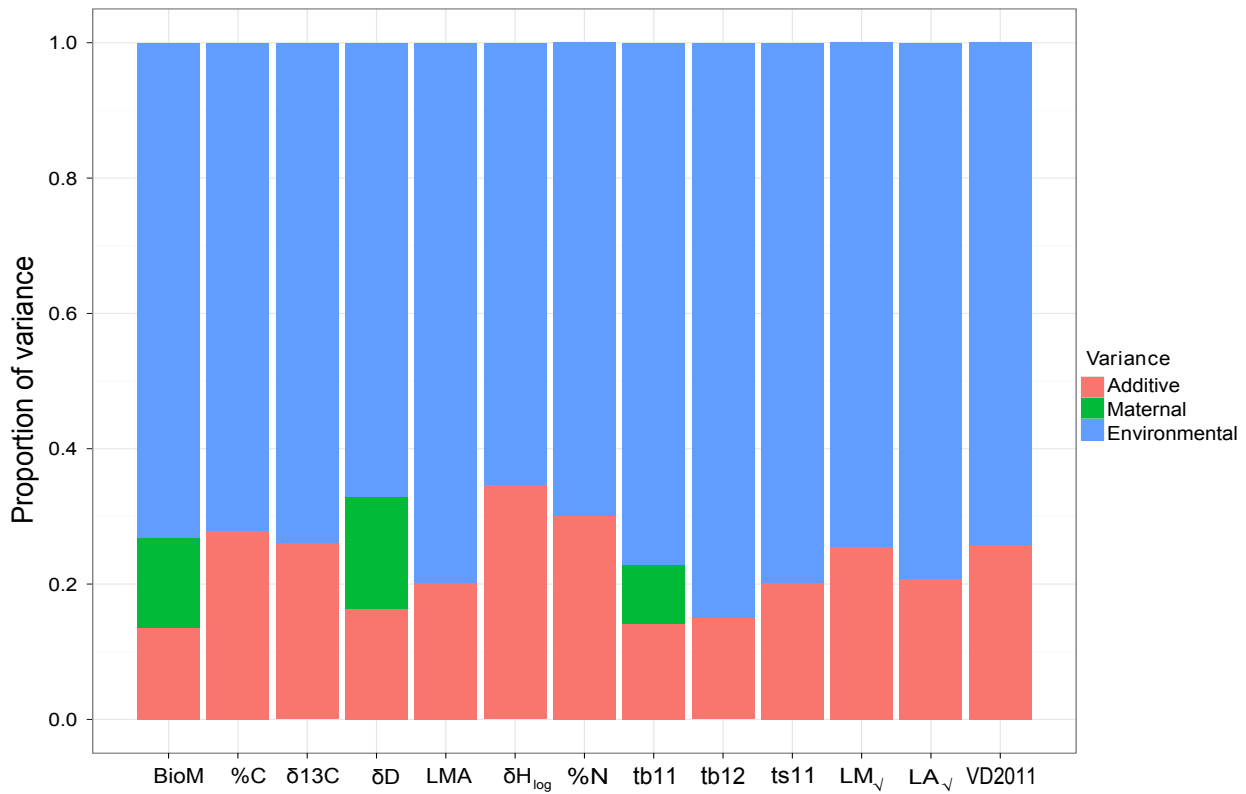


FIGURE 3.8 – Proportion of total phenotypic variance explained by environmental, additive and maternal variances for each trait. Variance components were estimated using the hybrid relatedness method and the model 3.12 for traits presenting significant maternal effects and the model 3.10 for the other traits.

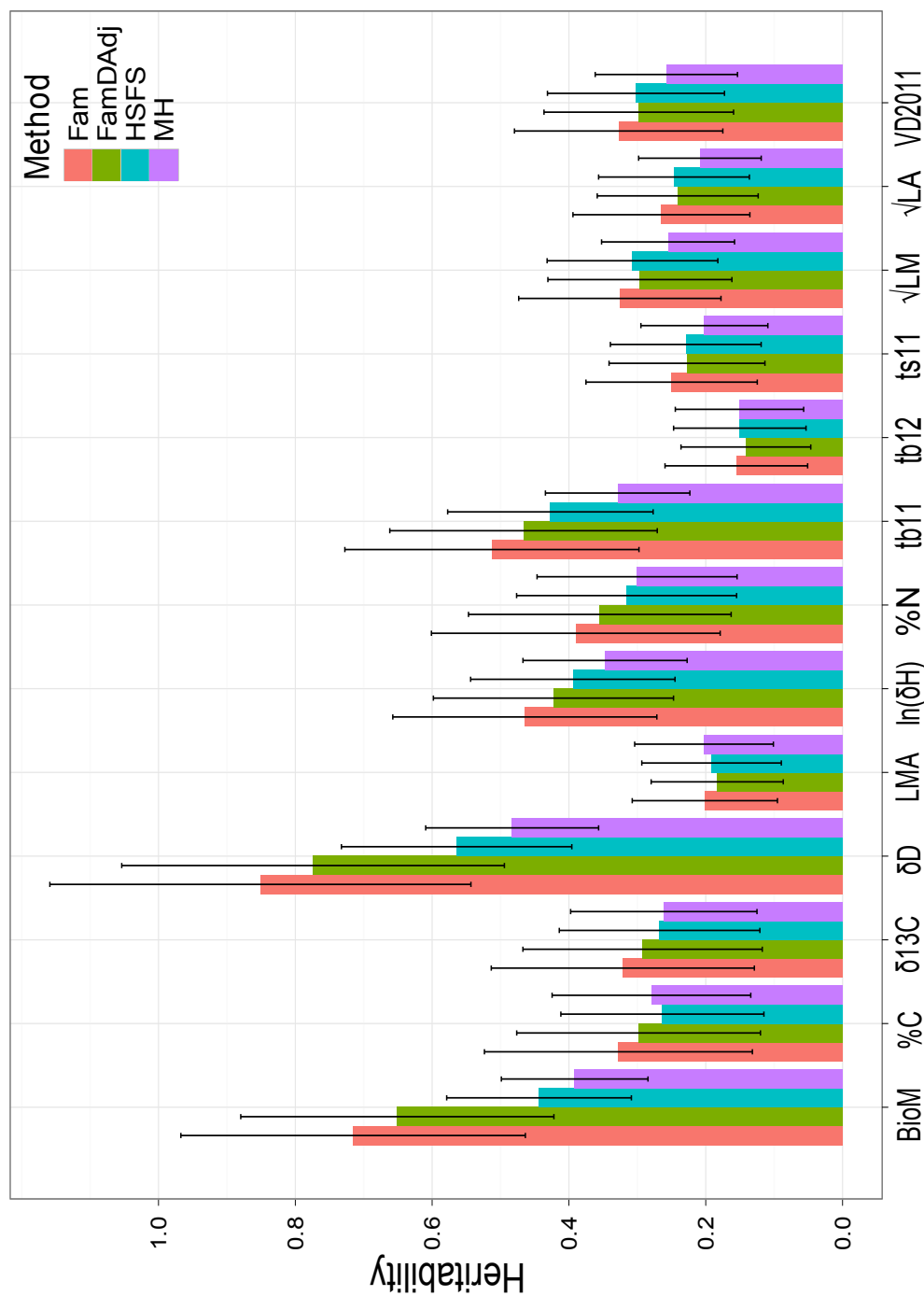


FIGURE 3.9 – Comparison of different methods to estimate the heritability. The "Fam" and "FamDAdj" methods use the family model (model 3.7). The "HSFS" and "MH" methods use the animal model (model 3.10 without maternal effects). The "FamDAdj" accounts for the $\tilde{\rho}_w$ and $\tilde{\rho}_b$. The "HSFS" method uses the incomplete pedigree information available, and the "MH" method hybrid relatedness information, mixing pedigree and average marker-based coefficients.

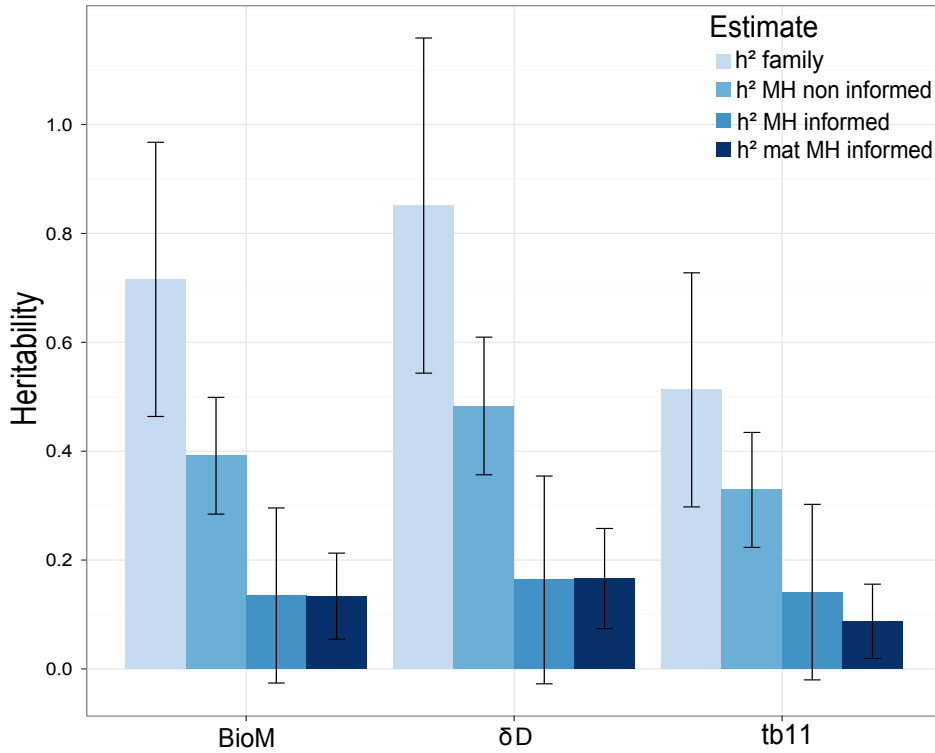


FIGURE 3.10 – Estimation of h^2 for the traits presenting significant maternal effects using the family method and the hybrid matrix method ("MH") with models where maternal effects were non-informed and informed. The estimation of h_{mat}^2 using the informed MH method is also represented. The bars around the estimates represent the confidence intervals at 95 %.

Discussion

DEPARTURE FROM RANDOM MATING FAINTLY AFFECTS h^2 ESTIMATES

Contrary to our expectation, the unbalanced mating system in the studied *F. sylvatica* populations only moderately increased the overall average marker-based relatedness within ($\tilde{\rho}_w = 0.276$) and between families ($\tilde{\rho}_b = 0.001$) compared to the assumptions in open-pollinated design ($\rho_w = 0.25$ and $\rho_b = 0$). For the 10 studied traits unaffected by maternal effects, the use of pedigree information in an animal model (methods "HSFS" and "MH"), instead of the family model, did not change h^2 estimates. Yet, the studied population showed significant departures from random mating, with pollen dispersal restricted by distance and unbalanced reproductive success among males. In fact, these mating patterns are probably counterbalanced by the high effective population size in Mont-Ventoux : Lander *et al.* 2011 estimated $N_e = 1.82 \times 10^3$ for the west region of Mont-Ventoux corresponding to the location of our study sites.

Another potential source of bias on h^2 estimates is the presence of inbreeding in parental population (Gauzere *et al.*, 2013b). Among the methods compared in this study, the method using only the one-generation pedigree ("HSFS" method) is expected to be more biased if parents are related (bi-parental inbreeding) than the method using also the average marker-based relatedness coefficients ("MH" method; Gauzere *et al.* (a) *in prep*). The similar h^2 estimates obtained with these two methods (Figure 3.9) thus suggest that the parents of the seedlings were faintly related and we do not expect that inbreeding or inbreeding depression in our progeny population significantly impacted the h^2 estimates.

More generally, in large, outcrossed, weakly inbred natural populations, highly unequal male fertilities are not likely to result in significant departures from half-sib assumptions. Thus, the inferences of the adaptive potential of fragmented or disturbed populations using family estimates may be strongly biased. Note however that these biases are not systematic : in cases where family trials significantly departs from both $\rho_w = 0.25$ and $\rho_b = 0$, these departures can have balancing effects on the estimation of h^2 : $\rho_w > 0.25$ tend to overestimate V_A and h^2 , while $\rho_b > 0$ tend to underestimate V_A and h^2 (Gauzere *et al.* (a) *in prep*). In these situations, neglecting ρ_b and correcting the h^2 calculations according to ρ_w only will conduct to bias h^2 (e.g. considering $h^2 = \frac{1/\rho_w \times V_F}{V_R + V_F}$; Gauzere *et al.* (a) *in prep*). Only the double correction used in equation 3.9 can balance both departures.

MATERNAL EFFECTS AFFECT h^2 ESTIMATES DERIVED FROM THE FAMILY AND ANIMAL METHODS

In our progeny trial, neglecting maternal effects was the main source of bias in h^2 estimates. Indeed, for the 3 traits presenting strong maternal effects, the family and animal methods produced highly divergent h^2 estimates, with an overestimation of h^2 in the family methods. Similarly, Hansen & Nielsen 2010 showed that in presence of moderate departure from $\rho_w = 0.25$, models using half-sib assumptions or pedigree information conduct to the same estimates except for traits presenting significant dominance effects. Thus, we can generalize that in presence of non-additive genetic effects, pairwise relatedness information must be used to accurately estimate h^2 .

Maternal effects can be viewed either as nuisance effects, wrongly inflating the additive genetic variance estimated in family trials, if these effects are non-permanent effects, or as part of the phenotypic variation possibly responding to selection, if these effects are permanent and transmissible effects (Bossdorf *et al.*, 2008). Strong maternal effects were detected for diameter growth and biomass traits. More surprisingly, maternal effects were also found for the first year of bud burst monitoring. The second year, this effect was no more detected suggesting that maternal effects were non-permanent effects on the expression of the dates of bud burst. The significant correlation between the average biomass production and the seed weight suggested that maternal effects were also non-permanent effects of $BioM$ phenotypic variation (seedlings with more resource available in the seeds produced higher biomass in the first years of growth). Thus, maternal effects observed in this study are likely to be nuisance effects. To our knowledge this study is the first to highlight significant maternal effects on the determinism of juvenile traits for the european beech and on bud burst phenology for the *Fagaceae* family. The maternal determinism of phenological traits notably questions the previous studies which have investigated the genetic determinism and adaptive potential for species within this wide phylogenetic family.

We also showed that maternal effects challenge the performances of the animal model using incomplete pairwise relatedness information. Indeed, animal models accounting for the maternal effects estimated h^2 with lower precision than models without maternal effects. Moreover, our previous simulation work (Gauzere *et al.* (a) *in prep*) showed that incomplete pairwise relatedness information conduct to significantly underestimate h^2 in such cases. Thus, the true h^2 value for the traits presenting maternal effects is probably higher than the one estimated with the model in equation 3.12. But this underestimation does not completely explain why $V_M + V_A$ values estimated with the animal model were lower than the global V_A estimated with the family model. One possible explanation is that the use of incomplete pedigree information also conducts to underestimate V_M . Unfortunately, the use of incomplete pedigree information in tree natural populations is the rule rather than the exception, notably because of the high migration rates (Ashley, 2010).

ARE h^2 ESTIMATES GOOD PROXY OF THE ADAPTIVE POTENTIAL ?

Heritabilities were significant but globally low (ranging from 0.13 to 0.35), highlighting low

to moderate genetic control for the functional traits investigated in this study. The theoretical and empirical literature show that traits closely linked to fitness should present low h^2 because the genetic variance of these traits should be rapidly eroded by selection (Stirling *et al.*, 2002). Notably, for trees, growth traits are expected to be more related to fitness than leaf morphological traits. Surprisingly, in this study we saw no differences in the h^2 values by class of functional traits : growth traits presented on average $h^2 = 0.21$ and morphological traits $h^2 = 0.22$.

Comparison of our estimates with other ex situ h^2 estimates in the literature is difficult because, first of all, these estimates greatly depend on the population studied and the environment in which the phenotypic and genetic variances are expressed (Charmantier & Garant, 2005), and second, to our knowledge, only one study has already estimated h^2 of phenological and ecophysiological traits for european beech populations in common garden conditions (Kramer, 2004). Based on several populations from different provenances in europe, they estimated a narrow-sense heritability, h_{ns}^2 , of 0.56 and 0.58 for two years of bud burst monitoring, and non significant broad-sense heritabilities, h_{bs}^2 , for *LMA* and *%N*. h_{ns}^2 were estimated assuming $V_F = 4 \times V_A$, and h_{bs}^2 using a clonal test. Note that using the family method, we estimated a $h_{ns}^2 = 0.51$ for the first year of bud burst monitoring, which was quite similar to the estimates of Kramer 2004. Overall, a review of several progeny tests of different temperate and boreal tree species has revealed that bud burst phenology was generally under strong genetic control (on average $h^2 = 0.61$ for bud burst ; Howe *et al.* 2003). In our progeny test the low h^2 estimates for the two years of bud burst phenology can be partly explain. Indeed, $h_{tb2 \rightarrow 3(2011)}^2$ was probably underestimated due to methodological biases and $h_{tb2 \rightarrow 3(2012)}^2$ was probably deflated due to the particular environmental conditions suffered in autumn 2011, which have increased the environmental variance expressed this second year of bud burst phenology ($CV_{tb2 \rightarrow 3(2011)} < CV_{tb2 \rightarrow 3(2012)}$; Table 3.10).

Despite the difficulties to compare h^2 derived from different common garden experiments, some studies have shown that for the same population, h^2 estimated in controlled conditions were comparable to h^2 estimated in natural conditions (Weigensberg & Roff, 1996; St Juliana & Janzen, 2007), and thus, should provide relevant estimations of both the significance and the magnitude of h^2 in nature. Bontemps *et al.* *in prep* have estimated in situ h^2 for several functional traits in population N1 using the method of Ritland (1996). Using the "MH"

method for the subset of seedlings belonging to this same population, we found equivalent h^2 estimates for $\delta^{13}C$ and LA (Appendix C). For the bud burst phenology, LMA and $\%N$ the confidence intervals for the ex situ and in situ estimates did not overlap (Appendix C). Overall, the in situ h^2 were higher than the ex situ estimates. This trend has also been highlighted by the review of Weigensberg & Roff 1996, contrary to the expectation that the natural h^2 should be overestimated in controlled conditions due to the reduction of the environmental variability. Note also that ex situ h^2 were measured on juvenile traits, while in situ they were measured on adult traits. The lower h^2 estimated for juveniles may be due to the fact that juveniles are more subject to non-additive effects, such as epigenetic or maternal effects.

CONCLUSION

Beech populations in the Mont-Ventoux presented unbalanced mating system leading to only slight departures from family assumptions in progeny trials. Thus, methods using family or animal models presented similar h^2 estimates. In presence of maternal effects, assuming that V_G only includes V_A conducted to overestimate h^2 . However, even in informed models (models specifying maternal effects) the incomplete knowledge of the relatedness relationships among individuals may prevent to accurately estimate h^2 . We showed that all the functional traits investigated in this study presented a significant V_A and h^2 , and thus were able to evolve in natural populations.

Overall, this study highlight the main importance to better understand the genetic structure and mating characteristics of the populations investigated using molecular markers. This first evaluation of the departures from family assumptions occurring in family trials is a useful stage to choose the more suitable method to estimate h^2 . Our results also suggested that methods using family assumptions can be used to accurate the adaptive potential of large outcrossed plant natural populations, except in presence of maternal effects.

Appendix

A. Protocol of bud burst monitoring in the common garden experiment ¹

Stage 1:
buds are dormant or swelling



Stage 4:
at least 50 % of the leaves are
emerging on the plant



Stage 2:
bud scales are broken



Stage 5:
leaves are spread out



Stage 3:
at least 15 % of the leaves are
emerging on the plant



1. Photographs taken by Fabrice Bonne, Frédéric Jean and Olivier Gilg.

B. Protocol of leaf senescence monitoring in the common garden experiment ²

Stage 1:
leaves are not fallen and
colored



Stage 3:
at least 50 % of the leaves
are colored or have fell



Stage 2:
at least 10 % of the leaves
are colored or have fell



2. Photographs taken by Fabrice Bonne, Frédéric Jean and Olivier Gilg.

C. In situ h^2 of functional traits studied in the PhD of A. Bontemps

Trait	Population	h^2 in situ	h^2 ex situ
<i>PSS</i>	N1	0.92* [0.54 ; 1.29]	0.44* [0.34 ; 0.55]
<i>%N</i>	N1	0.81* [0.45 ; 1.19]	0.30* [0.15 ; 0.45]
$\delta^{13}C$	N1	0.52* [0.20 ; 0.84]	0.26* [0.12 ; 0.40]
<i>LMA</i>	N1	-0.19 [-0.53 ; 0.16]	0.20* [0.10 ; 0.30]
<i>LA</i>	N1	0.55* [0.13 ; 0.96]	0.21* [0.12 ; 0.30]

Legend : PSS is the sum of the phenological notations recorded at each weekly monitoring used as a surrogate of the date of bud burst. "*" indicated significant h^2 .

DIFFÉRENCIATION GÉNÉTIQUE ET POTENTIEL DE RÉPONSE À LA SÉLECTION DES TRAITS FONCTIONNELS



Synthèse

Dans ce chapitre de thèse, je me suis intéressée à la réponse des traits fonctionnels à la sélection passée et leur potentiel pour répondre aux pressions de sélection futures. La longue tradition de tests de provenance chez les arbres, entrepris dans un premier temps pour identifier les provenances les plus performantes et les utiliser en plantation, a permis d'étudier l'adaptation des populations d'arbres à une large gamme d'environnements. Les méta-analyses montrent que les populations d'arbres présentent généralement de forts niveaux de divergence adaptative (sur la base de tests de *Fst-Qst*; Savolainen *et al.* 2007; Alberto *et al.* 2013), indiquant que ces populations se sont probablement adaptées localement à leur habitat depuis la recolonisation post-glaciaire. Cependant, l'échelle large à laquelle ont été menées ces études ne permet pas d'explorer certaines questions en biologie évolutive, comme quelle est l'échelle spatiale à laquelle l'adaptation locale a lieu et quels sont les facteurs qui peuvent la contraindre (e.g., flux de gènes; Hendry & Taylor 2004; Sambatti & Rice 2006). De plus, les quelques études qui ont étudié le potentiel adaptatif de populations d'arbres se sont généralement focalisées sur l'estimation des variances génétiques additives (V_A) et héritabilités (h^2) des traits fonctionnels (Scotti *et al.*, 2010), sans aller jusqu'à mesurer les covariances génétiques entre ces traits et comprendre comment les contraintes génétiques pouvaient affecter les taux d'évolution des traits chez les arbres.

La **section 4.1** de ce chapitre a eu pour objectif de tester la divergence génétique des traits fonctionnels entre populations de hêtre commun le long du gradient altitudinal du Mont-Ventoux malgré les forts flux de pollen immigrant reçus par les populations ($m = 56\%$; Chapitre 2). Les tests classiques de *Fst-Qst* étant inefficaces pour détecter des traces de différenciation adaptative sur un petit nombre de populations comme celui étudié dans cette thèse (O'Hara & Merila, 2005), j'ai utilisé une approche bayésienne proposée récemment par Ovaskainen *et al.* (2011) pour tester les traces de sélection uniformisante ou divergente sur la variation des traits quantitatifs entre populations, permettant de contourner ces problèmes méthodologiques. Ce travail a été réalisé en collaboration avec Oliver Brendel, chercheur à l'INRA de Nancy, et Hendrik Davi, chercheur à l'URFM.

Dans la **section 4.2** de ce même chapitre, j'ai cherché à mesurer le potentiel adaptatif multi-traits des populations en utilisant l'approche et les statistiques proposées par Kirkpa-

trick (2009). Cette approche a notamment permis d'évaluer l'évolutivité et la dimensionnalité des matrices de variance-covariance des traits (appelée \mathbf{G}) par population afin d'évaluer les contraintes *potentielles* à l'évolution future. Une suite naturelle à ce travail consisterait à mesurer les gradients de sélection en œuvre dans les trois populations naturelles, afin de (i) comparer la direction de la sélection avec la géométrie des contraintes évolutives définies par la matrice \mathbf{G} et (ii) mesurer plus directement la variation des pressions de sélection aux différentes altitudes sur ce gradient.

Globalement, les résultats de ce chapitre permettent d'observer que des patrons de différenciation adaptative entre populations ont pu émerger malgré les faibles distances séparant les populations (~ 1 km séparant N1 et N4) et la présence de forts flux de pollen immigrants. Pour un certain nombre des traits présentant une divergence génétique le long du gradient, on observe des niveaux de V_A hétérogènes entre populations, ce qui suggère que la sélection pourrait être à l'origine d'une perte de diversité génétique. L'approche multi-traits indique que les populations se trouvant aux deux extrêmes du gradient altitudinal (N1 et N4) présentent des évolutivités totales plus faibles que la population intermédiaire (N2). Les nombreuses covariances génétiques entre traits mesurées pour la population à plus basse altitude (N1) soulignent que les mesures de réponse à la sélection univariées peuvent être inappropriées pour évaluer le taux d'évolution des traits chez les populations d'arbres.

Ces résultats sont cependant à nuancer puisqu'ils ont été obtenus dans un dispositif en pépinière, par la mesure de traits sur des semis. Or, on ne connaît pas bien la relation qui existe entre les traits adultes et juvéniles chez les arbres et l'environnement non naturel dans lequel ont été réalisées les mesures pourrait affecter l'expression des variances et covariances génétiques entre traits. De plus, je n'apporte ici pas la preuve empirique rigoureuse que les populations de hêtre sont localement adaptées le long du gradient. Pour cette question, les manipulations de transplantations croisées seraient particulièrement intéressantes à mettre en œuvre dans la suite de ce travail de thèse puisqu'elle permettraient de s'affranchir d'un grand nombre de limitations par rapport aux expérimentations en jardin commun, en permettant de mesurer les traits dans leur environnement naturel (tout du moins climatique), tout en connaissant les covariances environnementales et génétiques entre individus. Les semis suivis durant mes trois années de thèse ont récemment été transplantés dans deux sites à différentes altitudes sur le Mont-Ventoux. Le suivi des ces arbres transplantés dans les années à venir

devrait apporter un complément très intéressant à l'étude de la différenciation génétique que j'ai réalisée ici.

L'échelle locale à laquelle nous avons étudié l'interaction entre différenciation génétique et flux de gènes est à la fois très intéressante pour comprendre à partir de quelle échelle spatiale les flux de gènes peuvent empêcher toute différenciation, et à la fois limitante car elle réduit notre capacité à pouvoir détecter la différenciation entre populations. En effet, la faible distance séparant le peuplement de hêtre aux deux extrêmes altitudinaux sur le Mont-Ventoux limite la caractérisation d'un grand nombre de populations sur le gradient climatique. Or, de nombreuses analyses à la fois en écologie et en biologie évolutive nécessitent d'étudier beaucoup de populations ($n > 10$) pour pouvoir faire des inférences sur les processus écologiques ou évolutifs qui régissent les différences inter-populations (e.g., différences dans les caractéristiques du régime de reproduction, divergence adaptative ou neutre des traits). Une approche plus pertinente pour étudier ce gradient climatique restreint aurait peut être été d'avoir une approche de "cline", i.e. plutôt que d'étudier 60 arbre-mères répartis en 3 populations, suivre 60 arbre-mères localisées de manière spatialement explicite le long d'un transect altitudinal. L'inconvénient de cette approche auraient cependant été de perdre l'information à l'échelle intra-population qui est importante pour les analyses de génétique quantitative telle que cette discipline est formalisée aujourd'hui (e.g., mesure de la variance génétique, des covariances génétiques locales).

4.1 Réponse plastique et adaptative des traits fonctionnels à la sécheresse et la température entre populations de *F. sylvatica* le long d'un gradient altitudinal

Title : Plastic and adaptive response of functional traits to drought and temperature between *F. sylvatica* populations along an elevational gradient

Authors : Gauzere J., Oddou-Muratorio S., Klein E. K., Brendel O., Davi H.

Introduction

Identifying the factors that promote or prevent adaptive divergence between populations is a main topic in evolutionary biology (Hendry & Taylor, 2004; Etterson & Shaw, 2001; Saint-Laurent *et al.*, 2003; Kawecki & Ebert, 2004). This issue has gained a renewed interest in the current context of climate change as it could help to evaluate the abilities of populations to respond to ongoing and predicted environmental variations (Davis & Shaw, 2001; Etterson, 2004). Evidences of population phenotypic divergence with spatial or temporal environmental changes have been widely documented (Keller *et al.*, 2013). However, the adaptive component of phenotypic divergence can only be inferred if (i) the relative contributions of genetic and plastic variations with environment can be dissociated and (ii) the action of natural selection and genetic drift on the population genetic responses can be distinguished.

The genetic response to environmental variation is classically investigated by phenotyping populations experiencing different environments in common environmental conditions, also called "ex situ" or common garden experiments (Conover & Schultz, 1995). For a given trait, the comparison of the patterns of phenotypic divergence observed in situ and genetic divergence observed ex situ allow to infer the role of phenotypic plasticity and that of genetic variation, and thus to understand the causes of phenotypic variability or uniformity in nature (Conover & Schultz, 1995). Moreover, if the phenotypic and genetic values show the same pattern of variation (genetic and environmental influences on phenotypic variation are

consistent), the trait is said to display cogradient variation (CoGV). On the contrary if the phenotypic and genetic values show opposite patterns of variation, the trait display "counter-gradient variation" (CnGV).

Regarding for the action of genetic drift, the relative contributions of neutral evolutionary processes and natural selection to genetic divergence are classically disentangled by comparing the Fst and Qst statistics (Merila & Crnokrak, 2001; McKay & Latta, 2002; Leinonen *et al.*, 2008). If the Qst statistics, which summarizes the genetic differentiation of the quantitative traits is significantly different from the Fst statistics, which summarizes the neutral genetic differentiation, then the genetic divergence between populations can be attributed to the effects of natural selection. The main limitation of $Fst - Qst$ method is that it is not adapted to the analysis of a small number of populations : at least 10 populations are necessary to get robust $Fst - Qst$ comparisons (O'Hara & Merila, 2005; Whitlock & Guillaume, 2009). Moreover, by using summarized statistics, this approach does not allow to fully use all pairwise population comparisons and thus tends to provide nonsignificant tests for populations presenting low level of phenotypic divergence. These limitations have been overcome in a new approach proposed by Ovaskainen *et al.* (2011), an alternative to the $Fst-Qst$ comparison method which notably allows to account for the random nature of the evolutionary processes and the population-to-population levels of genetic and trait divergences. This new method, although potentially powerful, has been implemented only recently in a user-friendly R package and thus seldom used for the moment (but see Karhunen *et al.* 2013, 2014).

Elevational gradients are popular experimental designs to investigate the adaptive divergence between natural populations since they provide wide and continuous variation of the environmental conditions over short spatial distances (Korner, 2007; Keller *et al.*, 2013). Studies on elevational gradient are generally less dependent on evolutionary confounding effects due to distinct regional evolutionary histories than studies on latitudinal gradients (Keller *et al.*, 2013). Moreover, if the environmental clines are steep enough relative to the distance of gene dispersal, elevational gradients are relevant sites to investigate the interplay between natural selection and gene flow on the adaptive responses of the populations. However, care should be taken to accurately interpret the phenotypic divergences along elevational gradients if the environmental changes due to the elevation changes, such as the temperature or atmospheric pressure, are confounded with other environmental changes that are not elevation

specific, such as the human land use, soil geology or wind exposure (Korner, 2007).

Tree species are among the most important terrestrial species group, in term of area covered and Earth's biomass (Petit & Hampe, 2006). Their wide distribution range involves that trees have adapted to a wide range of environmental conditions. However, the evolutionary potential of these species is still questioned because of their long generation time, late maturity and high gene flow that may prevent genetic adaptation (Petit & Hampe, 2006; Savolainen *et al.*, 2007). In the Mediterranean zone, the annual mean temperature is predicted to increase and the mean rainfall to decrease (Lindner *et al.*, 2010). Therefore, phenological, morphological and physiological traits responding to temperature and drought may play an important role in the maintenance and adaptation of tree species in Mediterranean ecosystems in the next future.

The variation of phenological traits along environmental gradients has been widely documented since the current climate change is already affecting the phenological timing of most animal and plant species (Root *et al.*, 2003). For plants, warmer temperatures found at low latitudes and altitudes enable longer vegetative seasons and thus increase the net carbon accumulation (Bertin, 2008). At the opposite, at high latitudes and altitudes, the cold temperatures tend to delay the timing of bud burst phenology due to the highest risk of frost damage on buds. The role of the bud burst phenology in response to drought is more uncertain. Some studies have notably suggested that earliest phenology may allow to avoid unfavourable climatic conditions and constitute, in association with other functional traits, a multivariate ecological strategy of drought escape (Geber & Dawson, 1990; Arntz & Delph, 2001; Franks *et al.*, 2007). Among the morphological and physiological traits, the water use efficiency and leaf size has already been shown to play an important role in the adaptation to drought in some plant species (Dudley, 1996; Donovan *et al.*, 2009). In these cases, plants display higher water use efficiency and thicker leaves in the drier environments to optimise the photosynthetic capacities and water transpiration. However, selection for drought tolerance has rarely been investigated for deciduous temperate tree species (but see Ramirez-Valiente *et al.* 2009).

In this study, we investigated the genetic differentiation of three populations of *Fagus sylvatica* along an elevational gradient in the south-east of France. The studied *F. sylvatica*

tica populations are under a Mediterranean climate, corresponding to the southern ecological conditions tolerated by this species. The adaptation to drought may thus be one of the main pressure differentiating the populations along the gradient. However, the short distance separating the populations (almost 1 km of ground distance) and the high immigrant pollen flow previously estimated (Gauzere *et al.*, 2013a) may prevent adaptive divergence (Lenormand, 2002; Hendry & Taylor, 2004). In addition to in situ measurements of the traits in the three populations, we analysed growth, morphological, physiological and phenological traits on 20 open-pollinated progenies sampled from each of the 3 natural populations (60 families in total) and which were grown in controlled environmental conditions. Our aims were to (i) investigate the patterns of phenotypic differentiation in the natural populations, (ii) characterize the pattern of genetic differentiation between populations and the plastic response of the traits to drought based on the ex situ trial, and (iii) evaluate the importance of natural selection and neutral evolutionary processes on the genetic differentiation measured.

Material and Method

SPECIES AND SITES

Fagus sylvatica L. (European beech) is a monoecious, anemophilous and predominantly out-crossed tree species (Merzeau *et al.*, 1994; Wang, 2003). This species is a shade-tolerant species requiring well drained and moderately deep soils and relatively high humidity rates; it represents one of the main essence of temperate forests in Europe. On the Mont-Ventoux, a relief in the south-east of France, beech populations are located at the southern limit of their ecological range, in a Mediterranean ecosystem. On the northern face of Mont-Ventoux, a beech forest ranges almost continuously from 750 to 1700 m a.s.l.

We studied three beech populations located at 1,020, 1,140 and 1340 m a.s.l. along this elevational gradient (Figure 4.1). Main competitors of beech in these sites are *Pinus sylvestris* and *Abies alba*. The northern side of the Mont-Ventoux presents almost linear variations with elevation of the mean temperature, humidity rate and rainfall from March to August. Populations at the lowest elevational levels experiment the warmest and driest climate (Table 4.1). Populations present high pollen immigration rates (on average 53.3 %) with moderate distances of pollen dispersal (on average 45.7 m) and non-negligible long-distance pollen dis-

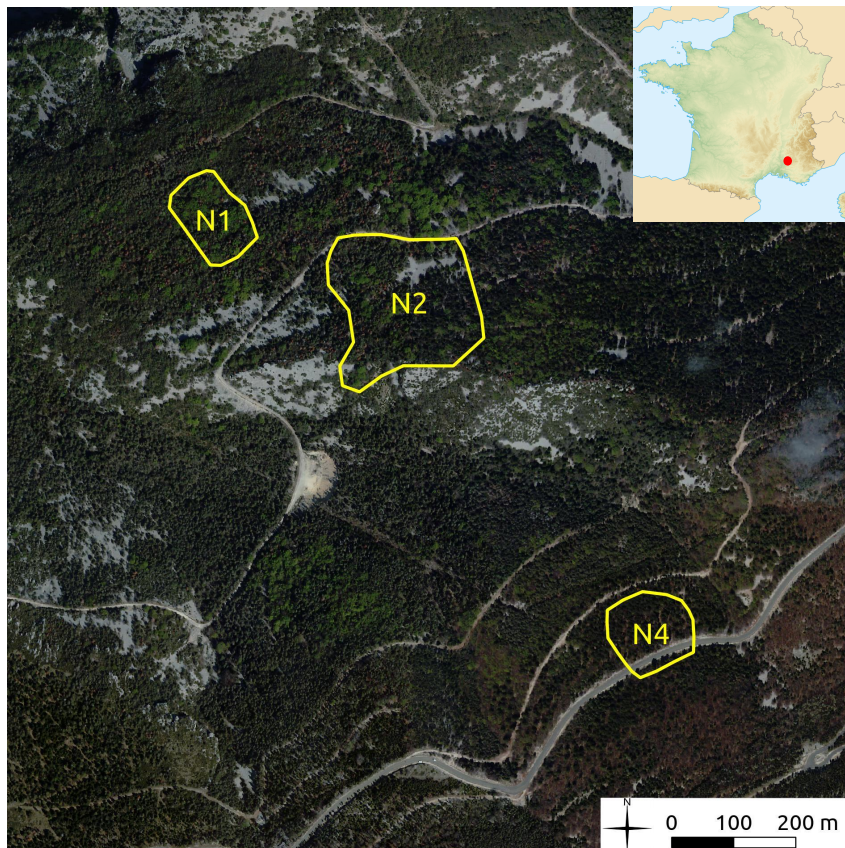


FIGURE 4.1 – Location of the three sampled populations of *F. sylvatica* on the north-face of the Mont-Ventoux, in the south-east of France.

	N1	N2	N4
Altitude (metres)	1020	1140	1340
Dimension (hectare)	1.30	2.20	0.80
Mean temperature (degree)	9.4	8.5	6.7
Mean rainfall (mm)	688	815	1046
Mean humidity rate (%)	71.7	72.2	73.0
Mean competitor density (/ha)	230.20	130.16	320.91
Mean number of clonal copies per mother-tree	2.55	6.81	1.82
Mean N_{ep}	12.3	15.1	29.2
Mean immigration rate	0.43	0.53	0.64
Number of reproductive adults (mothers included)	154	172	364

TABLE 4.1 – Location, climatic and ecological characteristics of the three studied populations. Mating system results are extracted from Gauzere *et al.* (2013a), with N_{ep} the effective number of pollen donors.

persal events (Gauzere *et al.*, 2013a). Seed dispersal also occurs at moderate distance (on average 23.5 m; Bontemps *et al.* 2013). Moreover, the short distances that separate populations N1 and N4 (~ 1 km) suggests that the three populations are highly connected by gene flow. Despite their large sizes and high immigration rates, the highly unequal male fertilities (Gauzere *et al.* (a) *in prep*) leading to reduced effective numbers of pollen donors (N_{ep}) may induce non-negligible genetic drift within the populations.

In august 2009, 20 open-pollinated progenies were sampled in each of the 3 populations (60 families in total). The mother-trees were chosen according to their high fertilities and random location within populations. All the potentially reproductive individuals within these populations were sampled and genotyped at 13 microsatellite markers (690 adults in total; Gauzere *et al.* 2013a).

EXPERIMENTAL DESIGN

The experiment was performed in the State nursery of Aix-en-Provence (43 ° 30'N 5 ° 24'E) with a complete block design (details of the transplantation are described in Gauzere *et al.* (b) *in prep*). After germination, an average of 91.25 seedlings per family were randomly planted in 50 blocks (1.8 seedlings per family in each block), distributed in 6 planks (Figure 4.2). Using a system of watering rails, half of the trial was watered and the other half submitted to water stress conditions during the driest months, i.e. from May to October. In each water condition, a subset of pots within one block was regularly weighted to calculate a point of permanent wilt (PW). In the watering conditions, seedlings were watered to maintain $PW = 1$. In the water stress conditions, seedlings were watered when PW was above a threshold of 2.5. The water stress applied was voluntarily weak to avoid extreme mortality in the trial. Each plank can be dissociated in two lanes, on both sides of the watering rails. This cutting allow to more finely characterize the water stress condition, which was not homogenous in the treatments, since the pots were not individually watered.

An average of 34.8 seedlings per family were genotyped at the same 13 microsatellites markers than the reproductive adults in the populations. Using a categorical paternity assignment method (the likelihood based method implemented in CERVUS; Marshall *et al.* 1998), we successfully assigned the paternity for an average of 16.6 offspring per family (Gauzere *et al.* (b) *in prep*). Only the maternal relatedness was known for the rest of the seedlings (i.e., those not genotyped and those genotyped but not assigned).

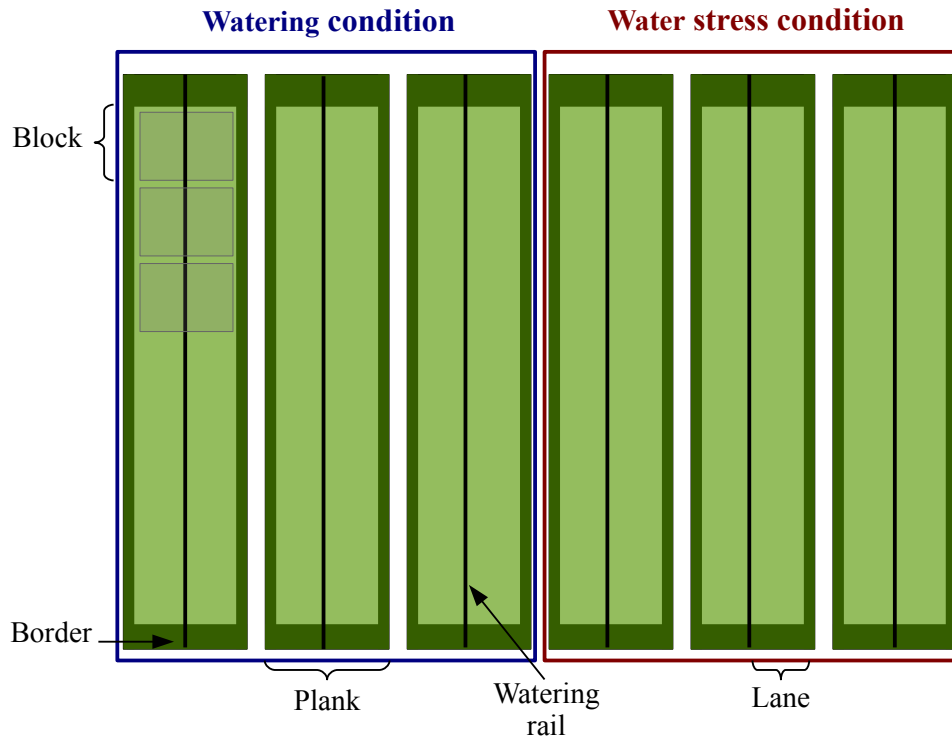


FIGURE 4.2 – Scheme of the split-block trial. It is composed of 6 planks distributed in two water content treatments. Each plank is surrounded by seedlings not phenotyped on the borders and located on the bottom of watering rail. Each plank is constituted of two lanes, on each side of the watering rail. Planks contain 8 to 9 complete blocks in which all families are randomised.

FUNCTIONAL TRAITS

We investigated the genetic differentiation and plastic response of the same phenotypic traits than previously described by Gauzere et al (b) *in prep.* Briefly, growth traits (δH , δD , $BioM$), bud burst phenology in 2011 and 2012 ($t_{b2 \rightarrow 3(2011)}$, $t_{b2 \rightarrow 3(2012)}$) and morphological traits (LM , LA , LMA) were measured on the whole trial. Leaf senescence ($t_{s1 \rightarrow 2(2011)}$) and the length of the vegetative season (VD_{2011}) were monitored for the watering condition only. Physiological traits ($\%N$, $\%C$, $\delta^{13}C$) were measured on leaves of a subset of 1574 individuals, distributed in the two water conditions (Table 4.2). Among these traits, LMA , $\delta^{13}C$ and $\%N$ are notably expected to respond to drought. Indeed, LMA is related to the photosynthetic capacities and stomatal conductance (Reich *et al.*, 2003), $\delta^{13}C$ is classically used as a surrogate of the water use efficiency (Farquhar *et al.*, 1989) and $\%N$ is considered related to the photosynthetic

Abbreviation	Trait	Unit	N_{water}	N_{stress}
δH	Total height increase	no unit	1490	1810
δD	Total diameter increase	no unit	2376	2309
$BioM$	Total biomass produced	cm ²	1490	1810
$t_{b2 \rightarrow 3(2011)}$	Bud burst date in 2011	Degree days	2574	2525
$t_{b2 \rightarrow 3(2012)}$	Bud burst date in 2012	Degree days	1420	903
$t_{s1 \rightarrow 2(2011)}$	Senescence date in 2011	Degree days	2378	0
VD_{2011}	Vegetation season duration in 2011	Degree days	2378	0
LM	Dry leaf mass	g	2481	2184
LA	Leaf Area	cm ²	2481	2184
LMA	Leaf Mass Area	g m ⁻²	2481	2184
$\%N$	Nitrogen content in leaves	%	1034	540
$\%C$	Carbone content in leaves	%	1034	540
$\delta^{13}C$	Water use efficiency	‰	1034	540

TABLE 4.2 – Phenotypic measurements in the two water content treatments. We phenotyped N_{water} offspring in the watered condition and N_{stress} in the water stress condition.

Trait	Data set 1			Data set 2			Data set 3		
	N1	N2	N4	N1	N2	N4	N1	N2	N4
$t_{b2 \rightarrow 3(2012,2013)}$	10	40	40						
$t_{s1 \rightarrow 2(2013)}$ VD_{2013}	19	55	37						
LMA $\%N$ $\delta^{13}C$				8	8	8	155	21	30
LM LA							155	21	30
$\%C$				8	8	8			

TABLE 4.3 – Number of phenotyped individuals in the three natural populations. Dataset 1 comes from a long-term phenology monitoring, dataset 2 from M. Ducrey (personal communication) and dataset 3 from A. Bontemps (Bontemps, 2012).

capacities since the RubisCO, a key enzyme in photosynthesis, represents 20 to 30 % of the leaf nitrogen content. In Gauzere et al. (b) *in prep*, we previously showed that the phenotypic variation was significantly affected by maternal effects for three traits : δD , $BioM$ and $t_{b2 \rightarrow 3(2011)}$.

For all these functional traits except the growth and biomass traits, we also get phenotypic measurements for a subset of adult trees in the natural populations. These data came from

a long-term project of phenology monitoring and two previous studies in the Mont-Ventoux (Table 4.3). Diameter growth has already been characterized along the elevational gradient using tree-ring measures (Cailleret & Davi, 2011).

QUANTITATIVE GENETIC MODEL AND DIFFERENTIATION ESTIMATES

In situ phenotypic differentiation was investigated using an ANOVA model with a single population fixed effect for each trait separately and then testing pairwise differences among populations with a Tukey test (function "TukeyHSD" in R). Common traits from data sets 2 and 3 were analysed independently since the leaves had not been sampled at the same time and with the same protocol.

The analysis of the trait Y measured in common garden was performed using the following linear mixed model :

$$Y_{i,f,g,h,m,n,k} = \mu + P_i + T_g + H_{f(g)} + L_{h(f)} + B_{m(f)} + E_{n(g)} + P.T_{ig} + a_k + \epsilon_{i,f,g,h,m,n,k} \quad (4.1)$$

where μ is the mean phenotype, P_i the effect of the population of origin, T_g the effect of the water treatment, $H_{f(g)}$ the effect of the plank, nested in the treatment effect, $L_{h(f)}$ the effect of the lane, nested in the planks, $B_{m(f)}$ the effect of the block, nested in the planks, and $E_{n(g)}$ the effect of presence or absence of the fertilizer, nested in the treatment. $P.T_{ig}$ is the interaction effect between the population and the treatment effects. All these provenance and environmental effects were considered as fixed effects, i.e. effects affect the mean of the phenotypic distributions. The random effects are the additive genetic values of the individuals k , $\{a_k\} \sim N(0, A \times V_A)$ with A the genetic relatedness matrix, and the residuals effects, $\{\epsilon_{i,f,g,h,m,n,k}\} \sim N(0, I_{de} \times V_R)$. V_A is the additive genetic variance and V_R the residual variance. The covariance structure of the additive genetic effects was defined by a incomplete pedigree information (paternal identity known for 33 % of the genotyped individuals, i.e. 12 % of all individuals; Gauzere et al. (a) in prep). This linear mixed model defined at the individual level is also called "animal model".

To analyse the traits presenting significant maternal effects, we included to the models 4.1

and 4.3 the maternal identity, m_j , as a random effect :

$$Y_{i,f,g,h,m,n,j,k} = \mu + P_i + T_g + H_{f(g)} + L_{h(f)} + B_{m(f)} + E_{n(g)} + P.T_{ig} + a_{j,k} + m_j + \epsilon_{i,f,g,h,m,n,j,k} \quad (4.2)$$

The population genetic differentiation and the plastic response of the traits were investigated by testing the significance of the population, P , and treatment, T , fixed effects in the model 4.1. We also investigated whether the plastic response of the traits differed between populations by testing the population-by-treatment interaction effect, $P.T$. In presence of significant or marginal $P.T$ effect we measured the population effects in each treatment modality independently with the model :

$$Y_{i,f,h,m,n,k} = \mu + P_i + H_f + L_h + B_{m(f)} + E_n + a_k + \epsilon_{i,f,h,m,n,k} \quad (4.3)$$

Three of the studied traits were transformed to improve the adequacy with the hypothesis of the linear mixed models : δH with a logarithm transformation, LM and LA with a square-root transformation (Gauzere et al (b) *in prep*). We analysed these models using a restricted maximum likelihood procedure with the "AsReml-R" package in R (Gilmour *et al.*, 2006). Fixed effects were tested with a conditional Wald F-test, using the function "wald.asreml". Tests were considered significant if their p_{value} was lower than 0.05 and marginally significant if $0.05 < p_{value} < 0.1$.

TEST FOR ADAPTIVE DIFFERENTIATION

We used the method recently developed by Ovaskainen *et al.* 2011 to detect signatures of natural selection based on differentiation of traits. The method applies to phenotypic data of offspring measured in a common garden, collected from different natural populations, and for which the one-generation pedigree is known. This methods largely relies on the hierarchical Bayesian framework that enables statistical inference from observations resulting from several hierarchical levels of randomness (O'Hara *et al.*, 2008). In this method, the molecular marker data and the phenotypic data (mono-trait or multi-trait) are analyzed under a neutral model that assumes that the genes coding for the trait(s) are as divergent as the neutral markers ($Q_{st} = F_{st}$ case). Testing the occurrence of disruptive selection or stabilizing selection then rely on a goodness-of-fit statistics S . We briefly present the key steps of the method below for a single quantitative trait.

The first level of the statistical model considers the population mean genetic additive values, assuming that they are normally distributed with the covariance between a pair of population means being proportional to the pairwise average coancestry between these populations (thus assuming neutrality for the trait). This assumption mimics the decomposition of the individual genetic values in the animal model :

$$\mathbf{A}^{\mathbf{P}} \sim N(0, 2V_A \times \theta^{\mathbf{P}}) \quad (4.4)$$

with $\mathbf{A}^{\mathbf{P}}$ be the vector of the population-level effects , V_A the additive genetic variance of the trait, $\theta^{\mathbf{P}}$ the matrix of pairwise population-level coancestry coefficients. The validity of this model depends on a low mutation rate.

The matrix $\theta^{\mathbf{P}}$ is estimated from the molecular marker data using an admixture F Model generalized from Gaggiotti & Foll (2010) (Karhunen & Ovaskainen, 2012). Although it can be integrated in the hierarchical model and analyzed jointly with the trait data, it is estimated in a previous step for practical reasons ("afm" of the package RAFM).

The second level of the statistical model considers the individual phenotypic values Z_i as the sum of the fixed effects (μ_i), the population-level additive effect (p_i), the individual-level additive effect (s_i) and the environmental effect (e_i). The population-level additive effect (p_i) is determined by the mean genetic additive values in the populations of the parents of individual i and the individual-level additive effects (s_i) are normally distributed with covariances between s_i and $s_{i'}$ equal to $2V_A$ times the relatedness between i and i' generated by their common parents in the one-generation pedigree. The environmental effects (e_i 's) are independent normal variables with variance V_E . The individual-level additive effects (s_i) also include the effect of inbreeding.

This second level thus corresponds to an animal model in which the additive value is decomposed into a population effect and an individual effect.

A Monte Carlo Markov Chain based on a Metropolis-Hastings algorithm is used to retrieve posterior distributions for the parameters (V_A and fixed effects) and the latent variables (in particular the $\mathbf{A}^{\mathbf{P}}$ vector). To test the presence of selection on the trait, a test statistics S is derived that computes whether the realized pattern of population mean genetic additive values is likely under the neutral model (equation 4.4) :

$$S = Pr(f_P(\mathbf{A}^P) < f_P(N(0, 2V_A \times \theta^{\mathbf{P}}))), \quad (4.5)$$

with f_P the probability density of the distribution of equation 4.4 and \mathbf{A}^P the realized pattern of population divergence. If the pattern of population divergence is not governed by selection then the the probability density f_P for \mathbf{A}^P will be similar to the one expected for a random variable that follows 4.4 (i.e., $S \sim 0.5$ indicates signal of neutral divergence). Otherwise, $S > 0.95$ indicates signal of divergent selection ; $S < 0.05$ indicates signal of stabilizing selection.

The method is implemented in R-packages using two-step procedures : first, the admixture F-model for neutral divergence among populations is implemented in the function "afm" of the package RAFM and provides $\theta^{\mathbf{P}}$ (Karhunen & Ovaskainen, 2012) ; second, and the phenotypic divergence is analyzed with the function "MH" of the package DRIFTSEL, which estimate μ , V_A , \mathbf{A}^P and S (Karhunen *et al.*, 2013).

For our dataset, we applied "afm" on the microsatellite genetic data of the adult trees in the natural populations. We also estimated the classical index of neutral genetic differentiation, F_{st} . We used DRIFTSEL for the quantitative traits of the seedlings located in the watered condition of the common garden. We used the fixed effects listed in the following model (identical to the model used in Gauzere et al. (b) *in prep*) :

$$P_{m,n,j,k} = \mu + B_m + L_h + E_n + p_{j,k} + s_{j,k} + \epsilon_{m,n,j,k} \quad (4.6)$$

where B_m is the block effect, L_h is the lane effect, E_n is the fertilizer effect and $p_{j,k}$ and $s_{j,k}$ are the population-level and individual-level additive effects. The model was informed by the pedigree information available from the paternity analysis conducted with Cervus 3.0.

These bayesian analyses were performed using 50 000 Monte Carlo Markov Chain (MCMC) iterations with a burn in of 10 000 iterations and an interval of 3 iterations to sample in the MCMC. We checked the convergence of the chains visually and by using the function "auto-corr" of the package CODA which evaluates the autocorrelation of the MCMC chains : low autocorrelation within the chains indicates that the space of the parameter values is efficiently explored and thus suggests no convergence problems.

Results

IN SITU PHENOTYPIC VARIATION

In situ, the dates of bud burst ($t_{b2\rightarrow3}$) and leaf senescence ($t_{s1\rightarrow2}$) and the vegetation season duration (VD) significantly differed between populations, with $t_{b2\rightarrow3}$ increasing with the altitude and $t_{s1\rightarrow2}$ and VD decreasing with elevation (Figure 4.3). The two years of bud burst monitoring showed the same phenotypic trend along the elevational gradient. Along the gradient, the phenotypic delay in phenology was almost equal to 12 days/1 kilometer for $t_{b2\rightarrow3(2012)}$, 6 days/1 kilometer for $t_{b2\rightarrow3(2013)}$, 5 days/1 kilometer for $t_{s1\rightarrow2(2013)}$ and 13 days/1 kilometer for $VD_{(2013)}$. For $\%N$, a consistent pattern of phenotypic differentiation was found by the analysis of the data sets 2 and 3, with a significantly lower $\%N$ for population N2 compared to the two others (only the results of the dataset 3 were represented; Figure 4.3). However, for LA , LMA and $\delta^{13}C$ only the analysis of the data set 3 revealed significant or marginal differences between populations : LA increased with the elevation, LMA decreased with the elevation and $\delta^{13}C$ tended to decrease with the elevation (Figure 4.3). There was no in situ phenotypic differentiation between populations for LM and $\%C$ (results not shown).

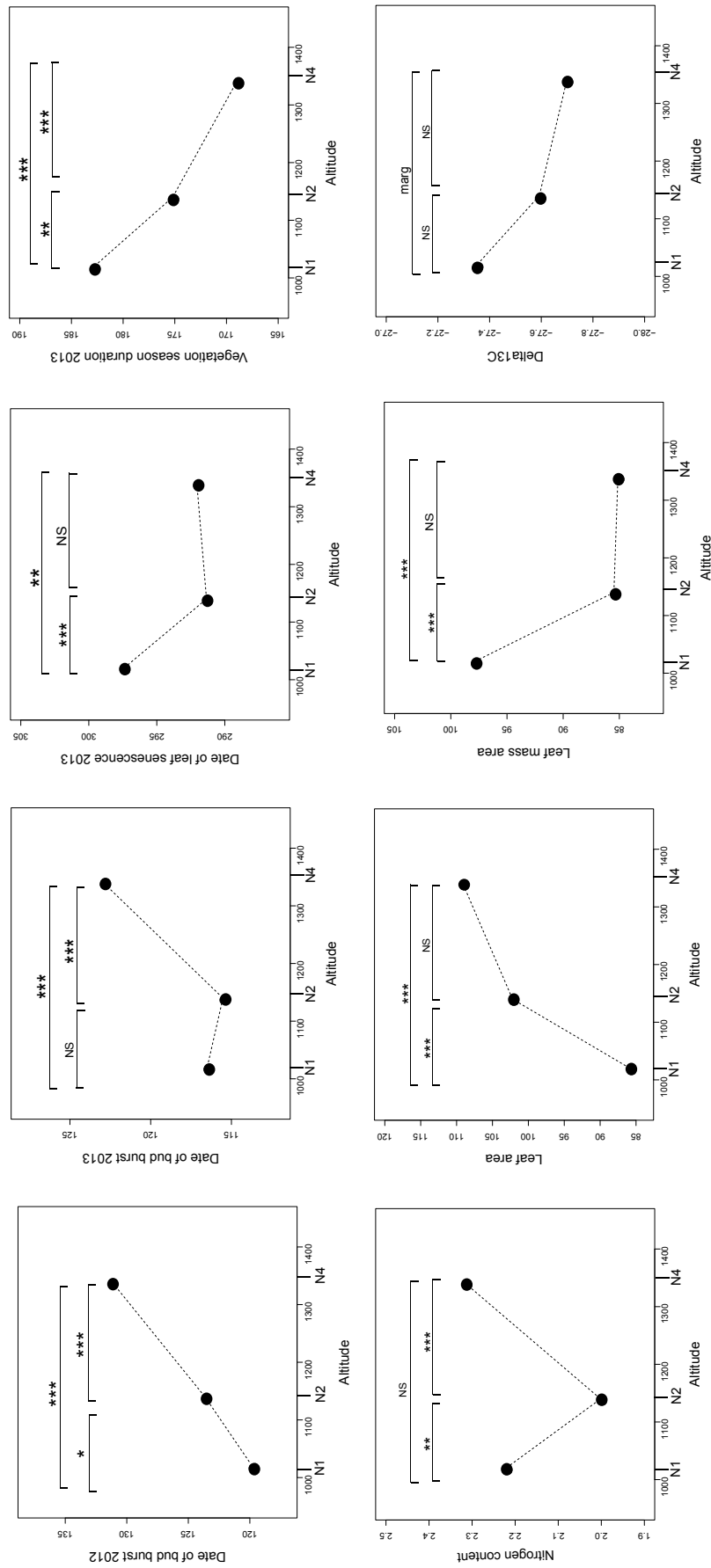


FIGURE 4.3 – In situ differentiation of the phenological, morphological and physiological traits. Date of passage by key stages were measured in julian days. The other traits had the same units than the traits measured ex situ. Only the results of the analysis of the data set 1 for the phenological traits and the data set 3 for the morphological and physiological traits were represented here. Legend : "marg" 0.1 < p_{value} < 0.05; "*" 0.01 < p_{value} < 0.05; "**" 0.001 < p_{value} < 0.01; "***" 0.0001 < p_{value} < 0.001; "NS" p_{value} > 0.05.

GENETIC DIFFERENTIATION DETECTED IN THE COMMON ENVIRONMENT

Significant genetic differentiation between the populations was found for δH , δD , $t_{b2 \rightarrow 3(2011)}$, $t_{b2 \rightarrow 3(2012)}$ and $\%C$. Marginal population differentiation was found for $t_{s1 \rightarrow 2(2011)}$ (Table 4.4). Population N2 presented highest height and diameter growth. Population N1 presented the highest $\%C$ (Figure 4.4). Phenological traits presented almost continuous pattern of genetic variation with elevation, with $t_{b2 \rightarrow 3(2011)}$, $t_{b2 \rightarrow 3(2012)}$ and $t_{s1 \rightarrow 2(2011)}$ decreasing with elevation (Figure 4.4). For the phenological traits we can see that the genetic differences between the populations were quantitatively low compared to the range of phenotypic variation found in situ : between populations N1 and N4 we estimated a difference of 23 degree days for $t_{b2 \rightarrow 3(2011)}$, that is about 1.5 julian days, instead of an average of 9 days in situ (Figures 4.4).

Trait	<i>P</i> effect	<i>T</i> effect	<i>P.T</i> effect	<i>P_{watered}</i>	<i>P_{stressed}</i>
δH_{ln}	0.0037 **	0.0502 .	0.4910	-	-
δD	< 0.0001 ***	< 0.0001 ***	0.0883	-	-
<i>BioM</i>	0.1599	< 0.0001 ***	0.8449	-	-
$t_{b2 \rightarrow 3(2011)}$	< 0.0001 ***	0.7234	0.3885	-	-
$t_{b2 \rightarrow 3(2012)}$	0.0078 **	0.0001 ***	0.9317	-	-
$t_{s1 \rightarrow 2(2011)}$	0.0767 .	-	-	-	-
<i>VD₂₀₁₁</i>	0.4866	-	-	-	-
$LM_{\sqrt{}}$	0.7691	< 0.0001 ***	0.0608 .	0.7738	0.6654
$LA_{\sqrt{}}$	0.4001	< 0.0001 ***	0.1145	-	-
<i>LMA</i>	0.2027	< 0.0001 ***	0.0728 .	0.4421	0.2185
$\%N$	0.1603	< 0.0001 ***	0.0107 *	0.7104	0.0405 *
$\%C$	< 0.0001 ***	< 0.0001 ***	0.0636 .	< 0.0001 ***	0.0007 ***
$\delta^{13}C$	0.3893	< 0.0001 ***	0.1137	-	-

TABLE 4.4 – Effect of the population of origin (*P*), treatment condition (*T*) and population-by-treatment interaction effect (*P.T*). *P_{watered}* and *P_{stressed}* are the test of the population effect in the watered and water stress conditions respectively using the model 4.3. *ln* and $\sqrt{}$ indicate that a transformation has been applied. Legend : "." $0.1 < p_{value} < 0.05$; "*" $0.01 < p_{value} < 0.05$; "***" $0.001 < p_{value} < 0.01$; "****" $p_{value} < 0.001$.

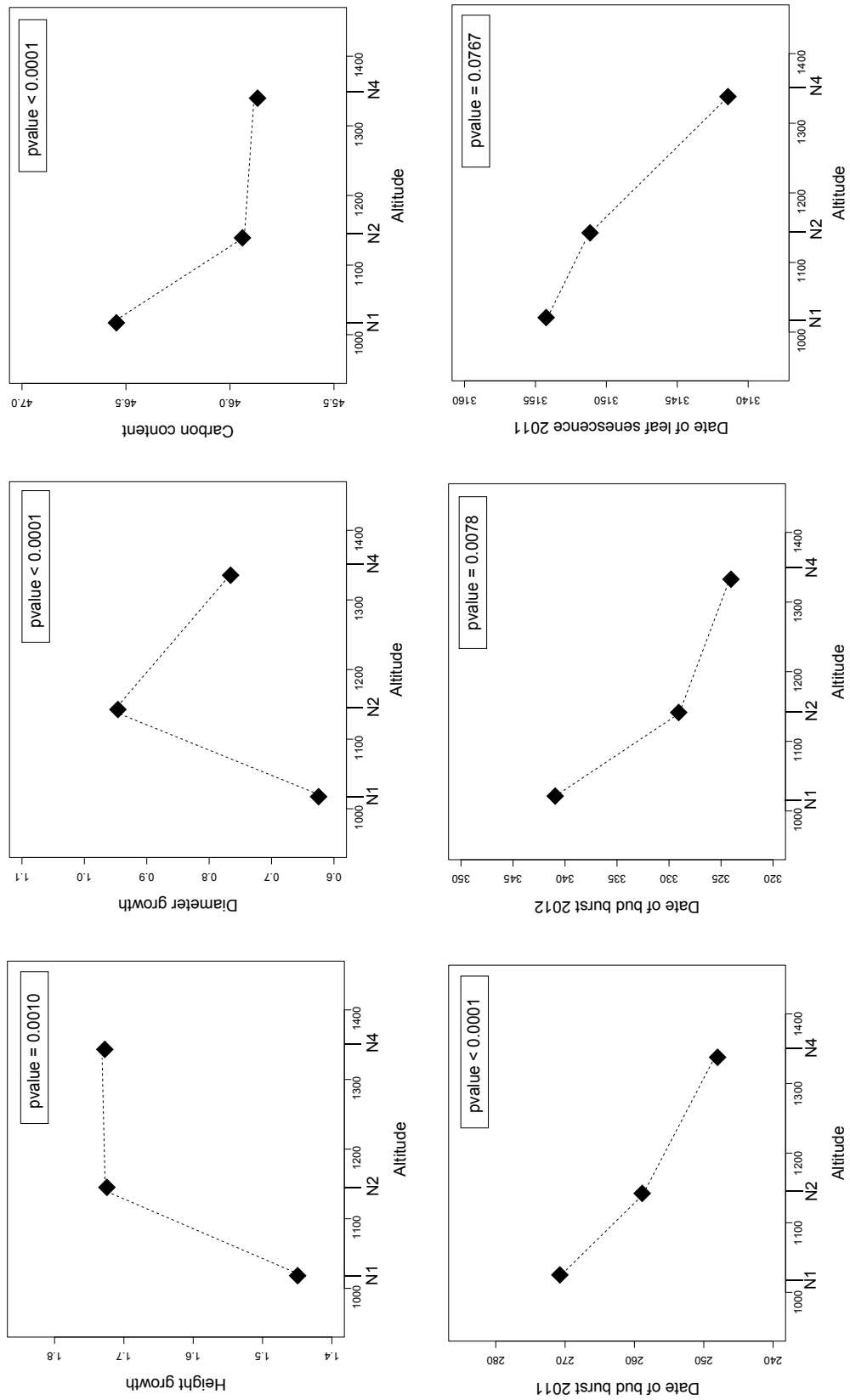


FIGURE 4.4 – Ex situ genetic differentiation of the phenological, growth, morphological and physiological traits between populations. *N.B.* : Phenological traits are represented in degree days

PLASTIC RESPONSE MEASURED EX SITU

As expected, many traits were significantly affected by the water treatment in the common garden experiment (9 traits in total) : $BioM$, δD , $t_{b2 \rightarrow 3(2012)}$, LM , LA , LMA , $\%C$, $\%N$ and $\delta^{13}C$. In the water stress condition, the biomass production and diameter growth were lower than in the watered condition. In the stress condition, the leaves presented lower LM , LA and LMA . The leaves also contained less nitrogen and more carbon and their $\delta^{13}C$ were less negative than in the watered condition. The date of bud burst in 2012 was earlier in the water stress condition than in the watered condition.

	$BioM$	δD	$t_{b2 \rightarrow 3(2012)}$	LM	LA	LMA	$\delta^{13}C$	$\%C$	$\%N$
W	13.690	0.994	336.463	0.174	30.494	57.258	-29.033	46.429	2.149
S	11.725	0.948	334.699	0.143	26.584	53.957	-28.502	47.814	1.977

TABLE 4.5 – Mean trait estimates in the watered (W) and water stress (S) conditions of the common garden experiment for the traits presenting significant treatment effect. Significance of the treatment effect is detailed in Table 4.4.

POPULATION-BY-TREATMENT EFFECT DETECTED EX SITU

The leaf nitrogen content, $\%N$, was the only trait presenting a significant population-by-treatment interaction effect. Marginal population-by-treatment effects were detected for $LM_{\sqrt{}}$,

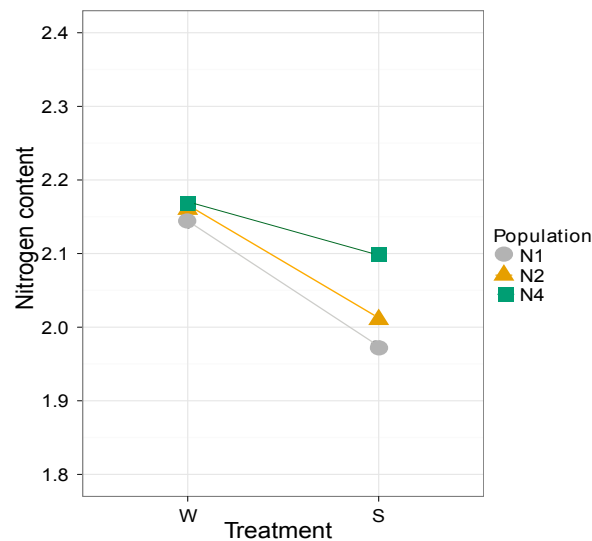


FIGURE 4.5 – Representation of the population means in the different water content treatment for the trait presenting significant population-by-treatment effect : $\%N$.

LMA, %*C* and δD . %*N* notably presented a significant population differentiation in the water stress condition but not in the watered condition (Table 4.4). This result can also be interpreted as a highest plastic response to drought for the populations at the bottom of the gradient (N1 and N2) ; population N4 presented nearly no plastic response (Figure 4.5).

SIGNAL OF NATURAL SELECTION

The MCMC chains of the "afm" and "MH" bayesian analyses presented very low autocorrelation structure suggesting that the convergence was easily achieved (results not shown). The admixture F-model of Karhunen & Ovaskainen (2012) estimated a posterior *Fst* value of 0.026 [0.023 ; 0.029]. The diagonal of the θ^P matrix representing the coancestry within the populations indicated that population N1 tend to have the lowest population size and populations N2 and N4 similar effective sizes (Figure 4.6). The off-diagonal element of θ^P showed that the two closest populations (N1 and N2) exchanged more gene flow than did population N4 with the two others ($\theta_{N1N2} > \theta_{N1N4} \sim \theta_{N2N4}$; Figure 4.6). However, the coancestry estimates were low (maximum $\theta_{N1} = 0.065$) indicating that effective population sizes were globally large.

$$\theta^P = \begin{pmatrix} 0.06516 & 0.024380 & 0.0162200 \\ 0.024380 & 0.03204 & 0.015190 \\ 0.0162200 & 0.015190 & 0.03824 \end{pmatrix}$$

FIGURE 4.6 – Matrix of the coancestry coefficients estimated by the admixture F-model of Karhunen & Ovaskainen (2012).

Trait	<i>S</i>	Trait	<i>S</i>
δH_{ln}	0.96	LM_{\checkmark}	0.27
δD	0.95	LA_{\checkmark}	0.42
<i>BioM</i>	1	<i>LMA</i>	0.58
$t_{b2 \rightarrow 3(2011)}$	0.99	% <i>N</i>	0.38
$t_{b2 \rightarrow 3(2012)}$	0.96	% <i>C</i>	0.95
$t_{s1 \rightarrow 2(2011)}$	1	$\delta^{13}C$	0.35
VD_{2011}	1		

TABLE 4.6 – Signal of selection (*S*) in three beech populations measured for each trait using the method of Ovaskainen *et al.* (2011). *S* coefficient above 0.95 indicates significant signal of divergent selection.

For 8 traits among the 13 in total, we detected a strong signal of natural selection on the pattern of phenotypic divergence between populations (i.e. $S > 0.95$) : δH_{ln} , δD , $BioM$, $t_{b2 \rightarrow 3(2011)}$, $t_{b2 \rightarrow 3(2012)}$, $t_{s1 \rightarrow 2(2011)}$, VD_{2011} and $\%C$. These traits were the same than previously detected by the test of the population effect in the quantitative genetic model, except $BioM$ and VD_{2011} . The approach of Ovaskainen *et al.* (2011) thus indicated a significant genetic divergence for these traits, with increased length of vegetative season and biomass production with elevation (Figure 4.7).

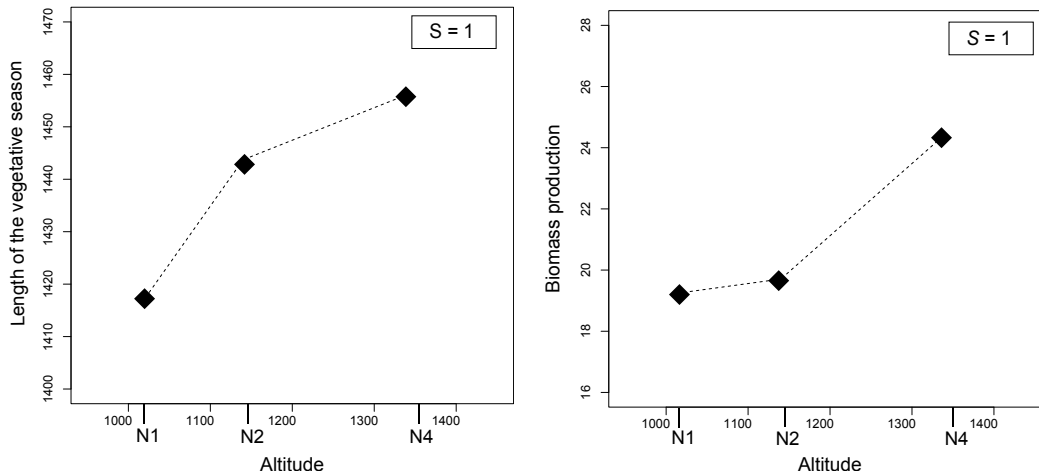


FIGURE 4.7 – Genetic differentiation of the two traits significantly detected under divergent selection by the method of Ovaskainen *et al.* (2011) ($S > 0.95$) and that were not detected as divergent with a classical quantitative genetic model (equation 4.1; results in Table 4.4). *N.B.* : Phenological traits are represented in degree days

Discussion

Investigations of the genetic and plastic variation of functional traits related to temperature and drought response in tree populations are rare. Some studies have already measured the genetic variation of ecophysiological traits in response to temperature variation in beech populations (Vitasse *et al.*, 2009; Bresson *et al.*, 2011). However, to our knowledge, this study is the first to investigate the differentiation of beech populations in drought limiting conditions and to use a proper quantitative genetic approach to accurately test if the genetic divergence between populations is adaptive.

GENETIC DIFFERENTIATION OCCURS DESPITE GENE FLOW

The low Fst value ($Fst = 0.026$) indicated that most of the genetic variation was contained within populations and only a moderate differentiation existed between the populations. Similarly, Lander *et al.* (2011) estimated an overall Fst of 0.026 analysing 51 sites on three regions of Mont-Ventoux using the same 13 microsatellites as those used in our study. This low neutral differentiation between populations can be due to the large population sizes ($Ne = 1.82 \times 10^3$; Lander *et al.* 2011) and high gene flow between populations (on average the immigrant pollen flow equals $m = 53.3\%$; Gauzere *et al.* 2013a).

Along the elevational gradient, genetic differentiation between populations was detected for 6 traits among the 13 studied : the height and diameter growth, bud burst, leaf senescence and leaf carbon content (Table 4.4). More precisely, growth performances were highest in the population at the middle of the gradient, bud burst and leaf senescence were earliest at higher elevations and leaf carbon content was highest at the bottom of the elevational gradient. The genetic differentiation of these traits was significantly stronger than neutral differentiation indicating an effect of natural selection. These results highlight that divergent selective pressures are strong enough to counteract the homogenizing action of gene flow on population differentiation. Patterns of genetic divergence at restricted spatial scale has been showed for many plant species (Linhart & Grant, 1996). But among these studies, patterns of microdifferentiation, i.e. at the scale of the meter to the kilometer, have mainly been found for annual plant species (Linhart & Grant, 1996; Gonzalo-Turpin & Hazard, 2009). To our knowledge patterns of adaptive divergence at the scale of 1 kilometer has never been reported for a temperate tree species (for tropical tree species see Brousseau *et al.* 2013).

Our study also highlighted the interest of the method of Ovaskainen *et al.* (2011) to detect adaptive divergence between a small number of populations. With our experimental design, the classical $Fst - Qst$ comparison method was not suitable to infer a signal of natural selection. Indeed, with only three populations we were not able to construct accurate confidence intervals around the Qst estimates. Moreover, we measured quantitatively low genetic divergences between populations : for instance, the Qst value estimated for the bud burst phenology was only of 0.06 (result not shown). Using a simulation approach, Karhunen *et al.* (2013) showed that for this range of genetic divergence the classical $Fst - Qst$ methods generally fail to detect a signal of natural selection. By using all the pairwise population comparison, the

method of Ovaskainen *et al.* (2011) succeeded in these cases where the $Fst - Qst$ methods would fail.

Note however that we have not used yet the approach of Ovaskainen *et al.* (2011) by informing the maternal effects in the quantitative genetic models. It probably explained why we found a discrepancy between the results of the test of the genetic differentiation with the informed quantitative genetic model (equation 4.2) and the S coefficient of Ovaskainen *et al.* (2011) for the total biomass production, since this trait was affected by a maternal effect ($pvalue$ population effect = 0.1599; Table 4.4; $S= 1$; Table 4.6; see section 3.3 for results on maternal effects). However, for the vegetative season duration, not affected by maternal effects, the discrepancy between the two approaches used is more tricky to explain ($pvalue$ population effect = 0.4866; Table 4.4; $S= 1$; Table 4.6). According to our previous results, i.e. a genetic divergence of both the bud burst and leaf senescence phenology, we can expect that the approach of Ovaskainen *et al.* (2011) accurately detect signal of natural selection for extended duration of vegetative season at high elevations. However, the hypothesis that DRIFTSEL is just more powerful to detect small effects as significant is challenged by the fact that VD_{2011} had not the smallest non significant $pvalue$ in Table 4.4 (the smallest $pvalue$ is found for $\%N$ but it does not appear as significant with DRIFTSEL).

PLASTIC VERSUS GENETIC RESPONSE OF FUNCTIONAL TRAITS TO DROUGHT

As expected, the morphological and physiological traits presented different average phenotypic value in the water treatment highlighting the high plastic response of leaf functional traits for beech trees (Bresson *et al.*, 2011). Overall, the water stress treatment the seedlings presented lowest investment in total above-ground growth, leaf sizes and leaf nitrogen content, but highest water use efficiency and leaf carbon content. More surprisingly, we observed that seedlings in the water stress treatment presented earliest bud burst. This plastic response could allow to start photosynthesizing and accumulating carbon before drought occurs, and thus would be an adaptative answer to drought occurrence.

The plastic response to drought measured *ex situ* and the pattern of phenotypic variation observed *in situ* can be compared to provide evidence that drought variation along the elevational gradient is one of the main environmental pressure shaping phenotypic variation. This

comparison is notably relevant for the traits presenting no genetic divergence between populations, since in these cases the in situ phenotypic variation must be explained by plasticity only. For LA and $\delta^{13}C$, the plastic responses to drought measured ex situ is consistent with the patterns of phenotypic variation in situ : the lowest LA and highest $\delta^{13}C$ were found at the bottom of the elevational gradient, i.e. in presence of the driest conditions. Therefore, for these two traits we can expect that the pattern of trait variation in situ was mainly driven by a plastic response to a drought environmental cline. However, for $\%N$ and LMA the patterns of phenotypic variation in situ can not be understood by the plastic response to drought only. Indeed, if drought was the only ecological driver then the LMA and $\%N$ should be the lowest in the driest population (N1). Thus, other environmental conditions than the humidity rate must significantly differ between populations. The competitor densities, which has already been shown to vary between populations (Gauzere *et al.*, 2013a), is likely to be a main ecological factors affecting the population phenotypic response in natural conditions (Bontemps, 2012).

For the traits measured both in situ and ex situ, we detected a pattern of cogradient variation (CoGV) for leaf senescence and diameter growth (Cailleret & Davi, 2011). In patterns of CoGV, the phenotypic plasticity tends to accentuate the phenotypic divergences generated by selection ; thus in these cases phenotypic plasticity is generally considered adaptive. However, this notion of "adaptive plasticity" must be taken with caution, since Chevin *et al.* (2013) recently showed that these patterns of plasticity, which helps to track the phenotypic optimum, are expected to slow down the action of natural selection. Patterns of countergradient variation (CnGV) were detected for bud burst phenology and the leaf carbon content. Note that the plastic response completely hides the genetic variation of $\%C$ in situ. Thus, patterns of CnGV are particularly interesting since they allow to understand apparent lack of phenotypic response to selection in nature (Conover & Schultz, 1995). The increasing number of articles reporting patterns of CnGV for many morphological and physiological traits in plant and animal species indicated that CnGV may be commun in nature (Conover *et al.*, 2009). Our results highlight the importance to understand causes of phenotypic similarities as well as causes of phenotypic divergences in nature.

Finally, the detection of different population reaction norms to drought for $\%N$ (as depicted by the significant population per treatment interaction ; Figure 4.5) suggests that adaptive

response to drought can occur via plasticity. Gene-by-environment interactions, i.e. genetic variance for plasticity, has already been reported for a variety of species, meaning that plasticity can evolve by responding to natural selection as any other trait (Scheiner, 1993; Pigliucci, 2005). Temporal variability of the environment is expected to select for more plasticity. Here, we showed that the populations at the bottom of the gradient, i.e. populations that suffer the summer droughts the more variable between years and that experience the driest environmental conditions, also presented the largest plastic responses. The difference in plastic response among populations could thus be explained by a selection for widely plastic genotypes in drought limiting conditions. However, we can wonder why this pattern is not observed for other functional traits related to drought response, such as *LMA* or $\delta^{13}C$. Moreover, as the leaf nitrogen content is a very integrative trait, its variation can be affected by many other biological processes than the photosynthesis capacity; it can notably depend on the capacity of nitrogen absorption. Thus, independently of the variation of other leaf morphological or physiological traits, it is quite difficult to discuss the response of %N in the water stress condition according to some adaptive response (i.e. to provide the evidence that this plastic response allows to increase the fitness in this particular environment). More generally, the lack of theoretical and empirical work on the evolution of plasticity still limits our understanding of the role of the evolution of plasticity in adaptation (Chevin *et al.*, 2013).

ADAPTATION OF BEECH POPULATIONS TO MEDITERRANEAN CLIMATE

The traits expected to be the most directly linked to drought response (i.e. leaf morphology and water use efficiency) were not genetically differentiated between beech populations along the elevational gradient of Mont-Ventoux. The leaf carbon content is the only physiological trait to present adaptive divergence. However, the variation of %C depends on too many ecophysiological processes (e.g. leaf structure, carbon storage), which precludes an easy interpretation with regards to adaptation to drought or another environmental pressure. A first explanation of the lack of adaptive divergence for morphological and physiological traits is that studied elevational gradient does not fully capture the spatial scale of variation for drought stress (or, at least, less appropriately than for the sum of temperatures). Nourtier *et al.* (2013) have notably shown that soil water storage capacity was highly variable within sites. The microheterogeneity of the soil may thus prevent the genetic differentiation between populations of leaf traits, which are moreover highly plastic. Secondly, we cannot rule out

that population N1 effectively evolved traits conferring it advantages under drought stress, but that we could not detect this evolution, either because it involves functional traits that we did not measure (such as hydraulic traits), or because it involves coordinated suite of traits (whereas here we search for a single-trait signal of adaptation). Finally, traits were measured in seedlings grown in a common garden outside of the *Fagus sylvatica* range ; we cannot rule out that these environmental conditions affect the expression of the additive genetic effects, and blur a signal of genetic adaptation that could be revealed inside the natural populations.

On the other hand, the genetic differentiation of juvenile growth traits constitute strong evidence of adaptation to limiting ecological conditions (Saenz-Romero *et al.*, 2006). In this study, the highest diameter growth was found in population N2 (Figure 4.4) and the highest height growth were found at the middle and the top of the gradient. This pattern is opposite to the one classically characterized for height growth along elevational gradient for beech by common garden experiments : populations at lowest elevation classically presented the highest height (Vitasse *et al.*, 2009). Our results could suggest that in a sub-Mediterranean ecosystem, beech populations at low elevation display a more conservative growth strategy, probably associated to increase drought tolerance. This conservative strategy should allow to invest more carbon in other performance traits, such as survival, vegetative or root biomass production. Unfortunately, our experiment did not allow mortality to be analysed (because of important frost damage in winter 2011) and the measure of vegetative production, via the *LMA*, did not indicate that individuals from population N1 invested more in this performance trait than populations at middle and high elevations.

CONCLUSION

Our study highlighted that (i) patterns of adaptive divergences can occur over very short spatial scale and despite high immigrant gene flow, (ii) morphological and physiological traits presented plastic response to drought, (iii) adaptation to Mediterranean climate occurs via conservative growth strategy. A better characterisation of adaptive differentiation of functional traits in response to drought would require to account for the fact that the drought variation is less continuous than the temperature variation along the gradient of Mont-Ventoux.

4.2 Capacités évolutives et contraintes multivariées à l'évolution

Introduction

Un des enjeux principaux de la biologie évolutive est de comprendre les mécanismes qui contraignent ou favorisent l'adaptation des populations à leur environnement. Dans le contexte du changement climatique actuel, il est devenu crucial de répondre à cette interrogation afin d'anticiper les conséquences qu'auront les changements environnementaux sur le maintien ou l'extinction locale des espèces. Un des modèles les plus classiquement utilisés pour prédire le taux d'évolution d'un trait est l'équation du sélectionneur : $\Delta z = V_A \times \beta$, avec Δz le changement moyen de la valeur du trait d'une génération à la suivante, V_A la variance génétique additive et β la force de la sélection exercée sur ce trait (Falconer & MacKay, 1996). Les variances génétiques additives des traits sont souvent standardisées par rapport à leur variances phénotypiques (V_P) afin de dériver une mesure comparable entre traits, populations et espèces appelée "héritabilité" : $h^2 = V_A/V_P$ (Falconer & MacKay, 1996). La réponse à la sélection s'écrit alors : $\Delta z = V_P \times h^2 \times \beta$. Cependant, de plus en plus d'incohérences sont relevées entre les taux d'évolution observés et ceux prédits par l'équation du sélectionneur (Merila *et al.*, 2001; Morrissey *et al.*, 2010). Une première explication possible à ces incohérences est que la mesure d'héritabilité n'est pas une mesure correcte pour évaluer le potentiel adaptatif d'une population ; une deuxième provient du fait de considérer l'évolution d'un trait indépendamment des autres traits pouvant répondre à la sélection.

De l'équation de réponse à la sélection, on peut définir un paramètre d'évolutivité¹ d'un trait comme étant : $e = \Delta z/\beta = V_A$. Classiquement, on utilise deux options pour standardiser e : soit par rapport à la variance phénotypique totale du trait, soit par rapport à la moyenne phénotypique du trait (m ; Hansen *et al.* 2011)

$$e_\sigma = V_A/V_P = h^2$$

$$e_\mu = V_A/m^2 = I_A$$

La variance additive standardisée par la moyenne du trait, I_A , est directement reliée au coefficient de variation génétique additive, CV_A , tel que définit par Houle (1992) : $I_A =$

1. "evolvability" en anglais

$(CV_A/100)^2$. CV_A a plus de sens que I_A en tant que mesure de l'évolutivité car sa valeur peut être interprétée comme le pourcentage de changement du trait sous une unité de force de sélection. Alors que h^2 et I_A sont censées être deux mesures de l'évolutivité, leurs variations sont quasi-indépendantes (Hansen *et al.*, 2011). La standardisation n'est en réalité pas une action anodine et notamment la standardisation de V_A par V_P n'a de sens que si d'autres composantes de variances, possiblement incluses dans V_P , ne varient pas avec V_A (Hansen *et al.*, 2011). En théorie on s'attend pourtant à ce que les variances génétiques épistatiques (V_{AA}) et variance génétique de dominance (V_D) soient reliées à V_A (Figure 4.8; Merila & Sheldon 1999), ce qui a pour conséquence de générer des relations complexes entre h^2 et l'évolutivité. De ce fait, de nombreux auteurs recommandent d'utiliser les estimations de I_A ou CV_A plutôt que h^2 pour caractériser l'évolutivité des traits (Houle, 1992; Hereford *et al.*, 2004; Hansen *et al.*, 2011).

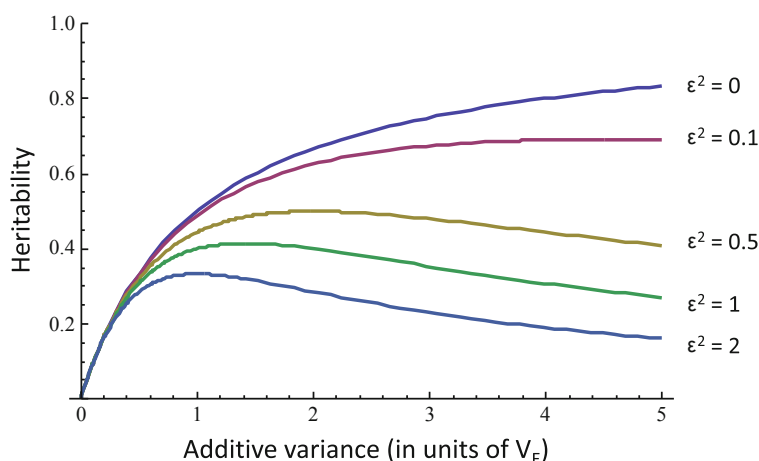


FIGURE 4.8 – Relation entre héritabilité et variance génétique additive dans un modèle multilinéaire d'interaction de gènes. Le paramètre ϵ^2 mesure l'intensité des effets épistatiques, étant donné que la variance des effets épistatique est égale à $V_{AA} = \epsilon^2 V_A^2 / 2$. Plus les effets épistatiques sont forts et plus la relation entre h^2 et V_A atteint un maximum puis est décroissante quand V_A augmente. Figure extrait de Hansen *et al.* (2011)

D'autre part, les différents traits affectant la valeur sélective n'étant généralement pas indépendants, leur évolution doit être considérée conjointement. Cette indépendance est d'autant plus vraie que la sélection a pour effet de réduire la variance génétique en générant des covariances négatives entre traits (effet Bulmer; Bulmer 1971). Si l'on considère les covariances génétiques, le changement moyen multivarié de valeurs de traits sur une génération s'écrit

alors : $\Delta\bar{z} = \mathbf{G} \times \beta$, avec $\Delta\bar{z} = \{\Delta z_1, \Delta z_2, \dots, \Delta z_n\}$ le vecteur des changements moyens de n traits, \mathbf{G} la matrice de variance-covariance additive génétique des traits et $\beta = \{\beta_1, \beta_2, \dots, \beta_n\}$ le vecteur des gradients de sélection directionnels des n traits (Lande & Arnold, 1983). Les gradients de sélection étant difficiles à mesurer en populations naturelles (car cela nécessite d'avoir des mesures de fitness et de traits indépendantes des effets environnementaux), les études s'intéressant au potentiel évolutif des populations s'attachent d'abord à caractériser la matrice \mathbf{G} . Classiquement, des tests de descendance en environnement contrôlé sont utilisés afin de découpler les effets génétiques et environnementaux dans l'estimation des variances et covariances composant \mathbf{G} .

L'approche empirique la plus commune pour étudier les contraintes évolutives est de caractériser les corrélations génétiques entre traits. L'avantage de la standardisation des covariances génétiques par les écarts-types génétique ($\sqrt{V_A}$) est une fois de plus d'avoir des variables plus parlantes biologiquement (e.g. non dépendantes de l'échelle de mesure) et comparables. Cependant, les corrélations génétiques sont de mauvais prédicteurs de la réponse évolutive puisque qu'elles varient en fonction de deux paramètres clefs de l'évolution : les variances et covariances entre traits

$$r_{z_1 z_2} = \frac{cov_{z_1 z_2}}{\sqrt{V_{z_1} \times V_{z_2}}}$$

En effet, on peut démontrer numériquement que deux matrices \mathbf{G} produisant des corrélations génétiques similaires entre traits peuvent avoir des réponses à la sélection différentes et, à l'inverse, que deux matrices \mathbf{G} avec des corrélations génétiques entre traits différentes peuvent générer la même réponse à la sélection (Conner, 2012).

Afin d'éviter ces erreurs d'interprétations et d'avoir une compréhension plus globale des contraintes multitraits, une approche consiste à étudier la dimensionnalité de la matrice \mathbf{G} (Hine & Blows, 2006; McGuigan, 2006; Kirkpatrick, 2009). Un paramètre important pour caractériser \mathbf{G} est notamment son nombre efficace de dimensions. Ce nombre efficace de dimensions traduit le fait que le nombre de combinaisons de traits qui peuvent évoluer sous effet de la sélection peut être inférieur au nombre de traits, même si ces derniers présentent une variance génétique significative. Ainsi, théoriquement, les covariances génétiques entre traits peuvent contraindre les populations à évoluer dans certaines directions privilégiées dans l'espace des traits, aussi appelées "lignes de moindre résistance", contenant le maximum de la

variance génétique (Figure 4.9 ; Schluter 1996). L'influence de ces lignes de moindre résistance dans l'évolution à long-terme des populations a déjà été démontrée (Schluter, 1996) mais leur rôle sur la réponse évolutive à court-terme reste peu connu.

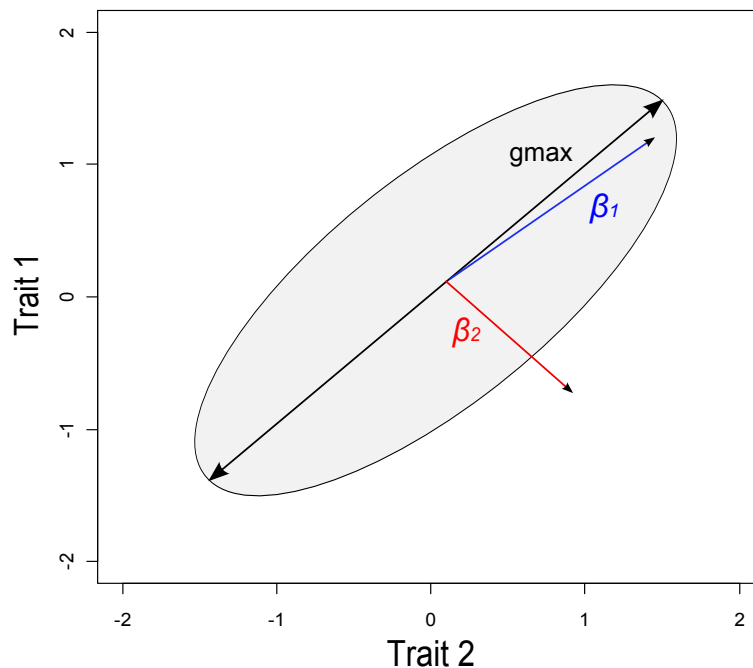


FIGURE 4.9 – Représentation géométrique du potentiel évolutif de deux traits dans une population et illustration du concept de "ligne de moindre résistance". *gmax* est l'axe contenant le plus de variance génétique additive et β_1 et β_2 deux vecteurs représentant des gradients de sélection directionnels. Une covariance génétique significative entre les deux traits se traduit par une représentation elliptique de \mathbf{G} . Ici, β_1 représenterait une pression de sélection qui s'alignerait quasiment sur la direction d'évolution privilégiée par les covariances génétiques et β_2 une pression de sélection qui à l'inverse irait à l'encontre de la covariance génétique entre traits.

Dans cette étude nous nous sommes intéressés aux capacités évolutives de trois populations de hêtre commun situées le long d'un gradient altitudinal. L'étude de 60 descendances maternelles issues de ces populations nous a auparavant permis d'évaluer que les traits potentiellement adaptatifs présentaient une héritabilité significative et que 6 traits parmi les 12 mesurés présentaient une différenciation vraisemblablement due à l'action de la sélection. À l'aide de ce même jeu de données nous avons cherché à savoir :

- Quel était le potentiel adaptatif des populations ?
- Existe-t-il des différences significatives de variances génétiques additives et des matrices \mathbf{G} différentes entre populations du fait de l'action de la sélection ?

- Et enfin est ce que les covariances génétiques entre traits peuvent favoriser ou ralentir la réponse évolutive des populations ?

Matériel et méthodes

MATÉRIEL D'ÉTUDE

Sur le versant nord du Mont-Ventoux, nous avons étudié le potentiel évolutif de 3 populations de hêtre commun situées le long d'un gradient altitudinal (Table 4.1; Gauzere et al. (c) *in prep*). Plus précisément, sur 20 arbres au sein de chacune de ces 3 populations ont été échantillonnés 60 familles maternelles issues de pollinisation libre en 2009. Ces familles maternelles ont été placées en test de descendance dans la pépinière d'état d'Aix les Milles selon un schéma en blocs complets randomisés (protocole décrit dans la section 3.3; Figure 4.2). Les 5475 semis ayant été plantés avec succès (soit 91,25 en moyenne par famille) ont été suivis sur deux années de croissance (2010 et 2011). Douze traits fonctionnels potentiellement adaptatifs ont été mesurés dans ce dispositif en environnement contrôlé. Ces traits sont décrits dans Gauzere et al. (b) *in prep* (section 3.3) et les effectifs phénotypés détaillés dans la Table 4.2 de l'article Gauzere et al. (c) *in prep* (section 4.1). Tous ces traits présentent une héritabilité significative à l'échelle du versant nord du Mont-Ventoux et 3 d'entre eux présentent des effets maternels significatifs ($BioM$, δD et $t_{b2 \rightarrow 3}$; section 3.3).

Un sous-ensemble de 2088 semis (en moyenne 34,8 par famille) a été génotypés pour 13 marqueurs microsatellites décrits dans Gauzere *et al.* (2013a). L'ensemble des adultes potentiellement reproducteurs au sein de chacune des 3 populations étudiées ont été génotypés pour ces mêmes marqueurs moléculaires afin de réaliser des assignations de paternité. Un père a pu être identifié pour 1000 descendants (section 3.3). Seule l'information des phénotypes et génotypes des individus contenus dans la modalité arrosée du dispositif en pépinière a été utilisée dans cette étude, sachant que l'expression des variances génétiques et phénotypiques peut être modifiée en environnement stressant (Charmantier & Garant, 2005).

MODÈLE DE GÉNÉTIQUE QUANTITATIVE UNIVARIÉ

Afin d'identifier si les traits potentiellement adaptatifs présentaient des variances génétiques différentes entre populations, nous avons utilisé le modèle mixte univarié suivant :

$$Y_{i,f,h,m,n,k} = \mu + P_i + H_f + L_{h(f)} + B_{m(f)} + E_n + a_k + \epsilon_{i,f,h,m,n,k} \quad (0)$$

avec P_i l'effet de la population d'origine, H_f l'effet de la planche, $L_{h(f)}$ l'effet de la bande hiérarchisée dans la planche, $B_{m(f)}$ l'effet du bloc hiérarchisé dans la planche et E_n l'effet de l'absence ou présence du fertilisateur dans les pots contenant les semis. Tous ces effets environnementaux et populationnels sont considérés comme des effets fixes. a_k et $\epsilon_{i,f,h,m,n,k}$ sont les composantes aléatoires du modèle, avec a_k l'effet génétique additif de l'individu k et $\epsilon_{i,f,h,m,n,k}$ les résidus du modèle. Les effets aléatoires sont supposés normalement distribués et les éléments a_k sont supposés être tirés dans $\{a_k\} \sim N(0, \mathbf{A} \times V_A)$, avec \mathbf{A} la matrice des apparentements et V_A la variance génétique additive, et les éléments $\epsilon_{i,f,h,m,n,k}$ tirés dans $\{\epsilon_{p,m,n,j,k}\} \sim N(0, \mathbf{I}_{de} \times V_R)$, avec \mathbf{I}_{de} la matrice identité et V_R la variance résiduelle du modèle.

En conservant ces effets fixes, nous avons complexifié la partie aléatoire du modèle (0), en considérant que (i) la variance des effets génétiques additifs, V_A , varie selon la population d'origine des individus, (ii) la variance des résidus, V_R , varie selon la population d'origine et (iii) V_A et V_R varient selon la population d'origine, en définissant les modèles suivants :

$$Y_{i,f,h,m,n,k} = \mu + P_i + H_f + L_{h(f)} + B_{m(f)} + E_n + a_{k(i)} + \epsilon_{i,f,h,m,n,k} \quad (i)$$

$$Y_{i,f,h,m,n,k} = \mu + P_i + H_f + L_{h(f)} + B_{m(f)} + E_n + a_k + \epsilon_{i,f,h,m,n,k(i)} \quad (ii)$$

$$Y_{i,f,h,m,n,k} = \mu + P_i + H_f + L_{h(f)} + B_{m(f)} + E_n + a_{k(i)} + \epsilon_{i,f,h,m,n,k(i)} \quad (iii)$$

Ne disposant pas de l'information nécessaire pour correctement dissocier V_A de la variance due aux effets maternels (V_M ; section 3.3), nous avons préféré négliger les effets maternels dans les modèles ci-dessus. Ces modèles ont été appliqués sur des variables transformées pour 3 traits : δH , LM et LA .

Ces modèles ont été analysés en utilisant la version R du logiciel AsReml (Gilmour *et al.*, 2006) et l'information sur les relations de parentés connues sous forme de pedigree. Les modèles hiérarchisés (0), (i), (ii) et (iii) ont été comparés avec un test de ratio de vraisemblance (Wilson *et al.*, 2010). Les tests ont été considérés significatifs lorsque leurs p_{valeur} étaient inférieures à 0.05 et marginalement significatifs si $0.05 < p_{valeur} < 0.1$.

MODÈLE DE GÉNÉTIQUE QUANTITATIVE MULTIVARIÉ

Nous avons estimé les matrices de variance-covariance génétiques additives des traits quantitatifs, aussi appelées matrices \mathbf{G} , pour chaque population indépendamment grâce au modèle suivant :

$$Y = \mu + L + B + E + a + \epsilon \quad (4.7)$$

avec Y le vecteur des observations phénotypiques standardisées, μ le vecteur des moyennes phénotypiques, L le vecteur des effets fixes bande, B le vecteur des effets fixes bloc, E le vecteur des effets fixes de la présence de fertilisateur. Pour les effets aléatoires, a est le vecteur des valeurs génétiques additives et ϵ le vecteur des effets résiduels. Les effets aléatoires sont supposés normalement distribués et les a sont supposés tirés dans $a \sim N(0, \mathbf{A} \otimes \mathbf{G})$, où \otimes est le produit de Kronecker, et ϵ tirés dans $\epsilon \sim N(0, \mathbf{I}_{de} \otimes \mathbf{E})$, où \mathbf{E} est la matrice de variance-covariance des résidus. Nous avons choisi d'appliquer ce modèle uniquement à un sous-ensemble des 12 traits mesurés au total pour faciliter l'interprétation des matrices \mathbf{G} . Les 5 traits analysés, qui sont δD , $t_{b2 \rightarrow 3(2011)}$, $t_{s1 \rightarrow 2(2011)}$, LMA et $\%N$ ont été choisis de manière à ce que chacune des grandes catégories fonctionnelles soit représentée et au vu des résultats précédemment obtenus sur la différenciation adaptative des traits (Gauzere et al. (c) *in prep*). La statistique du z-ratio (z-ratio = estimation / erreur standard) nous a permis de détecter les covariances génétiques entre traits significativement (z-ratio > 2) ou marginalement différentes de zéro (z-ratio > 1,7).

Pour étudier la dimensionnalité des matrices \mathbf{G} , nous avons suivi l'approche proposée par Kirkpatrick (2009), qui est une extension de la notion d'évolutivité proposée par Houle (1992) à plusieurs traits. Ainsi, nous avons calculé les covariances génétiques entre les traits i et j sans dimension comme :

$$\tilde{G}_{ij} = r_{ij} \sqrt{G_{ii} G_{jj}} / (\bar{z}_i \bar{z}_j) = cov(z_i, z_j) / (\bar{z}_i \bar{z}_j) \quad (4.8)$$

et les variances génétiques sans dimensions pour chaque trait comme :

$$\tilde{G}_{ii} = G_{ii} / (\bar{z}_i^2) = V_{Ai} / (\bar{z}_i^2) \quad (4.9)$$

avec r_{ij} la corrélation génétique entre les traits i et j , G_{ii} la variance génétique du trait i et \bar{z}_i la moyenne du trait i . À partir de la matrice de variance-covariance génétique normalisée,

$\tilde{\mathbf{G}}$, nous avons calculé trois indices statistiques décrivant le potentiel évolutif pour chaque population : n_D , e_{max} et ν_T (Kirkpatrick, 2009). Les valeurs propres de $\tilde{\mathbf{G}}$, λ_i , ont tout d'abord été utilisées pour calculer le nombre efficace de dimensions, n_D :

$$n_D = \sum_{i=1}^n \frac{\lambda_i}{\lambda_1}$$

où n est le nombre de traits analysés et λ_1 est la valeur propre la plus forte.

La racine carrée de la première valeur propre nous a permis de calculer le maximum de l'évolutivité (e_{max}), ce maximum étant obtenu lorsque la sélection et l'axe contenant la plus forte V_A (appelé \mathbf{g}_{max}) sont alignés : $e_{max} = \sqrt{\lambda_1}$ (Kirkpatrick, 2009). Enfin, la variance génétique totale (ν_T) a été calculée comme étant la somme des valeurs propres de $\tilde{\mathbf{G}}$: $\nu_T = \sum_{i=1}^n \lambda_i$ (Kirkpatrick, 2009). Graphiquement, on peut représenter la matrice $\tilde{\mathbf{G}}$ par une ellipsoïde plus ou moins aplatie et recouvrant l'espace des valeurs de traits selon le nombre efficace de dimensions et la variance génétique totale (Figure 4.10).

Les modèles multivariés ont été analysés avec AsReml-R (Gilmour *et al.*, 2006), toujours en utilisant l'information des apparentements paternels disponible sous forme de pedigree.

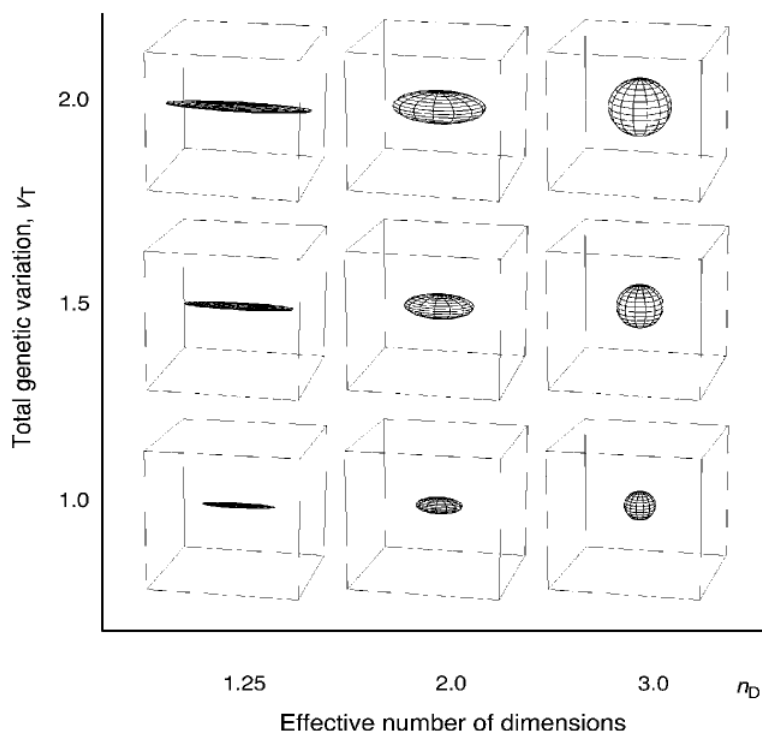


FIGURE 4.10 – Patrons de variation multivarié avec $n = 3$ traits par rapport au nombre de dimension efficace (n_D) et de la variance génétique totale (ν_T). Figure tirée de (Kirkpatrick, 2009).

L'analyse des valeurs propres et vecteurs propres des matrices $\tilde{\mathbf{G}}$ s'est faite en utilisant la fonction "eigen" du logiciel R.

Résultats

VARIATION DES V_A ENTRE POPULATIONS

L'analyse univariée des traits montre que 8 traits présentent des variances additives significativement différentes entre populations (Table 4.7). Notamment, les traits morphologiques et physiologiques foliaires présentent des variances génétiques réduites en population N1 et N4 (δD , $LM_{\sqrt{}}$, $LA_{\sqrt{}}$, LMA et $\%N$; Figure 4.11). Les patrons de variation de V_A pour les traits d'accroissement ne montrent pas de tendance systématique claire, avec pour le trait d'accroissement en diamètre des V_A réduites en N1 et N4 et pour le trait de production de biomasse une V_A réduite en N2 (Figure 4.11). Les traits phénologiques présentent des V_A réduites en population N4 (et N2 pour la durée de saison de végétation; Figure 4.11). Le fait de définir V_R par population améliore le modèle pour 8 traits (Table 4.7).

Trait	H0 $V_{AN1} = V_{AN2} = V_{AN4}$	H0 $V_{RN1} = V_{RN2} = V_{RN4}$	H0 $V_{AN1} = V_{AN2} = V_{AN4}$ $V_{RN1} = V_{RN2} = V_{RN4}$	Meilleur modèle
δH_{ln}	0,1591	0,0659	0,0320	(ii)
δD	< 0,0001	< 0,0001	< 0,0001	(iii)
$BioM$	< 0,0001	< 0,0001	< 0,0001	(iii)
$t_{b2 \rightarrow 3(2011)}$	1	1	0,11207	(0)
$t_{b2 \rightarrow 3(2012)}$	0,0492	0,1232	0,0255	(i)
$t_{s1 \rightarrow 2(2011)}$	0,5208	0,7908	0,1639	(0)
VD_{2011}	0,5398	0,2305	0,0130	(iii)
LM_{\checkmark}	0,0120	0,8021	0,0132	(iii)
LA_{\checkmark}	0,0144	0,4158	< 0,0001	(iii)
LMA	0,4343	0,5313	0,0360	(iii)
$\%N$	0,0265	0,0141	0,0132	(i)
$\%C$	0,2692	0,7672	0,0762	(0)
$\delta^{13}C$	0,0052	0,0006	0,0005	(ii)

TABLE 4.7 – Test des modèles hiérarchiques supposant différentes structurations des variances génétiques additives (V_A) et résiduelles (V_R) par le niveau du facteur population (N1, N2 et N4). Ici nous avons représenté les résultats du test de 3 hypothèses H0 et indiqué quel modèle expliquait le mieux la variance des traits. Pour rappel, le modèle (0) suppose des V_A et V_R indépendantes de la population, le modèle (i) suppose que V_A varie entre populations, le modèle (ii) suppose que V_R varie entre populations et le modèle (iii) suppose que V_A et V_R varient entre populations. ln et \checkmark indiquent les transformations qui ont été appliquées pour l'analyse des traits.

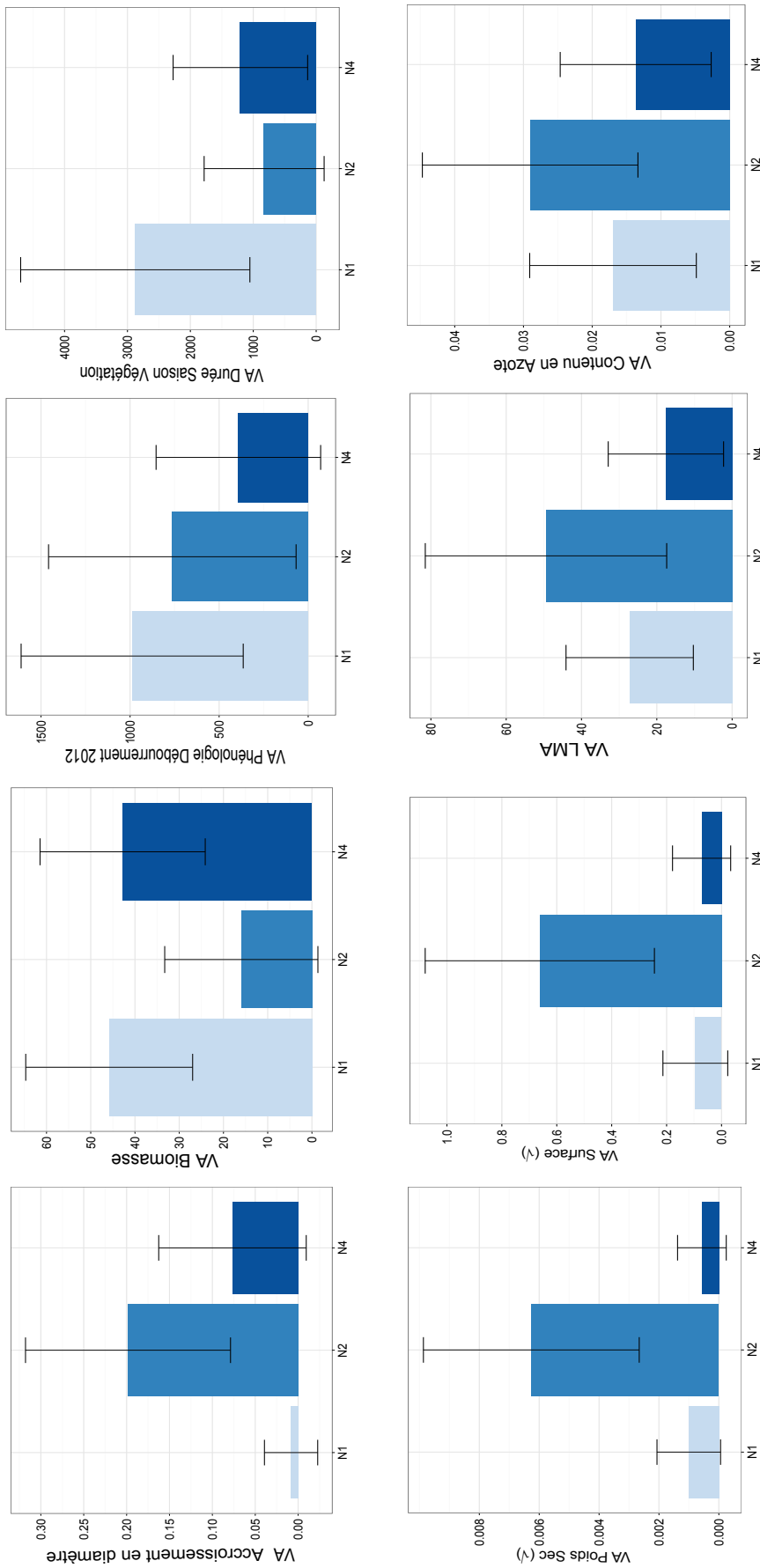


FIGURE 4.11 – Estimation des variances génétiques additives (V_A) par population pour les modèles univariés où l'hypothèse $H_0 V_{AN1} = V_{AN2} = V_{AN4}$ a pu être rejetée (voir résultats dans la Table 4.7). Les barres autour des estimations représentent les limites de l'intervalle de confiance à 95 %.

MATRICES G PAR POPULATIONS

Une tendance générale sur le patron de variation de l'évolutivité par classe fonctionnelle de trait se retrouve entre populations : l'évolutivité du trait d'accroissement en diamètre est bien plus forte ($CV_{A\delta D} = 2,95$ en moyenne) que celles des traits de débourrement végétatif et morphologique ($CV_{At_{b2\rightarrow3}} = 1,26$ et $CV_{ALMA} = 0,99$ en moyenne), elles-mêmes plus fortes que l'évolutivité du trait physiologique ($CV_{A\%N} = 0,64$ en moyenne ; Table 4.8). Le trait avec la plus faible évolutivité est le trait de sénescence foliaire ($CV_{At_{s1\rightarrow2}} = 0,13$ en moyenne). La faible évolutivité pour ce trait peut certainement s'expliquer en partie par la qualité de sa mesure phénotypique. En effet, du fait de la difficulté de caractériser des stades de sénescence foliaire, seuls 2 stades ont été utilisés pour décrire cette phénologie foliaire, ce qui a eu pour conséquence de très faiblement faire ressortir les variations individuelles de ce trait (trait avec un des plus faibles coefficients de variation phénotypique CV ; voir Table 3.10 Gauzere et al. (b) *in prep*).

	Population N1		Population N2		Population N4	
	I_A	CV_A	I_A	CV_A	I_A	CV_A
δD	3,811	1,952	18,943	4,352	6,524	2,554
$t_{b2\rightarrow3(2011)}$	3,009	1,735	1,442	1,201	0,731	0,855
$t_{s1\rightarrow2(2011)}$	0,051	0,226	0,005	0,071	0,009	0,092
LMA	1,027	1,014	1,449	1,204	0,572	0,756
$\%N$	0,421	0,649	0,416	0,645	0,396	0,629

TABLE 4.8 – Évolutivité ($\times 100$), I_A , et coefficient de variance génétique additive, CV_A ($CV_A = \sqrt{I_A} \times 100$), pour les traits étudiés dans l'analyse multivariée.

	Population N1			Population N2			Population N4		
	Estimation	IC ₉₅ inf	IC ₉₅ sup	Estimation	IC ₉₅ inf	IC ₉₅ sup	Estimation	IC ₉₅ inf	IC ₉₅ sup
δD	0,045	0,015	0,075	0,224	0,115	0,333	0,077	0,017	0,137
$t_{b2 \rightarrow 3(2011)}$	1890,000	1039,686	2750,214	908,000	220,509	1595,907	460,212	77,007	843,417
$t_{s1 \rightarrow 2(2011)}$	5150,000	2710,331	7580,731	509,000	-83,446	1102,116	859,594	90,835	1628,352
<i>LMA</i>	34,000	14,816	53,138	47,900	16,901	78,961	18,925	2,774	35,075
%N	0,019	0,002	0,037	0,019	0,001	0,037	0,018	0,001	0,035
$\delta D : t_{b2 \rightarrow 3(2011)}$	-2,870	-6,796	1,062	2,720	-3,415	8,862	1,206	-2,240	4,653
$\delta D : t_{s1 \rightarrow 2(2011)}$	-3,520	-10,340	3,295	0,142	-6,173	6,457	2,937	-2,060	7,934
$\delta D : LMA$	0,519	-0,065	1,104	0,133	-1,204	1,471	-0,414	-1,121	0,292
$\delta D : \%N$	-0,008	-0,024	0,009	-0,013	-0,047	0,021	0,015	-0,009	0,038
$t_{b2 \rightarrow 3(2011)} : t_{s1 \rightarrow 2(2011)}$	-46,600	-1088,529	995,398	-58,700	-557,086	439,659	52,143	-334,380	438,666
$t_{b2 \rightarrow 3(2011)} : LMA$	-161,000	-261,038	-60,826	-0,948	-103,589	101,694	-38,567	-98,121	20,988
$t_{b2 \rightarrow 3(2011)} : \%N$	2,790	-0,368	5,956	-0,403	-3,036	2,229	-0,124	-1,942	1,695
$t_{s1 \rightarrow 2(2011)} : LMA$	-58,200	-221,777	105,469	9,980	-90,104	110,067	-12,312	-91,694	67,070
$t_{s1 \rightarrow 2(2011)} : \%N$	4,840	-0,670	10,351	-0,540	-2,902	1,822	1,353	-1,359	4,064
%N : <i>LMA</i>	-0,446	-0,892	0,001	-0,843	-1,490	-0,195	-0,415	-0,842	0,011

TABLE 4.9 – Estimation des variances et covariances génétiques entre traits (matrices **G**) à partir du modèle multivarié 4.7 pour les trois populations de hêtre.

Peu de covariances génétiques significatives sur la base du z-ratio sont détectées au sein des populations (Table 4.10). L'essentiel des covariances entre traits sont retrouvées en population N1 et ne sont que marginalement significatives. Les intervalles de confiance assez larges autour de l'estimation des composantes de \mathbf{G} (Table 4.9) indiquent que nous manquons de puissance statistique pour détecter plus de covariances génétiques entre traits du fait du faible nombre de familles étudiées par population (20 familles/population). La seule covariance retrouvée dans toutes les populations indique qu'il existe une forte corrélation négative entre le LMA et $\%N$ ($r_{LMA:\%N} = -0.73$ en moyenne ; Table 4.10). Cette covariance pourrait être expliquée par un effet de dilution de l'azote plus important dans les feuilles larges et épaisses (et donc aux forts LMA). En population N1, on retrouve une forte corrélation négative entre

Population N1					
	$t_{b2 \rightarrow 3}$	LMA	δD	$\%N$	$t_{s1 \rightarrow 2}$
$t_{b2 \rightarrow 3}$	1				
LMA	-0,634	1			
δD	-0,310	0,419	1		
$\%N$	0,463	-0,551	-0,259	1	
$t_{s1 \rightarrow 2}$	-0,126	-0,139	-0,231	0,486	1

Population N2					
	$t_{b2 \rightarrow 3}$	LMA	δD	$\%N$	$t_{s1 \rightarrow 2}$
$t_{b2 \rightarrow 3}$	1				
LMA	-0,005	1			
δD	0,191	0,041	1		
$\%N$	-0,097	-0,882	-0,198	1	
$t_{s1 \rightarrow 2}$	-0,410	0,064	0,013	-0,173	1

Population N4					
N4	$t_{b2 \rightarrow 3}$	LMA	δD	$\%N$	$t_{s1 \rightarrow 2}$
$t_{b2 \rightarrow 3}$	1				
LMA	-0,413	1			
δD	0,202	-0,343	1		
$\%N$	-0,043	-0,709	0,390	1	
$t_{s1 \rightarrow 2}$	0,449	-0,097	0,360	0,343	1

TABLE 4.10 – Table des corrélations génétiques entre traits par population. Les cellules rouges et jaunes indiquent les corrélations génétiques significativement et marginalement différentes de zéro respectivement d'après le z-ratio.

LMA et la phénologie du débourrement ($r_{LMA:tb2\rightarrow3} = -0,63$ en N1 ; Table 4.10), ainsi que de nombreuses corrélations marginales, dont une corrélation négative entre δD et $t_{b2\rightarrow3}$ et positive entre δD et LMA . Ces corrélations traduisent le fait que les individus au débourrement végétatif précoce ont des accroissements en diamètre plus forts et de plus forts LMA , donc investissent plus dans le développement végétatif. La corrélation positive entre le contenu en azote des feuilles et la sénescence pourrait s'expliquer par une longue re-mobilisation de l'azote des feuilles quand celui-ci est en forte concentration avant la chute des feuille.

DIMENSIONNALITÉ DES MATRICES $\tilde{\mathbf{G}}$

Pour toutes les matrices $\tilde{\mathbf{G}}$ nous estimons un faible et relativement similaire nombre efficace de dimensions ($n_D < 2$; Table 4.12). Ce résultat se retrouve par le fait que la composante principale de $\tilde{\mathbf{G}}$ (g_{max}) contient toujours plus de la moitié de la variance génétique totale pour chacune des populations (Table 4.11).

L'analyse du premier vecteur propre des matrices $\tilde{\mathbf{G}}$ indique que pour toutes les populations, la ligne de moindre résistance est très largement associée au trait de croissance en diamètre ($r_{\delta D:g_{max}} = 0,67$ en N1, $r_{\delta D:g_{max}} = 0,99$ en N2 et $r_{\delta D:g_{max}} = 0,93$ en N4 ; Table 4.11). Ce résultat est cohérent avec le fait que δD présente la plus forte évolutivité dans toutes les populations (Table 4.8). Pour la population N1, la composante principale de $\tilde{\mathbf{G}}$ est éga-

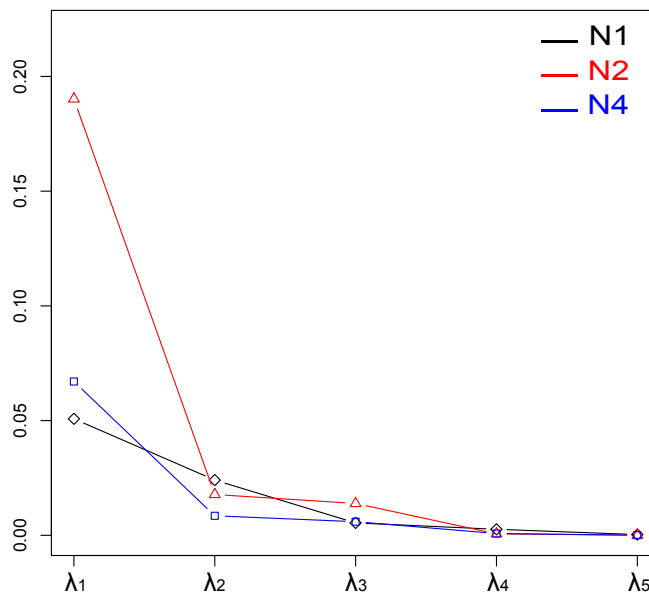


FIGURE 4.12 – Patron de décroissance des cinq premières valeurs propres (λ_i) de $\tilde{\mathbf{G}}$ pour chaque population.

Population	Pourcentage de variance expliqué par $gmax$	δD	$t_{b2 \rightarrow 3(2011)}$	$t_{s1 \rightarrow 2}$	LMA	$\%N$
N1	0,61	0,67	-0,05	-0,18	0,62	-0,45
N2	0,85	0,99	0,19	0,013	0,04	-0,20
N4	0,81	0,93	-0,21	-0,36	0,36	-0,40

TABLE 4.11 – Caractéristiques de la première valeur propre et premier vecteur propre des matrices $\tilde{\mathbf{G}}$ estimées par populations. $gmax$ est l'axe contenant le plus de variance génétique additive. Les cinq dernières colonnes de la table représentent les corrélations de chacun des traits avec la composante principale de $\tilde{\mathbf{G}}$ (i.e. le poids des traits dans $gmax$).

Population	n_D	e_{max}	ν_T
N1	1,64	0,23	0,08
N2	1,17	0,44	0,22
N4	1,23	0,26	0,08

TABLE 4.12 – Statistiques résumant le patron de variance génétique pour les 5 traits étudiées dans chacune des populations. n_D est le nombre efficace de dimensions de la matrice $\tilde{\mathbf{G}}$, e_{max} le maximum de l'évolutivité et ν_T la variance génétique totale.

lement associée au LMA et $\%N$ ($r_{LMA:gmax} = 0,62$ et $r_{\%N:gmax} = -0,45$; Table 4.11). $\tilde{\mathbf{G}}_{N2}$ est la matrice avec la plus faible dimensionnalité ($n_D = 1,17$), mais c'est aussi la population où on trouve les plus fortes V_A ($e_{max} = 0,44$ et $\nu_T = 0,22$; Table 4.12). À l'inverse, $\tilde{\mathbf{G}}_{N1}$, qui possède le plus fort n_D ($n_D = 1,64$), présente le moins de V_A ($e_{max} = 0,23$ et $\nu_T = 0,08$). $\tilde{\mathbf{G}}_{N4}$ présente quand à elle à la fois un faible nombre efficace de dimensions et peu de variance génétique comparativement aux autres populations.

Discussion

LA SÉLECTION NATURELLE ÉRODE LA VARIANCE GÉNÉTIQUE AU SEIN DES POPULATIONS

La mesure d'héritabilité a souvent été utilisée pour comparer la variation génétique de différents traits adaptatifs et interpréter ses variations en regard de la variabilité des régimes de sélection sur les traits (Houle, 1992; Stirling *et al.*, 2002). En théorie les allèles défavorables devraient être rapidement éliminés par la sélection (Merila & Sheldon, 1999). Globalement, il se dégage de ces analyses que les traits d'histoire de vie présentent des héritabilités plus faibles

que les traits moins reliés à la valeur sélective, comme les traits morphologiques (Stirling *et al.*, 2002). Cependant, lorsque les mesures d'évolutivités standardisées par la moyenne des traits sont comparées, i.e. I_A ou CV_A , le patron de variation inverse est observé : les traits les plus reliés à la fitness sont également les plus variables (Houle, 1992). Cette observation est souvent expliquée par le fait que les traits reliés à la fitness auraient un déterminisme génétique plus complexe que les autres traits (Merila & Sheldon, 1999). Ainsi, les faibles héritabilités mesurées pour ces traits ne traduiraient pas une plus faible V_A , mais une plus forte influence de l'environnement sur leur variation phénotypique (V_E ; Price & Schluter 1991). Dans notre étude l'évolutivité plus élevée est trouvée pour le trait d'accroissement en diamètre, un trait de performance (Violle *et al.*, 2007), qui s'avère être fortement relié aux fertilités chez les arbres (Oddou-Muratorio *et al.*, 2005). Cependant, la forte évolutivité de ce trait pourrait en partie s'expliquer par l'influence significative d'effets maternels qui ont pour effet de sur-estimer les valeurs de variance génétique additive (section 3.3). Ainsi, nos résultats ne permettent pas de vérifier les attendus empiriques de variation de I_A par catégorie fonctionnelle de traits.

Ces études soulèvent également la question de quels sont les facteurs qui érodent significativement la variance génétique au sein des populations. Dans les populations de faible taille efficace, la dérive génétique peut fortement réduire les niveaux de diversité génétique. Dans notre système d'étude nous avons vu que la dérive génétique n'était pas à l'origine de la différenciation génétique des traits entre populations. Le faible niveau de divergence génétique neutre entre population ($F_{st} = 0.026$) suggère que la dérive génétique réduit faiblement la diversité génétique au sein des populations, et surtout nos précédents résultats ne suggèrent pas que la population N2, présentant la plus forte évolutivité totale, ait une taille efficace supérieure à celle de N1 et N4 ($Nep_{N1} = 12,3$, $Nep_{N2} = 15,1$ et $Nep_{N4} = 29,2$; Chapitre 2; et sur l'apparentement intra-population : $\theta_{N1} = 0,065$, $\theta_{N2} = 0,032$ et $\theta_{N4} = 0,038$; section 4.1). De plus, parmi les 8 traits présentant des V_A différentes entre populations, 4 ont montré un signal de sélection divergente (δD , $BioM$, $t_{b2 \rightarrow 3(2012)}$ et VD_{2011}) et 3 autres présentent des réponses plastiques significativement ou marginalement différentes entre populations qui pourraient être dues à une action de la sélection sur la plasticité des traits ($\%N$, LM et LMA ; section 4.1). Il est donc vraisemblable que le facteur principal réduisant la V_A au sein des populations soit l'action de la sélection naturelle. Dans ce cas, les patrons de variation de V_A indiquent qu'à haute altitude les pressions de sélection s'exercent sur l'ensemble des traits phénologiques, physiologiques et morphologiques (sauf pour la production de biomasse) et à

basse altitude les pressions de sélection s'exercent plus sur les traits physiologiques et morphologiques (sauf également la production de biomasse). À part pour les traits phénologiques et la production de biomasse, la population N2 semblerait subir moins de pression de sélection. Les variances génétiques globalement plus élevées en N2 pourraient également être dues à des pressions de sélection plus hétérogènes au sein de cette population. En effet, bien qu'une analyse précise de l'hétérogénéité micro-environnementale n'ait pas été menée sur le gradient, la population N2 est celle qui semble présenter la plus forte hétérogénéité environnementale qui se manifeste par la structure du peuplement et la densité plus faible dans cette population (Chapitre 2).

Pour la majorité des traits, les modèles qui spécifient que V_R est structurée par population sont meilleurs que les modèles avec V_R global. Ce résultat indique qu'il est probable que V_R des traits inclus d'autres composantes génétiques de variances, comme des variances épistatiques ou de dominance, dépendant de V_A et pouvant varier entre populations. Ainsi, pour étudier l'évolutivité des traits il est ici d'autant plus pertinent de standardiser V_A par la moyenne phénotypique plutôt que par la variance phénotypique totale des traits.

LES COVARIANCES GÉNÉTIQUES LIMITENT LES COMBINAISONS DE TRAITS POUVANT ÉVOLUER SOUS EFFET DE LA SÉLECTION

Le nombre efficace de dimensions des matrices de variance-covariance génétiques pour chaque population indique que les possibilités d'action de la sélection sont réduites et que les matrices \mathbf{G} se résument par moins de deux directions privilégiées d'évolution des traits. Cette interprétation est donc assez pessimiste sur les capacités d'évolution des populations de hêtre sur le Mont-Ventoux à court terme. Ceci peut paraître étrange au vu du faible nombre de covariances génétiques entre traits détectées comme significativement ou marginalement différentes de zéro (Table 4.10). D'une part, il se pourrait que le fait que la statistique du n_D ne prenne pas en compte les incertitudes autour des estimations de $\tilde{\mathbf{G}}$ conduisent à sous-estimer le nombre de dimension efficace. D'autre part, la valeur de n_D ne dépend pas seulement des covariances génétiques entre traits mais aussi de la proportion d'évolutivité totale expliquée par chacun des traits. Ainsi, si un trait présente une très forte évolutivité comparés aux autres (ce qui est le cas dans notre étude avec δD) un très faible nombre efficace de dimensions permet de caractériser \mathbf{G} (e.g., la Figure 4.10 représente des matrices à très faible n_D en présence

d'une covariance nulle entre les 3 traits).

Dans le papier présentant pour la première fois la mesure sans dimension du nombre efficace de dimensions, n_D , Kirkpatrick (2009) avait utilisé ce paramètre pour étudier 5 jeux de données sur des espèces différentes; il avait alors remarqué que le nombre de dimension efficace des matrices était faible ($n_D < 2$) et relativement constant entre études. Cette statistique ayant été encore peu appliquée, il est difficile d'en savoir plus sur le patron général de variation de n_D , en revanche nous ne pouvons pas exclure que cette statistique ne soit pas la plus pertinente pour différencier le potentiel adaptatif de différentes populations et espèces. Les deux autres statistiques descriptives de $\tilde{\mathbf{G}}$, e_{max} et ν_T , présentent une plus forte gamme de variation entre populations. Les mesures de ν_T estimées ici pour chaque population varient autant que les estimations faites par Kirkpatrick (2009) entre populations de différentes espèces. Au travers de cette statistique, qui ne dépend pas des corrélations génétiques entre traits, nous retrouvons que la population N2 est celle où l'évolutivité totale est la plus forte, probablement car cette population subit moins de pressions de sélection ou des régimes de sélection plus hétérogènes.

L'analyse des valeurs et vecteurs propres de $\tilde{\mathbf{G}}$ révèle la forte influence du trait d'accroissement en diamètre dans la direction principale d'évolution. L'évolution des traits covariants avec δD pourrait donc être fortement conditionnée par cette ligne de moindre résistance dans l'espace des traits. De telles covariances n'ont été détectées qu'en population N1 avec les traits LMA et $t_{b2 \rightarrow 3}$. Il est très probable qu'en milieu chaud et sec (i.e. dans les conditions rencontrées en N1) le fait de débourrer précocement (avoir de plus faibles $t_{b2 \rightarrow 3}$) et d'avoir une forte masse surfacique foliaire augmente la valeur adaptative des individus. Ainsi, la covariance de ces traits avec δD devrait faciliter l'évolution de ces traits à basse altitude.

L'inconvénient de l'approche qui ne consiste qu'à analyser les composantes principales de $\tilde{\mathbf{G}}$ est qu'elle ne permet que d'évaluer les contraintes **potentielles** à l'évolution, et non pas les contraintes **efficaces**. En effet, seule la connaissance des gradients de sélection (β) au sein de chacune des populations permettrait de pouvoir déterminer si les covariance génétiques vont effectivement modifier l'action de la sélection. Une approche intéressante pour évaluer le rôle des covariances génétiques dans l'évolution en connaissant β est de mesurer la réponse à la sélection de manière classique ($\Delta \bar{z} = \mathbf{G} \times \beta$), puis en contraignant les covariances de \mathbf{G} à

être nulles ($\Delta \bar{z}_i = \mathbf{G}_{ii} \times \beta_i$) (Agrawal & Stinchcombe, 2009; Conner, 2012; Teplitsky *et al.*, 2014). Une perspective importante à ce travail consiste donc à aller mesurer les gradients de sélection β dans les différentes populations.

ANALYSE CRITIQUE DES MÉTHODES POUR ÉVALUER L'ÉVOLUTIVITÉ ET LES CONTRAINTES ÉVOLUTIVES

Même s'il est clair que l'héritabilité n'est pas une mesure idéale du potentiel évolutif, la mesure d'évolutivité I_A (ou CV_A), qui pour rappel implique une standardisation par la moyenne du trait, n'est pas non plus une mesure parfaite du potentiel évolutif et ne doit pas être utilisée sans un regard critique. Notamment, et de manière très problématique dans notre cas, cette mesure n'a pas beaucoup de sens lorsqu'elle est appliquée aux traits qui n'ont pas de zéro naturel, ce qui est le cas des traits correspondant à des dates de l'année. Dans notre cas, bien qu'il s'agisse de traits très impliqués dans l'adaptation locale des populations, les évolutivités des mesures de phénologies ne peuvent donc pas être comparées avec celles des autres traits. Les analyses sur la dimensionnalité des matrices $\tilde{\mathbf{G}}$ de cette partie devraient donc être reprises sans inclure les traits phénologiques. On voit ici une des contraintes majeure de la mesure d'évolutivité standardisée par la moyenne des traits qui ne permet pas d'être appliquée de manière pertinente à tous les traits adaptatifs.

Malgré des preuves empiriques que l'évolution puisse être contrainte par la variance génétique disponible et les covariances génétiques entre traits (Etterson & Shaw, 2001), les approches de génétique quantitative visant à étudier les limites de la sélection en populations naturelles sont de plus en plus remises en cause (Conner, 2012). En effet, en reprenant les données de la littérature, Agrawal & Stinchcombe (2009) ont récemment montré que globalement les covariances génétiques avaient peu ou pas d'effet sur les taux d'évolutions des traits. De manière plus problématique encore, d'autres études montrent qu'une sélection artificielle perpendiculaire aux directions de moindre résistance entre traits peuvent générer une réponse rapide (Frankino *et al.*, 2005). Ainsi, les analyses multivariées ne permettraient toujours pas de prédire de manière fiable le changement des traits et les facteurs qui contraindront l'évolution.

Une critique aux approches de génétique quantitative est qu'il n'existe pas de moyen de

prouver que la variance génétique de certains traits soit exactement nulle puisqu'il restera toujours de la variance autour des estimations (Conner, 2012; Kirkpatrick, 2009). Or, au travers du nombre efficace de dimensions de \mathbf{G} c'est à la question cruciale de savoir si certaines dimensions de l'analyse en composante principale ne contiennent aucune variance génétique et donc s'il existe certaines combinaisons de traits qui sont inaccessibles à l'évolution à laquelle on aimerait répondre. Théoriquement, une manière de tester cela serait de déterminer si certaines valeurs propres de \mathbf{G} sont nulles (Kirkpatrick, 2009). Cependant, les développements statistiques permettant de réaliser ce test se heurtent à de nombreuses difficultés et biais expérimentaux et il n'existe aucun moyen de certifier qu'une valeur propre est nulle. De plus, toutes les inférences qui sont faites sur les capacités évolutives des populations se basent sur l'hypothèse que \mathbf{G} n'évolue pas dans le temps, ou à des échelles de temps plus élevées que le temps nécessaire à la différenciation moyenne des traits. Cependant, \mathbf{G} peut théoriquement évoluer par l'action de tous les processus affectant les fréquences alléliques : la sélection, la migration, la mutation et la dérive (McGuigan, 2006). À l'heure actuelle, les incertitudes sur les capacités évolutives de \mathbf{G} sont un frein supplémentaire à notre compréhension des causes et conséquences des contraintes sur l'évolution des traits.

CONCLUSION

L'estimation de la variance génétique au sein des populations est intéressante à la fois pour comprendre l'histoire évolutive des populations et pour déterminer leur capacités à répondre à des pressions de sélection futures. Ici, la forte érosion de V_A sur des traits présentant des valeurs moyennes ou effets plastiques divergeant entre populations révèle que l'action de la sélection a vraisemblablement diminué les capacités adaptatives des populations aux deux extrémités du gradient altitudinal étudié (N1 et N4). L'analyse des composantes principales des matrices \mathbf{G} standardisées par la moyenne des traits ($\tilde{\mathbf{G}}$) indique que les covariances entre traits peuvent potentiellement limiter les combinaisons de traits pouvant évoluer sous effet de la sélection. Notamment, la forte évolutivité du trait d'accroissement en diamètre pourrait fortement impacter les directions d'évolution des traits qui lui sont corrélées dans la population à plus basse altitude sur le gradient.

EFFETS DES FLUX DE POLLEN SUR LES CAPACITÉS ADAPTATIVES DES POPULATIONS D'ARBRES



Synthèse

Dans ce chapitre, je me suis intéressée aux flux de gènes en tant que processus permettant de ré-alimenter les populations en diversité génétique. Malgré de nombreuses études théoriques ayant balisé les attendus sur les effets de la migration sur les capacités d'adaptation des populations, très peu d'études empiriques ont démontré que l'arrivée de migrants dans une population apportait des phénotypes "atypiques" ou "innovants" que la sélection pourrait ensuite favoriser.

Nous avons précédemment vu que les populations de hêtre sur le versant nord du Mont-Ventoux recevaient de fortes quantités de pollen immigrant (malgré des distances de dispersion relativement faibles ; Chapitre 2) et que de nombreux traits fonctionnels présentaient une différenciation adaptative entre ces populations (Chapitre 4.1). Ces résultats permettent donc de démontrer que les effets défavorables des flux de gènes sur l'adaptation génétique n'étaient pas prédominants. Ici, les données génétiques disponibles pour les descendants récoltés et les pères potentiels au sein des populations ont été utilisées afin d'assigner les descendants à deux catégories selon l'origine de leur père : une catégorie "Non-migrant" regroupant les individus issus de reproduction avec un père local et une catégorie "Migrant" regroupant les individus issus d'un événement de reproduction avec du pollen extérieur à la population. La comparaison des valeurs génétiques de ces deux catégories de descendants m'a ainsi permis de tester par une **approche directe** les effets de la migration pollinique sur les capacités d'adaptation des populations de hêtre.

Les résultats préliminaires exposés dans ce chapitre indiquent que la migration pollinique peut augmenter les variances génétiques des traits différenciés le long du gradient d'altitude. Cependant, les effets flux de pollen restent complexes puisque n'affectant pas tous les traits différenciés et toutes les populations. Une meilleure compréhension des effets bénéfiques des flux de pollen semble nécessiter de pouvoir mieux caractériser l'origine du pollen migrant, et notamment les événements de dispersion à longue distance.

Introduction

Le maintien de diversité génétique au sein des populations naturelles expérimentant des pressions de sélection reste une question centrale en biologie évolutive (Le Corre & Kremer, 2012). Une interprétation classique du "théorème fondamental de la sélection naturelle" de Fisher est que la variance génétique (V_A) des traits fortement reliés à la valeur sélective des individus devrait être faible ou nulle (Merila & Sheldon, 1999; Stirling *et al.*, 2002). Paradoxalement, on mesure des variances génétiques significatives pour la grande majorité des traits dans les populations naturelles, y compris pour les traits d'histoire de vie (Houle, 1992). Des approches théoriques ont montré que la mutation, la variation temporelle des pressions de sélection, les flux de gènes entre populations localement adaptées ou encore les contraintes génétiques entre traits peuvent expliquer le maintien de variance génétique au sein des populations et donc leur capacité à répondre à de nouvelles pressions de sélection (Barton, 2001; Johnson & Barton, 2005). Cependant, on ne connaît toujours pas le poids de ces différents processus sur l'augmentation de V_A dans les populations naturelles.

Ici, nous nous sommes plus particulièrement intéressés à l'effet des flux de gènes sur le maintien de la variance génétique intra-population et les capacités adaptatives des populations. Le long de clines environnementaux, on observe que les pressions de sélection spatialement hétérogènes peuvent générer des patrons de différenciation génétique et conduire à l'adaptation locale des populations (Savolainen *et al.*, 2007; Leimu & Fischer, 2008). Dans les cas où l'intensité des flux de gènes ne contraint pas l'adaptation locale (Lenormand, 2002), les flux de gènes sont supposés fortement contribuer à l'augmentation des variances génétiques (Bridle & Vines, 2007). Le modèle théorique de Barton (2001) a permis de montrer que cette augmentation des variances génétiques est attendue comme d'autant plus forte que les populations connectées par les flux de gènes sont différenciées génétiquement. Le contexte de changement climatique actuel, soulève la question des capacités d'évolution à de nouvelles conditions environnementales des espèces localement adaptées (Harte *et al.*, 2004). Les modèles climatiques prédisent un décalage des enveloppes thermiques dans l'espace (Loarie *et al.*, 2009). Dans ce cas particulier d'environnement variable à la fois dans l'espace et dans le temps, on peut penser que les flux de gènes entre populations localement adaptées pourraient favoriser la diffusion d'allèles favorables pré-adaptés aux conditions environnementales futures (Aitken *et al.*, 2008; Davis & Shaw, 2001). Les événements de dispersion à longue-distance

pourraient tout particulièrement favoriser ces diffusions d'allèles (Kremer *et al.*, 2012).

Les arbres constituent de bon modèles pour étudier l'interaction entre flux de gènes à longue distance et adaptation. En effet, les populations d'arbres présentent généralement de forts taux d'immigration (immigration par pollen égale à 44 % en moyenne chez les espèces anémophiles; Ashley 2010) favorisés par des distances de dispersion par graine et pollen élevées (Ashley, 2010). Chez les arbres, la dispersion à longue distance se fait généralement par pollen : chez les espèces anémophiles les distances de migration peuvent facilement atteindre la centaine de mètres et de rares événements de dispersion sur une dizaine de kilomètres chez certaines espèces (Petit & Hampe, 2006; Ashley, 2010). En outre, l'importance des événements de dispersion sur de très longues distances pourrait avoir été largement sous-évaluée du fait de la difficulté d'observer ou mesurer ces événements (Savolainen *et al.*, 2007). Les traits d'histoire de vie des arbres laissent donc penser que les effets favorables de la migration jouent un rôle important dans l'adaptation de ces espèces (Savolainen *et al.*, 2007; Kremer *et al.*, 2012). Les flux de gènes permettraient notamment d'expliquer les forts taux de diversité génétique mesurés au sein des populations d'arbre (Hamrick & Nason, 1986). Chez ces espèces à cycle de vie long pour lesquelles la manipulation expérimentale est difficile (e.g. croisements contrôlés), peu d'études empiriques ont testé les attendus théoriques en terme d'apport de V_A au sein de populations localement adaptées. À ma connaissance, seules deux études ont permis de démontrer que la migration pouvait affecter les valeurs génétiques des traits chez les arbres. Dans une première étude sur *Pinus sylvestris*, Nilsson (1995) a montré que, sur des traits différenciés le long d'un gradient latitudinal, les descendants récoltés après un événement de reproduction présentaient des moyennes génétiques décalées par rapport à la moyenne de leurs mères. Ces différences de moyennes génétiques permettaient de déduire que la migration pollinique provenait majoritairement de populations de plus hautes latitudes. Dans une étude plus récente sur 142 populations de *Pinus contorta*, Yeaman & Jarvis (2006) ont démontré que l'hétérogénéité régionale des conditions climatiques expliquait les niveaux de variance génétique mesurés au sein des populations. Cependant, ces deux études caractérisaient l'effet de la migration dans les populations naturelles sans réaliser de mesures directe de la migration et donc en utilisant des approches que l'on peut qualifier d'*indirectes*. Le problème de ces approches est qu'elles peuvent possiblement être affectées par des effets confondants mal pris en compte. Ces études sont donc encore très insuffisantes pour conclure l'importance des effets favorables de la migration en terme d'augmentation de V_A dans les

populations naturelles d'arbres, et notamment pour comprendre à partir de quelle échelle et quels niveaux de différenciation génétique ces effets peuvent être non-négligeables.

Les populations de hêtre commun (*Fagus sylvatica*) sur le versant nord du Mont-Ventoux sont des sites expérimentaux pertinents pour mesurer les effets de la migration sur les capacités d'adaptation des populations. En effet, les 3 populations étudiées sur ce gradient présentent une différenciation génétique pour de nombreux traits phénologiques, physiologiques et d'accroissement tout en étant très proches géographiquement (~ 1 km sépare les deux populations les plus extrêmes sur le gradient) et en recevant de forte quantité de pollen immigrant ($m = 56$ % en moyenne ; Gauzere *et al.* 2013a ; Chapitre 2). Les fenêtres phénologiques impactant la quantité de pollen immigrant reçue à l'échelle individuelle et les populations présentant des taux moyen d'immigration en augmentation avec l'altitude, nous avons précédemment conclu que le décalage des fenêtres phénologiques à l'échelle des populations pouvait générer des flux de pollen directionnels, du bas vers le haut du gradient, sur le versant nord du Mont-Ventoux (Gauzere *et al.*, 2013a). Si cette hypothèse est correcte, les flux de gènes favoriseraient tout particulièrement la diffusion vers le haut du gradient de gènes pré-adaptés aux conditions environnementales sèches et chaudes rencontrées en bas du gradient et permettraient aux

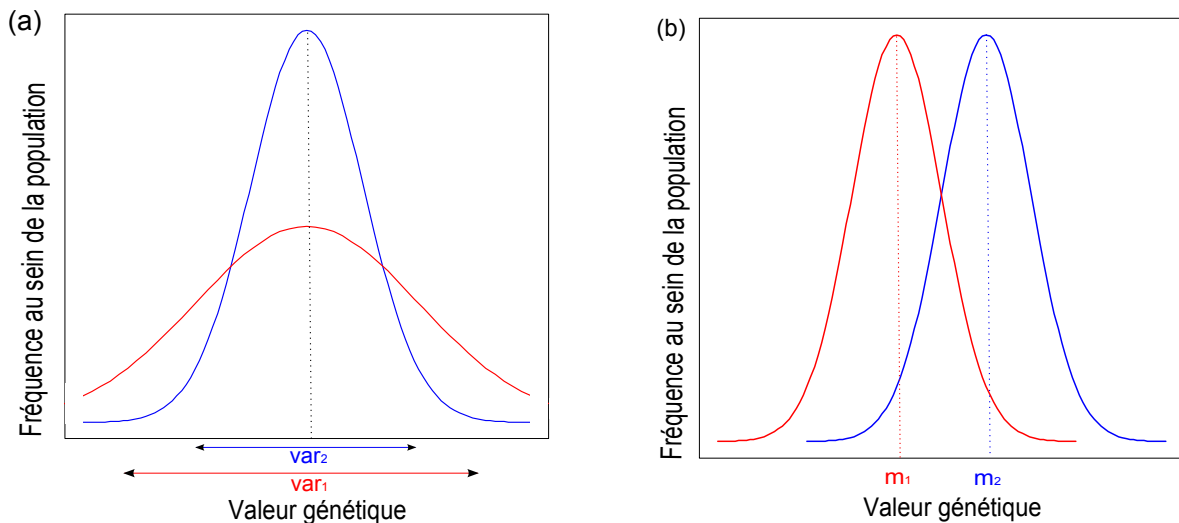


FIGURE 5.1 – Effets hypothétiques de la réception de pollen immigrant. m_1 , m_2 , var_1 et var_2 représentent les moyennes et variances génétiques de deux groupes d'individus provenant de deux groupes d'individus provenant de croisements avec du pollen local (bleu) ou migrant (rouge). Le cas (a) correspond à l'attendu théorique d'augmentation de variance génétique par flux de gènes entre populations localement adaptées. Le cas (b) correspond à l'attendu empirique de décalage des valeurs génétiques moyennes par flux de gènes directionnel entre populations localement adaptées.

populations de hautes altitudes de répondre plus rapidement aux changements climatiques prédits pour les prochaines décennies (Keller *et al.*, 2013).

L'objectif de cette partie de ma thèse a été de tester ces attendus sur l'impact des flux de pollen à la fois sur l'apport de variance génétique et sur le décalage des valeurs génétiques : (a) dans le cas de flux de gènes non-directionnels on s'attend à ce que les individus issus de croisements avec du pollen migrant soient plus variables génétiquement que les individus provenant de croisements avec du pollen local (Figure 5.1); (b) dans le cas de flux de gènes directionnels on s'attend à ce que la moyenne génétique des individus migrants soit décalée par rapport à celle des individus locaux (Figure 5.1). Pour cela, nous avons utilisé 60 familles maternelles récoltées en 2009 sur les 3 populations de hêtre et distingué différentes catégories selon si les descendants pouvaient être assignés ou non-assignés à un père local. De manière originale, nous avons ensuite testé si ces deux catégories de descendants présentaient différentes valeurs génétiques et/ou variances génétiques pour plusieurs traits fonctionnels.

Matériel et méthodes

MATÉRIEL D'ÉTUDE

Le matériel d'étude qui a été utilisé pour cette étude est le même que celui décrit dans les précédents chapitres de cette thèse : 5475 semis issus de pollinisation libre, récoltés sur 60 arbre-mères au sein de trois populations de hêtre commun (20 familles/population) situées à différents niveaux altitudinaux, et placés en test de descendance (voir la Table 4.1 pour la description des populations; Figure 4.2 pour la description du test de descendance; Table 4.2 pour le détail du phénotypage). Un sous-ensemble de 2088 semis (en moyenne 34,8 par famille) a été génotypé pour 13 marqueurs microsatellites décrits dans Gauzere *et al.* (2013a). L'ensemble des adultes potentiellement reproducteurs au sein de chacune des 3 populations étudiées ont été génotypés pour ces mêmes marqueurs moléculaires.

Pour $n= 10, 20$ et 20 arbres adultes en population N1, N2 et N4 respectivement nous disposons de données sur le débourrement végétatif en 2012 et 2013. Même si peu d'individus ont été mesurés sur ces suivis de phénologie, ces derniers ont été spécifiquement choisis sur des critères de précocité ou tardiveté du débourrement afin de couvrir complètement de la

gamme de variation du débourrement au sein des populations. Chez le hêtre commun, les phénologies du débourrement végétatif et de floraison sont très liées (Milleron *et al.*, 2012), celles-ci se produisant presque simultanément (Figure 5.2). L'étalement des dates de débourrement végétatif nous a donc permis de caractériser indirectement les fenêtres de floraison les différentes populations le long du gradient.

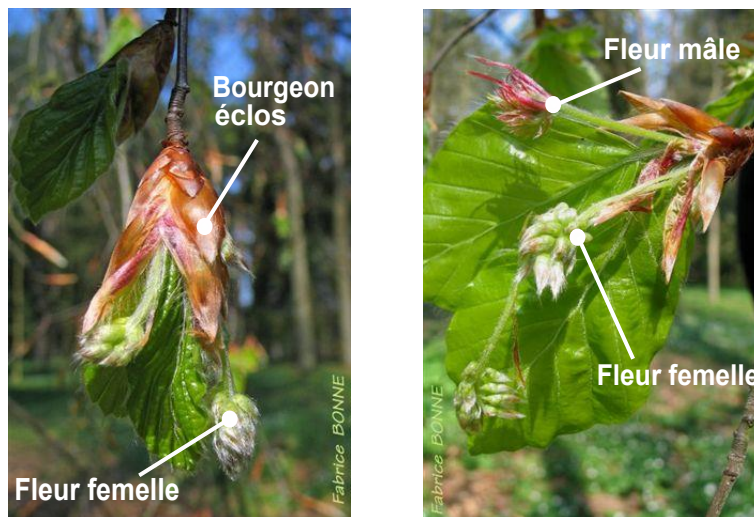


FIGURE 5.2 – Photographies de débourrement végétatif chez des arbres adultes de *F. sylvatica*. Photographies prises par Fabrice Bonne.

ASSIGNATION DE PATERNITÉ ET CARACTÉRISATION DES MIGRANTS

À partir du génotypage sur 13 marqueurs microsatellites de 2088 descendants et de l'ensemble des arbres adultes potentiellement reproducteurs au sein des populations, nous avons réalisé des assignations de paternité grâce au logiciel CERVUS version 3.0 (Marshall *et al.*, 1998). Nous avons utilisé les mêmes critères d'assignation que ceux décrits dans Gauzere *et al.* (b) *in prep*, mis à part le fait que nous avons autorisé les assignations de descendants à des pères n'étant pas issus de la même population d'origine que le descendant. Ces assignations de paternité nous ont permis d'identifier pour chaque semis génotypé une catégorie parmi "Non migrant", "Migrant" et "Inconnu". Les descendants assignés à un père se trouvant dans la population d'origine de leur mère ont été catégorisés "Non migrant". Les descendants ne trouvant aucun père dans les populations ont été catégorisés "Migrant". Enfin, les descendants n'ayant pas été génotypés sur assez de marqueurs (moins de 6), ayant été assignés à plusieurs pères ou à des pères ne se trouvant pas dans la population d'origine des mères ont été catégo-

risés "Inconnu". Ce choix des critères d'assignation et la définition des différentes catégories nous permet d'être conservatif sur les statuts "Migrant" et "Non migrant".

APPARENTEMENT AU SEIN DES NUAGES POLLINIQUES

L'information moléculaire disponible pour les 2088 descendants a également été utilisée pour estimer l'apparement paternel moyen au sein des nuages polliniques "migrant" et "non-migrant" par population. Connaissant le génotype des mères, nous avons pu extraire l'haplo-type paternel pour chaque descendant génotypé (voir la section 3.3). Ces haplotypes paternels ont été utilisés pour calculer les apparements paternels par paire d'individus, $\tilde{f}_{k,k'}$, à l'aide du logiciel SPAGEDI (Hardy & Vekemans, 2002). Puis, sur la base des catégories définies par les assignations de paternité, nous avons estimé la moyenne des apparements paternels pour chaque catégorie au sein des populations :

$$\tilde{\rho}_{p,h} = \sum_{k \neq k' \in \{population(p), categorie(h)\}} \frac{\tilde{\rho}_{k,k'}}{n_{ph}(n_{ph}-1)/2}$$

avec j et h les facteurs définissant l'appartenance aux groupes $p = \{N1, N2, N4\}$ et $h = \{"Migrant", "Non-migrant"\}$.

Afin de calculer l'apparement paternel moyen au sein du groupe "Non-migrant" les individus issus d'autofécondation ont été supprimés de l'analyse. Ainsi, le nombre efficace de pères, $Nep = 1/\tilde{\rho}_{p,h}$, a pu être calculé pour représenter la diversité des pères contribuant aux nuages polliniques "Migrant" et "Non-migrant" dans chaque population.

ANALYSE DES TRAITS PAR CATÉGORIE "MIGRANT"- "NON MIGRANT"

Afin d'évaluer l'effet de la dispersion pollinique à moyenne et longue distance sur les valeurs phénotypiques des descendants nous avons utilisé des modèles génétiques intégrant l'information sur la provenance migrant ou non migrante du pollen (M_p) par population :

$$Y_{g,f,h,m,n,p,k} = \mu + T_g + H_{f(g)} + L_{h(f)} + B_{m(f)} + E_{n(g)} + M_p + a_k + \epsilon_{g,f,h,m,n,p,k} \quad (0)$$

$$Y_{g,f,h,m,n,p,k} = \mu + T_g + H_{f(g)} + L_{h(f)} + B_{m(f)} + E_{n(g)} + M_p + a_{k(p)} + \epsilon_{g,f,h,m,n,p,k} \quad (M)$$

avec T_g l'effet du traitement, $H_{f(g)}$ l'effet de la planche hiérarchisé dans le traitement, $L_{h(f)}$ l'effet de la bande hiérarchisé dans la planche, $B_{m(f)}$ l'effet du bloc hiérarchisé dans la planche

et $E_{n(g)}$ l'effet de l'absence ou présence du fertilisateur dans les pots contenant les semis. Tous ces effets environnementaux et de l'origine du pollen sont considérés comme des effets fixes. a_k et $\epsilon_{g,f,h,m,n,p,k}$ sont les termes aléatoires des modèles, avec a_k l'effet génétique additif de l'individu k et $\epsilon_{g,f,h,m,n,p,k}$ les résidus du modèle. Les effets aléatoires sont supposés normalement distribués et les éléments a_k sont supposés être tirés dans $\{a_k\} \sim N(0, A \times V_A)$, avec A la matrice des apparentements et V_A la variance génétique additive, et les éléments $\epsilon_{i,f,h,m,n,k}$ tirés dans $\{\epsilon_{p,m,n,j,k}\} \sim N(0, I_{de} \times V_R)$, avec I_{de} la matrice identité et V_R la variance résiduelle du modèle. Dans le modèle (M), $a_{k(p)}$ a pour variance V_{Ap} , avec deux paramètres de V_A différents pour les catégories "Migrant" et "Non-migrant".

Le modèle (0) décrit un effet de l'origine du pollen sur la moyenne phénotypique des descendants à travers l'effet fixe M_p ; le modèle (M) décrit en plus un effet de l'origine du pollen sur la variance génétique additive du trait Y , par la présence de deux paramètres de variance V_A différents. Pour chaque trait, nous avons tout d'abord comparé les modèles hiérarchisés (0) et (M) avec un test de ratio de vraisemblance (Wilson *et al.*, 2010). Puis, à partir du meilleur modèle (sur un critère de $p_{valeur} < 0.05$), nous avons testé l'effet de la provenance d'origine du pollen sur la moyenne phénotypique des individus (test de l'effet fixe M_p). Les tests ont été considérés significatifs lorsque leurs p_{valeur} étaient inférieures à 0.05 et marginalement significatifs si $0.05 < p_{valeur} < 0.1$. Du fait du grand nombre de tests réalisés (autant que de traits \times populations), nous avons choisi de ne discuter que les tests ayant une $p_{valeur} < 0.02$.

Ces modèles ont été appliqués au sous-ensemble de semis qui ont été à la fois génotypés et identifiés dans une des deux catégories "Migrant" ou "Non-migrant". Afin d'utiliser toute l'information disponible, les individus présents dans les deux modalités du dispositif ont été analysés. Les modèles ont été analysés en utilisant la version R du logiciel AsReml (Gilmour *et al.*, 2006) et l'information sur les relations de parenté connues sous forme de pedigree. Pour les traits présentant des effets maternels (δD , $BioM$ et $t_{b2 \rightarrow 3(2011)}$), où l'information sur les apparentements paternels a un effet non-négligeable sur l'estimation de V_A (voir Gauzere *et al.* (b) *in prep*; Chapitre 3.3) nous n'avons utilisé que l'information sur les apparentements maternels afin de ne pas biaiser les analyses (l'information sur les apparentements paternels n'étant connue que pour les individus "Non-migrant" par construction).

Résultats

COMPOSITION DES NUAGES POLLINIQUES MIGRANT ET NON MIGRANT

La représentation des fenêtres de débourrement végétatif par population nous permet de retrouver le patron classique de variation des dates de débourrement avec l'altitude, i.e. des débournements plus précoces à basses altitudes (N1 et N2) qu'à haute altitude (N4; Figure 5.3). Globalement, nous observons que les populations les plus proches géographiquement, N1 et N2, sont aussi celles qui ont les fenêtres de débourrement les plus chevauchantes. Cependant, entre années les fenêtres phénologiques peuvent être plus ou moins chevauchantes entre populations (Figure 5.3). En 2012 notamment, les populations N1 et N4 se retrouvent complètement décalées phénologiquement.

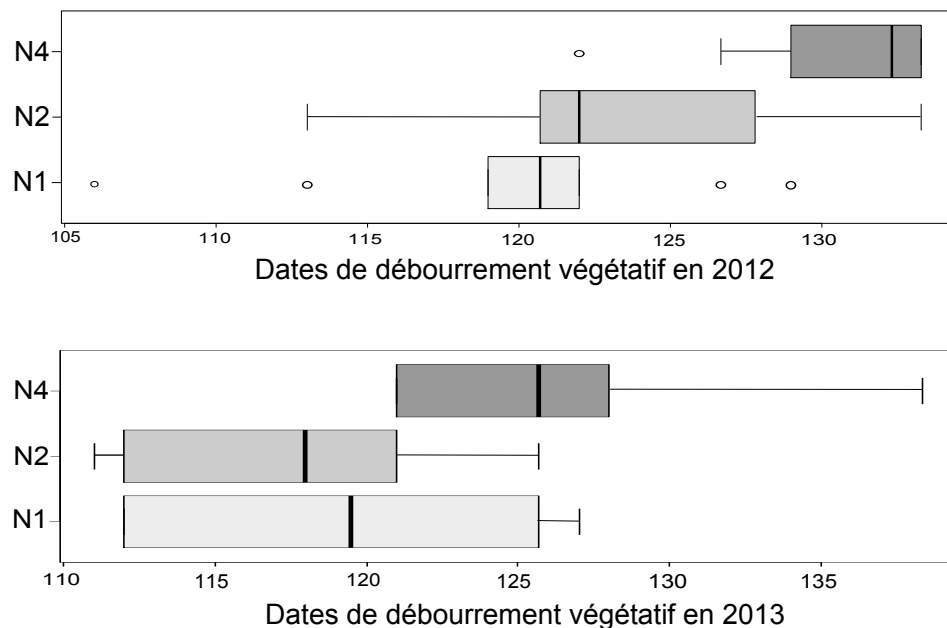


FIGURE 5.3 – Fenêtres de débourrement végétatif en population N1, N2 et N4 évaluées sur un sous-échantillon d'arbres adultes au sein de ces populations. Dates de débourrement représentées en jour julien.

Globalement, ces résultats laissent penser que les croisements entre individus le long du gradient altitudinal sont en général possibles. Au vu de la variation inter-annuelle des dates de débourrement, il aurait été préférable de pouvoir analyser les phénologies des arbres en populations naturelles en 2009, l'année où ont été échantillonnées les descendances maternelles. Malheureusement, l'année 2009 nous ne disposons pas d'un suivi phénologique assez important en population N2 afin de pouvoir mesurer des dates de débourrement pour les adultes

de cette population.

Les assignations de paternité ont permis de caractériser 1000 individus "Non migrant", 758 "Migrant" et 330 "Inconnus". Les individus au sein de ces catégories sont répartis de manière équilibrée dans les différentes modalités du traitement du dispositif en pépinière. La majorité des pères assignés aux descendants sont des pères qui se trouvent dans la même population que l'arbre-mère (en moyenne 89 % d'assignation intra-population ; Table 5.1). En comparaison, les cas où le père assigné provient d'une autre population que celle de l'arbre-mère (i.e., assignation inter-population) est seulement de 11 % en moyenne. Ces assignations inter-populations ne nous permettent pas de déduire de direction privilégiée de dispersion pollinique. En revanche, nous pouvons remarquer que les assignations inter-populations sont plus fortes pour les descendants de la population N2 que les deux autres (presque 8 % de plus ; Table 5.1).

Population Mère	Population Père		
	N1	N2	N4
N1	91,90	3,47	4,63
N2	7,25	83,19	9,57
N4	4,61	4,32	91,07

TABLE 5.1 – Proportion d'individus assignés intra-population et inter-population par population d'origine des arbre-mères.

L'estimation des apparentements paternels moyens par catégorie "Migrant"/"Non-migrant" et par population indique que les nuages polliniques migrants ont tendance à être moins apparentés, i.e. plus diversifiés, que les nuages polliniques non migrants (Figure 5.2). Cette différence entre les nuages polliniques "Migrant" et "Non migrant" n'est pas variable entre populations, ce qui paraît étonnant puisqu'on aurait pu s'attendre à ce que la localisation des populations sur le gradient altitudinal impacte la diversité du pollen pollinique reçue. En effet, si comme le suggèrent nos résultats précédents (Chapitre 2) la dispersion pollinique se réalise de manière privilégiée du bas vers le haut du gradient altitudinal, la population N1, localisée à basse altitude, aurait pu recevoir un nuage pollinique migrant moins diversifié que les populations N2 et N4, à plus hautes altitudes.

Catégorie	Population		
	N1	N2	N4
Non migrant	19,8	18,0	19,8
Migrant	20,3	20,7	21,7

TABLE 5.2 – Nombre efficace de père ($Nep = 1/\tilde{\rho}_{p,h}$) moyen au sein de chacune des catégories "Migrant" et "Non migrant" par population.

VARIANCES GÉNÉTIQUES CHEZ LES MIGRANTS ET NON-MIGRANTS

Globalement, peu d’effet de la provenance du pollen sur les variances génétiques des traits est détecté (Table 5.3). De plus, ces effets, lorsqu’ils sont significatifs pour certains traits, ne se retrouvent pas sur toutes les populations. Plus précisément, le seul trait qui présente un signal fort de l’effet de l’origine du pollen sur la variance génétique par rapport au seuil de

Trait	N1		N2		N4	
	Meilleur Modèle	p_{valeur}	Meilleur Modèle	p_{valeur}	Meilleur Modèle	p_{valeur}
$t_{b2 \rightarrow 3(2011)}$	(0)	-	(M)	0,002	(0)	-
%N	(M)	0,041	(M)	0,085	(0)	-

TABLE 5.3 – Effet de la provenance d’origine du pollen sur la variance génétique des descendants. Seuls les résultats pour les traits où le modèle (M) est significativement meilleur (avec une $p_{valeur} < 0.05$) que le modèle (0) pour au moins une population sont représentés. Legend : "-" $p_{valeur} > 0.10$; "." $0.1 < p_{valeur} < 0.05$; "*" $0.01 < p_{valeur} < 0.05$.

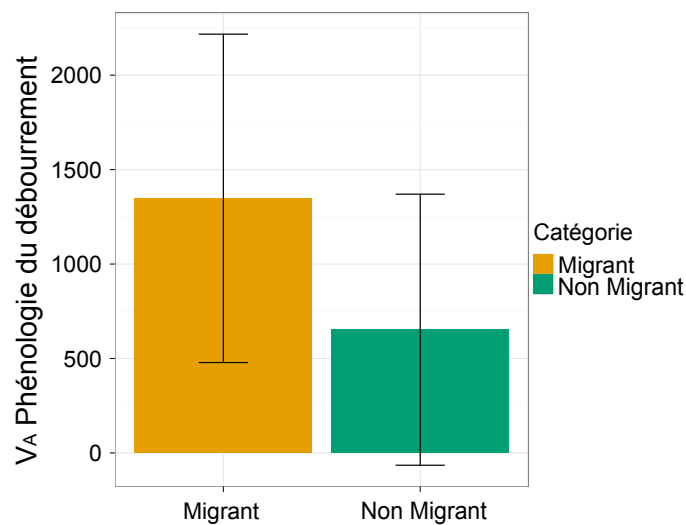


FIGURE 5.4 – Variance génétique additive (V_A) par catégorie "Migrant" et "Non migrant" pour le trait de débourrement végétatif en population N2.

$p_{valeur} < 0,02$ est la date de débourrement en population N2 (Table 5.3). Dans cette population, on peut voir que le groupe d'individus issus de croisements avec du pollen immigrant est plus variable génétiquement que le groupe d'individus issus de croisements avec un père local (Figure 5.4).

EFFET MOYEN D'APPARTENANCE AU POOL MIGRANT

L'effet de la migration pollinique sur les moyennes génétiques n'est détecté que pour 2 traits et non généralisable entre populations (Table 5.4). Avec un seuil de $p_{valeur} = 0,02$, seul l'effet de la migration sur la moyenne de l'accroissement en diamètre dans la population N2 peut être considéré significatif. Pour ce trait on peut voir que la moyenne génétique des individus migrants est plus forte que celle des individus non-migrants (Table 5.4).

Trait	N1			N2			N4		
	$p_{valeur}M_p$	NM	M	$p_{valeur}M_p$	NM	M	$p_{valeur}M_p$	NM	M
δD	0,0912 .	0,60	0,67	0,008 *	0,66	0,78	-	-	-
VD_{2011}	-	-	-	0,031 *	2883	2864	-	-	-

TABLE 5.4 – Test de l'effet de l'origine du pollen (M_p) sur la moyenne génétique des descendants. " $p_{valeur}M_p$ " est la statistique du test, "M" la valeur génétique moyenne des individus issus de croisement avec un père non-local et "NM" la valeur génétique des individus issus de croisement avec un père local. Seuls les résultats pour les traits présentant un effet significatif de M_p (avec une $p_{valeur} < 0.05$) dans une des populations sont représentés. Legend : "." $0.1 < p_{valeur} < 0.05$; "*" $p_{valeur} < 0.05$; "-" $p_{valeur} > 0.10$.

Discussion

Cette étude nous a permis de démontrer que les flux de pollen provenant de moyenne et longue distance pouvaient impacter les valeurs génétiques des traits fonctionnels. Plus précisément, pour la phénologie du débourrement ($t_{b2 \rightarrow 3}$), les individus issus de croisements avec du pollen migrant sont plus variables génétiquement que les individus issus de pollen local dans la population centrale N2. Pour l'accroissement en diamètre (δD), les individus migrants ne sont pas plus variables mais présentent une moyenne génétique différente de celles des non-migrants en population N2. Ces deux impacts de la migration pollinique conduisent au même effet global qui est d'augmenter la variance génétique disponible pour la sélection au sein de

la population N2.

Dans le chapitre précédent, nous avons vu que δD et $t_{b2 \rightarrow 3}$ présentaient une différenciation adaptative significative entre populations (section 4.1). La divergence génétique des traits dans le paysage est donc bien ici un pré-requis pour que la migration ait un effet sur les variances génétiques. Une première question qui émerge de ces résultats est pourquoi la migration n'affecte qu'un sous-ensemble des traits présentant une différenciation génétique sur le gradient (e.g. la sénescence foliaire, accroissement en hauteur, le contenu en carbone des feuilles sont également sous sélection divergente). Une première explication porte sur la puissance statistique dont nous disposons pour aller tester cet effet ; étant donnée la différenciation relativement modeste à l'échelle du versant, les tests réalisés ici (Tables 5.3 et 5.4) ont une probabilité assez élevée de ne pas détecter les différences génétiques entre les migrants et non-migrants qui elles-même doivent être quantitativement faibles. Une deuxième explication est que la divergence génétique de δD et $t_{b2 \rightarrow 3}$ est plus forte que celle des autres traits sous sélection (Barton, 2001). Nos résultats actuels ne nous permettent pas de pouvoir comparer quantitativement la divergence des traits entre populations et donc de tester cette hypothèse. Par la suite, il pourrait être intéressant de parvenir à classer les traits étudiés en utilisant une mesure de Qst . Cependant, les fortes incertitudes sur les mesures de variances génétique inter-population, dues au fait que nous ne travaillons que sur trois populations, pourraient fortement contraindre cette comparaison et l'interprétation plus fine de nos résultats. Une autre explication à ces résultats est que les traits sur lesquels on observe un effet de la migration pourraient être des traits directement reliés aux capacités de dispersion à longue distance chez les arbres. Pour la phénologie par exemple, le décalage des phénologies de débourrement et floraison le long du gradient pourrait conduire à ce que les arbres-mères de la population N2 ne se reproduisent qu'avec des mâles tardifs à basse altitude et des mâles précoces à haute altitude, ce qui aurait pour conséquence d'augmenter la variance génétique de ce trait. Toujours de manière hypothétique, si les traits d'accroissement juvéniles sont reliés au diamètre définitif des arbres adultes, alors on pourrait imaginer les diamètres des arbres, en affectant la fertilité de ces derniers (Oddou-Muratorio *et al.*, 2005), affectent également leurs capacités de dispersion à longue distance.

Un autre résultat important de cette étude est que les effets de la migration pollinique n'ont été détectés que sur les capacités d'adaptation de la population N2. Nous avons vu que

dans cette population, localisée au milieu du gradient altitudinal étudié, la proportion d'assignation inter-population était plus importante que dans les deux autres populations (Table 5.1). De plus, nous avons observé précédemment que les distances de dispersion pollinique avaient tendance à être plus fortes au sein de de cette population que dans les deux autres ($\delta_{N2} = 63$ m, $\delta_{N1} = 39$ m, $\delta_{N4} = 35$ m ; Chapitre 2). Une interprétation possible à ce résultat serait donc que l'effet de la migration dépend plus de la qualité (i.e. pollen de plus ou moins longue distance) que de la quantité de pollen immigrant intercepté par les populations (la population N2 ne présente en effet pas les plus forts taux d'immigration mais pourrait recevoir du pollen de plus longue distance ; Chapitre 2).

De manière similaire à l'étude de Nilsson (1995), l'origine du pollen immigrant peut a priori se déduire de la moyenne génétique des individus migrants et de la divergence génétique moyenne des populations sur le gradient. Pour le trait δD , les précédents résultats montraient que la population N2 présentait les plus fortes valeurs d'accroissement en diamètre sur le gradient (Figure 5.5). Ainsi, les flux de gènes provenant de plus basses ou hautes altitudes devraient tout deux apporter des allèles réduisant l'accroissement en diamètre. Or, nous observons que la valeur génétique des individus migrants pour δD est supérieure à celle des individus non-migrants. Ce résultat, qui s'oppose à notre attendu que les flux de pollen proviendraient majoritairement de plus basses altitudes du fait des décalages de floraison

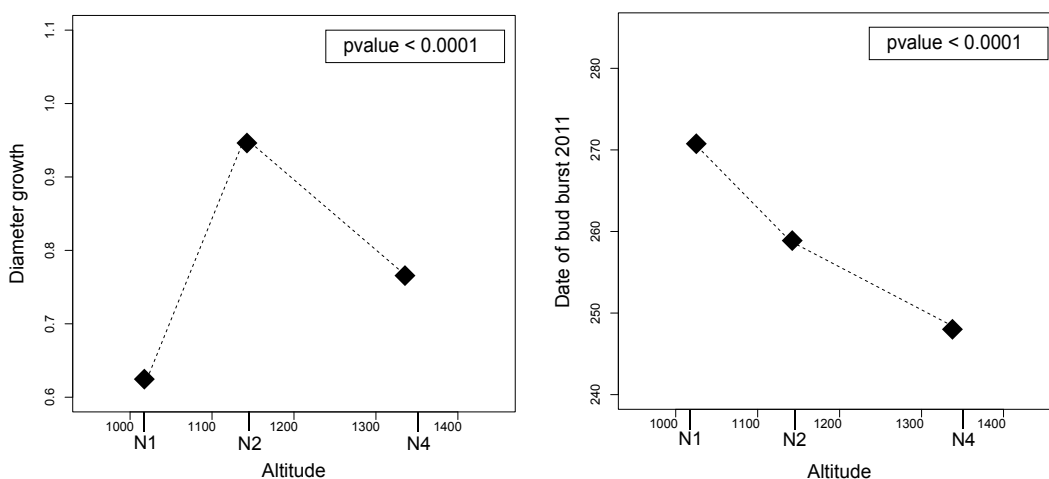


FIGURE 5.5 – Patron de différenciation génétique mesuré en jardin commun pour les traits d'accroissement en diamètre et phénologie du débourrement. Graphiques repris du Chapitre 4 section 4.1.

mâle et femelle, pourrait s'expliquer de deux manières. D'une part, les populations étudiées se trouvant dans une forêt quasi-continue de hêtre de 750 m à 1700 m d'altitude, l'origine du pollen migrant peut être variée et notamment provenir d'un même niveau altitudinal. D'autre part, il est possible que le suivi de trois populations sur le gradient ne nous permette pas de pouvoir définir assez finement le niveau altitudinal auquel les hêtres sont à leur optimum de croissance. Notamment, l'optimum écologique du hêtre et les meilleures capacités de croissance pourraient se trouver entre les populations N2 et N4, et dans ce cas les flux de pollen provenant d'altitudes légèrement plus hautes pourraient expliquer nos résultats. Ainsi, dans notre cas la différence de moyenne génétique du δD entre individus migrants et non-migrants ne permet donc pas de conclure sur l'origine du pollen. Les valeurs d'accroissement plus fortes pour les individus migrants pourraient également s'expliquer par des effets de vigueur hybride (Costa e Silva *et al.*, 2014). Mais la population N2 n'est pas la population où on s'attendrait à ce que ces effets soient les plus forts sachant qu'elle présente un nombre efficace de pères similaire aux autres populations et surtout un coefficient d'apparentement intra-population faible ($\theta_{N1} = 0,065$, $\theta_{N2} = 0,032$ et $\theta_{N4} = 0,038$; section 4.1).

Globalement, nos résultats soulignent la nécessité d'aller vers une meilleure caractérisation des flux réalisés à l'échelle du versant et notamment une meilleure caractérisation de l'origine des migrants afin de mieux comprendre l'effet de la migration sur les capacités évolutives en utilisant des approches directes. Premièrement, il est important de souligner que les assignations réalisées par Cervus et les définitions des catégories qui ont été utilisées pour ces premières analyses avaient été réalisées dans le but de construire des pedigrees avec le moins d'erreur possible et de correctement estimer les V_A et h^2 dans les chapitres précédents (Chapitres 3 et 4). En revanche, ces assignations étaient moins appropriées pour analyser l'effet de la migration; les assignations de paternité mériteraient donc d'être reprises pour mieux définir les catégories "Migrant" et "Non migrant", notamment afin de prendre en compte les individus assignés à des pères extérieurs à leur population dans la catégorie "Migrant". Deuxièmement, une distinction plus fine des individus migrants pourrait être faite afin mieux prendre en compte l'effet de la dispersion à longue distance dans ces analyses. En effet, notre catégorisation "Migrant" fait que les individus issus d'un croisement avec un père avoisinant la population des mères et un père plus lointain sont considérés de manière égale. Sur la base du modèle de croisement dans le voisinage développé par Klein *et al.* (2008) et des données microsatellites nous avons tout de même pu dériver une mesure quantitative de l'effet migrant

(ν_M) pour chaque descendant génotypé, traduisant la vraisemblance d'appartenir au pool migrant local ou global, qui nous permettra d'assigner avec plus de certitude des individus dans la catégorie "Migrant" :

$$\nu_M = \log \left(\frac{p(M/FA_{global})}{p(M/FA_{local})} \right)$$

avec $p(M/FA_{global})$ la vraisemblance d'appartenir au pool migrant sachant les fréquences alléliques globales sur tout le versant nord du Mont-Ventoux, et $p(M/FA_{local})$ la vraisemblance d'appartenir au pool migrant sachant les fréquences alléliques de la population d'origine de la mère. Cependant, il est probable que cette variable soit également peu informative ; en effet la faible différenciation génétique des populations aux marqueurs étudiés ($F_{st} = 0.026$) restreint probablement notre capacité à pouvoir distinguer l'origine des pères grâce aux marqueurs moléculaires.

Une autre approche qui peut être utilisée afin de comprendre nos résultats est de comparer les observations empiriques réalisées durant cette thèse aux prédictions faites par des modèles mécanistes représentant explicitement les processus étudiés. L'idée serait donc de parvenir à développer un modèle de dispersion pollinique à l'échelle du versant nord (e.g., en s'appuyant sur la connaissance de la densité de hêtre en chaque point du versant : données IFN, cartographies réalisées par l'URFM), gradient environnemental sur lequel les valeurs génétique des traits seraient distribuées spatialement (en extrapolant spatialement les résultats obtenus dans la section 4.2). Ce modèle alimenté par nos informations empiriques sur l'héritabilité et le niveau de divergence génétique des traits fonctionnels permettrait de prédire l'effet de la dispersion pollinique au sein des populations de hêtre. La mise en place de ce type d'approche nécessite alors de caractériser la dispersion sur le versant nord du Mont-Ventoux. Une première solution consisterait à utiliser les fonctions de dispersion du pollen estimées dans le Chapitre 2 et de vérifier si leur utilisation à l'échelle du versant aboutit à des taux de migration prédits cohérents avec ceux estimés dans ce même chapitre. Une deuxième solution consisterait à développer des méthodes plus mécanistes permettant de tracer les mouvements polliniques (basées sur des modèles utilisant des données météorologiques sur les vents et de la topographie ; Kremer *et al.* 2012). Ces réflexions sont actuellement abordées dans le cadre du projet ANR MECC.

Conclusion

En utilisant une approche originale de caractérisation des migrants, nous avons pu démontrer empiriquement que les flux de gènes pouvaient avoir un rôle favorable sur l'apport de variance génétique au sein des populations naturelles. De manière très intéressante, ces effets favorables de la migration par pollen sur les capacités d'adaptation ont pu être détectés sur un évènement de reproduction, à une échelle micro-locale et sur des populations présentant de faibles niveaux de divergence génétique avec leur voisinage. L'effet de la migration semble cependant complexe, les migrants pouvant apporter de la variance génétique totale et décaler les moyennes génétiques des descendants. Globalement, le peu d'effet de la migration détecté et uniquement au sein de la population la moins dense, suggère qu'une meilleure caractérisation de l'origine des migrants et de la dispersion pollinique à l'échelle du Mont-Ventoux sont nécessaires pour pouvoir tirer des conclusions sur l'effet de la migration dans les populations naturelles.

DISCUSSION GÉNÉRALE ET PERSPECTIVES



Quels sont les facteurs favorisant l'adaptation génétique des populations naturelles à un environnement changeant ? Quels sont ceux la limitant ? Quels sont ceux maintenant la diversité génétique dans les populations soumises à sélection ? Autant de questions centrales en biologie évolutive dont une des réponses est : les flux de gènes entre populations localement adaptées. Cette constatation souligne les effets très divers que peuvent avoir les flux de gènes sur l'adaptation. Dans le contexte actuel de changements climatiques rapides, une question majeure est celle du poids des effets favorables et défavorables des flux de gènes sur l'adaptation future des populations aux changements environnementaux. Bien que les modèles théoriques aient abordé ces questions dans la littérature (Kirkpatrick & Barton, 1997; Barton, 2001; Duputié *et al.*, 2012), les études empiriques se sont beaucoup focalisées sur la mesure des effets défavorables des flux de gènes sur l'adaptation locale, i.e. leur effet homogénéisant (e.g., Hendry & Taylor 2004; Sambatti & Rice 2006). Plus globalement, les études sur ces questions sont généralement basées sur des approches indirectes, i.e. ne mesurant pas directement la migration mais simplement la distance ou divergence génétique neutre (F_{st}) entre les populations, ou des manipulations d'évolution expérimentale (Swindell & Bouzat, 2006). Finalement, le rôle des effets favorables des flux de gènes sur l'adaptation dans les populations naturelles est très mal connu.

L'objectif principal de ma thèse a été de comprendre l'effet des flux de gènes par pollen sur les capacités adaptatives de trois populations de hêtre situées sur un gradient altitudinal. Plus précisément, nous avons évalué les effets de la dispersion pollinique sur les variances génétiques (V_A) des traits fonctionnels adaptatifs. Préalablement, nous avons démontré que notre système d'étude était pertinent pour explorer cette question.

En effet, nous avons tout d'abord démontré que **1)** les populations étudiées recevaient de forts flux de pollen immigrant ($m = 56\%$). Les caractéristiques des courbes de dispersion indiquent que la dispersion pollinique se réalise essentiellement dans le voisinage proche des arbres adultes ($\delta = 46$ m ; Chapitre 2), même si des événements de dispersion à longue-distance sont non-négligeables.

Nous avons ensuite démontré que **2)** tous les traits fonctionnels étudiés présentaient une base génétique significative qui pouvait donc être transmise d'une génération à l'autre (Chapitre 3).

Et enfin, **3)** nous avons détecté une divergence génétique significative entre les trois populations, vraisemblablement due à l'action de la sélection divergente pour la moitié des traits

fonctionnels étudiés. Ce résultat est obtenu malgré les faibles distances séparant les populations (Chapitre 4).

Forts de ces trois résultats, une approche originale a été mise en œuvre pour mesurer l'apport contemporain de variance génétique par la migration pollinique. Celle-ci a consisté à comparer les valeurs génétiques d'individus migrants et non-migrants par pollen. Ainsi, un effet des flux de pollen sur l'augmentation de la variance génétique de deux traits différenciés génétiquement le long du gradient a pu être détecté au sein de la population localisée au milieu du gradient altitudinal (N2; Chapitre 5). Ces résultats permettent de démontrer que les effets favorables de la migration par pollen sur les capacités d'adaptation peuvent être détectés sur un événement de reproduction, à des échelles micro-locales et sur des populations présentant de faibles niveaux de divergence génétique avec leur voisinage.

Plus globalement, les parties de cette thèse apportent différentes informations sur les capacités évolutives des populations de hêtre étudiées. Dans cette discussion générale, je synthétiserai donc ces résultats afin d'éclaircir la question des capacités d'adaptation des populations d'arbres aux changements environnementaux. Je discuterai également de points méthodologiques et d'éléments de perspectives soulevés par ce travail de thèse.

6.1 Vue d'ensemble des capacités adaptatives des populations d'arbres aux changements environnementaux

Un des concepts les plus difficiles à définir en biologie évolutive est celui de "potentiel adaptatif". Bien qu'il n'existe pas de définition universelle, une définition large et consensuelle du potentiel adaptatif d'une population est : "la capacité à maintenir de la diversité génétique malgré des variations environnementales dans l'espace et le temps" (Scotti, 2010). Selon cette définition, l'ensemble des résultats de cette thèse suggère que la population N2, localisée au milieu du gradient altitudinal étudié, présente un fort potentiel adaptatif. En effet, cette population présente à la fois une forte évolutivité totale ($\mu_T = 0,22$) et la migration pollinique au sein de cette population augmente les variances génétiques de deux traits adaptatifs importants, la phénologie du débourrement et l'accroissement en hauteur. À l'inverse, les populations aux deux extrêmes du gradient, N1 et N4, présentent de faibles évolutivités totales ($\mu_T = 0,08$), probablement à cause d'un effet plus fort de la sélection naturelle sur ces

populations, et aucun effet de la migration n'a pu être détecté sur la variance génétique des traits. Cependant, de nombreuses études suggèrent que l'existence de variance génétique pour les traits ne garantit pas la réponse de ces traits à la sélection, notamment si les changements évolutifs sont trop importants ou si l'interaction entre les traits contraint trop la réponse (Etterson & Shaw, 2001; Blows & Hoffmann, 2005). Même si l'étude du nombre efficace de dimensions souligne peu de différences entre les populations étudiées (section 4.2), la matrice \mathbf{G} de la population N1 présente plus de covariances génétiques significatives ou marginales entre traits que les autres populations. Ces covariances entre traits pourraient alors fortement affecter les taux d'évolution dans cette population. De plus, la population N1, qui se trouve au plus bas niveau altitudinal étudié, est dans une situation que l'on pourrait qualifier de "marge arrière" dans le sens où cette population marque le début de la présence du hêtre sur les pentes du Mont-Ventoux et subit des conditions écologiques proches de celles rencontrées dans la limite sud de son aire de distribution (climat Méditerranéen). Cette position pourrait fortement contraindre la réponse de cette population qui ne peut probablement pas recevoir de flux de gènes pré-adaptés de plus basses altitudes ou latitudes qui favoriseraient son évolution future. Même si à l'heure actuelle les flux de gènes immigrants en N1 ne contraignent vraisemblablement pas l'action de la sélection (section 4.1), la détérioration des conditions environnementales pourrait accentuer les effets défavorables des flux de gènes provenant de plus hautes altitudes. De plus, nos résultats sur le régime de reproduction des populations et sur l'apparentement génétique intra-population ont montré que la population N1 était également la population qui avait tendance à présenter les plus faibles nombres de pères efficaces et la plus forte consanguinité ($Nep_{N1} = 12$; Chapitre 2; $\theta_{N1} = 0.065$; Chapitre 4.1). Les effets combinés du manque d'effet favorable des flux de pollen, des covariances génétiques et des faibles variances génétiques dans la population N1 semblent donc indiquer que cette population est relativement plus vulnérable aux changements climatiques que les populations N2 et N4.

Les résultats de cette thèse nous permettent de comparer le potentiel adaptatif *relatif* entre populations. L'évaluation du potentiel adaptatif *absolu* des populations nécessiterait dans un premier temps de connaître les pressions de sélection à l'œuvre au sein de chacune des populations. Durant sa thèse, Aurore Bontemps avait utilisé une approche de génétique quantitative en population naturelle pour évaluer la réponse à la sélection de cette population N1 (Bontemps, 2012). Sur la base de l'estimation de gradients de sélection standardisés et

d'héritabilités, l'application de l'équation du sélectionneur montrait une forte réponse potentielle à la sélection pour les différents traits fonctionnels suivis (phénologie du débourrement, LM , $\delta^{13}C$ et $\%N$). L'intensité de cette réponse était notamment expliquée par la mesure de fortes pressions de sélection sur les traits comparativement à la distribution des gradients de sélection issue de la méta-analyse de Kingsolver *et al.* (2001). Il se dégageait donc globalement de cette thèse que la population N1 présentait des capacités adaptatives suffisantes pour répondre aux fortes pressions de sélection existantes. Cependant, Bontemps (2012) avait également détecté de nombreuses covariances génétiques entre traits au sein de la population N1 qui justifient la nécessité d'estimer des réponses à la sélection multi-traits plutôt qu'univariées. Ce travail en cours devrait permettre de mieux comprendre comment les covariances génétiques entre traits en population N1 affectent les taux d'évolution des traits. Pour aller plus loin, il serait notamment intéressant de mesurer les gradients de sélection dans les populations N2 et N4 et de les confronter aux matrices \mathbf{G} estimées durant ma thèse.

Une deuxième étape pour caractériser le potentiel adaptatif *absolu* des populations est de tenir compte des conséquences démographiques sur l'action de la sélection. Les études réalisées jusqu'à présent sur le hêtre du Mont-Ventoux se sont particulièrement intéressées aux mécanismes pouvant contraindre ou faciliter l'adaptation génétique des populations d'arbres dans le but de comprendre les capacités micro évolutives de cette espèce (cette thèse, Bontemps 2012). Or, les modèles théoriques de sauvetage évolutif¹ montrent que la probabilité d'extinction d'une population est également fonction de la taille des populations (Gonzalez *et al.*, 2013). La compréhension des capacités de survie des populations ne doit donc pas se limiter à la simple caractérisation du potentiel évolutif des traits, mais doit intégrer l'étude du lien complexe qu'il existe entre adaptation génétique et démographie (Kinnison & Hairston, 2007), notamment pour les populations de faibles tailles. Généralement, les espèces d'arbres présentent de fortes tailles de population qui devraient empêcher la chute démographique jusqu'à des seuils critiques où les probabilités d'extinction sont fortes. Néanmoins, le maintien de tailles de population viables dépendra fortement des capacités de reproduction et régénération des populations d'arbres. Le recrutement des semis, permettant la régénération des forêts, est connu pour être une phase clef de l'évolution des populations d'arbres, car, chez ces espèces, une forte purge démographique a lieu lors du stade juvénile (Petit & Hampe, 2006). Par ailleurs, chez des peuplements de hêtre en zone Méditerranéenne, un effet du changement

1. "evolutionary rescue" en anglais

climatique sur le déplacement des hêtraies a déjà observé (Penuelas & Boada, 2003). Ce déplacement se produit notamment via une isolation progressive des populations "arrières" dans lesquelles on mesure une accentuation de la défoliation et une diminution du taux de recrutement (Penuelas & Boada, 2003). Les changements de conditions climatiques ont donc bien un impact mesurable à des échelles de temps courtes sur la démographie et le fonctionnement des peuplements forestiers. Sur le Mont-Ventoux, les premiers signes de mortalité du hêtre ont été mesurés récemment (Hendrik Davi, communication personnelle).

Dans cette thèse je me suis particulièrement focalisée sur le test empirique des effets hypothétiques des flux de gènes par pollen sur l'adaptation locale. Cependant, la dispersion, et notamment le déplacement des individus, peut également avoir des effets sur la démographie des populations et peut fortement affecter la persistance des populations. En effet, la migration, en maintenant les tailles de populations, peut avoir des effets favorables sur leur persistance (Figure 6.1). L'augmentation de la persistance des populations peut ensuite augmenter les chances d'adaptation aux nouvelles conditions environnementales (Garant *et al.*, 2007). La compréhension des facteurs écologiques générant des dysfonctionnements du régime de reproduction, de la migration ou entraînant la mortalité semble cruciale pour pouvoir, en complément à l'évaluation du potentiel adaptatif, comprendre la persistance des populations d'arbres malgré les changements climatiques.

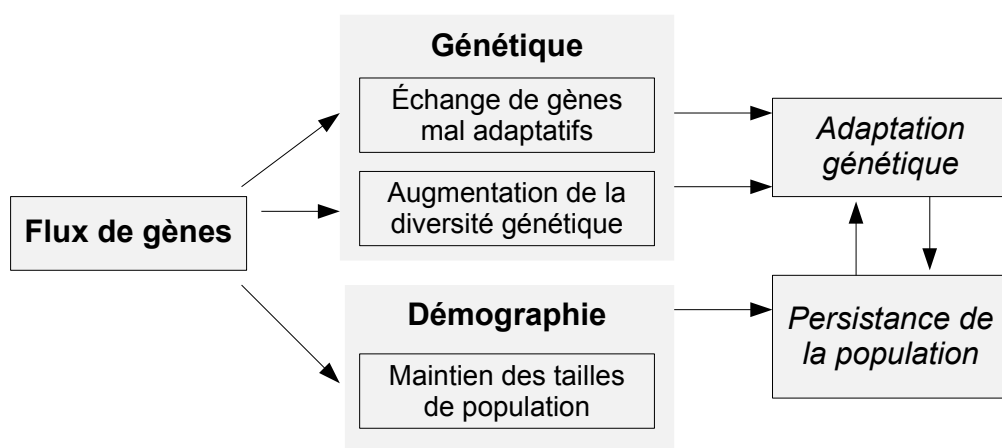


FIGURE 6.1 – Effets complexes de la migration par pollen et graines sur les capacités d'adaptation génétique et la démographie de populations connectées par des flux de gènes. Inspiré de Garant *et al.* (2007).

Une manière de caractériser les processus génétiques et démographiques impliqués dans l'évolution de populations dans l'espace et le temps est notamment d'utiliser des approches de modélisation. Chez les espèces d'arbres, pour lesquelles l'évolution des traits phénotypiques ne peut pas être mesurée empiriquement, ces approches ont l'intérêt de permettre la compréhension des facteurs génétiques, écologiques et environnementaux participant à la mise en place de l'adaptation locale et au maintien des populations. Chez les arbres, ces modèles "démogénétiques" ont déjà souligné l'importance des dynamiques démographiques dans la réponse aux changements environnementaux, notamment via la mortalité (Kuparinen *et al.*, 2010; Oddou-Muratorio & Davi, 2014). Ainsi, il a pu être démontré que, chez les populations d'arbres, la mortalité agit à la fois en accentuant les risques d'extinction locale et en offrant de nouvelles opportunités à l'adaptation génétique grâce au remplacement plus rapide des générations mal adaptées (Kuparinen *et al.*, 2010). Les modèles basés sur les processus soulignent également la nécessité de travailler sur des traits fonctionnels afin de mieux caractériser les aires de distributions actuelles et mieux prédire la réponse des populations aux changements climatiques (Chuine & Beaubien, 2001; Benito Garzon *et al.*, 2011). Les approches empiriques étudiant le déterminisme et la différenciation génétique de traits fonctionnels dans des tests de provenance ou tests de descendances (comme dans cette thèse) sont donc essentiels et doivent être liées à des approches de modélisation afin de déterminer les réponses de ces populations aux changements climatiques futurs (Benito Garzon *et al.*, 2011).

6.2 Influences de la gestion sur l'adaptation des forêts

Les forêts n'étant pas des systèmes complètement naturels dans le sens où elles sont généralement exploitées par l'homme, l'influence des activités anthropiques sur la réponse adaptative des forêts est une composante importante à prendre en compte lorsque l'on souhaite évaluer le potentiel de réponse des populations d'arbres aux changements climatiques. L'échelle locale à laquelle le potentiel évolutif des populations et les effets des flux de gènes ont été caractérisés durant cette thèse est pertinente car elle correspond à l'échelle à laquelle les peuplements forestiers sont gérés. Dans un article de revue auquel j'ai contribué, Lefèvre *et al.* (2013) discutent des aspects de gestion adaptative locale qui peuvent être mis en œuvre pour limiter la vulnérabilité des arbres aux changements climatiques (Annexe A). Afin de mesurer l'intérêt et les conséquences des pratiques de gestion adaptative, il semble crucial de

comprendre comment les pratiques de gestions classiques, telles que la diminution de la densité des peuplements par éclaircies ou l'élagage des arbres, affectent le potentiel évolutif des arbres (Figure 6.2). Dans cette thèse, nous avons vu que la densité des peuplements (conspécifique et totale) semblait être l'un des principaux déterminants écologiques affectant le régime de reproduction et notamment la réception de pollen migrant à longue distance (Chapitre 2 et Chapitre 5 ; effet déjà connu dans la littérature, e.g., Robledo-Arnuncio & Gil 2004). La réduction des densités, en population N1 et N4 notamment, permettrait peut être d'accentuer les effets bénéfiques de la dispersion pollinique à longue distance sur les capacités évolutives de ces populations. En revanche, cette réduction de densité, si elle passe par une réduction de densité conspécifique, peut également diminuer le nombre de pères efficaces contribuant à la reproduction dans les populations (Nep) et avoir des effets défavorables sur la variance génétique (V_A) et le niveau de consanguinité (F) via la taille efficace de la population (Ne ; Figure 6.2).

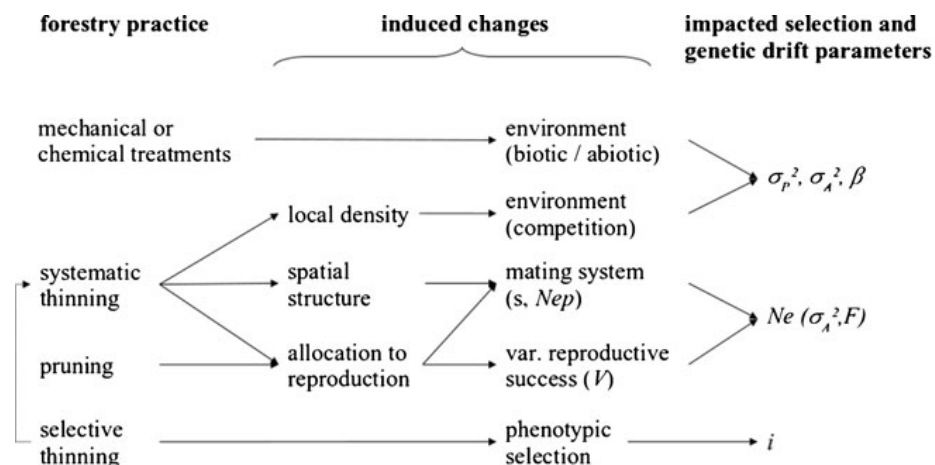


FIGURE 6.2 – Effets attendus des pratiques de gestion forestière sur les processus évolutifs (avec Nep le nombre efficace de pères, Ne la taille efficace de la population, σ_p^2 la variance phénotypique, σ_A^2 la variance génétique, β le gradient de sélection et F le coefficient de consanguinité). Tiré de Lefèvre *et al.* (2013).

Une autre pratique qui permettrait d'aider l'adaptation des populations à de nouvelles conditions climatiques est la pratique des flux de gènes assistés, qui consiste à gérer le mouvement des individus ou gamètes entre populations au sein de l'aire de distribution de l'espèce (Aitken & Whitlock, 2013). En effet, pour les espèces à aires de distribution étendues, on peut imaginer que les génotypes les mieux adaptés aux conditions climatiques nouvellement appliquées à une population se trouvent en fréquence élevée dans une autre localité subissant

déjà les conditions prédites. Plus globalement, cette pratique permettrait d'aller ré-injecter de la diversité génétique dans les populations "à risques" (e.g. populations de faibles tailles efficaces) et d'augmenter ainsi leur probabilité de persistance. Les pré-requis pour utiliser cette pratique de gestion sont l'existence de patrons d'adaptation locale et l'existence d'un minimum de connexion génétique par flux de gènes naturels entre les populations pour éviter des effets d'outbreeding depression (Aitken & Whitlock, 2013). Les populations d'arbres correspondent à ces critères puisqu'elles présentent généralement de forts niveaux de divergence adaptative malgré des forts flux de gènes entre populations (Savolainen *et al.*, 2007). Cependant, le poids des effets défavorables des flux de gènes assistés versus ceux de la mal adaptation causée par le changement climatique est difficilement mesurable. Dans cette thèse, nous avons montré que les effets favorables des flux de gènes pouvaient se mesurer à des échelles micro-locales, indiquant que les programmes de flux de gènes assistés pourraient être réalisés chez les arbres entre populations relativement proches géographiquement, en limitant ainsi les effets potentiels de l'outbreeding depression.

6.3 Le cadre Bayésien : un outil statistique pour propager l'incertitude

Dans les études en écologie et évolution, la faible taille des jeux de données disponibles et la complexité des questions qui sont abordées avec ces jeux de données font que les problèmes sont souvent divisés en parties utilisant différentes méthodes indépendantes entre elles. L'incertitude des données est alors mal prise en compte aux différentes étapes du processus. Le développement des statistiques bayésiennes et leur application récente dans différents domaines de la biologie ont permis de considérablement complexifier les modèles statistiques en permettant la formalisation de modèles hiérarchiques, i.e. des modèles où la vraisemblance des observations peut être facilement écrite par rapport à des niveaux intermédiaires de variables non observées reliées elles-mêmes aux paramètres à estimer. Au delà de la complexification, les approches bayésiennes ont surtout permis de proposer des méthodes qui propagent naturellement les erreurs à travers le modèle jusqu'à l'estimation des paramètres, qui sont considérés comme des variables aléatoires (O'Hara *et al.*, 2008). Les champs d'application des approches bayésiennes aux analyses en écologie et évolution sont larges et promettent à la fois de pouvoir mieux utiliser l'information contenue dans les données et d'améliorer la prise en compte de

l'incertitude, donc d'avoir des estimations et des inférences plus robustes.

Dans le domaine de l'écologie et de la génétique des populations, le développement bayésien du modèle de croisement spatialement explicite, utilisé dans le chapitre 2 de ma thèse (Klein *et al.*, 2008), offre un exemple de l'intérêt d'utiliser une approche bayésienne pour estimer les composantes du régime de reproduction conjointement à la dispersion pollinique. En effet, la structure hiérarchique de ce modèle avec la prise en compte et l'estimation des fertilités individuelles des adultes reproducteurs (variables non-observées) sur le régime de reproduction justifiait tout particulièrement l'intérêt d'une approche bayésienne. Dans le chapitre 2 de cette thèse, nous montrons comment ce modèle a pu être amélioré pour estimer les composantes individuelles du régime de reproduction. D'un point de vue pratique, la maximisation d'un modèle à autant de paramètres avec une approche classique serait difficile, voire impossible, selon la taille du jeu de données. Ce modèle pourrait encore être amélioré pour exploiter tout le potentiel offert par le cadre Bayésien, en définissant une variance des paramètres du régime de reproduction par population et en tirant les valeurs individuelles de ces paramètres dans des distributions à priori. De plus, l'influence des facteurs écologiques sur le régime de reproduction pourrait directement être prise en compte au sein du modèle sous forme de covariables affectant la variation du régime de reproduction à l'échelle individuelle. Ces avancées répondraient à un vrai besoin pour tester les différences de régime de reproduction ou fonction de dispersion entre populations en prenant bien en compte le niveau individuel de variation (Klein *et al.*, 2011; Chybicki & Burczyk, 2013).

En génétique quantitative, les approches bayésiennes sont également de plus en plus utilisées, que ce soit dans le cadre d'assignations de parenté pour prendre en compte différentes sources d'informations sur les relations de parenté entre individus (e.g. statut social des animaux, informations spatiales; Hadfield *et al.* 2006), utiliser le modèle animal sur des données non-gaussiennes (Hadfield, 2010) ou encore estimer les réponses à la sélection multi-traits en prenant en compte les incertitudes autour de l'estimation des matrices \mathbf{G} et des covariances génétiques entre traits et valeur sélective (Stinchcombe *et al.*, 2014). Dans le chapitre 3 de cette thèse, nous nous sommes questionnés sur l'utilisation d'information moléculaire incomplète dans les modèles de génétique quantitative. Le cadre Bayésien permettrait de créer un modèle sur l'apparentement génétique entre individus qui prendrait en compte à la fois des informations moléculaires et la connaissance a priori des pédigrées (par observation ou

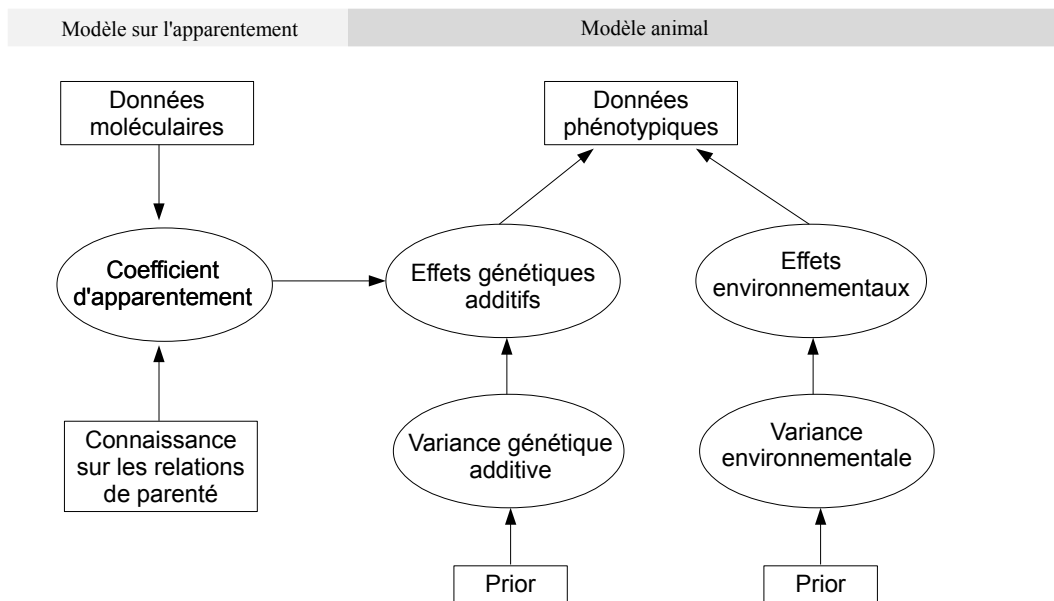


FIGURE 6.3 – DAG hypothétique d’un modèle de génétique quantitative reliant les deux approches classiquement utilisées pour mesurer la variance génétique d’un trait phénotypique : un modèle sur l’apparentement génétique et un modèle animal. Inspiré de O’Hara *et al.* (2008).

échantillonnage de familles) et de coupler ce modèle avec le modèle animal mesurant la variance génétique d’un trait (Figure 6.3). Alors que l’approche classique consisterait à réaliser deux types d’analyses, puis à les confronter, l’approche hiérarchique bayésienne permettrait de prendre en compte l’incertitude sur les apparentements génétiques des individus dont les phénotypes sont mesurés.

Enfin, toujours dans le domaine de la génétique quantitative, l’approche de Ovaskainen *et al.* (2011) donne un bel exemple du niveau de complexité et d’intégration que permettent d’obtenir les modèles hiérarchiques analysés par des approches bayésiennes. En effet, ce modèle, en intégrant à la fois un modèle sur l’évolution des marqueurs neutres (modèle F) et un modèle sur la divergence génétique des traits (modèle animal), permet d’utiliser toute l’information contenue dans les données et ainsi de gagner en puissance sur le test de détection de la sélection dans les cas où les approches classiques de $Fst - Qst$ échouent, i.e. les situations où peu de populations sont étudiées. Ces exemples et beaucoup d’autres soulignent les importantes avancées méthodologiques qu’a permises l’application des statistiques bayésiennes aux études s’intéressant à l’écologie et évolution des populations naturelles. Ces modèles sont d’autant plus intéressants qu’ils permettent d’inclure différentes sources d’informations (e.g.,

a priori, génétique, spatiale, environnementale), de prendre en compte l'incertitude liée à l'utilisation de jeux de données restreints et de propager cette incertitude dans des modèles hiérarchiques. Cependant, les approches bayésiennes n'ont pas toujours lieu d'être, et notamment à complexité égale et prior non informatifs, l'utilisation d'une approche bayésienne à la place d'une approche classique n'apportera aucun bénéfice, avec un temps de calcul rallongé : par exemple, l'utilisation d'un modèle bayésien (MCMCglmm ; Hadfield 2010) ou d'un modèle de maximum de vraisemblance restreint (REML ; Gilmour *et al.* 2006) pour étudier des traits gaussiens produit des résultats équivalents (Villemereuil *et al.*, 2013).

6.4 Vers une approche combinant la génétique quantitative et génomique des populations : comprendre le déterminisme génétique et l'évolution des traits écologiquement importants

Au cours des dernières décennies, le développement de marqueurs génétiques polymorphes et de méthodes permettant d'utiliser ces données moléculaires a joué un rôle majeur dans le domaine de l'écologie et de la biologie évolutive. En effet, comme nous avons pu le voir dans les différents chapitres de cette thèse, un faible nombre de marqueurs moléculaires très polymorphes, tels que les microsatellites, permettent de caractériser le régime de reproduction et de retracer les apparentements entre individus par des méthodes de reconstruction ou d'assignation de parenté (Blouin, 2003). Plus récemment, l'amélioration des méthodes de séquençage à haut débit a permis le marquage dense du génome sur des zones de polymorphisme nucléotidique, ou SNP², même chez des espèces non-modèles. Les approches de génomique des populations qui utilisent ces données promettent ainsi de pouvoir identifier le polymorphisme génétique à l'origine de la variation des traits reliés à la valeur sélective et les gènes impliqués dans l'adaptation locale en général.

Sur des gradient où les populations présentent peu de variation quantitative aux traits, comme c'est le cas sur le versant nord du Mont-Ventoux, l'adaptation locale implique probablement de faibles variations des fréquences alléliques aux gènes codants, qui rendent les

2. Single Nucleotide Polymorphism

méthodes classiques de détection de SNPs sous sélection peu efficaces (Le Corre & Kremer, 2012). Durant sa thèse, Hadrien Lalagüe a eu l'occasion de développer des marqueurs génomiques chez le hêtre et d'utiliser des approches de détection de trace de sélection sur SNPs sur les versants nord et sud du Mont-Ventoux (Lalagüe, 2013). Il a notamment permis d'identifier 3 SNPs outliers ainsi que de nombreuses paires de gènes sélectionnés conjointement, traces de sélection épistatique, entre les populations de hêtre sur le versant nord du Mont-Ventoux. Parmi ces gènes candidats potentiellement impliqués dans la réponse aux variations environnementales sur le gradient, certains seraient a priori impliqués dans la réponse aux stress abiotiques et dans la phénologie. Cependant, chez le hêtre, espèce non-modèle, les fonctions des gènes restent largement inconnues ce qui limite l'interprétation conjointe des approches phénotypiques (réalisées durant ma thèse) et génomiques (Lalagüe, 2013). De manière consistante, les approches de génomique et de génétique quantitative parviennent à détecter des signaux de sélection divergente sur le gradient altitudinal de faible taille étudié.

Bien que les approches utilisant les marqueurs génomiques pour étudier la micro évolution en soient encore à leurs débuts, spécialement pour les espèces non-modèles, les outils et résultats de la génomique des populations ont déjà considérablement révolutionné la génétique quantitative. En effet, les marqueurs génomiques permettent aujourd'hui de pouvoir estimer de manière fiable des coefficients d'apparentement (ou identité par état) entre individus. Jusqu'à présent, l'estimation des apparentements sur marqueurs était très peu utilisée par manque de performance quand trop peu de marqueurs étaient utilisés (Gay *et al.*, 2013). Les marqueurs SNP offrent aujourd'hui l'opportunité d'aller étudier la base génétique des traits quantitatifs sans avoir de connaissance sur le pedigree et de pouvoir partitionner la variance génétique en composantes additives et non-additives (Visscher *et al.*, 2006), ce qui constitue une avancée majeure sur l'application des méthodes de génétique quantitative aux populations naturelles. De plus, la génomique a apporté à la génétique quantitative une nouvelle compréhension du déterminisme génétique des traits. En effet, les approches de génomique ont déjà démontré l'influence des nombreux effets épistatiques, épigénétiques ou encore d'un nombre limité de gènes sur le déterminisme des traits chez de nombreuses espèces, remettant en cause le modèle infinitésimal utilisé en génétique quantitative et les hypothèses simplificatrices souvent faites sur le déterminisme non-additif des traits (Nelson *et al.*, 2013).

Dans ce contexte, on peut se demander quel est le devenir de la génétique quantitative, une

discipline qui a été théorisée il y a près d'un siècle maintenant. Il ne faut pas oublier qu'avec des modèles simples, la génétique quantitative a permis et permet toujours d'apporter des réponses sur les grands mécanismes qui régissent l'évolution des traits fonctionnels. Utiliser une approche de génomique déconnectée des traits, c'est également considérer tout polymorphisme génétique non-neutre comme adaptatif. Or, les approches multi-traits en génétique quantitative nous ont par exemple permis de savoir qu'un trait non directement affecté par la sélection pouvait évoluer par simple corrélation avec un trait soumis à sélection. Plus globalement, la génétique quantitative et génomique doivent évoluer conjointement si elles veulent permettre de comprendre comment le génotype affecte le phénotype. C'est la rencontre de ces deux disciplines qui permettra, d'une part, aux modèles de génétique quantitative de ne pas se retrouver dépassés par la connaissance actuelle sur le déterminisme génétique des traits et, d'autre part, à la génomique des populations de ne pas tomber dans le travers d'avoir une vision de l'action de la sélection complètement déconnectée de la réponse des traits.

BIBLIOGRAPHIE

- Agrawal, A.F. & Stinchcombe, J.R. (2009) How much do genetic covariances alter the rate of adaptation? *PROCEEDINGS OF THE ROYAL SOCIETY B-BIOLOGICAL SCIENCES*, **276**, 1183–1191.
- Aitken, S.N. & Whitlock, M.C. (2013) Assisted Gene Flow to Facilitate Local Adaptation to Climate Change. *ANNUAL REVIEW OF ECOLOGY, EVOLUTION, AND SYSTEMATICS*, **44**.
- Aitken, S.N., Yeaman, S., Holliday, J.A., Wang, T. & Curtis-McLane, S. (2008) Adaptation, migration or extirpation : climate change outcomes for tree populations. *EVOLUTIONARY APPLICATIONS*, **1**, 95–111.
- Alberto, F.J., Aitken, S.N., Alia, R., Gonzalez-Martinez, S.C., Hanninen, H., Kremer, A., Lefevre, F., Lenormand, T., Yeaman, S., Whetten, R. & Savolainen, O. (2013) Potential for evolutionary responses to climate change evidence from tree populations. *GLOBAL CHANGE BIOLOGY*, **19**, 1645–1661.
- Alleaume-Benharira, M., Pen, I. & Ronce, O. (2006) Geographical patterns of adaptation within a species' range : interactions between drift and gene flow. *JOURNAL OF EVOLUTIONARY BIOLOGY*, **19**, 203–215.
- Arntz, A. & Delph, L. (2001) Pattern and process : evidence for the evolution of photosynthetic traits in natural populations. *OECOLOGIA*, **127**, 455–467.
- Ashley, M.V. (2010) Plant parentage, pollination, and dispersal : How dna microsatellites have altered the landscape. *CRITICAL REVIEWS IN PLANT SCIENCES*, **29**, 148–161.

- Asuka, Y., Tani, N., Tsumura, Y. & Tomaru, N. (2004) Development and characterization of microsatellite markers for *Fagus crenata* Blume. *MOLECULAR ECOLOGY NOTES*, **4**, 101–103.
- Atkin, F.C., Dieters, M.J. & Stringer, J.K. (2009) Impact of depth of pedigree and inclusion of historical data on the estimation of additive variance and breeding values in a sugarcane breeding program. *THEORETICAL AND APPLIED GENETICS*, **119**, 555–565.
- Barton, N. (2001) *Integrating Ecology and Evolution in a Spatial Context : Adaptation at the edge of a species range*. pp. 365-392. Blackwell Science, New York.
- Benito Garzon, M., Alia, R., Robson, T.M. & Zavala, M.A. (2011) Intra-specific variability and plasticity influence potential tree species distributions under climate change. *GLOBAL ECOLOGY AND BIOGEOGRAPHY*, **20**, 766–778.
- Berteaux, D., Reale, D., McAdam, A. & Boutin, S. (2004) Keeping pace with fast climate change : Can arctic life count on evolution? *INTEGRATIVE AND COMPARATIVE BIOLOGY*, **44**, 140–151.
- Bertin, R.I. (2008) Plant phenology and distribution in relation to recent climate change. *JOURNAL OF THE TORREY BOTANICAL SOCIETY*, **135**, 126–146.
- Blouin, M. (2003) DNA-based methods for pedigree reconstruction and kinship analysis in natural populations. *TRENDS IN ECOLOGY & EVOLUTION*, **18**, 503–511.
- Blows, M. & Hoffmann, A. (2005) A reassessment of genetic limits to evolutionary change. *ECOLOGY*, **86**, 1371–1384.
- Bontemps, A. (2012) *Potentiel évolutif d'une population de hêtre commun sur le Mont Ventoux*. Ph.D. thesis, University Paul Cézanne, Aix-Marseille III, Aix en Provence (Ecole Doctorale des Sciences de l'Environnement).
- Bontemps, A., Klein, E.K. & Oddou-Muratorio, S. (2013) Shift of spatial patterns during early recruitment in *Fagus sylvatica* : Evidence from seed dispersal estimates based on genotypic data. *FOREST ECOLOGY AND MANAGEMENT*, **305**, 67–76.
- Bossdorf, O., Richards, C.L. & Pigliucci, M. (2008) Epigenetics for ecologists. *ECOLOGY LETTERS*, **11**, 106–115.

-
- Brendel, O. (2014) Is the coefficient of variation a valid measure for variability of stable isotope abundances in biological materials? *RAPID COMMUNICATIONS IN MASS SPECTROMETRY*, **28**, 370–376.
- Bresson, C.C., Vitasse, Y., Kremer, A. & Delzon, S. (2011) To what extent is altitudinal variation of functional traits driven by genetic adaptation in European oak and beech? *TREE PHYSIOLOGY*, **31**, 1164–1174.
- Bridle, J.R. & Vines, T.H. (2007) Limits to evolution at range margins : when and why does adaptation fail? *TRENDS IN ECOLOGY & EVOLUTION*, **22**, 140–147.
- Brousseau, L., Bonal, D., Cigna, J. & Scotti, I. (2013) Highly local environmental variability promotes intrapopulation divergence of quantitative traits : an example from tropical rain forest trees. *ANNALS OF BOTANY*, **112**, 1169–1179.
- Bulmer, M. (1971) The effect of selection on genetic variability. *AMERICAN SOCIETY OF NATURALISTS*, **105**, 201–211.
- Bush, D., Kain, D., Matheson, C. & Kanowski, P. (2011) Marker-based adjustment of the additive relationship matrix for estimation of genetic parameters-an example using eucalyptus cladocalyx. *TREE GENETICS AND GENOME*, **7**, 23–35.
- Cailleret, M. & Davi, H. (2011) Effects of climate on diameter growth of co-occurring *Fagus sylvatica* and *Abies alba* along an altitudinal gradient. *TREES*, **25**, 265–276.
- Charmantier, A. & Blondel, J. (2003) A contrast in extra-pair paternity levels on mainland and island populations of mediterranean blue tits. *ETHOLOGY*, **109**, 351–363.
- Charmantier, A. & Garant, D. (2005) Environmental quality and evolutionary potential : lessons from wild populations. *PROCEEDINGS - BIOLOGICAL SCIENCES*, **272**, 1415–1425.
- Charmantier, A., McCleery, R.H., Cole, L.R., Perrins, C., Kruuk, L.E.B. & Sheldon, B.C. (2008) Adaptive phenotypic plasticity in response to climate change in a wild bird population. *SCIENCE*, **320**, 800–803.
- Charmantier, A. & Réale, D. (2005) How do misassigned paternities affect the estimation of heritability in the wild? *MOLECULAR ECOLOGY*, **14**, 2839–2850.

- Chen, I.C., Hill, J.K., Ohlemueller, R., Roy, D.B. & Thomas, C.D. (2011) Rapid Range Shifts of Species Associated with High Levels of Climate Warming. *SCIENCE*, **333**, 1024–1026.
- Chevin, L.M., Collins, S. & Lefevre, F. (2013) Phenotypic plasticity and evolutionary demographic responses to climate change : taking theory out to the field. *FUNCTIONAL ECOLOGY*, **27**, 966–979.
- Chuine, I. & Beaubien, E. (2001) Phenology is a major determinant of tree species range. *ECOLOGY LETTERS*, **4**, 500–510.
- Chybicki, I.J. & Burczyk, J. (2013) Seeing the forest through the trees : comprehensive inference on individual mating patterns in a mixed stand of *Quercus robur* and *Q-petraea*. *ANNALS OF BOTANY*, **112**, 561–574.
- Conner, J.K. (2001) How strong is natural selection? *TRENDS IN ECOLOGY AND EVOLUTION*, **16**, 215–217.
- Conner, J.K. (2012) Quantitative Genetic Approaches To Evolutionary Constraint : How Useful? *EVOLUTION*, **66**, 3313–3320.
- Conover, D.O., Duffy, T.A. & Hice, L.A. (2009) The covariance between genetic and environmental influences across ecological gradients. *THE YEAR IN EVOLUTIONARY BIOLOGY*, **1168**, 100–129.
- Conover, D. & Schultz, E. (1995) Phenotypic similarity and the evolutionary significance of countergradient variation. *TRENDS IN ECOLOGY & EVOLUTION*, **10**, 248–252.
- Corlett, R.T. & Westcott, D.A. (2013) Will plant movements keep up with climate change? *TRENDS IN ECOLOGY & EVOLUTION*, **28**, 482–488.
- Costa e Silva, J., Potts, B.M. & Lopez, G.A. (2014) Heterosis May Result in Selection Favouring the Products of Long-Distance Pollen Dispersal in *Eucalyptus*. *PLOS ONE*, **9**.
- Csillery, K., Johnson, T., Beraldi, D., Clutton-Brock, T., Coltman, D., Hansson, B., Spong, G. & Pemberton, J.M. (2006) Performance of marker-based relatedness estimators in natural populations of outbred vertebrates. *GENETICS*, **173**, 2091–2101.
- Davi, H. and Gillmann, M., Ibanez, T., Cailleret, M., Bontemps, A., Fady, B. & Lefèvre, F. (2011) Diversity of leaf unfolding dynamics among tree species : New insights from a study

- along an altitudinal gradient. *AGRICULTURAL AND FOREST METEOROLOGY*, **151**, 1504–1513.
- Davis, M. & Shaw, R. (2001) Range shifts and adaptive responses to Quaternary climate change. *SCIENCE*, **292**, 673–679.
- de Lucas, A.I., Robledo-Arnuncio, J.J., Hidalgo, E. & González-Martínez, S.C. (2008) Mating system and pollen gene flow in mediterranean maritime pine. *HEREDITY*, **100**, 390–399.
- DiBattista, J.D., Feldheim, K.A., Garant, D., Gruber, S.H. & Hendry, A.P. (2009) Evolutionary potential of a large marine vertebrate : quantitative genetic parameters in a wild population. *EVOLUTION*, **63**, 1051–1067.
- Donovan, L.A., Ludwig, F., Rosenthal, D.M., Rieseberg, L.H. & Dudley, S.A. (2009) Phenotypic selection on leaf ecophysiological traits in *Helianthus*. *NEW PHYTOLOGIST*, **183**, 868–879.
- Dudley, S. (1996) Differing selection on plant physiological traits in response to environmental water availability : A test of adaptive hypotheses. *EVOLUTION*, **50**, 92–102.
- Duputié, A., Massol, F., Chuine, I., Kirkpatrick, M. & Ronce, O. (2012) How do genetic correlations affect species range shifts in a changing environment? *ECOLOGY LETTERS*, **15**, 251–259.
- El-Kassaby, Y.A., Cappa, E.P., Liewlaksaneeyanawin, C., Klapste, J. & Lstiburek, M. (2011) Breeding without breeding : is a complete pedigree necessary for efficient breeding? *PLoS One*, **6**, e25737.
- ElKassaby, Y. & Jaquish, B. (1996) Population density and mating pattern in western larch. *JOURNAL OF HEREDITY*, **87**, 438–443.
- Etterson, J.R. & Shaw, R.G. (2001) Constraint to adaptive evolution in response to global warming. *SCIENCE*, **294**, 151–154.
- Etterson, J. (2004) Evolutionary potential of *Chamaecrista fasciculata* in relation to climate change. 1. Clinal patterns of selection along an environmental gradient in the great plains. *EVOLUTION*, **58**, 1446–1458.
- Falconer, D. & MacKay, T. (1996) *Introduction to Quantitative Genetics*. Ed 4. Longmans Green, Harlow, Essex, UK.

- FAO (2013) State of mediterranean forests. Technical report, <http://www.fao.org>.
- Farquhar, G., Ehleringer, J. & Hubick, K. (1989) Carbon Isotope Discrimination And Photosynthesis. *ANNUAL REVIEW OF PLANT PHYSIOLOGY AND PLANT MOLECULAR BIOLOGY*, **40**, 503–537.
- Fernandez, J. & Toro, M. (2006) A new method to estimate relatedness from molecular markers. *MOLECULAR ECOLOGY*, **15**, 1657–1667.
- Frankino, W., Zwaan, B., Stern, D. & Brakefield, P. (2005) Natural selection and developmental constraints in the evolution of allometries. *SCIENCE*, **307**, 718–720.
- Franks, S.J., Sim, S. & Weis, A.E. (2007) Rapid evolution of flowering time by an annual plant in response to a climate fluctuation. *PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA*, **104**, 1278–1282.
- Frentiu, F.D., Clegg, S.M., Chittock, J., Burke, T., Blows, M.W. & Owens, I.P.F. (2008) Pedigree-free animal models : the relatedness matrix reloaded. *PROCEEDINGS OF THE ROYAL SOCIETY B-BIOLOGICAL SCIENCES*, **275**, 639–647.
- Gaggiotti, O.E. & Foll, M. (2010) Quantifying population structure using the F-model. *MOLECULAR ECOLOGY RESOURCES*, **10**, 821–830.
- Garant, D. & Kruuk, L. (2005) How to use molecular marker data to measure evolutionary parameters in wild populations. *MOLECULAR ECOLOGY*, **14**, 1843–1859.
- Garant, D., Forde, S.E. & Hendry, A.P. (2007) The multifarious effects of dispersal and gene flow on contemporary adaptation. *FUNCTIONAL ECOLOGY*, **21**, 434–443.
- Gaspar, M., Lucas, A.I., de Alia, R., Paiva, J., Hidalgo, E., Louzada, J., Almeida, H. & Gonzalez-Martinez, S. (2009) Use of molecular markers for estimating breeding parameters : a case study in a pinus pinaster ait. progeny trial. *TREE GENETICS AND GENOME*, **5**, 609–616.
- Gauzere, J., Klein, E.K. & Oddou-Muratorio, S. (2013a) Ecological determinants of mating system within and between three fagus sylvatica populations along an elevational gradient. *MOLECULAR ECOLOGY*, **22**, 5001–5015.

-
- Gauzere, J., Oddou-Muratorio, S., Pichot, C., Lefèvre, F. & Klein, E. (2013b) Biases in quantitative genetic analyses using open-pollinated progeny tests from natural tree populations. *ACTA BOTANICA GALLICA*, **160** (3-4), 227–238.
- Gay, L., Siol, M. & Ronfort, J. (2013) Pedigree-free estimates of heritability in the wild : Promising prospects for selfing populations. *PLOS ONE*, **8**.
- Geber, M. & Dawson, T. (1990) Genetic-Variation In And Covariation Between Leaf Gas-Exchange, Morphology, And Development In Polygonum-Arenastrum, An Annual Plant. *OECOLOGIA*, **85**, 153–158.
- Geber, M. & Griffen, L. (2003) Inheritance and natural selection on functional traits. *INTERNATIONAL JOURNAL OF PLANT SCIENCES*, **164**, S21–S42.
- Gerhart, L.M. & Ward, J.K. (2010) Plant responses to low [CO₂] of the past. *NEW PHYTOLOGIST*, **188**, 674–695.
- Ghalambor, C.K., McKay, J.K., Carroll, S.P. & Reznick, D.N. (2007) Adaptive versus non-adaptive phenotypic plasticity and the potential for contemporary adaptation in new environments. *FUNCTIONAL ECOLOGY*, **21**, 394–407.
- Gilmour, A., Gogel, B., Cullis, B. & Thompson, R. (2006) *Asreml User Guide Release 2.0*. Hemel Hempsted : VSN International Ltd.
- Gonzalez, A., Ronce, O., Ferriere, R. & Hochberg, M.E. (2013) Evolutionary rescue : an emerging focus at the intersection between ecology and evolution. *PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B-BIOLOGICAL SCIENCES*, **368**.
- Gonzalo-Turpin, H. & Hazard, L. (2009) Local adaptation occurs along altitudinal gradient despite the existence of gene flow in the alpine plant species *Festuca eskia*. *JOURNAL OF ECOLOGY*, **97**, 742–751.
- Hadfield, J.D., Richardson, D.S. & Burke, T. (2006) Towards unbiased parentage assignment : combining genetic, behavioural and spatial data in a Bayesian framework. *MOLECULAR ECOLOGY*, **15**, 3715–3730.
- Hadfield, J.D. (2010) MCMC Methods for Multi-Response Generalized Linear Mixed Models : The MCMCglmm R Package. *JOURNAL OF STATISTICAL SOFTWARE*, **33**, 1–22.

- Hampe, A. & Petit, R. (2005) Conserving biodiversity under climate change : the rear edge matters. *ECOLOGY LETTERS*, **8**, 461–467.
- Hamrick, J.L., Godt, M.J.W. & Sherman-Broyles, S.L. (1992) Factors influencing levels of genetic diversity in woody plant species. *NEW FORESTS*, **6**, 95–124.
- Hamrick, J.L. & Nason, J.D. (1986) *Dynamics in Ecological Space and Time Chapter 6 : Consequence of dispersal in plants*.
- Hansen, M.M., Olivieri, I., Waller, D.M., Nielsen, E.E. & Grp, G.W. (2012) Monitoring adaptive genetic responses to environmental change. *MOLECULAR ECOLOGY*, **21**, 1311–1329.
- Hansen, O. & Nielsen, U. (2010) Microsatellites used to establish full pedigree in a half-sib trial and correlation between number of male strobili and paternal success. *ANNALS OF FOREST SCIENCE*, **67**, 703.
- Hansen, T.F., Pélabon, C. & Houle, D. (2011) Heritability is not evolvability. *EVOLUTIONARY BIOLOGY*, pp. DOI 10.1007/s11692-011-9127-6.
- Hardy, O.J., González-Martínez, S.C., Fréville, H., Boquien, G., Mignot, A., Colas, B. & Olivieri, I. (2004) Fine-scale genetic structure and gene dispersal in *centaurea corymbosa* (asteraceae) i. pattern of pollen dispersal. *JOURNAL OF EVOLUTIONARY BIOLOGY*, **17**, 795–806.
- Hardy, O. & Vekemans, X. (2002) SPAGEDi : a versatile computer program to analyse spatial genetic structure at the individual or population levels. *MOLECULAR ECOLOGY NOTES*, **2**, 618–620.
- Harrison, H.B., Saenz-Agudelo, P., Planes, S., Jones, G.P. & Berumen, M.L. (2013) Relative accuracy of three common methods of parentage analysis in natural populations. *MOLECULAR ECOLOGY*, **22**, 1158–1170.
- Harte, J., Ostling, A., Green, J. & Kinzig, A. (2004) Biodiversity conservation - Climate change and extinction risk. *NATURE*, **430**.
- Hendry, A. & Kinnison, M. (1999) Perspective : The pace of modern life : Measuring rates of contemporary microevolution. *EVOLUTION*, **53**, 1637–1653.

-
- Hendry, A. & Taylor, E. (2004) How much of the variation in adaptive divergence can be explained by gene flow? - An evaluation using lake-stream stickleback pairs. *EVOLUTION*, **58**, 2319–2331.
- Hereford, J., Hansen, T. & Houle, D. (2004) Comparing strengths of directional selection : How strong is strong? *EVOLUTION*, **58**, 2133–2143.
- Heschel, M. & Riginos, C. (2005) Mechanisms of selection for drought stress tolerance and avoidance in *Impatiens capensis* (Balsaminacea). *AMERICAN JOURNAL OF BOTANY*, **92**, 37–44.
- Hine, E. & Blows, M. (2006) Determining the effective dimensionality of the genetic variance-covariance matrix. *GENETICS*, **173**, 1135–1144.
- Houle, D. (1992) Comparing Evolvability And Variability Of Quantitative Traits. *GENETICS*, **130**, 195–204.
- Howe, G., Aitken, S., Neale, D., Jermstad, K., Wheeler, N. & Chen, T. (2003) From genotype to phenotype : unraveling the complexities of cold adaptation in forest trees. *CANADIAN JOURNAL OF BOTANY-REVUE CANADIENNE DE BOTANIQUE*, **81**, 1247–1266.
- IGN (2013) Inventaire forestier : La forêt en chiffres et en cartes.
- Ingvarsson, P. & Whitlock, M. (2000) Heterosis increases the effective migration rate. *PROCEEDINGS OF THE ROYAL SOCIETY B-BIOLOGICAL SCIENCES*, **267**, 1321–1326.
- IPCC (2013) *Climate Change 2013 : The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, UK, and New York, NY, USA.
- Johnson, T. & Barton, N. (2005) Theoretical models of selection and mutation on quantitative traits. *PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B-BIOLOGICAL SCIENCES*, **360**, 1411–1425.
- Jump, A.S., Matyas, C. & Penuelas, J. (2009) The altitude-for-latitude disparity in the range retractions of woody species. *TRENDS IN ECOLOGY & EVOLUTION*, **24**, 694–701.
- Jump, A. & Penuelas, J. (2005) Running to stand still : adaptation and the response of plants to rapid climate change. *ECOLOGY LETTERS*, **8**, 1010–1020.

- Karhunen, M., Merila, J., Leinonen, T., Cano, J.M. & Ovaskainen, O. (2013) driftsel : an R package for detecting signals of natural selection in quantitative traits. *MOLECULAR ECOLOGY RESOURCES*, **13**, 746–754.
- Karhunen, M., Ovaskainen, O., Herczeg, G. & Merila, J. (2014) Bringing Habitat Information Into Statistical Tests Of Local Adaptation In Quantitative Traits : A Case Study Of Nine-Spined SticklebackS. *EVOLUTION*, **68**, 559–568.
- Karhunen, M. & Ovaskainen, O. (2012) Estimating Population-Level Coancestry Coefficients by an Admixture F Model. *GENETICS*, **192**, 609.
- Kawecki, T. & Ebert, D. (2004) Conceptual issues in local adaptation. *ECOLOGY LETTERS*, **7**, 1225–1241.
- Keller, I., Alexander, J.M., Holderegger, R. & Edwards, P.J. (2013) Widespread phenotypic and genetic divergence along altitudinal gradients in animals. *JOURNAL OF EVOLUTIONARY BIOLOGY*, **26**, 2527–2543.
- Kingsolver, J.G., Hoekstra, H.E., Hoekstra, J.M., Berrigan, D., Vignieri, S.N., Hill, C.E., Hoang, A., Gibert, P. & Beerli, P. (2001) The strength of phenotypic selection in natural populations. *AMERICAN NATURALIST*, **157**, 245–261.
- Kinnison, M.T. & Hairston, Jr., N.G. (2007) Eco-evolutionary conservation biology : contemporary evolution and the dynamics of persistence. *FUNCTIONAL ECOLOGY*, **21**, 444–454.
- Kirkpatrick, M. & Barton, N. (1997) Evolution of a species range. *AMERICAN NATURALIST*, **150**, 1–23.
- Kirkpatrick, M. (2009) Patterns of quantitative genetic variation in multiple dimensions. *GENETICA*, **136**, 271–284.
- Kitano, J., Bolnick, D.I., Beauchamp, D.A., Mazur, M.M., Mori, S., Nakano, T. & Peichel, C.L. (2008) Reverse evolution of armor plates in the threespine stickleback. *CURRENT BIOLOGY*, **18**, 769–774.
- Klein, E.K., Desassis, N. & Oddou-Muratorio, S. (2008) Pollen flow in the wildservice 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., Carpentier, F.H. & Oddou-Muratorio, S. (2011) Estimating the variance of male fecundity from genotypes of progeny arrays : evaluation of the bayesian forward approach. *METHODS IN ECOLOGY & EVOLUTION*, **2**, 349–361.
- Korner, C. (2007) The use of ‘altitude’ in ecological research. *TRENDS IN ECOLOGY & EVOLUTION*, **22**, 569–574.
- Kozłowski, J. (1992) Optimal Allocation Of Resources To Growth And Reproduction - Implications For Age And Size At Maturity. *TRENDS IN ECOLOGY & EVOLUTION*, **7**, 15–19.
- Kramer, K. (2004) Effects of silvicultural regimes on dynamics of genetic and ecological diversity of european beech forests. *Final report of the 5th framework project DynaBeech, QLK5-CT-1999-1210*, pp. 1–269.
- Kremer, A., Ronce, O., Robledo-Arnuncio, J.J., Guillaume, F., Bohrer, G., Nathan, R., Bridle, J.R., Gomulkiewicz, R., Klein, E.K., Ritland, K., Kuparinen, A., Gerber, S. & Schueler, S. (2012) Long-distance gene flow and adaptation of forest trees to rapid climate change. *ECOLOGY LETTERS*, **15**, 378–392.
- Kruuk, L.E.B. (2004) Estimating genetic parameters in natural populations using the "animal model". *PROCEEDINGS OF THE ROYAL SOCIETY SERIES B-BIOLOGICAL SCIENCES*, **359**, 873–890.
- Kumar, S. & Richardson, T. (2005) Inferring relatedness and heritability using molecular markers in radiata pine. *MOLECULAR BREEDING*, **15**, 55–64.
- Kuparinen, A., Savolainen, O. & Schurr, F.M. (2010) Increased mortality can promote evolutionary adaptation of forest trees to climate change. *FOREST ECOLOGY AND MANAGEMENT*, **259**, 1003–1008.
- Lalague, H. (2013) *Genetic response of tree population to spatial climatic variation : an experimental genomic and simulation approach in Fagus sylvatica populations along altitudinal gradients*. Ph.D. thesis, Ecole doctorale SIBAGHE, Université Montpellier 2.
- Lande, R. (1992) Neutral theory of quantitative genetic variance in an island model with local extinction and colonization. *EVOLUTION*, **46**, 381–389.

- Lande, R. & Arnold, S. (1983) The Measurement Of Selection On Correlated Characters. *EVOLUTION*, **37**, 1210–1226.
- Lander, T.A., Oddou-Muratorio, S., Prouillet-Leplat, H. & Klein, E.K. (2011) Reconstruction of a beech population bottleneck using archival demographic information and bayesian analysis of genetic data. *MOLECULAR ECOLOGY*, **20**, 5182–5196.
- Le Corre, V. & Kremer, A. (2012) The genetic differentiation at quantitative trait loci under local adaptation. *MOLECULAR ECOLOGY*, **21**, 1548–1566.
- Lefèvre, F., Boivin, T., Bontemps, A., Courbet, F., Davi, H., Durand-Gillmann, M., Fdy, B., gauzere, J., Gidoin, C., Karam, M., Lalague, H., Oddou-Muratorio, S. & Pichot, C. (2013) Considering evolutionary processes in adaptive forestry. *ANNALS OF FOREST SCIENCE*, **71** (7), 723–739.
- Lefevre, S., Wagner, S., Petit, R.J. & de Lafontaine, G. (2012) Multiplexed microsatellite markers for genetic studies of beech. *MOLECULAR ECOLOGY RESOURCES*, **12**, 484–491.
- Legarra, A., Aguilar, I. & Misztal, I. (2009) A relationship matrix including full pedigree and genomic information. *JOURNAL OF DAIRY SCIENCE*, **92**, 4656–4663.
- Leimu, R. & Fischer, M. (2008) A Meta-Analysis of Local Adaptation in Plants. *PLOS ONE*, **3**.
- Leinonen, T., O'hara, R.B., Cano, J.M. & Merila, J. (2008) Comparative studies of quantitative trait and neutral marker divergence : a meta-analysis. *JOURNAL OF EVOLUTIONARY BIOLOGY*, **21**, 1–17.
- Lenormand, T. (2002) Gene flow and the limits to natural selection. *TRENDS IN ECOLOGY & EVOLUTION*, **17**, 183–189.
- Lindner, M., Maroschek, M., Netherer, S., Kremer, A., Barbati, A., Garcia-Gonzalo, J., Seidl, R., Delzon, S., Corona, P., Kolstrom, M., Lexer, M.J. & Marchetti, M. (2010) Climate change impacts, adaptive capacity, and vulnerability of European forest ecosystems. *FOREST ECOLOGY AND MANAGEMENT*, **259**, 698–709.
- Linhart, Y. & Grant, M. (1996) Evolutionary significance of local genetic differentiation in plants. *ANNUAL REVIEW OF ECOLOGY AND SYSTEMATICS*, **27**, 237–277.

-
- Loarie, S.R., Duffy, P.B., Hamilton, H., Asner, G.P., Field, C.B. & Ackerly, D.D. (2009) The velocity of climate change. *NATURE*, **462**, 1052–U111.
- Loiselle, B., Sork, V.L., Nason, J.D. & Graham, C. (1995) Spatial genetic structure of a tropical understory shrub, *Psychotria officinalis* (Rubiaceae). *AMERICAN JOURNAL OF BOTANY*, **82**, 1420–1425.
- Lopez, S., Rousset, F., Shaw, F.H., Shaw, R.G. & Ronce, O. (2008) Migration load in plants : role of pollen and seed dispersal in heterogeneous landscapes. *JOURNAL OF EVOLUTIONARY BIOLOGY*, **21**, 294–309.
- Lynch, L. & Walsh, B. (1998) *Genetics and analysis of Quantitative Traits*.
- Marshall, T., Slate, J., Kruuk, L. & Pemberton, J. (1998) Statistical confidence for likelihood-based paternity inference in natural populations. *MOLECULAR ECOLOGY*, **7**, 639–655.
- McDowell, N., Pockman, W.T., Allen, C.D., Breshears, D.D., Cobb, N., Kolb, T., Plaut, J., Sperry, J., West, A., Williams, D.G. & Yepez, E.A. (2008) Mechanisms of plant survival and mortality during drought : why do some plants survive while others succumb to drought ? *NEW PHYTOLOGIST*, **178**, 719–739.
- McGuigan, K. (2006) Studying phenotypic evolution using multivariate quantitative genetics. *MOLECULAR ECOLOGY*, **15**, 883–896.
- McKay, J.K. & Latta, R. (2002) Adaptive population divergence : markers, qtl and traits. *TRENDS in ECOLOGY and EVOLUTION*, **17**, 285–291.
- Merila, J. & Crnokrak, P. (2001) Comparison of genetic differentiation at marker loci and quantitative traits. *JOURNAL OF EVOLUTIONARY BIOLOGY*, **14**, 892–903.
- Merila, J. & Sheldon, B. (1999) Genetic architecture of fitness and nonfitness traits : empirical patterns and development of ideas. *HEREDITY*, **83**, 103–109.
- Merila, J., Sheldon, B. & Kruuk, L. (2001) Explaining stasis : microevolutionary studies in natural populations. *GENETICA*, **112**, 199–222.
- Merzeau, D., Comps, B., Thiébaud, B. & Letouzey, J. (1994) Estimation of *Fagus sylvatica* mating system parameters in natural populations. *ANNALS OF FOREST SCIENCE*, **51**, 163–173.

- Michelot, A., Breda, N., Damesin, C. & Dufrene, E. (2012) Differing growth responses to climatic variations and soil water deficits of *Fagus sylvatica*, *Quercus petraea* and *Pinus sylvestris* in a temperate forest. *FOREST ECOLOGY AND MANAGEMENT*, **265**, 161–171.
- Milleron, M., Lopez de Heredia, U., Lorenzo, Z., Perea, R., Dounavi, A., Alonso, J., Gil, L. & Nanos, N. (2012) Effect of canopy closure on pollen dispersal in a wind-pollinated species (*fagus sylvatica* l.). *PLANT ECOLOGY*, **213**, 1715–1728.
- Morrissey, M.B., Kruuk, L.E.B. & Wilson, A.J. (2010) The danger of applying the breeder's equation in observational studies of natural populations. *JOURNAL OF EVOLUTIONARY BIOLOGY*, **23**, 2277–2288.
- Nathan, R., Horvitz, N., He, Y., Kuparinen, A., Schurr, F.M. & Katul, G.G. (2011) Spread of North American wind-dispersed trees in future environments. *ECOLOGY LETTERS*, **14**, 211–219.
- Nelson, R.M., Pettersson, M.E. & Carlborg, O. (2013) A century after Fisher : time for a new paradigm in quantitative genetics. *TRENDS IN GENETICS*, **29**, 669–676.
- Nilsson, J. (1995) Genetic-Variation In The Natural Pollen Cloud Of *Pinus-Sylvestris* - A Study Based On Progeny Testing. *SCANDINAVIAN JOURNAL OF FOREST RESEARCH*, **10**, 140–148.
- Nourtier, M., Chanzy, A., Cailleret, M., Yingge, X., Huc, R. & Davi, H. (2013) Transpiration of silver fir (*abies alba* mill.) during and after drought in relation to soil properties in a mediterranean mountain area. *ANNALS OF FOREST SCIENCE*, pp. DOI 10.1007/s13595-012-0229-9.
- Oddou-Muratorio, S., Klein, E. & Austerlitz, F. (2005) Pollen flow in the wildservice 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. & Davi, H. (2014) Simulating local adaptation to climate of forest trees with a Physio-Demo-Genetics model. *EVOLUTIONARY APPLICATIONS*, **7**, 453–467.
- O'Hara, R.B., Cano, J.M., Ovaskainen, O., Teplitsky, C. & Alho, J.S. (2008) Bayesian approaches in evolutionary quantitative genetics. *JOURNAL OF EVOLUTIONARY BIOLOGY*, **21**, 949–957.

-
- O'Hara, R. & Merila, J. (2005) Bias and precision in Q(ST) estimates : Problems and some solutions. *GENETICS*, **171**, 1331–1339.
- Ovaskainen, O., Karhunen, M., Zheng, C., Arias, J.M.C. & Merila, J. (2011) A New Method to Uncover Signatures of Divergent and Stabilizing Selection in Quantitative Traits. *GENETICS*, **189**, 621–U729.
- Parmesan, C. (2006) Ecological and evolutionary responses to recent climate change. *ANNUAL REVIEW OF ECOLOGY EVOLUTION AND SYSTEMATICS*, **37**, 637–669.
- Pastorelli, R., Smulders, M., Van't Westende, W., Vosman, B., Giannini, R., Vettori, C. & Vendramin, G. (2003) Characterization of microsatellite markers in *Fagus sylvatica* L. and *Fagus orientalis* Lipsky. *MOLECULAR ECOLOGY NOTES*, **3**, 76–78.
- Pease, C., Lande, R. & Bull, J. (1989) A Model Of Population-Growth, Dispersal And Evolution In A Changing Environment. *ECOLOGY*, **70**, 1657–1664.
- Pemberton, J., Coltman, D., Smith, J. & Pilkington, J. (1999) Molecular analysis of a promiscuous, fluctuating mating system. *BIOLOGICAL JOURNAL OF THE LINNEAN SOCIETY*, **68**, 289–301.
- Penuelas, J. & Boada, M. (2003) A global change-induced biome shift in the Montseny mountains (NE Spain). *GLOBAL CHANGE BIOLOGY*, **9**, 131–140.
- Petit, R. & Hampe, A. (2006) Some evolutionary consequences of being a tree. *ANNUAL REVIEW OF ECOLOGY EVOLUTION AND SYSTEMATICS*, **37**, 187–214.
- Petit, R.J., Hu, F.S. & Dick, C.W. (2008) Forests of the past : A window to future changes. *SCIENCE*, **320**, 1450–1452.
- Pigliucci, M. (2005) Evolution of phenotypic plasticity : where are we going now? *TRENDS IN ECOLOGY & EVOLUTION*, **20**, 481–486.
- Price, T. & Schluter, D. (1991) On The Low Heritability Of Life-History Traits. *EVOLUTION*, **45**, 853–861.
- Pujol, B. & Pannell, J.R. (2008) Reduced responses to selection after species range expansion. *SCIENCE*, **321**, 96.

- Raesaenen, K. & Kruuk, L.E.B. (2007) Maternal effects and evolution at ecological time-scales. *FUNCTIONAL ECOLOGY*, **21**, 408–421.
- Ramirez-Valiente, J.A., Lorenzo, Z., Soto, A., Valladares, F., Gil, L. & Aranda, I. (2009) Elucidating the role of genetic drift and natural selection in cork oak differentiation regarding drought tolerance. *MOLECULAR ECOLOGY*, **18**, 3803–3815.
- Reich, P., Wright, I., Cavender-Bares, J., Craine, J., Oleksyn, J., Westoby, M. & Walters, M. (2003) The evolution of plant functional variation : Traits, spectra, and strategies. *INTERNATIONAL JOURNAL OF PLANT SCIENCES*, **164**, S143–S164.
- Reusch, T.B.H. & Wood, T.E. (2007) Molecular ecology of global change. *MOLECULAR ECOLOGY*, **16**, 3973–3992.
- Ritland, K. (1996) Estimators for pairwise relatedness and individual inbreeding coefficients. *GENETICAL RESEARCH*, **67**, 175–185.
- Ritland, K. (1996) Marker-based method for inferences about quantitative inheritance in natural populations. *EVOLUTION*, **50**, 1062–1073.
- Roach, D. & Wiulff, R. (1987) Maternal effects in plants. *ANNUAL REVIEW OF ECOLOGY EVOLUTION AND SYSTEMATICS*, **18**, 209–235.
- Robledo-Arnuncio, J.J., Austerlitz, F. & Smouse, P.E. (2007) POLDISP : a software package for indirect estimation of contemporary pollen dispersal. *MOLECULAR ECOLOGY NOTES*, **7**, 763–766.
- Robledo-Arnuncio, J . J .and Alia, R. & Gil, L.. (2004) Increased selfing and correlated paternity in a small population of a predominantly outcrossing conifer, pinus sylvestris. *MOLECULAR ECOLOGY*, **13**, 2567–2577.
- Root, T., Price, J., Hall, K., Schneider, S., Rosenzweig, C. & Pounds, J. (2003) Fingerprints of global warming on wild animals and plants. *NATURE*, **421**, 57–60.
- Saenz-Romero, C., Guzman-Reyna, R. & Rehfeldt, G. (2006) Altitudinal genetic variation among Pinus oocarpa populations in Michoacan, Mexico - Implications for seed zoning, conservation, tree breeding and global warming. *FOREST ECOLOGY AND MANAGEMENT*, **229**, 340–350.

-
- Saint-Laurent, R., Legault, M. & Bernatchez, L. (2003) Divergent selection maintains adaptive differentiation despite high gene flow between sympatric rainbow smelt ecotypes (*Osmerus mordax* Mitchill). *MOLECULAR ECOLOGY*, **12**, 315–330.
- Sambatti, J. & Rice, K. (2006) Local adaptation, patterns of selection, and gene flow in the Californian serpentine sunflower (*Helianthus exilis*). *EVOLUTION*, **60**, 696–710.
- Savolainen, O., Pyhajarvi, T. & Knurr, T. (2007) Gene flow and local adaptation in trees. *ANNUAL REVIEW OF ECOLOGY EVOLUTION AND SYSTEMATICS*, **38**, 595–619.
- Scheiner, S. (1993) Genetics And Evolution Of Phenotypic Plasticity. *ANNUAL REVIEW OF ECOLOGY AND SYSTEMATICS*, **24**, 35–68.
- Schluter, D. (1996) Adaptive radiation along genetic lines of least resistance. *EVOLUTION*, **50**, 1766–1774.
- Scotti, I. (2010) Adaptive potential in forest tree populations : what is it, and how can we measure it? *ANNALS OF FOREST SCIENCE*, **67**.
- Scotti, I., Calvo-Vialettes, L., Scotti-Saintagne, C., Citterio, M., Degen, B. & Bonal, D. (2010) Genetic variation for growth, morphological, and physiological traits in a wild population of the neotropical shade-tolerant rainforest tree *sextonia rubra* (mez) van der werff (lauraceae). *TREE GENETICS AND GENOME*, **6**, 319–329.
- Senneke, S., MacNeil, M. & Van Vleck, L. (2004) Effects of sire misidentification on estimates of genetic parameters for birth and weaning weights in hereford cattle. *JOURNAL OF ANIMAL SCIENCE*, **82**, 2307–2312.
- Smouse, P.E., Dyer, R.J., Westfall, R.D. & Sork, V.L. (2001) Two-generation analysis of pollen flow across a landscape. i. male gamete heterogeneity among females. *EVOLUTION*, **55**, 260–271.
- Sork, V., Nason, J., Campbell, D. & Fernandez, J. (1999) Landscape approaches to historical and contemporary gene flow in plants. *TRENDS IN ECOLOGY & EVOLUTION*, **14**, 219–224.
- Spitze, K. (1993) Population structure in *daphnia obtusa* : Quantitative genetic and allozymic variation. *GENETICS Society of America*, **135**, 367–374.

- Squillace, A. (1974) Average genetic correlations among offspring from open-pollinated forest trees. *SILVAE GENETICA*, **23**, 149–156.
- St Juliana, J.R. & Janzen, F.J. (2007) Can natural phenotypic variances be estimated reliably under homogeneous laboratory conditions? *JOURNAL OF EVOLUTIONARY BIOLOGY*, **20**, 1406–1414.
- Stinchcombe, J.R., Simonsen, A.K. & Blows, M.W. (2014) Estimating Uncertainty In Multivariate Responses To Selection. *EVOLUTION*, **68**, 1188–1196.
- Stirling, D., Reale, D. & Roff, D. (2002) Selection, structure and the heritability of behaviour. *JOURNAL OF EVOLUTIONARY BIOLOGY*, **15**, 277–289.
- Swindell, W. & Bouzat, J. (2006) Gene flow and adaptive potential in *Drosophila melanogaster*. *CONSERVATION GENETICS*, **7**, 79–89.
- Teplitsky, C., Tarka, M., Moller, A.P., Nakagawa, S., Balbontin, J., Burke, T.A., Doutrelant, C., Gregoire, A., Hansson, B., Hasselquist, D., Gustafsson, L., de Lope, F., Marzal, A., Mills, J.A., Wheelwright, N.T., Yarrall, J.W. & Charmantier, A. (2014) Assessing Multivariate Constraints to Evolution across Ten Long-Term Avian Studies. *PLOS ONE*, **9**.
- Thomas, C., Cameron, A., Green, R., Bakkenes, M., Beaumont, L., Collingham, Y., Erasmus, B., de Siqueira, M., Grainger, A., Hannah, L., Hughes, L., Huntley, B., van Jaarsveld, A., Midgley, G., Miles, L., Ortega-Huerta, M., Peterson, A., Phillips, O. & Williams, S. (2004) Extinction risk from climate change. *NATURE*, **427**, 145–148.
- Thomas, S., Coltman, D. & Pemberton, J. (2002) The use of marker-based relationship information to estimate the heritability of body weight in a natural population : a cautionary tale. *JOURNAL OF EVOLUTIONARY BIOLOGY*, **15**, 92–99.
- Thomas, S., Pemberton, J. & Hill, W. (2000) Estimating variance components in natural populations using inferred relationships. *HEREDITY*, **84**, 427–436.
- Vargashernandez, J. & Adams, W. (1991) Genetic-variation of wood density components in young coastal douglas-fir - implications for tree breeding. *CANADIAN JOURNAL OF FOREST RESEARCH*, **21**, 1801–1807.
- Verrier, E., Brabant, P. & Gallais, A. (2001) *Faits et concepts de base en génétique quantitative*. Polycopié INA Paris-Grignon.

-
- Villemereuil, P., Gimenez, O. & Doligez, B. (2013) Comparing parent-offspring regression with frequentist and Bayesian animal models to estimate heritability in wild populations : a simulation study for Gaussian and binary traits. *METHODS IN ECOLOGY & EVOLUTION*, **4**, 260–275.
- Violle, C., Navas, M.L., Vile, D., Kazakou, E., Fortunel, C., Hummel, I. & Garnier, E. (2007) Let the concept of trait be functional! *OIKOS*, **116**, 882–892.
- Visscher, P.M., Medland, S.E., Ferreira, M.A.R., Morley, K.I., Zhu, G., Cornes, B.K., Montgomery, G.W. & Martin, N.G. (2006) Assumption-free estimation of heritability from genome-wide identity-by-descent sharing between full siblings. *PLOS GENETICS*, **2**, 316–325.
- Vitasse, Y., Delzon, S., Bresson, C.C., Michalet, R. & Kremer, A. (2009) Altitudinal differentiation in growth and phenology among populations of temperate-zone tree species growing in a common garden. *CANADIAN JOURNAL OF FOREST RESEARCH-REVUE CANADIENNE DE RECHERCHE FORESTIERE*, **39**, 1259–1269.
- Vornam, B., Decarli, N. & Gailing, O. (2004) Spatial distribution of genetic variation in a natural beech stand (*Fagus sylvatica* L.) based on microsatellite markers. *CONSERVATION GENETICS*, **5**, 561–570.
- Wang, K.S. (2003) Mating system in isolated stands of european beech (*fagus sylvatica* l.). *FOREST GENETICS*, **10**, 159–164.
- Weigensberg, I. & Roff, D. (1996) Natural heritabilities : Can they be reliably estimated in the laboratory? *EVOLUTION*, **50**, 2149–2157.
- Whitlock, M. (1999) Neutral additive genetic variance in a metapopulation. *GENETICAL RESEARCH*, **74**, 215–221.
- Whitlock, M.C. & Guillaume, F. (2009) Testing for Spatially Divergent Selection : Comparing Q(ST) to F-ST. *GENETICS*, **183**, 1055–1063.
- Wilson, A., Coltman, D., Pemberton, J., Overall, A., Byrne, K. & Kruuk, L. (2005) Maternal genetic effects set the potential for evolution in a free-living vertebrate population. *JOURNAL OF EVOLUTIONARY BIOLOGY*, **18**, 405–414.

- Wilson, A.J., Reale, D., Clements, M.N., Morrissey, M.M., Postma, E., Walling, C.A., Kruuk, L.E.B. & Nussey, D.H. (2010) An ecologist's guide to the animal model. *JOURNAL OF ANIMAL ECOLOGY*, **79**, 13–26.
- Wright, I., Reich, P., Westoby, M., Ackerly, D., Baruch, Z., Bongers, F., Cavender-Bares, J., Chapin, T., Cornelissen, J., Diemer, M., Flexas, J., Garnier, E., Groom, P., Gulias, J., Hikosaka, K., Lamont, B., Lee, T., Lee, W., Lusk, C., Midgley, J., Navas, M., Niinemets, U., Oleksyn, J., Osada, N., Poorter, H., Poot, P., Prior, L., Pyankov, V., Roumet, C., Thomas, S., Tjoelker, M., Veneklaas, E. & Villar, R. (2004) The worldwide leaf economics spectrum. *NATURE*, **428**, 821–827.
- Wright, S. (1951) The genetical structure of populations. *ANNALS OF EUGENICS*, **15**, 323–354.
- Yang, J., Benyamin, B., McEvoy, B.P., Gordon, S., Henders, A.K., Nyholt, D.R., Madden, P.A., Heath, A.C., Martin, N.G., Montgomery, G.W., Goddard, M.E. & Visscher, P.M. (2010) Common SNPs explain a large proportion of the heritability for human height. *NATURE GENETICS*, **42**, 565–U131.
- Yeaman, S. & Jarvis, A. (2006) Regional heterogeneity and gene flow maintain variance in a quantitative trait within populations of lodgepole pine. *PROCEEDINGS OF THE ROYAL SOCIETY B-BIOLOGICAL SCIENCES*, **273**, 1587–1593.

ANNEXE A

Considering evolutionary processes in adaptive forestry

François Lefèvre · Thomas Boivin · Aurore Bontemps · François Courbet ·
Hendrik Davi · Marion Durand-Gillmann · Bruno Fady · Julie Gauzere ·
Cindy Gidoïn · Marie-Joe Karam · Hadrien Lalagüe ·
Sylvie Oddou-Muratorio · Christian Pichot

Received: 16 November 2012 / Accepted: 7 February 2013

© The Author(s) 2013. This article is published with open access at Springerlink.com

Abstract

• **Context** Managing forests under climate change requires adaptation. The adaptive capacity of forest tree populations is huge but not limitless. Integrating evolutionary considerations into adaptive forestry practice will enhance the capacity of managed forests to respond to climate-driven changes.

• **Aims** Focusing on natural regeneration systems, we propose a general framework that can be used in various and complex local situations by forest managers, in combination with their own expertise, to integrate evolutionary considerations into decision making for the emergence of an evolution-oriented forestry.

• **Methods** We develop a simple process-based analytical grid, using few processes and parameters, to analyse the impact of forestry practice on the evolution and evolvability of tree populations.

• **Results** We review qualitative and, whenever possible, quantitative expectations on the intensity of evolutionary drivers in forest trees. Then, we review the effects of actual

and potential forestry practice on the evolutionary processes. We illustrate the complexity of interactions in two study cases: the evolutionary consequences for forest trees of biotic interactions and of highly heterogeneous environment.

• **Conclusion** Evolution-oriented forestry may contribute adapting forests to climate change. It requires combining short-term and long-term objectives. We propose future lines of research and experimentation.

Keywords Genetic resources · Silviculture · Adaptation · Climate change · Evolution-oriented forest management

1 Introduction

Ecosystem functioning depends on the adaptation of living organisms to their physicochemical environment. In particular, the maladaptation of trees to local conditions can provoke ecosystem dysfunctions such as forest dieback or failure of regeneration, and it can also affect biotic interactions between trees and associated species. Multi-site common-garden experiments, which allow modelling the reaction norm of current tree populations to climatic parameters, suggest that climate change will lead to a high risk of maladaptation of tree species, at least in some parts of the current distribution range (Savolainen et al. 2007; St Clair and Howe 2007). To maintain forest services under climate change, tree stands will have to respond within one to ten generations to (1) more frequent and more intense extreme climatic events, (2) changing mean climatic parameters and (3) other related changes such as parasite outbreaks (IPCC 2007). Consequently, ecological services of the forests will depend on the intensity and velocity of the evolution of tree populations in response to climate change (Rehfeldt et al. 2001). Thus, adaptation should be considered in a dynamic perspective, as a bouquet of evolutionary processes that

Handling Editor: Erwin Dreyer

Contribution of the co-authors FL coordinated the review and wrote the text. All co-authors contributed to the revisions of preliminary drafts. In addition, more specific contributions were as follows:

Part 2: AB, BF, JG, SOM, CP

Part 3: FC, HD, BF, SOM, CP

Part 4: MJK, HL

Box 1: TB, MDG, CG

Box 2: FC, CP

F. Lefèvre (✉) · T. Boivin · A. Bontemps · F. Courbet · H. Davi ·
M. Durand-Gillmann · B. Fady · J. Gauzere · C. Gidoïn ·
M.-J. Karam · H. Lalagüe · S. Oddou-Muratorio · C. Pichot
INRA, UR629 Ecologie des Forêts Méditerranéennes, URFM,
Domaine Saint Paul, Site Agroparc, 84914 Avignon Cedex 9,
France
e-mail: francois.lefevre@avignon.inra.fr

change populations and communities to fit their environment. Among these processes, genetic adaptation, i.e. genetic change of a population responding to selection, can be rapid and contribute to the ecological success of species facing climate change including forest trees (Aitken et al. 2008; Hoffmann and Sgrò 2011). A recent review stressed the high potential of evolutionary response to climate change in trees (Alberto et al. 2013). However, evidence of lack of adaptation does also exist, e.g. niche limits and empty niches, including for tree species that have large population size and produce huge quantities of seeds (Bradshaw 1991).

During the last century, foresters have succeeded in adapting forest genetic resources to bioclimatic conditions very different from their native range, obtaining good survival, growth and reproduction in the new environments. Emblematic examples are the worldwide transfer of *Pinus radiata* (Yan et al. 2006) and the south to north translocation of *Picea abies* (Skrøppa et al. 2010). This adaptation was achieved in very few generations of trees, and it proceeds from plasticity and/or evolution. For each adaptive trait, the phenotypic plasticity and the capacity of evolution depend on the genetic content and the environment of the population, which can both evolve (Pigliucci 2008). Within each population, genetic changes of mean trait value, plasticity and evolvability result from the combination of random and selectively oriented processes that can be affected by forestry practice. Whether immediate response to selection can hamper future evolutions, e.g. due to erosion of the genetic diversity, remains an open question. Evidence from breeding experience shows that the genetic responsiveness of populations submitted to continuous selection can be maintained through time for some traits: Evolvability was maintained over more than 100 generations of selection for protein and oil content in the Illinois maize breeding population (Moose et al. 2004). No such long-term empirical evidence is available for trees. However, local adaptation that commonly emerged in most tree species over the course of post-glacial recolonisation provides another illustration of achieved evolution (Savolainen et al. 2007). Noticeably, this local adaptation did not completely erode within-population genetic variation of adaptive traits (Mimura and Aitken 2007; Alberto et al. 2013). The long-term maintenance of evolvability also depends on the genetic architecture of the traits under selection, and in the case of polygenic inheritance, Kremer and Le Corre (2012) showed that evolutionary changes first result from the selection of the fittest combinations of gene alleles before it reduces the allelic diversity at individual gene loci.

However, adaptation is not limitless. Futuyma (2010) reviewed the factors that can limit adaptation from the short term to the phylogenetic time scale. Focusing on an ecological rather than geological time scale, we can retain here seven

constraints to evolutionary changes. Firstly, developmental constraints result from functional interactions among traits involved in the elaboration of the performance. We use here 'performance' as a generic term, referring either to fitness components in an ecological perception or to forestry objectives like wood quantity or quality in an agronomic perspective, or to any combination of these traits. Secondly, genetic constraints result from the genetic architecture of traits, with complex epistatic interactions between several genes on one trait or pleiotropic effects of one single gene on several traits. Actually, forestry practices have little (but not null) impact on these first two limiting factors. Then, Futuyma (2010) identified four limiting factors of adaptation on which forestry practice may have direct or indirect impact: lack of genetic diversity, demographic stochasticity (counter-acting directional selection), random genetic drift and asymmetric gene flow (e.g. at niche limits). In addition to these, Kuparinen et al. (2010) identified another limiting factor potentially affected by forestry practice: low mortality.

Deciphering the factors that determine adaptation in the real forest, from the genes to the traits and from the traits to the performance, is complex. Each environment cannot be reduced to only one parameter, e.g. altitude combines temperature, soil, rainfall, biotic factors etc. Similarly, each performance, e.g. survival in stress conditions, can be achieved by different combinations of functional trait values. Finally, each value of a functional trait can be obtained by different combinations of gene alleles and interactions. As a consequence, one can hardly attribute a fixed intrinsic adaptive value to each physiological trait or to each gene allele. This complexity is also a chance for adaptation because it provides flexibility and there are multiple biological pathways to reach an ecological solution.

Forest management can enhance forest adaptation to climate change in three ways. Firstly, a full-control strategy consists in replacing the local population by a presumably better fit population. This is achieved through plantation of so-called forest reproductive material, which either comes from a breeding program or from a selected seed stand. This strategy allows for drastic stepwise evolutions, but it requires minimizing uncertainties about the ecological integration of the alien resource in the new site under future climates. Secondly, a driving strategy consists in guiding, i.e. supporting and accelerating, natural evolutionary processes using the local genetic resource, ecologically integrated within its current environment. This is achieved through natural regeneration. This strategy only produces progressive changes, limited by the evolutionary potential of the local resource, but it is flexible and relaxes the ecological uncertainty related to introduction of alien material. Thirdly, a combined strategy would follow the driving strategy after enrichment of the local resource with a certain amount of alien material in order to increase the evolutionary potential and to accelerate evolution

while limiting the ecological uncertainty due to introduction. Since the first approach has already been treated elsewhere and deserves a complete treatment, e.g. see St Clair and Howe (2007) for a concrete experience in *Pseudotsuga menziesii*, we focus here on the second and third strategies. Three main questions emerge in this context: (1) How fast can tree populations respond to changes? (2) Will the populations keep their capacity to adapt to both continuous and unpredictable changes? (3) How can forestry practice affect, positively or negatively, the properties of adaptation and adaptability through time? Due to the complexity of evolutionary mechanisms interacting with highly diverse local conditions and climate change scenarios, the first two questions can only receive case-specific answers. Here, we call evolution-oriented forestry a particular form of adaptive forestry that integrates the enhancement of evolutionary processes among its possible objectives, and we propose a process-based approach to investigate the impact of silviculture on the evolution and evolvability of tree populations facing climate change.

In a first part, we describe the basic evolutionary mechanisms using a limited number of parameters in order to define a simple analytical grid. We show how these few parameters can help understand complex situations. The analytical grid also provides a mechanistic interpretation of the evolutionary constraints mentioned above. Secondly, we use this analytical grid to evaluate the potential effects of current forestry practice and to suggest other silvicultural options that could preserve as much as possible the objectives of forestry while driving the tree populations into faster evolutionary changes. Finally, we review the possible genetic tools available for monitoring adaptive changes and evolutionary processes and conclude with future perspectives for experimental management and research.

2 The basic mechanisms driving genetic changes as an analytical grid

Genetic diversity is continuously changing: Each sexual reproduction event generates new and unique genotypic combinations, some of which are then eliminated by selection and random processes. The strength of selection and genetic drift can be efficiently approached with a limited number of parameters. A wide range of evolutionary scenarios can be obtained when considering the interactions between selection, genetic drift, gene flow and plasticity. Due to the short term considered here, we neglect the effect of mutation.

2.1 Single and multitrait response to selection

Selection is the elimination of the less fit, due to low reproduction or mortality. A selection pressure on one

trait can change the mean of the population (directional selection), or its variance (stabilizing or disruptive selection), or both. Although it does not integrate all the biological processes that effectively operate, the model of quantitative genetics (Falconer 1960), which assumes that a large number of genes interact with the environment to determine the phenotypic variation of each trait, has long proven a remarkable predictive value throughout all the practical achievements in plant and animal breeding programmes since mid-twentieth century. This model predicts the rate of change of trait mean per generation under direct directional selection as:

$$R = i.h.\sigma_A$$

or, expressed in phenotypic standard deviation units,

$$R' = R/\sigma_P = i.h^2$$

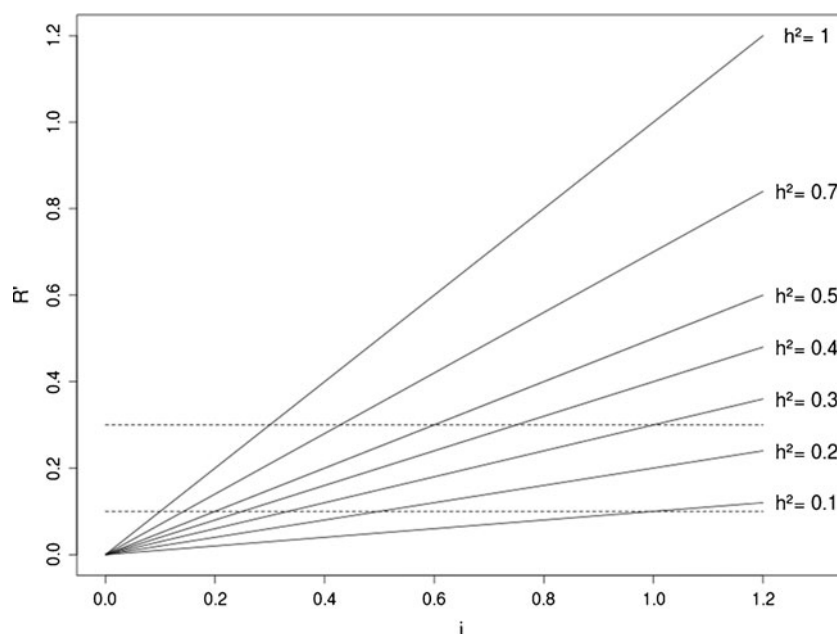
where

- i* Is the intensity of selection, i.e. mean differential between the adults that contribute and those that do not contribute to the next generation expressed in standard deviation units (*i* directly relates to the proportion of selected individuals)
- h* Is the square root of the narrow sense heritability, which is the ratio between the additive genetic variance and the phenotypic variance ($h^2 = \sigma_A^2/\sigma_P^2$)
- σ_A^2 Is the additive genetic variance
- σ_P^2 Is the phenotypic variance

Natural selection does not operate directly on the traits but on the global fitness of the individuals. Therefore, the response of each trait to natural selection is a type of indirect selection, and the previous equation is slightly modified as: $R = \beta.\sigma_A^2$ where β is the partial regression coefficient of the fitness on the trait, or selection gradient. This formulation extends to the multitrait case where the response on each trait integrates its correlations with other selected traits (Lande and Arnold 1983). Finally, selection is summarized with only four basic parameters: σ_P^2 , σ_A^2 , *i* (direct selection) or β (natural selection).

Based on a review of empirical studies in all kinds of organisms, Gingerich (2009) found relatively high rates of evolution (*R'*) in the magnitude of 0.1 to 0.3 phenotypic standard deviation per generation, similar for long-term evolution and micro-evolution. Figure 1 shows the expected rate of change of a single trait under direct selection with different values of heritability and selection intensity. As explained in the next section, forest management can modify *i* and, therefore, proportionally change *R'*.

Fig. 1 Expected rate of change per generation (R') for a single selected trait under direct selection, expressed in phenotypic standard deviation units, for different values of heritability (h^2) and different selection intensities (i). Assuming a Gaussian distribution of the trait, elimination of 25, 50 or 75 % of the population corresponds to values of $i=0.42$, 0.80 or 1.27, respectively. The range of empirical values of R' found in the literature by Gingerich (2009) lies between the dotted lines



A trait will respond to selection if three conditions are simultaneously met (Endler 1986): (1) there is phenotypic variation within the population, (2) this variation is heritable and (3) this variation is correlated with the variation of fitness. Each term in the equation above is not fixed, but it varies depending both on the environment (Charmantier and Garant 2005) and the genetic background of the population (Carter et al. 2005). Natural selection in trees is a complex process because the selection pressure varies between years and it can drastically change between life stages both in direction (e.g. shift in selective forces between a dense seedling patch under the canopy and the adult stage) and in intensity (e.g. mortality rate is much higher at juvenile stage while selection on reproductive success only occurs at adult stage). Thanks to this variability of the parameters, the response to selection does not systematically correlates with a total exhaustion of the genetic variation within populations: The partitioning of the genetic variance of adaptive traits into between- and within-population components (Q_{st} approach) has revealed that a large part of the genetic variance of functional traits is maintained within population (Alberto et al. 2013, for a review), and an absence of within population genetic variance for a quantitative trait has exceptionally been reported in trees (Sáenz-Romero et al. 2006). Interestingly, the genetic architecture of a trait, i.e. the system of genes involved in the variation of the trait and their interactions with other traits, simultaneously determines long-term persistence of evolvability by a capacity to release cryptic variation in a new environmental or genetic context (Le Rouzic et al. 2007) as well as a potential limit to selection in case of detrimental genetic correlations (Walsh and Blows 2009). Functional constraints, resulting in environmental correlations between

traits, can also limit the response to selection. Not included into this predictive model, epigenetic effects, i.e. environmentally determined heritable modification of gene expression, can also contribute to adaptation to sudden changes (Bossdorf et al. 2008; Skrøppa et al. 2010).

2.2 Random changes due to genetic drift and mating system

Genetic drift accounts for the reduction of genetic diversity that occurs in small populations, in absence of selection, mutation or migration, due to the variation of allele frequencies after random sampling from one generation to the next. Furthermore, small populations are prone to increased inbreeding due to the higher probability of mating between relatives. Inbreeding has a twofold effect: It reduces fitness whenever inbreeding depression is present, and it retains non-random association of gene alleles (linkage disequilibrium) at higher rate which represents a reduction in the diversity of genotypic combinations. Non-random mating system can also affect inbreeding: The mating system varies among individuals and populations, including selfing rate (most of tree species are not dioecious) and diversity of pollen donors, depending on the relative fecundity and spatial distribution of reproducing trees.

Under the assumptions of Wright–Fisher's model population (Wright 1931), the reduction of gene diversity and the increase of inbreeding are driven by one single parameter, population size. Using this model as a reference, the effective population size (N_e) of a real population of size N that has a per-generation rate of reduction of gene diversity (ΔH_e) or increase of inbreeding (ΔF) is such that: $\Delta H_e = -1/2N_e$ or $\Delta F = 1/2N_e$. It can be shown that, in the absence of dominance, N_e also measures the per-generation rate of reduction of

additive variance: $\Delta\sigma_A^2 = -1/2N_e$. N_e is defined on the rate of change of gene diversity or inbreeding, not on the actual population size N . N_e is most often not directly estimable in natural populations (unless longitudinal estimates of gene diversity, H_e or σ_A^2 , or inbreeding, F), but its changes can be predicted and decreasing N_e means intensifying the intensity of genetic drift. When the actual population only departs from the theoretical model by relaxing the assumption of Poissonian distribution of reproductive success, it can be shown that: $N_e = (4N-2)/(V+2)$ where V is the actual variance in reproductive success, i.e. N_e decreases substantially in proportion to this variance.

Tree populations are generally assumed to have large effective population size (Petit and Hampe 2006), in part because they are outcrossing and disperse their genes over long distances in particular through pollen (Ashley 2010); thus, they should not be too much affected by genetic drift. In their review, Schoen and Brown (1991) found N_e estimates for tree species in the range of other outbreeding plants: mean values around 3,000 for *Pseudotsuga menziesii* and >8,000 for *Pinus sylvestris* and *Picea abies*, with high variations among populations within each species. However, locally, seed and pollen contributions to reproduction are highly uneven among individuals (Burczyk et al. 2002; Krouchi et al. 2004; Oddou-Muratorio et al. 2005), and the great majority of the pollen disperses only in the close neighbourhood, which can greatly reduce the effective population size. In their review, Smouse and Sork (2004) found that the effective pollen pool size N_{ep} , defined as the inverse of the probability that a female draws two offsprings from the same father, ranges from 2 to 200 in tree populations. N_{ep} can be very small in some populations: Fragmented populations of wind pollinated species of *Quercus humboldtii* (Fernandez-M and Sork 2005) and *Quercus alba* (Smouse et al. 2001) exhibit estimates of N_{ep} around 6 and 8, respectively. By contrast, in continuous forest populations, several examples estimate high values of N_{ep} : Robledo-Arnuncio et al. (2004) estimated a $N_{ep} > 70$ in a Spanish population of *Pinus sylvestris*. Fragmented populations, isolated populations and populations at low density have a higher risk of extinction due to the erosion of diversity by genetic drift (Goodell et al. 1997; Hardy et al. 2004; Robledo-Arnuncio et al. 2004; Aguilar et al. 2008).

2.3 Interactions between selection, drift, gene flow and phenotypic plasticity

A well-known interaction between drift and selection is the vortex of extinction (Gilpin and Soule 1986): when there is a genetic load in the population, a rapid decrease in population size leads to increased genetic drift and increased inbreeding, resulting in reduced mean fitness that further reduces

population size, which over time will result in extinction in a geometric decline. However, there is no experimental evidence to our knowledge that this kind of extinction vortex ever occurred in trees. Alternatively, resistant genotypes that emerge in the population increase their contribution to the next generation and can restore population growth if not eliminated at random, a process known as evolutionary rescue (Gomulkiewicz and Holt 1995). Whether populations can be rescued depends on population size, genetic diversity and the degree of maladaptation to the new environment.

In the case when the environment changes both in space and time, gene flow can bring into the population pre-adapted genes (Pease et al. 1989; review by Kremer et al. 2012). Kuparinen et al. (2010) showed that pollen and seed dispersal at longer distance speed up the adaptation process. In tree populations, it is expected that pollen-mediated rather than seed mediated gene flow will contribute to this processes, with average pollination distances commonly being hundreds of meters (Ashley 2010), and maximum distance of 100 km measured in *Pinus sylvestris* airborne transported pollen (Robledo-Arnuncio 2011). In this species, along a latitudinal gradient, Nilsson (1995) showed that long-distance pollen migration brings into the population new phenotypes with a phenology specific of other locations. Local individuals with extreme phenotypic values, in particular for flowering phenology, are keen to catch more long-distance pollen grains because they are better synchronized with source population of interest and also because they are less saturated by local pollen.

Abundant theoretical literature exists on the role of phenotypic plasticity in evolution; recent reviews of predictions on the interaction between selection and plasticity in the context of climate change can be found in Chevin et al. (2010, 2012), the second review also provides key references of empirical studies of plasticity in trees. Maladaptive plasticity is obviously detrimental to adaptation. When adaptive plasticity is not genetically variable, it is expected to slow down the genetic response to directional selection in each generation but also to allow the phenotypes to track the environmental change more closely. The sustainability of this process of adaptation depends on the fitness cost of plasticity. When adaptive plasticity varies genetically, i.e. there is $G \times E$ interaction and plasticity can evolve, the amount of genetic variance of the plastic trait depends on the environment: If the new environment increases the genetic variance, then plasticity tends to accelerate the genetic response to selection and plasticity is itself selected for (Lande 2009). Considering a steep spatial environmental gradient, where gene flow interacts with selection and plasticity, the evolution of plasticity is expected to allow the population to explore a larger range, and marginal habitats are expected to show higher plasticity (Chevin and Lande 2011). A further degree of complexity arises when considering the interplay between abiotic environment, biotic interactions and evolutionary processes (Box 1).

3 Potential effects of forestry practice on the rate of evolution of trees

3.1 A global framework to assess the evolutionary impact of silviculture

Considering the evolutionary impact of silviculture is thus an additional requisite to adaptive forestry. While facing climate change and uncertain future, forestry practice should simultaneously accelerate genetic adaptation by helping tree populations to track the known environmental changes and preserve the genetic diversity as a reservoir of future options to respond to the next unknown changes. This is a classical challenge in breeding programs: obtaining a rapid genetic gain while keeping the potential for long-term response to selection. Breeders solve this problem by combining selection and recombination. Genetic adaptation occurs if better performing genotypes emerge during reproduction and if the best performing genotypes spread in the population before extinction. Therefore, we can assign three objectives to evolution-oriented forestry practice: (1) increase the chance of emergence of ‘innovative’ genetic combinations, (2) facilitate the spread of the best adapted genotypes and (3) preserve genetic diversity for long-term response to selection. Acting simultaneously on the demography and the environmental conditions, silviculture has an impact on the parameters of selection and genetic drift. Due to the multiplicity and complexity of the processes involved and to the huge diversity of biological and environmental features among forests, it is more reliable to predict the effects of silviculture on the parameters of evolution rather than on the final state of the genetic diversity. Thus, silviculture should aim to limit the intensity of genetic drift, increase the diversity of mating pairs, avoid counter-selection and maintain selection pressure in the right direction, reduce inbreeding etc. Here, we briefly consider the evolutionary benefits, or risks, associated to current or future forestry practice. A more quantitative prediction of their balanced effects could rely on simulation studies using process-based models that explicitly integrate demographic, genetic and biophysical processes and allow to simulate interventions (Kramer et al. 2008).

3.2 Foreseen evolutionary impact of common forestry practice

A general consequence of silviculture on the drivers of evolution is the change of environmental conditions: competition and, possibly, other biotic or abiotic environmental factors. As mentioned above, the environmental conditions determines not only the phenotypic variance (σ_P^2) of the traits but also the genetic variance (σ_A^2), in case of G×E interactions, and the selection gradient that relates trait value to fitness (β) (Fig. 2). Furthermore, increased canopy closure

also affects pollen dispersal and mating system parameters: selfing rate (s), effective pollen pool size (N_{ep}) and variance of reproductive success (V) (García et al. 2005; Milleron et al. 2012).

Beyond these general effects, thinning, whether systematic or selective, will affect the spatial clustering of genetically related reproducing trees and their allocation to reproduction and, therefore, the mating system (Fig. 2). By removing neighbouring-related individuals, thinning may reduce spatial genetic structure and, consequently, inbreeding in the regeneration (Dounavi et al. 2002). Sagnard et al. (2011) showed that, when the density of seed trees is low (<16 trees ha^{-1}), a clustered distribution of seed trees will produce less spatial genetic structure in the seedlings than a random or dispersed distribution. Robledo-Arnuncio et al. (2004), in *Pinus sylvestris*, and Restoux et al. (2008), in *Abies alba*, found at seed stage that low densities of reproductive trees simultaneously increase the probability of selfing (from <5 to >20 %) and the diversity of the outcrossing pollen (N_{ep}), which can be mechanistically explained by a higher rate of long distance pollen pollination (Klein et al. 2006). Therefore, lower seed trees density can result in higher genetic diversity at seedling stage, whenever pollen is not limiting and purge of inbreeding occurs early, but it can be a risk when pollen is limiting, which must be considered in the context of climate change.

When the local population size is limited, a genetic drift effect may result from the reduction of the number of reproducing trees, leading to a loss of rare alleles (among which currently deleterious genes). Such effect was observed in old-growth forests of *Pinus strobus* comparing pre- and post-harvest stands (>100 and <30 trees, respectively) (Buchert et al. 1997; Rajora et al. 2000), as well as in *Picea rubens* old growth forests (Mosseler et al. 2003). Konnert and Hussendörfer (2001) compared 16 even-aged and nine uneven-aged management systems in *A. alba*, several forests for each group and also the two management systems within the same forest: They found a slightly higher number of rare alleles but smaller number of different gametic combinations in the uneven-aged forests. However, there are too few studies of this type to draw general conclusions on an eventual intrinsic difference between these two management systems regarding their effects on the genetic diversity. The drift effect is not a priori limited to the managed population but it may also affect secondary tree species as observed for *Abies amabilis*, *Tsuga heterophylla*, *Thuja plicata* and *Pinus monticola* in *Pseudotsuga menziesii* forests by El-Kassaby and Benowicz (2000).

In naturally regenerated stands, reproductive trees usually result from successive selective thinnings for their phenotypic value such as stem vigour, health consideration, lack of defects like thick branches or forked or twisted trunk, as

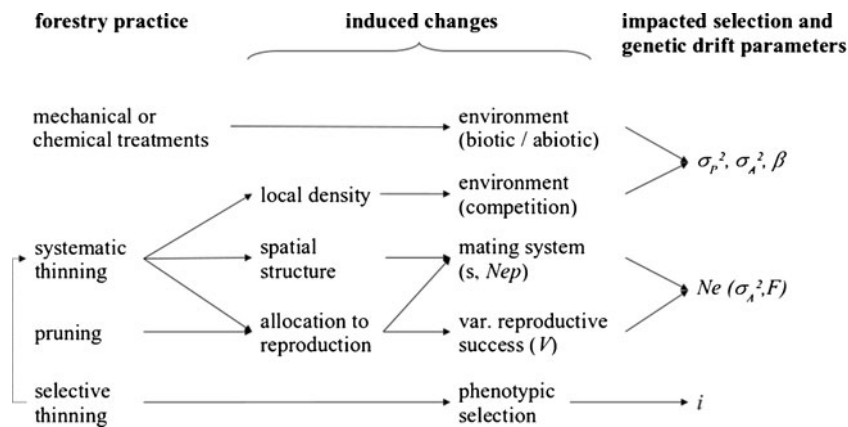


Fig. 2 Expected effects of forestry practice on evolutionary processes: schematic pathway from the forestry management in natural regeneration systems to the selection and genetic drift parameters. See text for explanation of the paths (σ_p^2 phenotypic variance, σ_A^2 additive genetic

variance, β selection gradient for natural selection, i selection intensity for direct selection, s selfing rate, N_{ep} effective pollen pool size, V variance of reproductive success, N_e effective population size, F inbreeding)

well as for their spatial distribution in order to reduce competition. In *Fagus sylvatica*, selective thinning favouring the most vigorous trees was found to increase the heterozygosity by 4 to 9 % (Lauber et al. 1997; Dounavi et al. 2002), even when selection occurs at a very early stage (Thiebaut et al. 1992). However, this effect was not detected in other studies on *A. alba* (Hussendörfer and Konnert 2000) or *Pinus contorta* (McDonald et al. 2001). We must remind here that an increase in heterozygosity during maturation of forest stands is frequently observed and that natural selection of seedlings can start at a very early stage in overstocked young regeneration (Pichot et al. 2006). Selective thinning favouring the best growing trees may act as selection for competing ability: Assuming a Gaussian distribution of tree height, the elimination of the 25 % (resp. 50 %) smallest individuals represents a selection intensity $i=0.42$ (resp. 0.80) on this trait. Selective thinning may also integrate a direct selection on other traits chosen by the forester (Fig. 2). We need to understand more clearly which functional traits are indirectly selected for and how these traits relate to future fitness in the context of changing climate. One main question that still needs to be addressed is how far silviculture will intensify selection for juvenile vigour and, if this is the case, how far juvenile vigour is genetically positively or negatively correlated with drought resistance.

The intensity of genetic drift, through the variance in reproductive success, and the mating system not only depends on the number and spatial distribution of reproductive trees but also on their allocation to reproduction. As indicated in Fig. 2, this allocation is influenced by local stand density as well as other practices like pruning (Ayari et al. 2012). Thinning and pruning also affect the plasticity, which will interact with selection processes as previously

mentioned: The question here is whether the plastic response induced by the silviculture is adaptive or maladaptive regarding future climate change.

Multiple forestry practices may also be combined and interact to modify the microenvironmental conditions of development, generally as to reduce stress and competition in order to favour growth (Forrester et al. 2012). Thus, after an extreme drought event in 1976, stand decline was reduced in *Picea abies* stands that had previously been thinned in 1971 (Misson et al. 2003): Compared to the control plot, heavy thinning had a more beneficial impact than moderate thinning (thinning from 36 to 14 or to 20 $m^2 ha^{-1}$ basal area, respectively). We are not aware of any study of the effects this could have on plasticity (e.g. reduced acclimation to future stress) and selection. As a case study, we present the analysis of evolution and potential effects of silviculture in the situation of highly heterogeneous environment within the forest (Box 2).

3.3 Evolutionary benefits and risks expected from some silvicultural recommendations related to climate change

New forestry practice is progressively implemented to reduce ecological and economic risks related to climate change (Legay and Mortier 2005; Yousefpour et al. 2012). From an evolutionary point of view, reducing environmental stress has a twofold effect. On the one hand, it reduces the damages and therefore contributes to increase the effective population size (N_e), which is a crucial issue when population size is already small or is expected to decrease drastically due to severe damages. But, on the other hand, it also slows down the genetic improvement in the next generation by reducing selection intensity (i), and it does not exploit potential adaptive plasticity (no acclimation to future stress), which is an important issue to consider in large populations.

We briefly consider here some of these practices, or changes in practices, from the evolutionary point of view and illustrate their possible balanced effects. In all cases, whenever applying new practice, it is crucial to keep precise records of what is done, how and when, in order to facilitate future evaluation ex post, in particular after marked climatic events.

Shortening rotation reduces the probability of the risk, e.g. to extreme climatic events, but it can also increase the vulnerability to the risk if shorter rotations select for higher juvenile vigour and if juvenile vigour is genetically negatively correlated with stress resistance. Both of these conditions still have to be investigated. The answer will probably depend on the species, on the environment and on the management system considered.

Reducing the density of stands is envisaged to reduce the effective drought stress supported by the trees. However, this immediate positive effect may be partly balanced by a long-term detrimental effect on selection by inducing a maladaptive phenotypic response and by reducing selective mortality (see first section). This risk is reduced if time is left for sufficient natural selection to proceed before thinning. More generally, from the evolutionary point of view, interventions occurring at juvenile stage raise the question of age–age correlations. Apart from the temporal changes in environmental conditions, life stages differ in their physiology and development. During the complex and temporally changing selection process in trees, juvenile–adult genetic correlations contribute to determine how far a selection pressure (or release of selection) during the juvenile stage will genetically affect the adult population. This issue can hardly be addressed in situ. Partial answer, here again, comes from the breeding experience and early selection schemes. Studies on growth and wood density in different *Pinus* species revealed that genetic age–age correlation >0.8 is generally achieved from the age of 10–12 years (Hannrup and Ekberg 1998; Gwaze et al. 2000; Wu et al. 2007; Bouffier et al. 2008). It varies greatly with environmental conditions, and in *Pinus radiata*, this level of correlation can be reached as soon as 2–5 years in certain sites (Gwaze et al. 2000; Wu et al. 2007). Matheson et al. (2002) showed that genetic age–age correlations in *Pinus radiata* also vary with the genetic background with higher correlations in presence of inbreeding. Thus, silviculture may have an effect on age–age correlations through its effects on the environmental conditions and on the genetic background. Further investigations on age–age correlations of functional traits are deeply needed.

In the case of massive dieback, sanitary logging is necessary to reduce the spread of primary or secondary parasites, and it can also be necessary for fire prevention or for the protection of forest users. However, excessive elimination of surviving trees could result in the elimination of resistance to the pathogen (Burke 2011).

3.4 Evolution-oriented forestry, why not?

We imagined some specifically evolution-oriented forestry practices in the case of natural regeneration management system (Table 1). These interventions should not be directly considered as recommendations or guidelines as such; we rather propose them as case studies to illustrate innovative adaptive forestry that would take into account short- and long-term evolutionary potential, still to be associated with other clues. This is not an exhaustive list, and any combination of the proposed interventions can be envisaged.

In order to reduce the intensity of genetic drift (increase N_e) in small populations, silviculture may be oriented towards reduced variance of fecundity (V) between trees: Reducing V will not only increase N_e within each annual seed production, as mentioned above, but it will also reduce the fluctuation of effective contributions across years and thus increase pluri-annual N_e estimates (Krouchi et al. 2004). This would be another objective assigned to thinning and pruning. It requires a balance between keeping sufficient number of seedling trees and sufficient spacing between them (optimisation should be made on a case-by-case basis, depending on dispersal capacities). Due to the tree \times year interaction effect on the variation in fecundity generally observed in trees, a general recommendation would be to cumulate reproduction during several years. Actually, current practice may already be optimal for this purpose. A negative side effect is to slow down the elimination of detrimental alleles (Couvét and Ronfort 1994) and reduce the response to selection. A compromise between preserving the genetic diversity for the future (reduce genetic drift) and accelerating the immediate response to selection could be to equalize the mating success per patch, in particular when the environment is spatially heterogeneous (Box 2).

In order to reshuffle the local genetic diversity and increase genetic recombination, silviculture could enhance local gene flow, either through artificial dispersal of local seeds or by assisting pollen dispersal. With the same objective, isolated seed trees should be considered with care: From one side, they may have a higher selfing rate but, on the other side, they can capture long distance pollen flow. If selfed seeds are eliminated at an early stage of development (e.g. empty seeds in some conifer species), the fertile seeds hamper a large genetic diversity. In the case of heterogeneous environment within the forest, areas for wood production and areas for evolution could be spatially dissociated while maintaining gene flow between these entities (Box 2).

In an environmental cline, typically an altitudinal cline, a strategy might be to accelerate the migration of the population towards more favourable areas. The velocity of migration depends on the effective dispersal, and effective dispersal is highly dependent on the local conditions for

Table 1 Some examples of evolution-oriented forestry practice, including re-orientation of usual interventions (no supplementary cost) and additional interventions

Forestry practice	Expected benefits	Associated costs and risks
N_e -oriented regulation of the density and spatial distribution to equalize reproductive success between trees in small populations	Reduce the variance in reproductive success to reduce genetic drift Reduce spatial genetic structure in the seedlings and inbreeding in next generation	No supplementary cost Risk to slow down the elimination of detrimental genes, prefer equalization of mating success per patch (compatible with the next line)
In heterogeneous environment, dissociate areas of production and areas of evolution (selection patches in harsh areas) and allow gene flow between these entities	Increase the reproductive contribution of the trees that have survived to drastic selection pressure	Limited supplementary cost Requires preliminary simulation studies to estimate benefits in various contexts (strength and spatial structure of the environmental heterogeneity)
Save the lone tree, which cumulates long distance dispersal (in allo-pollinated seeds) and can be adapted to marginal conditions; collect seeds for local assisted regeneration	Diversify the mating pairs to favour the emergence of new genotypic combinations Promote adaptation to marginal conditions	Limited supplementary cost Requires a protocol for assisted regeneration Risk of inbreeding if self-pollinated seeds are not purged at a very early stage (e.g. seed abortion)
Assisted local seed dispersal (e.g. collecting, possibly over several years, mixing and replanting seeds within the stand) or pollen dispersal (e.g. air flow used in seed orchards)	Enhance local gene flow to diversify the mating pairs and favour the emergence of new genotypic combinations	Potentially significant supplementary cost Requires preliminary studies to estimate benefits in various contexts (genetic diversity and spatial structure)
Enhance local migration capacity by favouring seed dispersal and germination at distance from the main stand	Reduce inbreeding Speed-up colonisation of locally favourable habitats in an environmental gradient	Requires a protocol for assisted regeneration Potentially significant supplementary cost
Genetic enrichment by introduction of a limited amount of seeds or pollen from presumably pre-adapted allochthonous origins	Introduce pre-adapted genotypes Increase local genetic diversity	Potentially significant supplementary cost Risk of gene swamping and reduction of effective population size (N_e) if local population is small and if introduced material has low genetic diversity Risk of unforeseen local maladaptation
Marker-assisted selective thinning (futurist)	Increase selection intensity on target major genes while retaining genetic diversity in the rest of the genome	High supplementary cost Requires accurate genetic knowledge and high-throughput genotyping capacities

See text and Box 2 for further explanation

seedling establishment conditions (Amm et al. 2012). We can imagine to enhance seed germination and seedling growth by preparing the soil or controlling competition and predation at distance from the core of the population, in the direction wanted for migration. For zoochorous species, we can also imagine to attract seed dispersers along the wanted migration route (Oddou-Muratorio et al. 2004; García et al. 2009; Schleuning et al. 2011).

Genetic enrichment of the local genetic resource by the introduction of a limited amount of allochthonous material from a putatively pre-adapted origin, through seed or pollen introduction, could present a twofold benefit of introducing gene alleles of interest and increasing the global genetic diversity. To avoid gene swamping effect and reduction of the effective population size, it is essential to use a large genetic base of the introduced material (Lefèvre 2004). For long-lived organisms, it is also important to anticipate a possible trade-off between adaptation to long-term climatic

trend and adaptation to current conditions and/or to annual fluctuations, such as vulnerability to late frost of early flushing genotypes.

As a complete utopy, we can imagine future access to intensive genotypic data on each adult tree. Inspired from the marker-assisted selection strategies used in plant and animal breeding, marker-assisted selective thinning could combine the objectives of increasing the adaptive change for the target traits that are unambiguously identified while preserving the maximum diversity in the rest of the genome.

4 Genetic monitoring and study tools

In a recent publication, Hansen et al. (2012) reviewed the different tests and approaches for genetic monitoring of adaptive changes using phenotypic or molecular tools. Their focus was on the capacity to demonstrate the adaptive

response and rule out alternative hypotheses that might explain the genetic change. In this section, we investigate how far recent genetic monitoring methods and tools can help to rationalize evolution-oriented forestry.

4.1 What is new in molecular and phenotypic tools?

A comprehensive review of the genetic markers and their use in trees was published by Prat et al. (2006). Almost all kinds of markers have been developed on one or several tree species, and they were mainly used to infer on the neutral genetic diversity and neutral processes (drift, mating system, dispersal). With the classical genome sequencing projects, which started in the 1990s for trees, a gap in terms of available tools had progressively appeared between a very limited number of model tree species and the other species. In the last 3 years, recent advances in DNA sequencing have revolutionized the field of genomics making it possible to generate a large amount of sequences and markers in time- and cost-effective way. Nowadays, thanks to the emergence and evolution of the so-called next generation sequencing techniques and related bioinformatics (Metzker 2009; Kircher and Kelso 2010), full-genome sequencing, reduced-representation sequencing and targeted sequencing are in progress not only for model species but also for non-model species where molecular monitoring is becoming worth considering from scratch. Evidence of genetic changes at molecular level in trees has recently emerged from the direct study of genome wide DNA polymorphisms: evidence of the correlations between genotype frequencies and environmental gradients (Eckert et al. 2010) or climate-related traits (Grivet et al. 2011), sometimes completed with functional information on the detected genes (Holliday et al. 2010). See also Alberto et al. (2013) for a recent review of single nucleotide polymorphisms associated to climate related traits in trees. However, as Rockman (2012) very wisely stated, the extraordinary potential of these approaches should not be misleading: Due to the genetic and environmental sensitivity of the response to selection as previously discussed, we should not expect to find many single gene alleles having large, constant and uniform effect in all populations. If such nucleotides are detected, we should not reduce the genetic variation to them because most of the genetic diversity of interest for adaptation to climate change will remain cryptic. These approaches will nevertheless be very useful in providing genetic indicators of the selection pressure.

The phenotypic approach of adaptation has also evolved in two directions. Firstly, the physiologists have produced proxies of physiological functions that can be measured in large sample size (hundreds of individuals). The relation between the measured trait and the actual function is generally indirect and requires careful interpretation, e.g. carbon

isotope discrimination or ring density used as proxies of the response to drought (Osório and Pereira 1994; Tene et al. 2011). It is probably worth reminding that splitting an integrated phenotypic trait into simpler functional components does not resolve the complexity, e.g. functional components do not necessarily have higher heritability or simpler genetic determinism than the integrated trait, because new interactions and regulations appear at finer scale. Secondly, methodologies combining phenotypic and genotypic information have been developed to estimate genetic parameters (variances and correlations) in situ, i.e. in the natural environment, at any life stage and without requiring controlled pedigrees (Ritland 1996). To estimate selection gradients (β) in trees, a common approach is to use performance traits like survival, growth or reproductive traits as proxies of fitness and study the impact of functional traits on these performance traits in controlled ex situ progeny tests. Using this approach in *Quercus suber*, Ramirez-Valiente et al. (2011) detected significant heritability but non-significant selection gradient for carbon isotope discrimination, contrasting with very low heritability and significant selection gradient for specific leaf area. An alternative approach of selection gradients through the assessment of actual reproductive success in situ was recently developed (Oddou-Muratorio et al. 2005; Burczyk et al. 2006; Klein et al. 2011). This method, based on the mixed-mating neighbourhood model, consists in estimating the reproductive success of individual adult trees using spatial genetic data of seedlings and their potential parents and then in relating this reproductive success to phenotypic traits. Bontemps (2012) used this method in a marginal population of *F. sylvatica*: In this case, the author found significant heritability and significant selection gradient for carbon isotope discrimination, contrasting with non-significant heritability and non-significant selection gradient for specific leaf area.

4.2 Possible monitoring

Various sets of state and pressure indicators have been proposed for the monitoring of the genetic diversity in forest trees, based on direct genetic assessment or indirect observations through the demography and ecology of the populations (Namkoong et al. 1996; Brown et al. 1997; Koski et al. 1997; Lefèvre and Kajba 2001; Aravanopoulos 2011). Table 2 briefly reviews the possible uses of molecular and phenotypic tools, and the requirements of these uses, for genetic monitoring of adaptation. We distinguish two main objectives for genetic monitoring: (1) quantification and characterization of the genetic diversity and (2) monitoring of the drivers of genetic changes.

The genes controlling the variance of most adaptive traits (quantitative trait loci, QTL) are expected to be numerous

Table 2 Objectives and requirements of genetic monitoring using molecular and phenotypic tools

Monitoring objectives	Requirements		
	Molecular tools only	Phenotypic tools only	Combined approach
Quantification and characterization of the genetic diversity			
Quantify the global genetic diversity and characterize its organisation	Neutral markers	–	–
Decipher the genetic architecture of adaptive traits (QTLs, variances and correlations)	Known pedigrees, validated candidate genes polymorphisms	Known pedigrees, common garden experiments (ex situ)	QTL mapping or QTL association studies, ex situ experiments and in situ methods
Monitoring the drivers of genetic changes			
Monitor recent changes in demography and genetic drift intensity	Neutral markers, multi-generation sampling	–	–
Monitor the mating system and hybridization	Neutral markers, multi-cohort sampling	–	–
Characterize pollen and seed dispersal functions	Neutral markers, seed or/and seedling samples	Pollen and seed traps	Combined approach possible
Detect signatures of past selection (β)	Genome wide markers and candidate genes polymorphisms	–	–
Monitor current selection gradient (i, β)	Validated candidate genes polymorphisms, multi-generation sampling	Known pedigrees, path analysis traits to performance, ex situ experiments	Selection gradient studies, in situ methods

See text for further explanation

QTL quantitative trait loci

with small individual effect. This is confirmed by empirical results, even though molecular tools only detect a small fraction of these QTLs. The QTLs and their genetic effects vary depending on the environment and the genetic background. Therefore, defining the adaptive genetic diversity as the whole set of QTLs potentially affecting fitness components, there is no strict frontier between neutral and adaptive genetic diversity: A gene polymorphism that is neutral in one context may become adaptive in another environmental or genetic context and vice versa. Thus, considering climate change and its uncertainties, quantification and characterization of both currently adaptive genetic diversity and currently neutral genetic diversity are needed. Classically, the overall genetic diversity is assessed with neutral markers (Buchert et al. 1997). Quantification and characterization of the current adaptive genetic diversity is assessed using phenotypic and molecular tools. Using only the phenotypic tools, the assessment of the genetic matrix of variances and covariances of adaptive traits requires known pedigrees and ex situ controlled experiments (preferable in trees for which deep pedigrees are not available in natural populations like in other organisms). Candidate gene polymorphisms, which include direct functional polymorphisms as well as linked markers, provide direct or indirect information on the diversity of target genes. When markers and phenotypic tools are combined, several approaches can be

conducted: QTL mapping in known pedigrees and controlled conditions, QTL association to traits in controlled conditions and in situ studies of QTL association to the environment (Neale and Savolainen 2004).

All kinds of molecular tools of known heredity may provide information on the neutral drivers of genetic changes: genetic drift, mating system and dispersal. More interesting is to investigate the changes in these drivers, which generally requires transgenerational sampling. Genome-wide markers as well as candidate genes polymorphisms are now commonly used to detect signature of past selection events by testing a departure of the diversity pattern from the neutral expectation. Monitoring ongoing selection processes requires to relate directly the trait to the fitness or at least to a performance trait (selection gradient β). With phenotypic tools alone, trait to performance mapping can be assessed ex situ with known pedigrees. When molecular and phenotypic tools are combined, in situ selection gradient studies can be performed.

5 Conclusion

The adaptive capacity of tree populations is potentially huge, and silviculture can have a significant impact on the rate of phenotypic and genetic change per generation: The

rate of change might probably be increased or reduced by a factor two depending on management interventions, which should not be neglected. The concept of evolution-oriented forestry that we introduced here does not pretend to allow for sufficient change in all cases. It should be considered as an option, with different associated benefits, risks and costs than those associated to the plantation strategy. Both strategies can also be combined. In any case, it is crucial to consider the potential evolutionary impact of silviculture when designing an adaptive forestry strategy. We proposed a simple framework to analyse and foresee the effects of forestry practice, and we identified a limited number of evolutionary processes and parameters that could be affected. Only few quantitative predictions can be made today, basically when evolutionary drivers are considered individually, and most expectations remain qualitative. Qualitative expectations can be used to draw research hypotheses. Quantitative predictions are needed to assess more precisely the cost effectiveness of forestry practice under various climate change scenarios. Quantitative comparison of evolution-oriented forestry with other baseline management options will require further process-based modelling and simulation studies for different forest types and species, different biotic and abiotic environments and different climate change scenarios. For a better understanding of the limits of the response to selection, we suggest to couple demogenetic models with biophysical models or host–parasite models.

One challenge for forestry decision-making under climate change is to reach a compromise between short-term and long-term objectives, e.g. speed-up the response to current selection pressure while preserving diversity and evolvability for uncertain future. A safe guideline is to favour natural selection for certainly adaptive traits, like drought resistance in the areas where more severe drought is expected, avoid random genetic erosion and increase genetic mixing. Over-selection for undue traits should also be avoided. As we have shown, each forestry practice has an effect on several evolutionary drivers (a geneticist would say pleiotropic effects) and interaction effects of different practices on a single evolutionary driver also exist (epistasis in genetic terms). To understand the global impact of forestry practice, in parallel to modelling approaches, long-term silvicultural options should be experimented. These new experiments will provide to the next generation of foresters very informative results, complementary to those obtained from comparative studies of existing situations.

Local decision should rely on a case by case approach taking into account each particular situation. Following quantitative genetics theory, we expect that in most cases genotypic diversity rather than allelic diversity constrains evolutionary changes. Can we obtain higher rate of change by driving recombination and selection within the current

population? Do we need to introduce allochthonous pre-adapted genotypes to accelerate the emergence and spread of adequate allelic combinations? To address these questions and make a diagnostic, combined genophenotypic approaches in situ are very promising and should be further developed. Molecular tools initially developed for model species now become available for non-model species. Similarly, methodologies and knowledge about the genetic architecture of traits and phenotype construction (e.g. genetic and environmental correlations, age–age correlations) should be generalized to non-model species to help better understanding the gene to trait and gene to fitness mapping.

Box 1—Raising-up complexity (1): the interplay between biotic interactions and evolutionary processes in forest trees

Trees interact with diverse mutualistic and antagonistic insect species. Mutualistic insects may be essential for plant reproduction during pollination, while antagonistic insects may be damageable to plants by consuming and removing plant parts and by selectively feeding on their reproductive tissues (Crawley 1989). The strength of such negative effects depends on the timing, the type and the amount of damage, as well as the stage of the plant's life cycle at which the damage occurs (Marquis 1992). By directly affecting tree reproduction or survival, the demographic and evolutionary consequences of the feeding activities of seed-specialized and tree-killing insect species are thus likely to differ from most forms of herbivory which only result in partial removal of tissues from individual plants (Hulme 1998).

Since seed predation leads to the eradication of individuals in a population, it plays a crucial role in plant population dynamics with possible genetic drift effects when population size is limited and potentially acts as a selective force driving the evolution of particular plant traits such as flowering synchrony, flowering phenology, inflorescence characteristics, flower size, flower longevity and mast seeding (Janzen 1971; Brody 1997; Fenner et al. 2002; Cariveau et al. 2004; Rose et al. 2005; Strauss and Whittall 2006). Many tree species suffer from large seed losses due to pre-dispersal seed predation, which can have significant effects on recruitment and plant population growth rate (Maron and Crone 2006; Kolb et al. 2007). However, the effect of such parasites on the long-term fitness of their host plant appears controversial (Crawley 1989; Horvitz and Schemske 2002), mainly due to the lack of data addressing this issue, especially on perennial plants. There is still a crucial need for studies examining the genetic consequences of massive seed losses within a host tree population, especially in a context in which pre-dispersal seed predation shows significant variation between trees. Indeed, seed loss due to seed-specialized chalcid wasps may vary from less than 1 to 100 % between trees (Roques 1981; Rappaport et al. 1993). At the tree population level, this raises the question of how such local variation in pre-dispersal seed predation may increase or, reversely, decrease the variance of effective seed set among trees, which influences the effective population size.

Evidence of insects directly acting as selective agents on forest trees is still lacking in the literature. Despite the extremely high tree mortality rates recorded during population outbreaks of the bark beetle *Dendroctonus ponderosae* Hopkins, many trees escape or survive bark beetle attacks, regardless of their vigour, age and/or size (Ott et al. 2011). Little is known about the heritability of tree traits involved in survival to bark beetle attacks such as resin acids (Baradat et al. 1978) and resin flow, viscosity and rate of crystallization (Nebeker et

al. 1992), in the exception of monoterpene production, which has been shown to be under strong genetic control (Ott et al. 2011).

Understanding how abiotic and biotic disturbances and tree dynamics are interdependent is also crucial for predicting the overall impact of parasitism on tree evolution. Indeed, severe abiotic changes such as droughts and/or heat waves may affect trees and parasites, as well as their interactions (Jactel et al. 2012). Successive drought episodes can affect directly tree survival (Allen et al. 2010), or indirectly when higher temperatures and lower tree resistance trigger severe forest insect outbreaks (OFEFP 2005; Netherer and Schopf 2010; Durand-Gillmann et al. 2012). The interdependence between climate, biotic factors and tree dynamics remains complex to predict. Drought induced changes in tree nutritional quality (water, carbohydrates and nitrogen contents) or in tree defence mechanisms can limit the development and the damages of parasites (Rouault et al. 2006; Jactel et al. 2012; Forkner et al. 2004). Extreme droughts may even be directly involved in the collapse of herbivorous populations at wide scales (Yames and Boecklen 2005). But drought can also affect negatively tree physiology and decrease the effectiveness of tree resistance mechanisms to pathogens and parasites (McIntyre et al. 1996).

Fire ecology provides interesting additional examples of the complexity of integrating interdependencies between trees, biotic and abiotic factors. Bark beetle outbreaks and forest fires have indeed jointly increased in extent and severity during the last decades, raising concerns about their possible interactions (Parker et al. 2006; Simard et al. 2011). Bark beetle outbreaks may increase the probability and intensity of active crown fire because they create great quantities of dead and ladder fuels (Brown 1975; McCullough et al. 1998). However, Simard et al. (2011) suggest that active crown fire are less probable in the short-term after outbreaks due to insect-driven stand thinning, while the probability of passive crown fire does not change in the short term but greatly increases in the decades following an outbreak. Thus, bark beetles are likely to indirectly affect non-attacked trees through subsequent enhanced fire risks. This clearly illustrates the critical need to integrate the possible interplay between the abiotic environment, biotic interactions and trees dynamics when designing forest management strategies.

Box 2—Raising-up complexity (2): evolution and silviculture in a highly spatially heterogeneous environment

In trees, the selection process is complex due to the long life-cycle and the high within-stand spatial environmental heterogeneity. Firstly, for long-lived and sessile organisms, different selection pressures may occur successively from the juvenile stage to the adult stage, e.g. selection for competing ability in a young dense regeneration vs. selection for stress resistance in the adult stage. Secondly, in heterogeneous environment, the phenotypic correlation between parents and offspring do not only depend on genetic control of the phenotype but also on the difference in environmental effects between the parents' and offsprings' sites. Finally, the environmental heterogeneity induces spatial variation in selection pressure, eventually leading to different selection pressure between parents and offsprings. Thus, evolution-oriented forestry should take environmental heterogeneity into account.

As a general objective, evolution-oriented silviculture should aim at favouring the mating success of the best growing trees located in the patches where the highest desired selection pressure occurs. In homogeneous conditions where the selection pressure is uniform, this objective is directly achieved by classical selective thinning. In heterogeneous conditions, this objective could be achieved if enough trees are selected for seeding within each patch where high selective

pressure occurs, even though these trees may have lower growth than neighbours growing in more favourable conditions.

In such context, assessing and mapping the environmental heterogeneity among patches is essential to avoid confusion between micro-environmental and genetic effects on the performance of the trees. It is crucial not only to assess the individual performance but also the patch conditions with synthetic indicators, easy to use and independent of the competition, such as site index, species composition of vegetation etc. Local variations of site index can be assessed through spatial patterns of tree height; therefore, it would be easier to identify and mark very early the 'good' phenotypes in highly selective areas, before a substantial reduction in number of trees happens due to other criteria.

Genetic improvement of the whole population will occur if these highly selected trees effectively contribute to the regeneration at stand level. If pollen or seed dispersal limits their effective contribution, considering here fecundity as included into the global dispersal process, it might be necessary to assist natural regeneration by local seed transfer from low to high selection patches within the stand.

Acknowledgments We particularly thank two anonymous reviewers and the editor for their relevant and useful comments on a previous version of the manuscript. The opinions expressed here benefitted from background experimental and modelling research funded by the EVOLTREE network of excellence (www.evoltree.eu), the ERA-Net Biodiversa LINKTREE project, the FRB VARIADAPT project, the ANR MACBI project, the Agropolis Foundation BIOFIS project as well as Région Provence Alpes Côte d'Azur and Département Santé des Forêts.

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

References

- Aguilar R, Quesada M, Ashworth L, Herrerias-Diego Y, Lobo J (2008) Genetic consequences of habitat fragmentation in plant populations: susceptible signals in plant traits and methodological approaches. *Mol Ecol* 17:5177–5188
- Aitken SN, Yeaman S, Holliday JA, Wang T (2008) Adaptation, migration or extirpation: climate change outcomes for tree populations. *Evol Appl* 1:95–111
- Alberto F, Aitken S, Alia A, Gonzáles-Martinez S, Hanninen H, Kremer A, Lefèvre F, Lenormand T, Yeaman S, Whetten R, Savolainen O (2013) Evolutionary response to climate change—evidence from tree populations. *Global Change Biol* (in press)
- Allen CD, Macalady AK, Chenchouni H, Bachelet D, McDowell N, Vennetier M, Kitzberger T, Rigling A, Breshears D, Hogg EH(Ted), Gonzalez P, Fensham R, Zhang Z, Castro J, Demidova N, Lim JH, Allard G, Running SW, Semerci A, Cobb N (2010) A global overview of drought and heat-induced tree mortality reveals emerging climate change risks for forests. *For Ecol Manag* 259:660–684
- Amm A, Pichot C, Dreyfus P, Davi H, Fady B (2012) Improving the estimation of landscape scale seed dispersal by integrating seedling recruitment. *Ann For Sci* 69:845–856
- Aravanopoulos FA (2011) Genetic monitoring in natural perennial plant populations. *Botany* 89:75–81

- Ashley M (2010) Plant parentage, pollination, and dispersal: how DNA microsatellites have altered the landscape. *Crit Rev Plant Sci* 29:148–161
- Ayari A, Zubizarreta-Gerendiain A, Tome M, Tome J, Garchi S, Henchi B (2012) Stand, tree and crown variables affecting cone crop and seed yield of Aleppo pine forests in different bioclimatic regions of Tunisia. *Forest Syst* 21:128–140
- Baradat P, Marpeau A, Bernard-Dagan C (1978) Variation of terpenes within and between populations of maritime pine. In: Rudin D (ed) *Biochemical genetics of forest trees*. Swedish University of Agricultural Sciences, Umea
- Bontemps A (2012) Potentiel évolutif d'une population de hêtre commun sur le Mont Ventoux. Ph.D. dissertation, University Paul Cézanne, Aix-Marseille III, Aix en Provence
- Bossdorf O, Richards CL, Pigliucci M (2008) Epigenetics for ecologists. *Ecol Lett* 11:106–115
- Bouffier L, Charlot C, Raffin A, Rozenberg P, Kremer A (2008) Can wood density be efficiently selected at early stage in maritime pine (*Pinus pinaster* Ait.)? *Ann For Sci* 65:106p2–106p8
- Bradshaw AD (1991) Genostasis and the limits to evolution. *Philos T R Soc B* 333:289–305
- Brody AK (1997) Effects of pollinators, herbivores, and seed predators on flowering phenology. *Ecology* 78:1624–1631
- Brown JK (1975) Fire cycles and community dynamics in lodgepole pine forests. In: Baumgartner DM (ed) *Symposium Proceedings: Management of lodgepole pine ecosystems*. Washington State University Cooperative Extension Service, Pullman, pp 430–456
- Brown A, Young A, Burdon J, Christidis L, Clarke G, Coates D, Sherwin W (1997) Genetic indicators for state of the environment reporting, Australia: State of the Environment Technical Paper Series (Environmental indicators). Department of the Environment, Sport and Territories, Canberra
- Buchert GP, Rajora OP, Hood JV, Dancik BP (1997) Effects of harvesting on genetic diversity in old-growth eastern white pine in Ontario, Canada. *Conserv Biol* 11:747–758
- Burczyk J, Adams WT, Moran GF, Griffins AR (2002) Complex patterns of mating revealed in a *Eucalyptus regnans* seed orchard using allozyme markers and the neighbourhood model. *Mol Ecol* 11:2379–2391
- Burczyk J, Adams WT, Birkes DS, Chybicki IJ (2006) Using genetic markers to directly estimate gene flow and reproductive success parameters in plants on the basis of naturally regenerated seedlings. *Genetics* 173:363–372
- Burke KL (2011) The effects of logging and disease on American chestnut. *For Ecol Manag* 261:1027–1033
- Cariveau D, Rebecca EI, Brody AK, Garcia-Mayeya LS, Ohe A (2004) Direct and indirect effects of pollinators and seed predators to selection on plant and floral traits. *Oikos* 104:15–26
- Carter AJR, Hermisson J, Hansen TF (2005) The role of epistatic gene interactions in the response to selection and the evolution of evolvability. *Theor Popul Biol* 68:179–196
- Charmantier A, Garant D (2005) Environmental quality and evolutionary potential: lessons from wild populations. *P Roy Soc B - Biol Sci* 272:1415–1425
- Chevin L-M, Collins S, Lefèvre F (2012) Phenotypic plasticity and evolutionary demographic responses to climate change: taking theory out to the field. *Funct Ecol*. doi:10.1111/j.1365-2435.2012.02043.x
- Chevin LM, Lande R (2011) Adaptation to marginal habitats by evolution of increased phenotypic plasticity. *J Evolution Biol* 24:1462–1476
- Chevin L-M, Lande R, Mace GM (2010) Adaptation, plasticity, and extinction in a changing environment: towards a predictive theory. *PLoS Biology* 8:e1000357
- Couvet D, Ronfort J (1994) Mutation load depending on variance in reproductive success and mating system. In: Loeschke V, Tomiuk J, Jain SK (eds) *Conservation genetics*. Birkhauser, Basel, pp 55–68
- Crawley MJ (1989) Insect herbivores and plant population dynamics. *Annu Rev Entomol* 34:531–564
- Dounavi K, Steiner W, Maurer W (2002) Effects of different silvicultural treatments on the genetic structure of European beech populations (*Fagus sylvatica* L.). *Manag For Ecosyst* 4:81–90
- Durand-Gillmann M, Cailleret M, Boivin T, Nageleisen LM, Davi H (2012) Individual vulnerability of Silver fir (*Abies alba* Mill.) to parasitism by two contrasting biotic agents: mistletoe (*Viscum album* L. ssp. *abietis*) and bark beetles (*Coleoptera: Curculionidae: Scolytinae*) during a decline process. *Ann For Sci*. doi:10.1007/s13595-012-0251-y
- Eckert AJ, Bower AD, Gonzalez-Martinez SC, Wegrzyn J, Coop G, Neale DB (2010) Back to nature: ecological genomics of loblolly pine (*Pinus taeda*, *Pinaceae*). *Mol Ecol* 19:3789–3805
- El-Kassaby YA, Benowicz A (2000) Effects of commercial thinning on genetic, plant species and structural diversity in second growth Douglas-fir (*Pseudotsuga menziesii* (Mirb.) Franco) stands. *For Genet* 7:193–203
- Endler JA (1986) Natural selection in the wild. *Monographs in population biology* 21. Princeton University Press, Princeton
- Falconer DS (1960) *Introduction to quantitative genetics*. Oliver and Boyd, Edinburgh
- Fenner M, Cresswell JE, Hurley RA, Baldwin T (2002) Relationship between capitulum size and predispersal seed predation by insect larvae in common *Asteraceae*. *Oecologia* 130:72–77
- Fernandez-M JF, Sork VL (2005) Mating patterns of a subdivided population of the Andean oak (*Quercus humboldtii* Bonpl., *Fagaceae*). *J Hered* 96:635–643
- Forkner RE, Marquis RJ, Lill JT (2004) Feeny revisited: condensed tannins as anti-herbivore defences in leaf-chewing. *Herbivore communities of Quercus*. *Ecol Entomol* 29:174–187
- Forrester DI, Collopy JJ, Beadle CL, Warren CR, Baker TG (2012) Effect of thinning, pruning and nitrogen fertiliser application on transpiration, photosynthesis and water-use efficiency in a young *Eucalyptus nitens* plantation. *For Ecol Manag* 266:286–300
- Futuyma DJ (2010) Evolutionary constraint and ecological consequences. *Evolution* 64:1865–1884
- García C, Arroyo JM, Godoy JA, Jordano P (2005) Mating patterns, pollen dispersal, and the ecological maternal neighbourhood in a *Prunus mahaleb* L. population. *Mol Ecol* 14:1821–1830
- García C, Jordano P, Arroyo JM, Godoy JA (2009) Maternal genetic correlations in the seed rain: effects of frugivore activity in heterogeneous landscapes. *J Ecol* 97:1424–1435
- Gilpin ME, Soule ME (1986) Minimum viable populations: processes of species extinction. In: Soule ME (ed) *Conservation biology: the science of scarcity and diversity*. Sinauer Associates, Sunderland, pp 19–34
- Gingerich PD (2009) Rates of evolution. *Annu Rev Ecol Evol S* 40:657–675
- Gomulkiewicz R, Holt RD (1995) When does evolution by natural selection prevent extinction? *Evolution* 49:201–207
- Goodell K, Elam DE, Nason JD, Ellstrand NC (1997) Gene flow among small populations of a self-incompatible plant: an interaction between demography and genetics. *Am J Bot* 84:1362–1371
- Grivet D, Sebastiani F, Alía R, Bataillon T, Torre S, Zabal-Aguirre M, Vendramin GG, González-Martínez SC (2011) Molecular footprints of local adaptation in two Mediterranean conifers. *Mol Biol Evol* 28:101–116
- Gwaze DP, Bridgewater FE, Byram TD, Woolliams JA, Williams CG (2000) Predicting age-age correlations in tree-breeding programs: a case study in *Pinus taeda* L. *Theor Appl Genet* 100:199–206
- Hannrup B, Ekberg (1998) Age-age correlations for tracheid length and wood density in *Pinus sylvestris*. *Can J For Res* 28:1373–1379

- Hansen MM, Olivieri I, Waller DM, Nielsen EE, the GeM Working Group (2012) Monitoring adaptive genetic responses to environmental change. *Mol Ecol* 21:1311–1329
- Hardy O, Gonzales-Martinez SC, Colas B, Fréville H, Mignot A, Olivieri I (2004) Fine-scale genetic structure and gene dispersal in *Centaurea corymbosa* (Asteraceae). II. Correlated paternity within and among sibships. *Genetics* 168:1601–1614
- Hoffmann AA, Sgrò CM (2011) Climate change and evolutionary adaptation. *Nature* 470:479–485
- Holliday JA, Ritland K, Aitken SN (2010) Widespread, ecologically relevant genetic markers developed from association mapping of climate-related traits in Sitka spruce (*Picea sitchensis*). *New Phytol* 188:501–514
- Horvitz CC, Schemske DW (2002) Effects of plant size, leaf herbivory, local competition and fruit production on survival, growth and future reproduction of a neotropical herb. *J Ecol* 90:279–290
- Hulme PE (1998) Post-dispersal seed predation: consequences for plant demography and evolution. *Perspect Plant Ecol* 1:32–46
- Hussendörfer E, Konnert M (2000) Impact of forest management on genetic variation of Silver fir and European beech populations. *For Landsc Res* 75:187–204
- IPCC (2007) Climate change 2007: AR4 synthesis report. An assessment of the Intergovernmental Panel on Climate Change, adopted at IPCC Plenary XXVII, Valencia, Spain, 12–17/11/2007
- Jactel H, Petit RJ, Desprez-Loustau M-L, Delzon S, Piou D, Battisti A, Koricheva J (2012) Drought effects on damage by forest insects and pathogens: a meta-analysis. *Global Change Biol* 195:267–276
- Janzen DH (1971) Seed predation by animals. *Annu Rev Ecol S* 2:465–492
- Kircher M, Kelso J (2010) High-throughput DNA sequencing? Concepts and limitations. *BioEssays* 32:524–536
- Klein EK, Carpentier FH, Oddou-Muratorio S (2011) Estimating the variance of male fecundity from genotypes of progeny arrays: evaluation of the Bayesian forward approach. *Methods Ecol Evol* 2:349–361
- Klein EK, Lavigne C, Gouyon PH (2006) Mixing of propagules from discrete sources at long distance: comparing a dispersal tail to an exponential. *BMC Ecol* 6:3
- Kolb A, Ehrlén J, Eriksson O (2007) Ecological and evolutionary consequences of spatial and temporal variation in pre-dispersal seed predation. *Perspect Plant Ecol* 9:79–100
- Konnert M, Hussendörfer E (2001) Genetic variation of silver fir (*Abies alba*) in unevenaged forests (Plenter forest) in comparison with evenaged forests (Altersklassenwald). *Forest Sci* 70:307–320
- Koski V, Skrøppa T, Paule L, Wolf H, Turok J (1997) Technical guidelines for genetic conservation of Norway spruce (*Picea abies* (L.) Karst.). International Plant Genetic Resources Institute, Rome
- Kramer K, Buiteveld J, Forstreuter M, Geburek T, Leonardi S, Menozzi P, Povillon F, Schelhaas MJ, Teissier du Cros E, Vendramin GG, van der Werf DC (2008) Bridging the gap between ecophysiological and genetic knowledge to assess the adaptive potential of European beech. *Ecol Model* 216:333–353
- Kremer A, Le Corre V (2012) Decoupling of differentiation between traits and their underlying genes in response to divergent selection. *Heredity* 108:375–385
- Kremer A, Ronce O, Robledo-Arnuncio JJ, Guillaume F, Bohrer G, Nathan R, Bridle JR, Gomulkiewicz R, Klein EK, Ritland K, Kuparinen A, Gerber S, Schueler S (2012) Long-distance gene flow and adaptation of forest trees to rapid climate change. *Ecol Lett* 15:378–392
- Krouchi F, Derridj A, Lefèvre F (2004) Year and tree effect on reproductive organisation of *Cedrus atlantica* in a natural forest. *For Ecol Manag* 197:181–189
- Kuparinen A, Savolainen O, Schurr FM (2010) Increased mortality can promote evolutionary adaptation of forest trees to climate change. *For Ecol Manag* 259:1003–1008
- Lande R (2009) Adaptation to an extraordinary environment by evolution of phenotypic plasticity and genetic assimilation. *J Evolution Biol* 22:1435–1446
- Lande R, Arnold SJ (1983) The measurement of selection on correlated characters. *Evolution* 37:1210–1226
- Lauber U, Rotach P, Hussendörfer E (1997) The influence of silvicultural interventions on the genetic structure of a young beech stand (*Fagus sylvatica* L.). *Schweiz Z Forstwes* 148:847–862
- Lefèvre F (2004) Human impacts on forest genetic resources in the temperate zone: an updated review. *For Ecol Manag* 197:257–271
- Lefèvre F, Kajba D (2001) Indicators for monitoring genetic diversity. In: Lefèvre F, Barsoum N, Heinze B, Kajba D, Rotach P, de Vries S, Turok J (eds) *In situ* conservation of *Populus nigra*. International Plant Genetic Resources Institute, Rome, pp 36–46
- Legay M, Mortier F (2005) La forêt face au changement climatique: adapter la gestion forestière—Prise en compte dans les documents d'orientation de la gestion forestière. Synthèse de l'atelier ONF/INRA du 20 octobre 2005. Office National des Forêts, Paris. Les Dossiers Forestiers n°16
- Le Rouzic A, Siegel PB, Carlberg O (2007) Phenotypic evolution from genetic polymorphisms in a radial network architecture. *BMC Biol* 5:50. doi:10.1186/1741-7007-5-50
- Maron JL, Crone E (2006) Herbivory: effects on plant abundance, distribution and population growth. *P Roy Soc B* 273:2575–2584
- Marquis RJ (1992) The selective impact of herbivores. In: Fritz RS, Simms EL (eds) Plant resistance to herbivores and pathogens. University of Chicago Press, Chicago, pp 301–325
- Matheson AC, Wu HX, Spencer DJ, Raymond CA, Griffin AR (2002) Inbreeding in *Pinus radiata* III. The effect of inbreeding on age–age correlation and early selection efficiency. *Silvae Genet* 51:115–122
- McCullough DG, Werner RA, Neumann D (1998) Fire and insects in northern and boreal ecosystems of North America. *Annu Rev Entomol* 43:107–127
- McDonald SE, Thomas BR, Cherniawsky DM, Purdy BG (2001) Managing genetic resources of lodgepole pine in west-central Alberta: patterns of isozyme variation in natural populations and effects of forest management. *For Ecol Manag* 152:45–58
- McIntyre GA, Jacobi WR, Ramaley AW (1996) Factors affecting *Cytospora* canker occurrence on aspen. *J Arboric* 22:229–233
- Metzker ML (2009) Sequencing technologies—the next generation. *Nat Rev Genet* 11:31–46
- Milleron M, Lopez de Heredia U, Lorenzo Z, Perea R, Dounavi A, Alonso J, Gil L, Nanos N (2012) Effect of canopy closure on pollen dispersal in a wind-pollinated species (*Fagus sylvatica* L.). *Plant Ecol* 213:1715–1728
- Mimura M, Aitken SN (2007) Adaptive gradients and isolation-by-distance with postglacial migration in *Picea sitchensis*. *Heredity* 99:224–232
- Misson L, Nicault A, Guiot J (2003) Effects of different thinning intensities on drought response in Norway spruce (*Picea abies* (L.) Karst.). *For Ecol Manage* 183:47–60
- Moose SP, Dudley JW, Rocheford TR (2004) Maize selection passes the century mark: a unique resource for 21st century genomics. *Trends Plant Sci* 9:358–364
- Mosseler A, Major JE, Rajora OP (2003) Old-growth red spruce forests as reservoirs of genetic diversity and reproductive fitness. *Theor Appl Genet* 106:151–161
- Namkoong G, Boyle T, Gregorius HR, Joly H, Savolainen O, Ratnam W, Young A (1996) Testing criteria and indicators for assessing the sustainability of forest management: genetic criteria and indicators. Center for International Forestry Research CIFOR, Bogor. Working Paper N°10
- Neale D, Savolainen O (2004) Association genetics of complex traits in conifers. *Trends Plant Sci* 9:325–330

- Nebeker TE, Hodges JD, Blanche CA, Honea CR, Tisdale RA (1992) Variation in the constitutive defensive system of Loblolly pine in relation to bark beetle attack. *Forest Sci* 38:457–466
- Netherer S, Schopf A (2010) Potential effects of climate change on insect herbivores—general aspects and a specific example (Pine processionary moth, *Thaumetopoea pityocampa*). *For Ecol Manag* 259:831–838
- Nilsson JE (1995) Genetic variation in the natural pollen cloud of *Pinus sylvestris*: a study based on progeny testing. *Scand J Forest Res* 10:140–148
- Oddou-Muratorio S, Demesure-Musch B, Pélissier R, Gouyon PH (2004) Impacts of gene flow and logging history on the local genetic structure of a scattered tree species, *Sorbus torminalis* L. Crantz. *Mol Ecol* 13:3689–3702
- Oddou-Muratorio S, Klein EK, Austerlitz F (2005) Pollen flow in the wildservice tree, *Sorbus torminalis* (L.) Crantz. II. Pollen dispersal and heterogeneity in mating success inferred from parent–offspring analysis. *Mol Ecol* 14:4441–4452
- OFEFP, Rapport forestier (2005) Faits et chiffres sur l'état de la forêt suisse. Office fédéral de l'environnement, des forêts et du paysage, Berne
- Osório J, Pereira JS (1994) Genotypic differences in water-use efficiency and ^{13}C discrimination in *Eucalyptus globulus*. *Tree Physiol* 14:871–882
- Ott DS, Yanchuk AD, Huber DPW, Wallin KF (2011) Genetic variation of lodgepole pine, *Pinus contorta* var. *latifolia*, chemical and physical defenses that affect mountain pine beetle, *Dendroctonus ponderosae*, attack and tree mortality. *J Chem Ecol* 37:1002–1012
- Parker TJ, Clancy KM, Mathiasen RL (2006) Interactions among fire, insects and pathogens in coniferous forests of the interior western United States and Canada. *Agri For Entomol* 8:167–189
- Pease CM, Lande R, Bull JJ (1989) A model of population growth, dispersal and evolution in a changing environment. *Ecology* 70:1657–1664
- Petit RJ, Hampe A (2006) Some evolutionary consequences of being a tree. *Annu Rev Ecol Evol S* 37:187–214
- Pichot C, Bastien C, Courbet F, Demesure-Musch B, Dreyfus P, Fady B, Frascaria-Lacoste N, Gerber S, Lefèvre F, Morand-Prieur ME, Oddou S, Teissier du Cros E, Valadon A (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 n°6, Bureau des Ressources Génétiques, Paris, pp 277–297
- Pigliucci M (2008) Is evolvability evolvable? *Nat Rev Genet* 9:75–82
- Prat D, Faivre-Rampant P, Prado E (2006) Analyse du génome et gestion des ressources génétiques forestières. INRA ed Quae, Paris
- Rajora OP, Rahman MH, Buchert GP, Dancik BP (2000) Microsatellite DNA analysis of genetic effects of harvesting in old-growth eastern white pine (*Pinus strobus*) in Ontario, Canada. *Mol Ecol* 9:339–348
- Ramirez-Valiente JA, Valladares F, Huertas AD, Granados S, Aranda I (2011) Factors affecting cork oak growth under dry conditions: local adaptation and contrasting additive genetic variance within populations. *Tree Genet Genomes* 7:285–295
- Rappaport NG, Mori S, Roques A (1993) Estimating effect of *Megastigmus spermotrophus* (Hymenoptera: Torymidae) on Douglas-fir seed production: the new paradigm. *J Econ Entomol* 86:845–849
- Rehfeldt GE, Wykoff WR, Ying CC (2001) Physiologic plasticity, evolution, and impacts of a changing climate on. *Clim Chang* 50:355–376
- Restoux G, Silva ED, Sagnard F, Torre F, Klein E, Fady B (2008) Life at the margin: the mating system of Mediterranean conifers. *Web Ecology* 8:94–102
- Ritland K (1996) A marker-based method for inferences about quantitative inheritance in natural population. *Evolution* 50:1062–1073
- Robledo-Arnuncio JJ (2011) Wind pollination over mesoscale distances: an investigation with Scots pine. *New Phytol* 190:222–233
- Robledo-Arnuncio JJ, Alía R, Gil L (2004) Increased selfing and correlated paternity in a small population of a predominantly outcrossing conifer, *Pinus sylvestris*. *Mol Ecol* 13:2567–2577
- Rockman MV (2012) The QTN program and the alleles that matter for evolution: all that's gold does not glitter. *Evolution* 66:1–17
- Roques A (1981) Biologie et répartition de *Megastigmus spermotrophus* Wachtl. (Hymenoptera: Chalcidoidea Torymidae) et des autres insectes liés aux cônes dans les peuplements forestiers et vergers à graines français de sapin de Douglas *Pseudotsuga menziesii* [Mirb.] Franco. *Acta Oecol* 2:161–180
- Rose KE, Louda SM, Rees M (2005) Demographic and evolutionary impacts of native and invasive insect herbivores on *Cirsium canescens*. *Ecology* 86:453–465
- Rouault G, Candau JN, Lieutier F, Nageleisen L-M, Martin JC, Warzée N (2006) Effects of drought and heat on forest insect populations in relation to the 2003 drought in Western Europe. *Ann For Sci* 63:613–624
- Sáenz-Romero C, Guzmán-Reyna R, Rehfeldt G (2006) Altitudinal genetic variation among *Pinus oocarpa* populations in Michoacan, Mexico. Implications for seed zoning, conservation, tree breeding and global warming. *For Ecol Manag* 229:340–350
- Sagnard F, Oddou-Muratorio S, Pichot C, Vendramin GG, Fady B (2011) Effects of seed dispersal, adult tree and seedling density on the spatial genetic structure of regeneration at fine temporal and spatial scales. *Tree Genet Genomes* 7:37–48
- Savolainen O, Pyhäjärvi T, Knürr T (2007) Gene flow and local adaptation in trees. *Annu Rev Ecol Evol S* 38:595–619
- Schleuning M, Farwig N, Peters MK, Bergsdorf T, Bleher B, Brandl R, Dalitz H, Fischer G, Freund W, Gikungu MW, Hagen M, Garcia FH, Kagezi GH, Kaib M, Kraemer M, Lung T, Naumann CM, Schaab G, Templin M, Uster D, Wagele JW, Bohning-Gaese K (2011) Forest fragmentation and selective logging have inconsistent effects on multiple animal-mediated ecosystem processes in a tropical forest. *PLoS One* 6:e27785
- Schoen D, Brown HD (1991) Intraspecific variation in population gene diversity and effective population size correlates with the mating system in plants. *P Natl Acad Sci USA* 88:4494–4497
- Silvertown JW (1980) The evolutionary ecology of mast seeding in trees. *Biol J Linn Soc* 14:235–250
- Simard M, Romme WH, Griffin JM, Turner MG (2011) Do mountain pine beetle outbreaks change the probability of active crown fire in lodgepole pine forests? *Ecol Monogr* 81:3–24
- Skrøppa T, Tollefsrud M, Sperisen C, Johnsen Ø (2010) Rapid change in adaptive performance from one generation to the next in *Picea abies*—central European trees in a Nordic environment. *Tree Genet Genomes* 6:93–99
- Smouse P, Dryer RJ, Westfall RD, Sork VL (2001) Two-generation analysis of pollen flow across a landscape. I. Male gamete heterogeneity among females. *Evolution* 55:260–271
- Smouse P, Sork VL (2004) Measuring pollen flow in forest trees: an exposition of alternative approaches. *For Ecol Manag* 197:21–38
- St Clair JB, Howe GT (2007) Genetic maladaptation of coastal Douglas-fir seedlings to future climates. *Global Change Biol* 13:1441–1454
- Strauss SY, Whittall JB (2006) Non-pollinator agents of selection on floral traits. In: Harder LD, Barrett SCH (eds) *Ecology and evolution of flowers*. Oxford University Press, Oxford, pp 120–138
- Tene A, Tobin B, Dyckmans J, Ray D, Black K, Nieuwenhuis M (2011) Assessment of tree response to drought: validation of a methodology to identify and test proxies for monitoring past environmental changes in trees. *Tree Physiol* 31:309–322
- Thiebaut B, Comps B, Leroux A (1992) Relation hauteur-génotype dans une régénération naturelle de hêtre (*Fagus sylvatica*) équiète et âgée de 18 ans. *Ann Sci For* 49:321–335

- Walsh B, Blows MW (2009) Abundant genetic variation+strong selection=multivariate genetic constraints: a geometric view of adaptation. *Annu Rev Ecol Evol S* 40:41–59
- Wright S (1931) Evolution in Mendelian populations. *Genetics* 16:97–158
- Wu HX, Powell MB, Yang JL, Ivkovic M, McRae TA (2007) Efficiency of early selection for rotation-aged wood quality traits in radiata pine. *Ann For Sci* 64:1–9
- Yan H, Bi HQ, Li R, Eldridge R, Wu Z, Li Y, Simpson J (2006) Assessing climatic suitability of *Pinus radiata* (D. Don) for summer rainfall environment of southwest China. *For Ecol Manag* 234:199–208
- Yarnes C, Boecklen WJ (2005) Abiotic factors promote plant heterogeneity and influence herbivore performance and mortality in Gambel's Oak (*Quercus gambelii*, Nutt.). *Entomol Exp Appl* 114:87–95
- Yousefpour R, Jacobsen JB, Thorsen BJ, Meilby H, Hanewinkel M, Oehler K (2012) A review of decision-making approaches to handle uncertainty and risk in adaptive forest management under climate change. *Ann For Sci* 69:1–15

RÉSUMÉ DE THÈSE

La rapidité du changement climatique observé et prédit soulève la question de la vitesse à laquelle les espèces pourront s'adapter au climat futur. Les populations exposées aux changements de conditions environnementales peuvent s'adapter sur place (sans migration) grâce, dans un premier temps, à la réponse plastique des individus, puis, à long terme, par la réponse à la sélection (adaptation génétique). En situation d'environnement variable à la fois dans l'espace et le temps, les flux de gènes peuvent faciliter la diffusion d'allèles bénéfiques entre populations. L'objectif de cette thèse est d'étudier expérimentalement l'interaction entre adaptation et dispersion pollinique à longue distance sur un gradient altitudinal de hêtre commun (*Fagus sylvatica*). Cette étude a été conduite à partir de trois populations de hêtre sur le versant Nord du Mont-Ventoux (de 900 m à 1400 m d'altitude). Douze traits fonctionnels (phénologiques, physiologiques et morphologiques) potentiellement adaptatifs ont été mesurés sur 60 descendances maternelles issues de ces populations (20 mères/population, 100 individus/mère) et placées en pépinière. Les résultats de cette thèse montrent que (i) les forts taux de pollen immigrant reçus par les populations ($m = 56\%$) n'empêchent pas l'émergence de patrons de différenciation adaptative, (ii) la sélection érode la variance génétique des traits, (iii) malgré l'existence de variance génétique pour tous les traits, les covariances génétiques entre traits peuvent affecter les taux et directions d'évolution future, et (iv) les flux de pollen contribuent à augmenter la variance génétique totale au sein des populations. Finalement, les populations à moyenne et hautes altitudes pourraient avoir les capacités de s'adapter au climat futur.

The rapidity of the climate change observed and predicted raises the issue of how fast species can adapt to future. In face of environmental change, populations can adapt without migrating thanks to their plastic response in a first time, and to their response to selection in the long term (genetic adaptation). In conditions of environment variable both in space and time, gene flow can enhance the spread of beneficial alleles between populations. The objective of this thesis is to experimentally measure the adaptive differentiation and the long-distance pollen dispersal along an elevational gradient of beech populations (*Fagus sylvatica*). In this study we investigated three beech populations located on the northern side of the Mont-Ventoux (from 900m to 1400m of elevation). Twelve potential adaptive functional traits (phenological, physiological and morphological) have been measured on 60 maternal families collected from these populations (20 mothers/population, 100 offspring/mother) and planted in nursery. The results of this thesis show that that (i) the strong pollen immigrant rate received by the populations ($m = 56\%$) did not their prevent adaptive differentiation, (ii) the action of natural selection erodes the genetic variance of the traits, (iii) despite significant genetic variance for all the traits, the genetic covariance between traits should affect the rate and direction of evolution in the future, and (iv) gene flow contribute to increase the total genetic variance within the populations. Finally, populations at medium and high elevations may have the potential to adapt to future climate.