



HAL
open science

Flux de gènes et arbres forestiers

Sophie Gerber

► **To cite this version:**

Sophie Gerber. Flux de gènes et arbres forestiers. Sciences du Vivant [q-bio]. Université des Sciences et Technologies (Bordeaux 1), 2006. tel-02813912

HAL Id: tel-02813912

<https://hal.inrae.fr/tel-02813912>

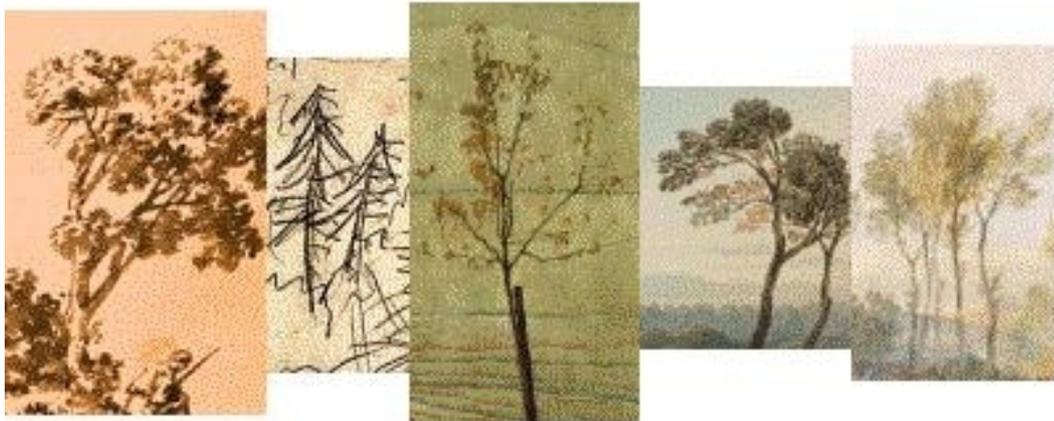
Submitted on 23 Aug 2021

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.

UNIVERSITÉ BORDEAUX I
HABILITATION À DIRIGER DES RECHERCHES

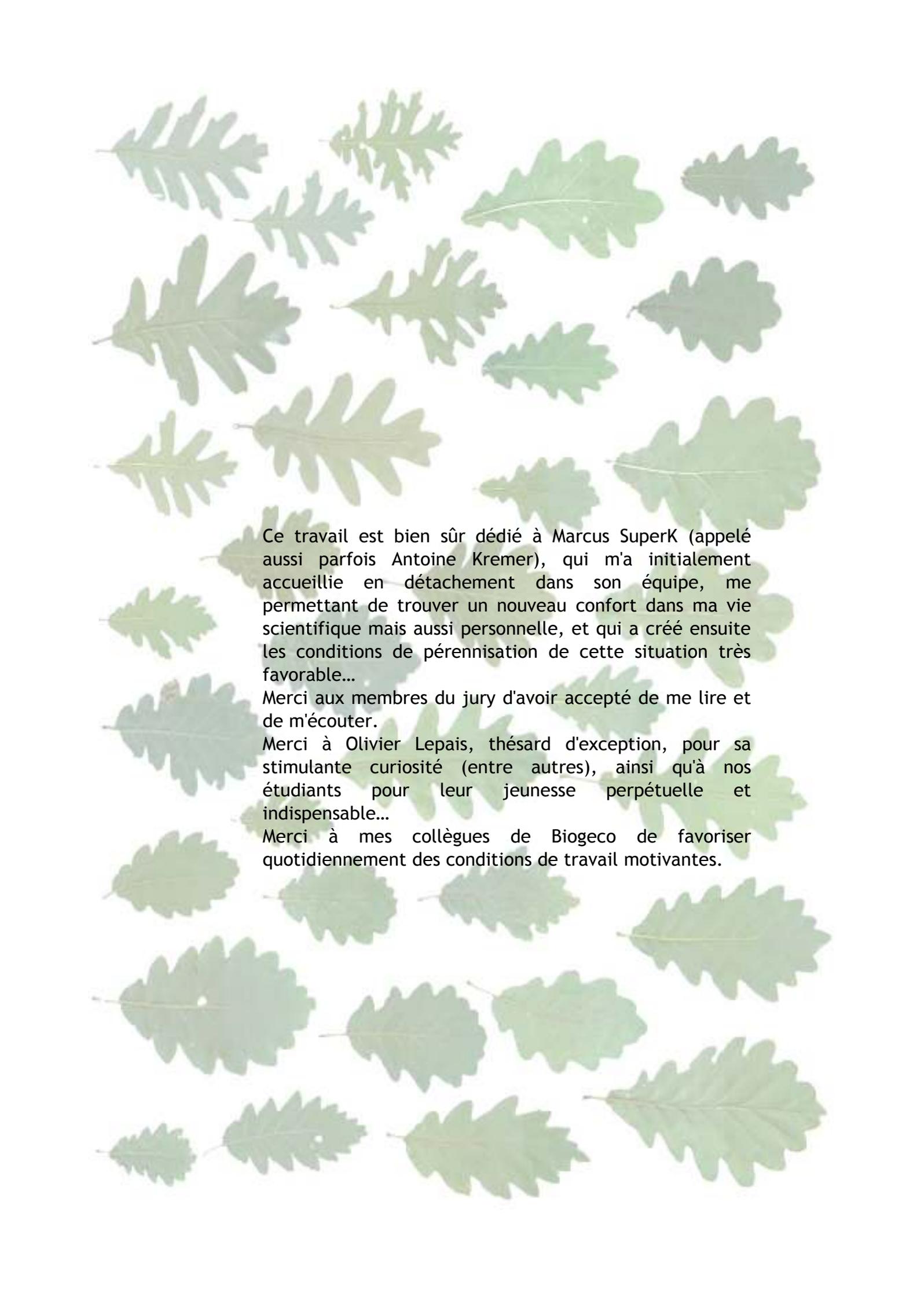
Flux de gènes et arbres forestiers



Sophie GERBER-DANJON, chargée de recherche INRA

Habilitation présentée le 14 décembre 2006 devant le jury composé de :

Didier Alard	Professeur des Universités, Bordeaux	examineur
Brigitte Crouau-Roy	Professeur des Universités, Toulouse	examinatrice
François Lefèvre	Directeur de recherche INRA, Avignon	rapporteur
Xavier Vekemans	Professeur des Universités, Lille	rapporteur
Frédérique Viard	Chargée de Recherche CNRS, Roscoff	rapporteure



Ce travail est bien sûr dédié à Marcus SuperK (appelé aussi parfois Antoine Kremer), qui m'a initialement accueillie en détachement dans son équipe, me permettant de trouver un nouveau confort dans ma vie scientifique mais aussi personnelle, et qui a créé ensuite les conditions de pérennisation de cette situation très favorable...

Merci aux membres du jury d'avoir accepté de me lire et de m'écouter.

Merci à Olivier Lepais, thésard d'exception, pour sa stimulante curiosité (entre autres), ainsi qu'à nos étudiants pour leur jeunesse perpétuelle et indispensable...

Merci à mes collègues de Biogeco de favoriser quotidiennement des conditions de travail motivantes.

SOMMAIRE

I. SYNTHÈSE DES ACTIVITÉS SCIENTIFIQUES.....	4
1. Introduction.....	4
1.1. Génétique de la qualité de la graine du soja.....	4
1.2. Méta-analyse de locus à effet quantitatif (QTL).....	5
2. Flux de gènes et arbres forestiers.....	6
2.1. Contexte.....	6
2.1.1. Les flux de gènes pourquoi, comment.....	6
2.1.2. Une brève histoire de la recherche de paternité.....	7
2.1.3. Arbres et flux de gènes.....	8
2.2. Développements méthodologiques.....	9
2.2.1. Mesurer la qualité d'un système de marquage : les probabilités d'exclusion..	9
2.2.1.1. Principe.....	9
2.2.1.2. Comparaison des marqueurs microsatellites et AFLP.....	10
2.2.2. Recherche de paternité et de parenté.....	11
2.3. Populations naturelles et flux de gènes.....	12
2.3.1. Structure génétique fine, flux de graines et fratries.....	12
2.3.2. Nombre de pères.....	13
2.4. Populations améliorées et flux de gènes.....	15
2.4.1. Test variétal et marqueurs chez le pin maritime.....	15
2.4.2. Fonctionnement des vergers à graines.....	16
2.4.3. Apparentement chez les vignes cultivées.....	19
2.5. Paternité et parenté chez les chênes européens : le projet OAKFLOW.....	19
2.5.1. Principe.....	20
2.5.2. Premiers résultats.....	20
2.5.3. Perspectives.....	21
2.6. Génétique, chênes et paysages.....	21
3. Perspectives : flux de gènes intra- et interspécifiques.....	23
3.1. Flux de gènes dans le complexe d'espèces des chênes blancs.....	23
3.2. Approches méthodologiques.....	26
3.2.1. Flux de gènes instantanés intra- et inter-population.....	26
3.2.2. Méthodes de reconstitution des apparentements sans pedigree.....	27
3.3. Interactions biotiques et biodiversité.....	28
3.3.1. Dispersion des chênes sous l'effet des geais.....	28
3.3.2. Dispersion des graines : génétique et pathologie.....	29
3.4. Une école-chercheur "philosophie et histoire des sciences biologiques".....	30
4. En guise de conclusion.....	32
5. Références.....	33
II. LISTE DES PUBLICATIONS.....	37
III. CURRICULUM VITAE.....	44

I. SYNTHÈSE DES ACTIVITÉS SCIENTIFIQUES

1. Introduction

L'idée selon laquelle l'hérédité est liée à des éléments discrets est issue des expériences de Mendel (prêtre-botaniste autrichien, 1822-1884) qui montre en 1865 que des caractères contrastés de deux lignées de pois réapparaissent dans des proportions bien définies à l'issue de l'intercroisement de leurs descendants, sans mélange. L'importance de cette découverte n'apparaît clairement que vers 1900, et l'usage associé des probabilités s'impose rapidement. Parallèlement, Galton (anthropologue anglais, 1822-1911) étudie l'hérédité de caractères continus chez l'homme de façon quantitative, suivi en cela par Pearson (mathématicien anglais, 1857-1936). Deux écoles de pensée s'opposent alors, l'une fondée sur les lois de Mendel, et l'autre sur des méthodes biométriques. Fisher (statisticien anglais, 1890-1962) suggère la compatibilité des deux approches en 1918 (Bailey 1961). Les variations discontinues et continues transmises par l'hérédité ont en effet en commun les propriétés de ségrégation, de dominance, d'interaction et de linkage (Thoday 1961). Le rapprochement des génétiques quantitatives et mendéliennes était déjà d'actualité.

En choisissant de m'orienter vers les métiers de la recherche à l'issue de mes études d'ingénieur agronome, je souhaitais combiner le domaine végétal et la génétique. La brève histoire de la génétique proposée plus haut introduit les deux outils principaux de mes activités de recherche : la génétique mendélienne, par l'usage permanent des marqueurs génétiques associés aux lois de probabilité, à la cartographie génétique ; la génétique quantitative, à la base de ma formation (Gallais 1990), à travers l'étude des caractères à distribution continue, et le recours régulier aux statistiques, à la biométrie. Mon parcours scientifique m'a d'abord amenée à travailler sur l'association entre marqueurs et caractères quantitatifs. Cette thématique a émergé dans le domaine de la génétique en 1988 avec une publication (Paterson *et al.* 1988) qui décrivait la première recherche de locus à effet quantitatif (QTL) chez la tomate. J'ai commencé ma thèse fin 1989, elle visait à mettre ce thème en œuvre chez le pin maritime. Ce travail de thèse a initié la thématique appelée actuellement "protéomique", qui a connu depuis un développement considérable chez cette espèce.

Je présenterai brièvement mes travaux antérieurs (après la thèse), qui concernent la génétique de la qualité de la graine de soja et la méta-analyse de locus à effet quantitatif. Mon propos sera ensuite consacré aux flux de gènes, à leur méthodologie et à leurs applications, sujet qui constitue le cœur de mes recherches actuelles et qui oriente mes projets de recherche futurs.

1.1. Génétique de la qualité de la graine du soja

Le soja (*Glycine max*) produit une graine riche en huile et en protéines. La qualité de cette graine pour la fabrication de produits destinés à l'alimentation humaine ou animale est essentiellement liée à sa teneur en protéines. Les sélectionneurs cherchent à augmenter cette teneur mais se heurtent à la corrélation négative, classique, existant entre la production de protéines et le rendement. En analysant les processus génétiques et physiologiques qui président à la mise en place de la teneur en protéines on espère pouvoir contourner l'incidence de cette corrélation négative.

Pour ce faire, une centaine de lignées recombinantes de soja (6^{ème} génération d'autofécondation) issues du croisement de deux lignées d'origines génétiques contrastées, toutes deux riches en protéines, ont été mesurées pour différents caractères quantitatifs impliqués dans la mise en place de la qualité de la graine, notamment la fixation symbiotique de l'azote (Fabre, 1998). J'ai suggéré d'analyser en parallèle la composition en protéines des graines de ces lignées.

J'ai mis au point l'utilisation de l'électrophorèse sur gel capillaire pour séparer et quantifier les composants de la graine selon leur poids moléculaire. J'ai analysé deux à trois mélanges de graines pour chaque lignée. Les surfaces relatives des pics séparés ont été comparées aux caractères quantitatifs. Des corrélations significatives (au seuil de 1 %) entre caractères et pics ont été repérées, en nombre plus important que ce que le simple hasard aurait produit, compte tenu du nombre de tests réalisés. En particulier, deux pics apparaissaient associés de façon inverse au rendement en graines et au taux de protéines des graines. Un autre pic était associé aux paramètres de fixation symbiotique.

Ce travail sur le soja confirme l'intérêt d'utiliser la quantification des protéines pour l'étude de caractères cibles de la sélection. Il ouvre des perspectives pour l'analyse plus fine des composants protéiques impliqués dans l'établissement de la qualité de la graine qui pourrait être combinée à l'étude de la régulation de ces protéines et à des approches classiques marqueurs/QTL (Gerber *et al.* 2000a).

1.2. Méta-analyse de locus à effet quantitatif (QTL)

Parallèlement à mon travail sur le soja, j'ai mené avec Bruno Goffinet (INRA, Laboratoire de Biométrie et Intelligence Artificielle, Toulouse) une collaboration portant sur la méta-analyse de locus à effet quantitatif (QTL). Une méta-analyse consiste à rassembler des résultats issus d'études différentes pour les analyser conjointement.

Cette collaboration nous a permis de proposer une méthode pour combiner les positions de QTL obtenues suite à des analyses indépendantes de caractères similaires dans une espèce donnée. Il s'agit de savoir si les QTL détectés dans les différentes expériences reflètent l'existence d'un, de deux, trois ou plus de véritables QTL. La méthode permet de calculer pour les différents modèles (modèle à une position, deux, etc.) un critère (critère d'Akaike modifié) qui mesure la qualité de chaque modèle et va permettre de choisir le meilleur. Des simulations ont été réalisées pour tester la méthode dans différentes situations de répartition des QTL et d'écart-type des positions. Quand peu de véritables QTL sont présents, la méthode permet de réduire significativement l'intervalle de confiance de la position du QTL.

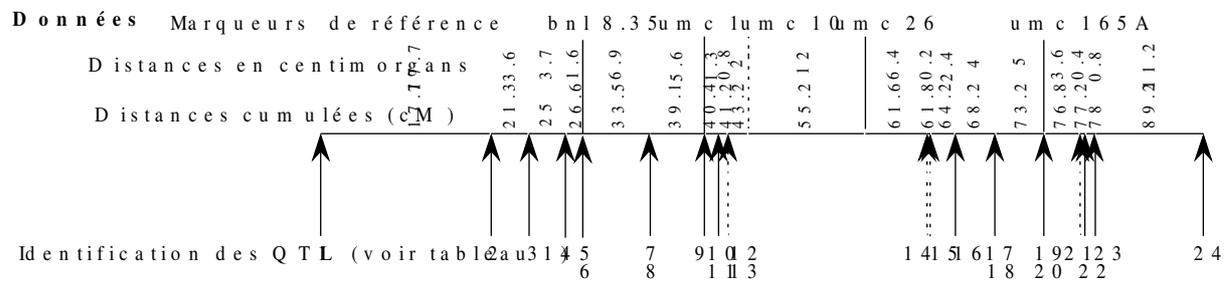
Nous avons appliqué notre méthode à des données relevées dans la base de données internationale "génomique maïs", accessible sur internet. En revenant aux publications originales, nous avons répertorié 24 QTL liés au rendement sur le groupe de liaison n°3 de cette espèce, décrits dans le tableau 1. Suite à notre analyse, ces 24 localisations pourraient représenter deux QTL dont nous avons estimé la position (figure 1).

Un logiciel mettant en œuvre notre méthode a été mis au point récemment (Arcade *et al.* 2004), et la méthode appliquée à la floraison du maïs grâce à sa syntenie avec le riz (Chardon *et al.* 2004, 2005). Ce type d'approche devrait permettre d'affiner les analyses de positionnement de QTL, notamment dans l'optique de la recherche de gènes candidats (Goffinet et Gerber 2000).

Tableau 1 : Description des QTL en relation avec le rendement sur le groupe de liaison n° 3 du génome du maïs.

Identification sur la figure 1	Caractère	Ecart-type estimé $\sigma_{E,i}$ de la position (cM)	Référence
1	Poids du grain	^a	(Veldboom et Lee 1996)
2	Diamètre de l'épi	16	(Beavis <i>et al.</i> 1994)
3	Hauteur de la plante	8	(Beavis <i>et al.</i> 1994)
4	"Poids test"	^a	(Ajmone-Marsan <i>et al.</i> 1995)
5	Hauteur de la plante	9.6	(Beavis <i>et al.</i> 1994)
6	Nombre de rangs	6.5	(Austin and Lee. 1996)
7	Hauteur de la plante	12	(Beavis <i>et al.</i> 1991)
8	Hauteur de la plante	12	(Beavis <i>et al.</i> 1991)
9	Hauteur de la plante	6.7	(Schön <i>et al.</i> 1993)
10	Poids du grain	9.8	(Austin et Lee. 1996)
11	Poids du grain	15.3	(Veldboom et Lee 1994)
12	Poids du grain	^a	(Maize Database, CIMMYT 1994)
13	Hauteur de la plante	^a	(Maize Database, CIMMYT 1994)
14	"Poids test"	8	(Beavis <i>et al.</i> 1994)
15	Hauteur de la plante	^a	(Maize Database, CIMMYT 1994)
16	Diamètre de l'épi	5.1	(Veldboom et Lee 1994)
17	Nombre d'épis / plante	8.5	(Veldboom et Lee 1994)
18	Nombre d'épis / plante	^a	(Veldboom et Lee 1996)
19	Diamètre de l'épi	7.6	(Austin et Lee 1996)
20	Poids du grain	7.6	(Austin et Lee 1996)
21	Hauteur de la plante	^a	(Maize Database, CIMMYT 1994)
22	Nombre d'épis / plante	5.5	(Austin et Lee 1996)
23	Hauteur de la plante	14.9	(Schön <i>et al.</i> 1994)
24	Longueur de l'épi	6.8	(Veldboom et Lee 1994)

^a pas d'information disponible dans la référence



Interprétation

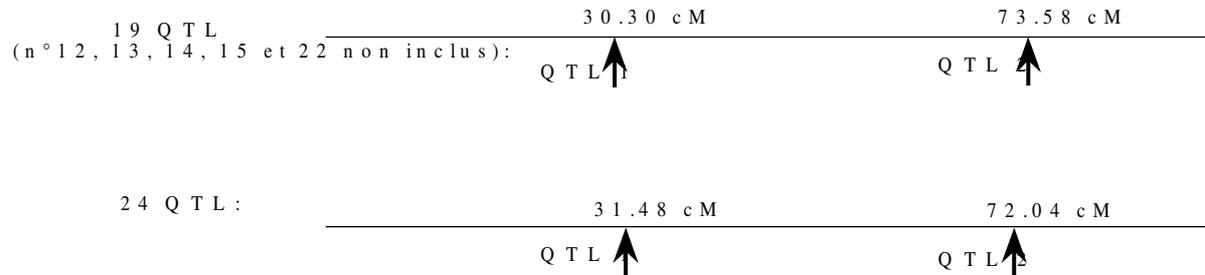


Figure 1 : QTL en relation avec le rendement sur le groupe de liaison n°3 du génome du maïs, localisation et interprétation.

2. Flux de gènes et arbres forestiers

2.1. Contexte

2.1.1. Les flux de gènes pourquoi, comment

Commençons par tenter une définition des flux de gènes. Ceux-ci peuvent simplement qualifier le mouvement des gènes dans les populations, sous forme de gamètes, de zygotes ou d'individus adultes - pour des espèces non fixées. Ces flux sont responsables de la modification de la distribution spatiale des gènes (Neigel 1997).

Dans le monde végétal, la dispersion du pollen et des graines, à l'origine des flux de gènes, constitue l'un des moteurs de l'évolution de la diversité. Les premiers flux historiquement étudiés ont été les **flux de gènes potentiels**, qui s'intéressent à la dispersion du pollen et des graines depuis un arbre source indépendamment de leur devenir. Le suivi physique de ces dispersions, par marquage avant libération du pollen et des graines (en utilisant par exemple un colorant) et piégeage, la mesure des distances parcourues, l'ajustement de courbes empiriques à ces données étaient les principaux résultats de ces études.

La découverte de marqueurs génétiques, les premiers ayant été les isozymes dans les années 1960, suivis dans les années 1980 de marqueurs de l'ADN, a notamment donné accès aux **flux efficaces**, différents des flux potentiels car conduisant à la naissance d'un nouvel individu (succès de la pollinisation ou de l'établissement des graines) : la dispersion, en Arctique, de pollens de pin et d'épicéa transportés sur 3000 km (Campbell *et al.* 1999), n'intéresse pas le généticien puisqu'elle n'a aucun avenir évolutif ! À l'origine, l'étude des flux de gènes avec des marqueurs se faisait *via* l'estimation de **flux de gènes historiques**, par des mesures indirectes utilisant la théorie de la génétique des populations. Cette approche consiste à réaliser un échantillonnage de la population étudiée, à génotyper les individus de l'échantillon et à caractériser la distribution spatiale des génotypes en utilisant des paramètres particuliers, comme par exemple des paramètres de différenciation. En appliquant un modèle de génétique des populations à ces données, il s'agit alors de déduire quel niveau de flux de gènes aurait généré une distribution de paramètres identiques. En procédant de cette façon, le paramètre statistique de différenciation F_{ST} est classiquement estimé et selon le modèle en îles de Wright, à l'équilibre, celui-ci est lié aux paramètres N (taille de la population) et m (taux de migrants) par la relation $F_{ST} = 1/4Nm$, Nm étant le nombre de migrants, un flux de gènes est ainsi indirectement estimé.

Les modèles alternatifs au modèle en île de Wright sont le modèle "stepping-stone" (pas japonais), dans lequel les migrations se passent uniquement entre populations adjacentes (Kimura 1953) et un modèle continu, d'isolement par la distance (Malécot 1950), dans lequel les probabilités de migrations entre populations sont proportionnelles à la distance qui les sépare. La réalité biologique se situe entre ces trois modèles.

La limitation de l'approche indirecte est liée au fait que ces flux cumulent une série d'événements hétérogènes de la vie de la population, en confondant l'effet des différentes pressions évolutives (dérive et migration notamment), ils reposent sur des hypothèses nombreuses, tel que l'équilibre, rarement atteint, et les résultats obtenus diffèrent selon les paramètres utilisés (Bossart et Prowell 1998, Ouborg *et al.* 1999). Le F_{ST} fournit une image globale de l'effet cumulé et général des flux de gènes et de son

rôle dans l'évolution : il garde un intérêt historique et comparatif car des données se sont accumulées sur de nombreuses espèces grâce à cette approche (Neigel 2002). Depuis que des marqueurs hautement polymorphes sont disponibles (par exemple les marqueurs microsatellites, qui sont dans l'ADN des répétitions généralement dinucléotidiques et qui varient d'un individu à l'autre par le nombre de motifs répétés), il est possible de retracer de manière rétrospective les **flux contemporains**. L'utilisation conjointe de plusieurs marqueurs moléculaires de ce type donne accès à une combinaison génotypique quasiment exclusive à chaque individu. De ce fait, les relations de parenté entre individus peuvent être reconstituées en comparant ces combinaisons, et l'observation directe des flux de pollen (recherche de paternité) ou des flux de graines, est alors possible.

2.1.2. Une brève histoire de la recherche de paternité

La recherche de paternité concerne au départ bien évidemment l'homme et a évolué avec les techniques disponibles.

La qualité d'un système de marquage génétique, qui autorise à répondre correctement à une recherche de paternité, se mesure par sa capacité moyenne à exclure une relation donnée (en l'occurrence ici, père-descendant) entre individus. Celle-ci n'a cessé d'augmenter avec les nouvelles techniques d'identifications disponibles.

La recherche de paternité débute dans les années 1920 avec les groupes sanguins. Ils ont été découverts en 1901 à l'Université de Vienne par Karl Landsteiner (Prix Nobel en 1930), qui essayait de comprendre pourquoi les transfusions sanguines sauvaient les patients mais étaient parfois responsables de leur mort. Le système ABO, qui détermine les protéines présentes à la surface des cellules sanguines, ne permet d'exclure comme pères potentiels chez les hommes (pour un descendant donné) que 30 % de la population masculine.

Dans les années 1930, on utilise des tests sérologiques basés sur les groupes Rhésus (découverts dans les années 1920 par le même Karl Landsteiner à l'Institut Rockefeller à New-York), les systèmes antigènes Kell et Duffy, qui ne permettent d'éliminer comme pères potentiels que 40 % de la population masculine.

Dans les années 1970, les groupes HLA ("Human Leukocyte Antigens"), des protéines qui interviennent dans le système immunitaire pour détecter les éléments étrangers à l'individu, permettent d'améliorer significativement la technique, pour atteindre 80 à 90 % d'exclusion. Ces protéines sont très variables d'un individu à l'autre, ce qui explique la meilleure résolution obtenue en recherche de paternité. Mais si deux individus apparentés sont comparés, ce système ne pourra souvent pas trancher.

Dans les années 1980, des tests ADN basés sur les RFLP ("Restriction Fragment Length Polymorphism") sont disponibles. Grâce à leur grand polymorphisme chez l'homme, des probabilités d'exclusion de 99,99 % sont alors possibles.

Depuis les années 1990, l'ADN peut être plus facilement testé grâce à la PCR (Polymerase Chain Reaction), qui requiert très peu de matériel biologique, et est à la base des méthodes standards en recherche de paternité ou dans les études criminelles. Les séquences microsatellites, répétitions en tandem de courtes séquences de paires de bases, sont extrêmement variables d'un individu à l'autre avec des probabilités d'exclusion supérieures à 99,99 %. En recherche de paternité, on utilise l'information de 16 locus différents aux États-Unis d'Amérique. En France les empreintes

génétiques légales portent sur 8 locus principaux (8 chromosomes différents) et 10 locus supplémentaires (Journal Officiel du 6 mars 2002).

Dans les bases de données bibliographiques à ma disposition, la première référence traitant de la recherche de paternité grâce aux groupes sanguins concerne l'homme et remonte à 1949 (Boyd, 1949). La recherche de paternité sera ensuite largement utilisée chez les animaux domestiques également avec leurs groupes sanguins (nombreuses publications dans les années 1970-1980). La première référence végétale à cette technique concerne justement les arbres forestiers, où la paternité d'un hybride supposé de pins est testée avec des isozymes (Morris *et al.* 1980). Il faut ensuite attendre 1984 pour voir une étude concernant une plante annuelle, le radis, (*Raphanus sativus*) avec les mêmes outils biochimiques (Ellstrand 1984, Ellstrand et Marshall 1985). À la fin des années 80, les publications témoignent du développement plus important de cette thématique chez les plantes. Un histogramme de la fréquence du mot "paternity" dans les articles publiés entre 1949 et 2005 montre en effet une nette augmentation à partir des années 90, grâce aux marqueurs génétiques : 82 % du total des articles parus le sont à partir de 1990 dans la base CAB (pour 16 années sur 36 soit 44 % de la période totale) et 56 % dans Medline (pour 16 années sur 56 soit 29 % de la période), les articles en biologie humaine étant bien sûr plus courants sur ce sujet, mais également affectés par l'apparition des techniques de biologie moléculaire (figure 2). Le mot "parentage" apparaît pour la première fois dans les bases de données en 1950, dans le même contexte que le mot "paternity", chez l'homme et avec les groupes sanguins (Wiener 1950). Ce mot est nettement plus courant dans la base CAB (ce qui s'explique par l'importance de ce concept dans les variétés améliorées, végétales et animales) que dans la base Medline. L'impact des techniques moléculaires est néanmoins sensible puisque 57 % des références dans CAB sont publiées à partir de 1990, et 70 % dans Medline (sur 44 % et 29 % de la période totale de publication des articles respectivement).

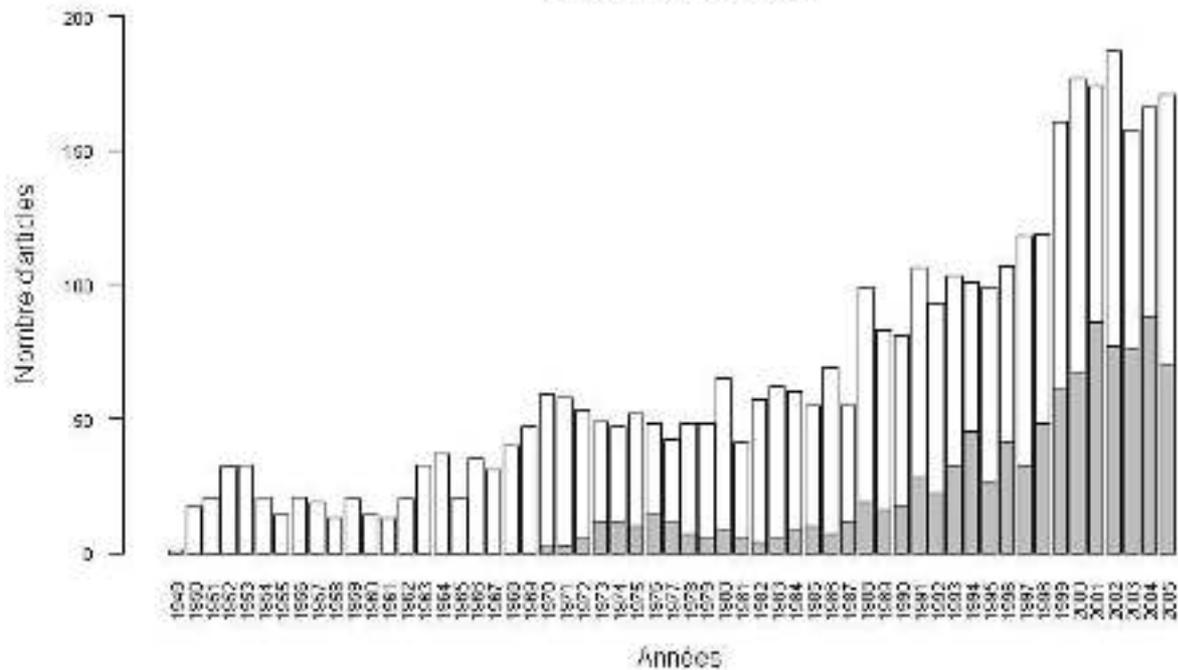
Mon travail de ces dernières années a concerné ces deux notions de paternité et de parenté (maternité et paternité), par la mise en place d'outils méthodologiques, le développement d'un logiciel pour les mettre en œuvre, et l'application de ces techniques à différentes espèces, notamment des arbres forestiers.

2.1.3. Arbres et flux de gènes

Les arbres forestiers se caractérisent par leur niveau élevé de polymorphisme et restent des espèces peu domestiquées, la sylviculture ayant été introduite vers le XIV^{ème} siècle : création de la maîtrise des Eaux et Forêts par Philippe IV le Bel en 1291, et mise en place de la première politique forestière complète par Colbert en 1669 (Lanier 1994). Avec l'accélération des interventions humaines, il apparaît cependant nécessaire aujourd'hui d'apprécier et de prévoir l'évolution de la variabilité sous l'action anthropique et biologique et ainsi d'introduire un raisonnement génétique dans les opérations sylvicoles, pour préserver *in-* ou *ex-situ* un niveau de diversité maximum.

L'équipe de génétique de l'UMR Biogeco (Biodiversité, Gènes et Communautés, INRA-Université Bordeaux I) développe notamment des recherches sur la structure et l'évolution de la diversité génétique dans les écosystèmes forestiers tempérés et tropicaux. Les chênes blancs européens font en particulier l'objet d'études intensives sur ces aspects depuis une vingtaine d'années. Ainsi, l'inventaire et la description de la diversité dans ce complexe d'espèces ont été faits à diverses échelles de plus en plus

Fréquence du mot 'paternity' dans les publications entre 1949 et 2005



Fréquence du mot 'parentage' dans les publications entre 1950 et 2005

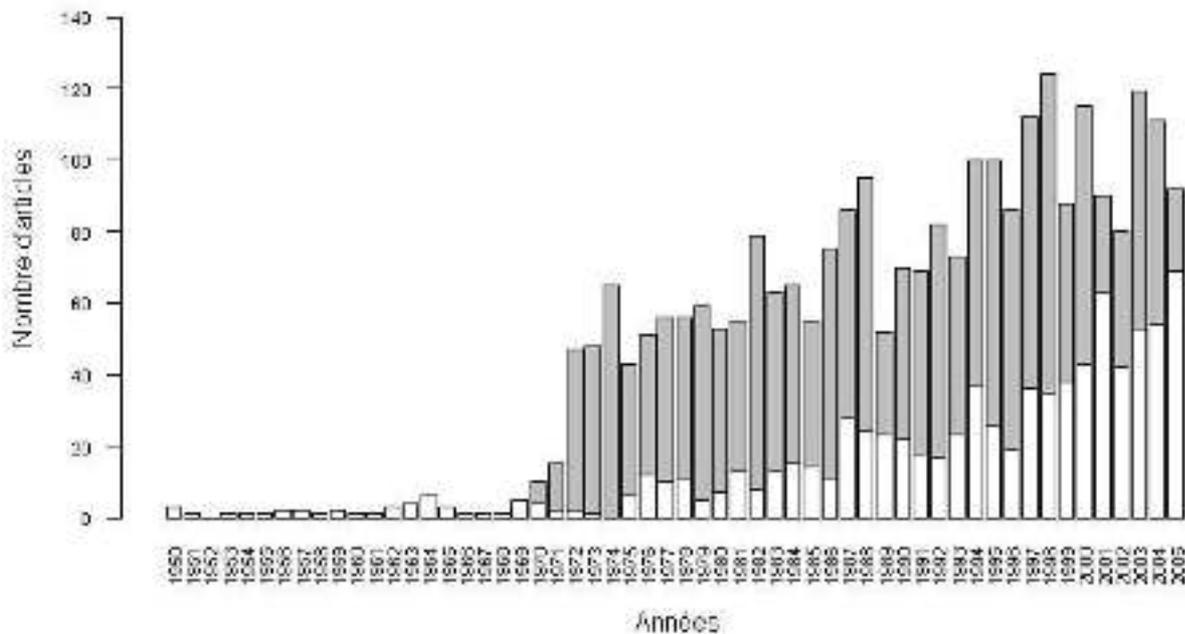


Figure 2 : Fréquence des mots "paternity" et "parentage" dans les articles des bases Medline (en blanc, depuis 1949) et CAB (en gris, depuis 1970) publiés entre 1949 et 2005 ('paternity' : 3772 publications référencées au total dans Medline et 993 dans CAB) ('parentage' : 844 publications référencées au total dans Medline et 2639 dans CAB).

Tableau 2 : probabilités d'exclusion pour un locus.

Un locus codominant

Pour un locus à n alleles (allele i en fréquence p_i) on écrit :
$$a_k = \sum_n^{i=1} p_i^k$$

Exclusion de paternité Exclusion d'un seul parent Exclusion d'une paire de parents

$$1 - 2a_2 + a_3 + 2a_4 - 3a_5 - 2a_2^2 + 3a_2 a_3 \quad 1 - 4a_2 + 2a_2^2 + 4a_3 - 3a_4 \quad 1 + 4a_4 - 4a_5 - 3a_6 - 8a_2^2 + 8a_2 a_3 + 2a_3^2$$

Un locus dominant

Deux phénotypes [+] et [-] au locus, p fréquence de l'allèle dominant +

Exclusion de paternité Exclusion d'un seul parent Exclusion d'une paire de parents

$$p(1-p)^4$$

Toujours nulle

$$p(2-p)(1-p)^4$$

finer : espèce, région, forêt, population. Le travail se concentre maintenant au niveau individuel. En effet, grâce aux marqueurs moléculaires très polymorphes, qui permettent une identification individuelle des arbres, de nouvelles perspectives se sont ouvertes pour l'étude de la dynamique de la diversité.

L'étude fine des flux contemporains, initiée au laboratoire dans le cadre de la thèse de Réjane Streiff, et concrétisée notamment par la description spatiale des événements polliniques autour de 13 arbres mères d'une parcelle de chênes (pédonculés et sessiles) de 5 hectares, dans laquelle tous les arbres adultes avaient été génotypés (Streiff *et al.* 1999), a été l'objet de mes travaux dès mon arrivée dans l'équipe. Les premières approches ont été d'ordre méthodologique, afin d'élaborer un outil de recherche de paternité et de parenté à partir de marqueurs codominants, dominants ou hérités de façon uniparentale. Les techniques mises au point ont alors pu être appliquées à diverses questions biologiques.

2.2. Développements méthodologiques

2.2.1. Mesurer la qualité d'un système de marquage : les probabilités d'exclusion

2.2.1.1. Principe

Tout système de marquage génétique peut être utilisé pour vérifier des relations de type parent-descendant, mais les caractéristiques du type de marqueur utilisé sont déterminantes dans la fiabilité de la réponse fournie. Pour mesurer la qualité d'un système de marquage pour la recherche de parenté, on calcule classiquement des probabilités d'exclusion. Une probabilité d'exclusion peut être définie comme la capacité moyenne de n'importe quel système de marquage génétique à exclure une relation donnée entre individus. Cette probabilité dépend du génotype des individus à comparer, de la fréquence dans la population des allèles aux marqueurs et du nombre de locus indépendants testés (Sandberg 1994 cité par Jamieson et Taylor 1997). Plus le nombre d'allèles est grand au marqueur et plus les fréquences des allèles sont équilibrées, plus la qualité du marqueur sera élevée pour répondre aux questions d'apparentement. Un marqueur présentant un grand nombre efficaces d'allèles (nombre d'allèles rapporté à leurs fréquences relatives) présentera des probabilités d'exclusion élevées.

Trois probabilités d'exclusion peuvent être calculées. La plus commune est l'exclusion de paternité. La deuxième est l'exclusion d'un parent potentiel unique comparé à un descendant supposé (sans information sur le deuxième parent, contrairement au cas précédent) et la troisième est l'exclusion d'un couple père/mère potentiel comparé à un descendant supposé. Pour un locus donné, la probabilité d'exclusion pour une paire de parents est toujours la plus élevée, suivie par la probabilité d'exclusion de paternité, la probabilité d'exclusion d'un seul parent étant toujours la plus faible.

Pour réaliser le calcul de ces probabilités, tous les cas possibles d'exclusion sachant les génotypes sont envisagés et sommés. Les probabilités associées à des marqueurs codominants sont classiquement données dans la littérature (Jamieson & Taylor 1997), nous avons recalculé celles associées à des marqueurs dominants, plus rarement utilisés pour ce genre d'étude. Ces différentes probabilités sont données dans le tableau 2 pour un locus codominant ou dominant.

La probabilité d'exclusion pour K locus indépendants s'écrit :

$$P = 1 - \prod_{i=1}^K (1 - P_i)$$

Un système de marquage avec une probabilité d'exclusion de paternité maximale (100 %), permettra d'exclure pour un descendant donné tous les pères possibles de la population au vu de leurs génotypes, à l'exception de son vrai père.

Notons au passage que si l'on s'intéresse à une relation de parenté moins forte que la relation parent-enfant, par exemple pour la plus forte d'entre elles en suivant la hiérarchie, la relation pleins-frères/sœurs, les probabilités d'exclusion sont nulles. Supposons, cas le plus extrême, qu'à un locus particulier, l'un des frère/sœur présente le génotype AB et l'autre le génotype CD, alors il est toujours possible de supposer que leurs deux parents communs sont de génotypes AC et BD. En l'absence de toute autre information et en considérant les individus deux par deux, il est impossible d'exclure une relation de pleine fraternité entre deux individus pour un locus et *a fortiori* pour plus d'un locus.

2.2.1.2. Comparaison des marqueurs microsatellites et AFLP

Le tableau 2 fait apparaître une probabilité nulle d'exclusion pour un seul parent pour les marqueurs dominants : aucun individu ne peut être exclu comme parent d'un descendant quelconque sur la base de son génotype à un marqueur dominant. Les marqueurs codominants sont plus performants et parmi eux, les marqueurs microsatellites, qui sont généralement très polymorphes, sont un outil de choix. Néanmoins, le développement de marqueurs microsatellites pour une espèce peut représenter un travail difficile, et nous avons voulu savoir si des marqueurs de type AFLP (Amplified Fragment Length Polymorphism), qui sont dominants mais qui sont nombreux, pouvaient offrir une alternative valable dans le cas où les microsatellites n'étaient pas disponibles.

Pour ce faire, nous avons utilisé des données du laboratoire concernant 89 chênes blancs (*Quercus petraea*, *Quercus robur*) de la parcelle 26 (5,76 ha) de la "Petite Charnie" (département de la Sarthe). Leur génotype a été établi pour 6 locus microsatellite par Streiff *et al.* (1998). Ils présentent chacun entre 15 et 31 allèles, de fréquences comprises entre 0,002 et 0,24. Ces mêmes arbres ont été typés pour 159 bandes AFLP polymorphes par Mariette (2001). L'histogramme de la répartition des 159 bandes selon la fréquence de l'allèle dominant (présence de la bande) est donné sur la figure 3. Les probabilités d'exclusion de 30 bandes AFLP théoriques présentant la même fréquence (entre 0 et 1) de l'allèle dominant sont données sur la figure 4 : les fréquences comprises entre 0,1 et 0,4 donnent lieu à des probabilités maximales.

Les figures 5 et 6 comparent les probabilités d'exclusion cumulées des deux systèmes de marquage, microsatellites et AFLP. Les probabilités totales sont proches de 1 pour les deux systèmes, la probabilité d'exclusion d'un couple de parents étant toujours plus élevée que les autres. Parmi les 159 bandes AFLP, 45 bandes (28,3 %) ont une fréquence de l'allèle dominant comprise entre 0,1 et 0,4. Lorsque les probabilités d'exclusion sont recalculées avec ces seules bandes, on constate que des valeurs élevées sont atteintes. Ce sous-ensemble de bandes est presque aussi informatif que les 159 bandes réunies : toutes les bandes ne participent pas de manière égale à la probabilité d'exclusion totale. Ces résultats suggèrent que des marqueurs dominants de type AFLP peuvent être utilisés pour l'analyse de parenté.

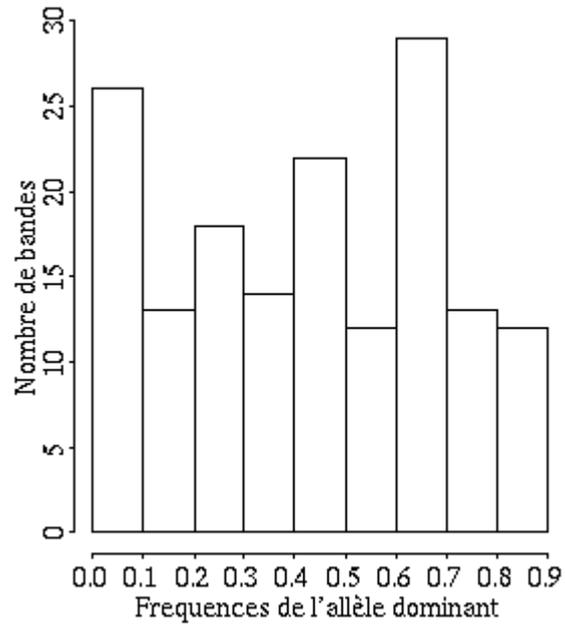


Figure 3 : Répartition des 159 bandes AFLP polymorphes dans l'échantillon des 89 chênes.

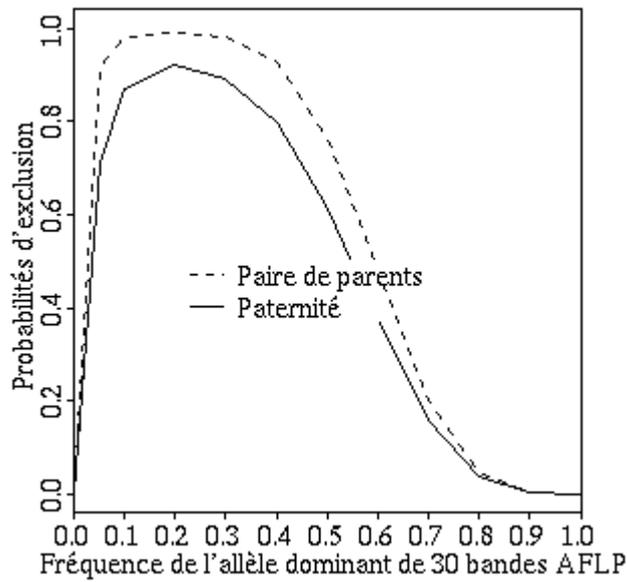


Figure 4 : Probabilités d'exclusion générées par 30 bandes AFLP de même fréquence pour l'allèle dominant.

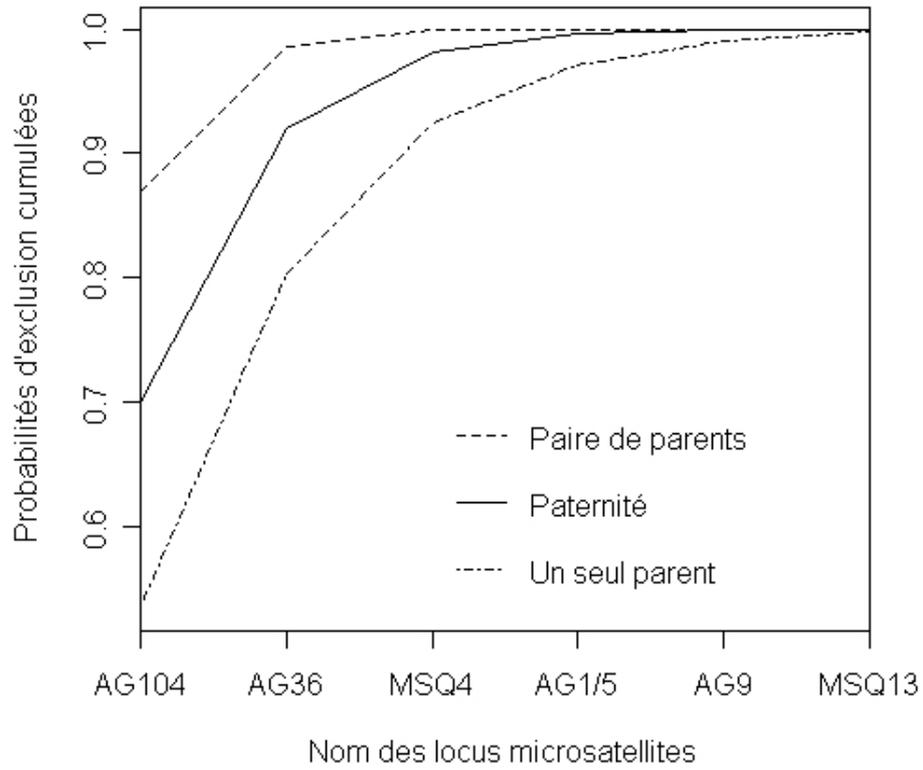


Figure 5 : Probabilités d'exclusion cumulées des 6 locus microsatellites.

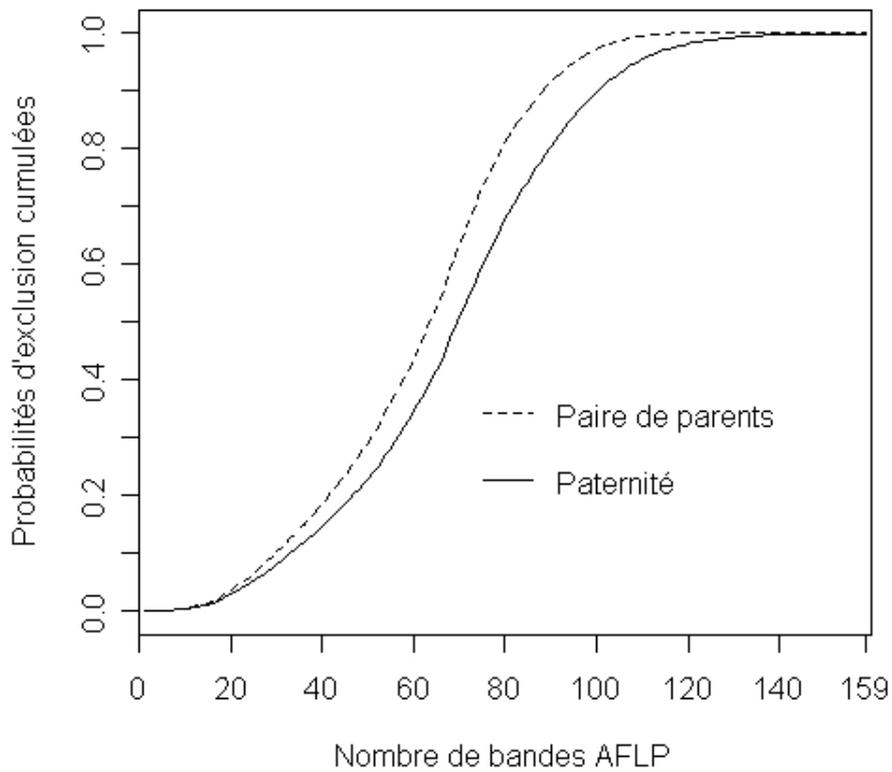


Figure 6 : Probabilités d'exclusion cumulées des 159 bandes AFLP.

2.2.2. Recherche de paternité et de parenté

Pour réaliser la reconstitution d'apparentement entre individus sur la base des informations fournies par des marqueurs génétiques polymorphes, les travaux que j'ai initiés ont été de nature méthodologique, ils ont été ensuite testés grâce à des simulations, et ont été appliqués à des données expérimentales. J'ai mis au point, en collaboration avec un informaticien de l'INRA (Patrick Chabrier, Toulouse), un logiciel dénommé FaMoz (un acronyme pour "father/mother"), qui permet à la communauté scientifique d'utiliser mes programmes facilement. Le logiciel est mis librement à la disposition des chercheurs sur Internet (Gerber *et al.* 2003) :

<http://www.pierroton.inra.fr/genetics/labo/Software/Famoz/index.html>

Ce logiciel était d'abord destiné aux membres du projet financé par le Bureau des Ressources Génétiques, *via* la Direction de l'Espace Rural et de la Forêt, que je coordonnais (2000-2002), et également aux partenaires du projet européen OAKFLOW (2001-2005), coordonné par Antoine Kremer (voir 2.5.).

Le logiciel est basé sur des calculs classiques de rapport de vraisemblance dans le cas parent-descendant développés par Meagher et Thompson (1986) pour les marqueurs codominants et que j'ai adaptés au cas des marqueurs dominants et cytoplasmiques (Gerber *et al.* 2000b). Un taux d'erreur de génotypage, qui permet d'éviter des incompatibilités de génotypes (et donc un rapport de vraisemblance nul) alors que les apparentements sont réels, suite à des génotypages incorrects, peut être utilisé dans les calculs, selon la méthode de Marshall *et al.* (1998). De la même façon, un écart à la panmixie est utilisable. Pour chaque père ou parent potentiel, une note, d'autant plus élevée que le parent en question est plus probablement un vrai parent selon les données génétiques, est calculée. Pour déterminer le seuil au-delà duquel un parent va être retenu comme parent réel, des simulations sont faites à partir des données de l'utilisateur. La note peut être le rapport de vraisemblance lui-même, une valeur delta correspondant à l'écart entre le rapport du meilleur parent et celui des autres parents potentiels ou la combinaison des deux. Les performances de ces deux types de notes ou de leur combinaison peuvent être comparées avec des simulations qui mesurent le pourcentage de bonnes décisions dans une étude de flux virtuelle. Les flux de gènes peuvent également être estimés grâce à des simulations. Dans une revue des méthodes disponibles, Jones et Ardren (2003) soulignent que FaMoz est sensible à l'estimation de la taille de la population reproductrice, ce paramètre a été supprimé dans les versions plus récentes du logiciel suite à une suggestion de calcul d'Olivier Hardy (Université Libre de Bruxelles). Il apparaît dans cette revue que FaMoz est le seul logiciel à proposer l'usage de marqueurs dominants et cytoplasmiques, et possède le plus de fonctions (Jones et Ardren 2003, tableau 2). Néanmoins, le fichier d'entrée est jugé difficile !...

J'illustrerai les usages du logiciel FaMoz en donnant quelques exemples tirés de publications.

Le logiciel a été utilisé dans deux articles parus dans le magazine Science. Le premier s'intéresse à un palmier tropical dans une forêt secondaire du Costa Rica. Grâce à des marqueurs AFLP (dominants), les auteurs ont montré que les arbres dans les forêts secondaires étaient nettement plus apparentés que ceux de la forêt primaire à proximité. L'analyse de parenté faite avec FaMoz a ainsi identifié que, parmi 66 arbres de la forêt primaire, seuls deux arbres avaient donné naissance à 56 % des descendants de la forêt secondaire, 23 d'entre eux ayant laissé 44 % de descendants restants et 41

arbres n'ayant eu aucun descendant. La diversité de la forêt secondaire subit un effet de fondation fort et est ainsi très diminuée (Sezen *et al.* 2005). Le second article démontre des flux de graines inédits, atteignant plusieurs kilomètres, dans des populations écossaises très fragmentées de frênes, dus à un transport par le vent ou l'eau. Les flux de graines apparaissent ainsi aussi importants que les flux de pollen, sinon plus, avec des queues de distribution pouvant atteindre des dizaines de kilomètres (Bacles *et al.* 2006)

La méthode a également été mise en œuvre pour rechercher les parents d'olives récoltées dans un verger d'oliviers contenant des variétés commerciales (Mookerjee *et al.* 2005). L'étude de parenté d'une silène, espèce pérenne menacée de Finlande, avec des marqueurs dominants, a montré que les flux de pollen étaient importants mais que les flux de graines restaient limités, conduisant à une structuration génétique spatiale significative (Tero *et al.* 2005). Une petite étude de parenté avec le logiciel sur le peuplier noir (*Populus nigra*), espèce endémique, a permis de démontrer l'existence d'hybrides avec les peupliers cultivés (*Populus deltoïdes*), espèce introduite, le long d'un cours d'eau en République Tchèque (Pospíšková et Šálková 2006)

Dans le monde animal, le logiciel semble avoir du succès chez les poissons : une analyse de degré d'apparentement et de parenté a été réalisée chez l'esturgeon blanc avec des marqueurs AFLP et FaMoz (Rodzen *et al.* 2004). La dispersion de larves de poisson clown a été également étudiée à l'aide du logiciel avec des marqueurs microsatellites. La dispersion observée à l'issue de cette analyse est l'une des plus faibles jamais obtenue pour un poisson à phase larvaire pélagique (Jones *et al.* 2005). Chez l'ombrine tropicale, élevée au sud des États-Unis, l'utilisation conjointe de marqueurs microsatellites nucléaires et du logiciel a permis de distinguer des poissons sauvages de poissons d'élevage, pour identifier les fraudes lors de vente de poissons annoncés comme élevés, et pourtant pêchés dans les ressources naturelles (Renshaw *et al.* 2006).

J'ai également pu valoriser mes travaux sur l'analyse de parenté grâce à des études consacrées à des espèces variées, au travers de collaborations. Les questions posées concernaient des populations naturelles ainsi que des populations d'amélioration, modelées par l'homme : les flux de gènes constituent un mécanisme important à étudier.

2.3. Populations naturelles et flux de gènes

Grâce aux méthodes d'étude d'apparentement, nous avons pu décrire avec précision la structure génétique d'une forêt naturelle, qui témoigne de l'histoire des flux de gènes d'une génération à l'autre, et nous avons tenté de déterminer le nombre de pères intervenant dans des lots de graines récoltés sur une plante mère, l'un des paramètres déterminant l'évolution de la diversité.

2.3.1. Structure génétique fine, flux de graines et fratries

Une forêt naturelle de pin maritime du centre de l'Espagne a fait l'objet d'une étude génétique fine, dans le cadre de la thèse de Santiago González-Martínez (INIA, Espagne). L'ensemble des arbres d'une parcelle de 50 m de diamètre a été échantillonné (76 individus) ainsi que tous les plants mesurant plus de 20 cm de haut dans un cercle concentrique de 25 m (132 individus). L'ensemble de ces arbres a été génotypé pour 3 locus microsatellites nucléaires. Les structures génétiques spatiales des deux cohortes et les liens parent-enfant les reliant ont été analysées en collaboration.

Une structuration génétique spatiale a été mise en évidence parmi les jeunes plants, plus apparentés à l'intérieur de taches de 10 m qu'au-delà. Cette structure disparaît chez les arbres adultes : la maturation de la forêt s'accompagne d'une élimination des individus apparentés à courte distance, suite à une diminution de densité associée à une mortalité locale forte (compétition avec les herbes et les buissons). Selon le niveau de signification retenu pour les liens parent-enfant, un seul parent a pu être attribué à un sous-ensemble de 12 à 40 % des descendants, alors qu'un couple parental a été retrouvé pour 3 à 9 % d'entre eux. Le flux de gènes extérieur (le pourcentage de gamètes ayant généré les jeunes plants en provenance de l'extérieur de la parcelle) est donc compris entre 70 et 90 % dans ce dispositif. Les distances parent-descendant observées se révèlent significativement inférieures (excès de distances inférieures à 15 m) à ce qui aurait été obtenu au hasard (en assignant de façon aléatoire un parent du dispositif aux jeunes plants) : les descendants sont plus proches de leur parent que d'un individu adulte quelconque de la parcelle (figure 7). En revanche, les directions de dispersion parent-descendant ne diffèrent pas de l'aléatoire (González-Martínez *et al.* 2002).

Ces mêmes données ont été utilisées pour reconstituer les fratries (demi- et pleins-frères), avec une méthode basée sur des comparaisons de rapport de vraisemblance que j'ai développée (et implémentée dans le logiciel FaMoz). La cohorte des arbres adultes est composée de 19,1 % de paires de demi-frères et de 1,8 % de paires de pleins-frères. Dans la régénération, ces chiffres sont de 19,2 et 2,4 %, donc très proches. L'apparentement moyen dans les deux cohortes est de 0,028 et 0,030, soit sensiblement égal et faible, témoignant notamment de l'absence d'autofécondation. De plus, parmi quatre groupes de régénérations un seul présentait une structure un peu différente (avec plus de 6 % de pleins-frères) (González-Martínez *et al.* 2003).

La méthode utilisée dans FaMoz dispose à présent de concurrents sérieux, utilisant des méthodes mathématiques plus élaborées (Almudevar et Field 1999, Painter 1997, Emery *et al.* 2001, Fiumera *et al.* 2001, Letcher et King 2001, Smith *et al.* 2001, Beyer et Bernie 2003, Butler *et al.* 2004, Wang 2004, Fernández et Toro 2006, voir 3.2.2.). La thèse de Johan Fogelqvist (Université d'Uppsala, Suède) devrait faire l'objet de développements dans ce domaine, notamment pour exploiter les données issues du projet européen OAKFLOW (étudier la structure de parenté des individus pour lesquels aucun parent n'a pu être retrouvé et en déduire le nombre de reproducteurs).

Cette population de pins maritimes au cœur de l'Espagne présente donc une dispersion de graines limitée, qui entraîne la mise en place d'une régénération structurée spatialement d'un point de vue génétique. Cette structuration sera néanmoins affectée par des événements de post-dispersion qui vont faire disparaître les liens à courte distance entre arbres. Cette image de la structure génétique suggère que pour la conservation de la diversité ou pour la reforestation, les récoltes de graines doivent être

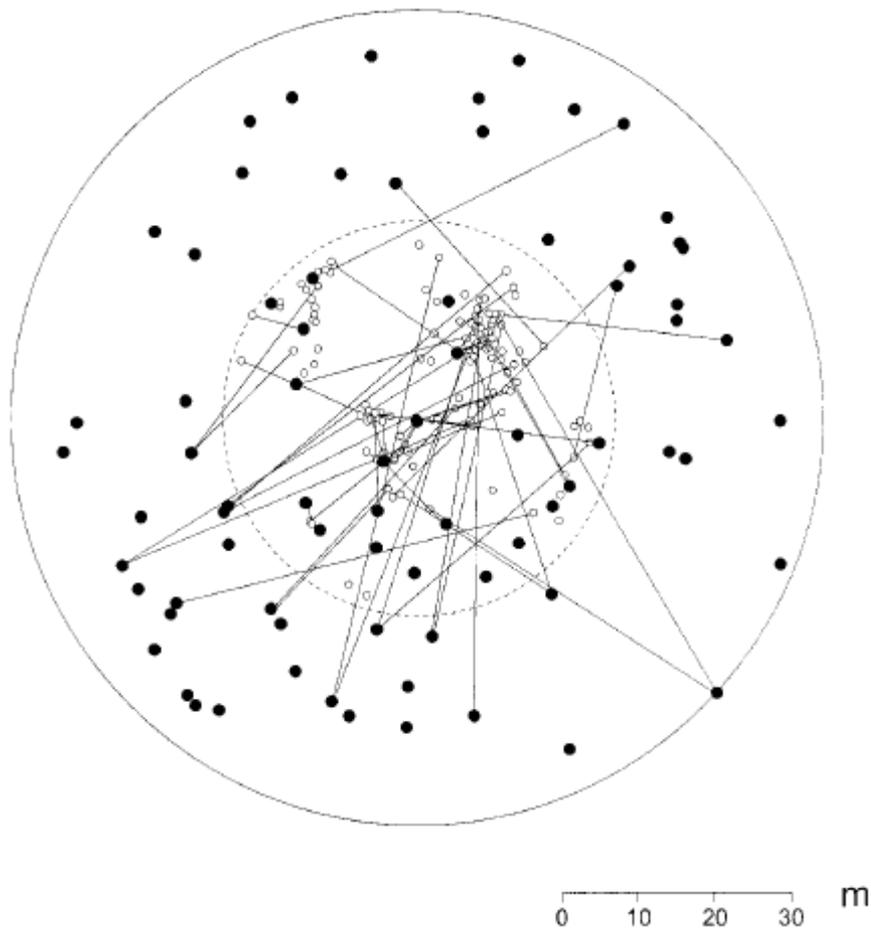


Figure 7 : descendants (cercles blancs) reliés à leurs parents (cercles noirs) retrouvés suite à la recherche de parenté dans le dispositif de pins maritimes en Espagne.

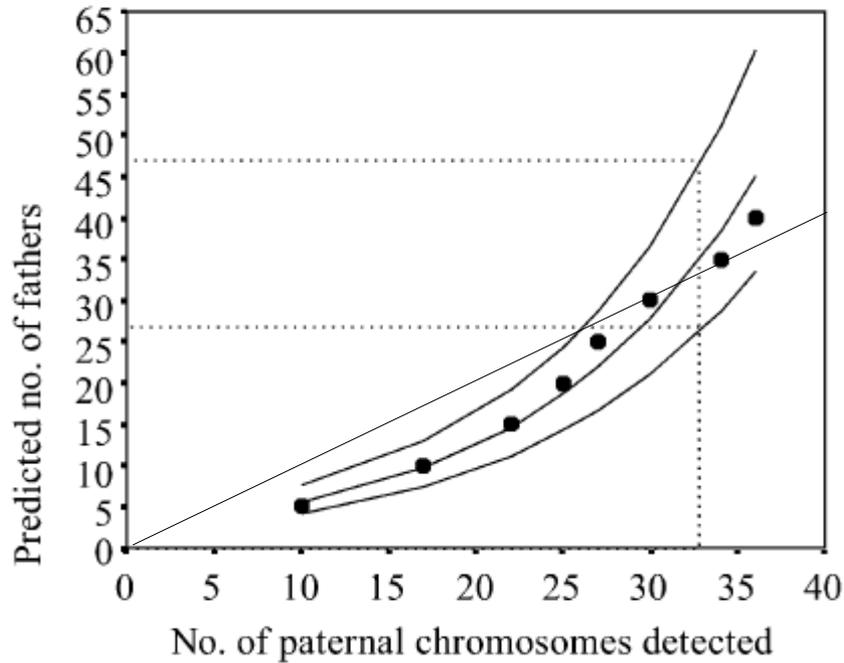


Figure 8 : nombre de pères différents détectés et réels dans des familles de demi-frères simulées, et intervalle de confiance à 95% de la valeur prédite. Les pointillés correspondent à la valeur détectée dans l'échantillon génotypé.

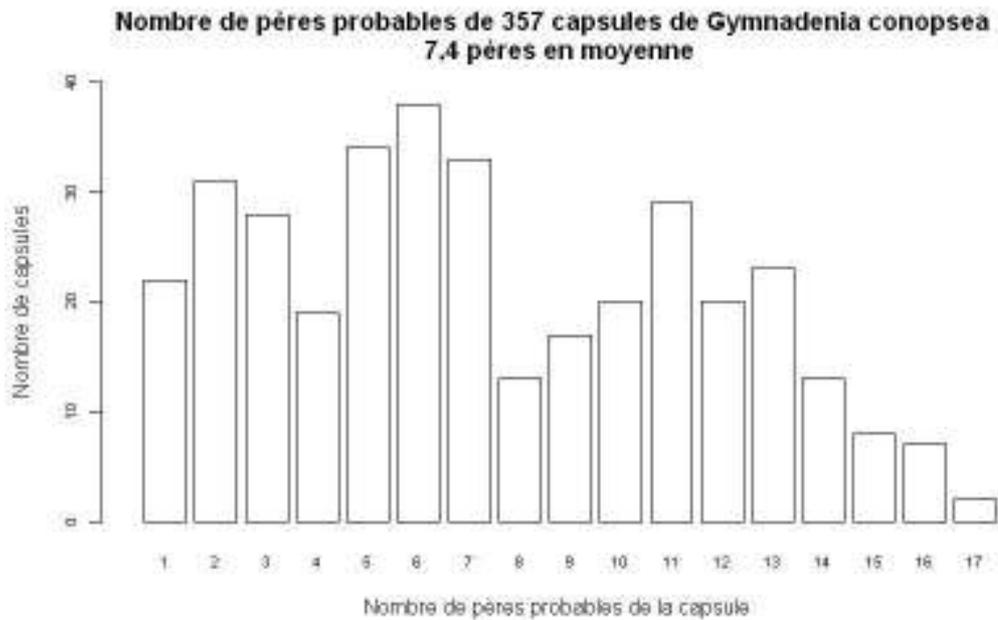


Figure 9 : nombre de pères probables de chacune des 357 capsules de *Gymnadenia conopsea* retrouvé suite au calcul des vraisemblances de chaque père possible de la population parentale.

faites sur des arbres suffisamment distants pour éviter de réduire la base génétique de l'échantillon.

2.3.2. Nombre de pères

Le nombre efficace de reproducteurs dans une population, c'est à dire le nombre de parents différents pondérés par le nombre de descendants qu'ils laissent, va déterminer l'évolution de cette population, et notamment les effets de dérive. Nous avons tenté une estimation du nombre de pères intervenant dans la génération de descendants par deux approches très différentes, basées sur des études d'apparentement, chez les chênes et chez une orchidée.

Le premier travail s'est fait dans le cadre de la thèse de Christian Lexer (Université de Vienne, Autriche), je lui ai proposé une analyse basée sur des simulations. Pour mettre en place une régénération artificielle en forêt, on récolte des graines sur un certain nombre d'arbres mères. Ces régénérations sont appelées à rester en place sur des périodes de temps longues, et devront faire face à de multiples changements, il est donc indispensable de leur assurer une base génétique la plus large possible, pour une stabilité optimale. Il est alors fondamental de s'appuyer sur un lot de graines issues d'un grand nombre d'arbres mères, et la question de savoir combien de pères ont participé à la reproduction dans ces lots de graines est également déterminante : c'est ce point que ce travail se proposait d'examiner. En utilisant l'information fournie par quelques marqueurs microsatellites polymorphes, fortement liés sur la carte génétique, l'idée était de compter le nombre d'haplotypes différents parmi le génotype des gamètes mâles ayant généré une famille de demi-frères récoltée sur un arbre mère de génotype connu, et d'en déduire le nombre de pères.

Un total de 9 marqueurs microsatellites nucléaires a été utilisé : 3 locus liés sur le groupe de liaison n°2, 2 locus liés plus faiblement sur le groupe n°7 et 4 locus indépendants. Des descendances maternelles au nombre de 43 ont été génotypées à ces locus (40 descendants), ainsi que leur mère. Pour estimer les fréquences alléliques de la population, 40 plantules ont également été génotypées. Pour mettre au point la méthode, nous avons tout d'abord simulé 10000 individus pendant 500 générations en utilisant les fréquences estimées à partir des plantules et les liaisons entre locus de la carte génétique de l'espèce. Huit descendances maternelles de 40 individus ont été créées à partir de cette population, en croisant une mère avec 5 à 40 pères, par pas de 5. Le nombre de pères, déduit du nombre d'haplotypes différents observés a été calculé pour chaque descendance et comparé au nombre utilisé dans la simulation, pour valider notre capacité à déduire ce nombre par cette méthode.

L'utilisation de locus liés améliore nettement l'estimation. La tendance est à la sous-estimation du nombre de pères réels (figure 8). Dans l'échantillon génotypé, au nombre de pères inconnu, le nombre de gamètes paternels différents observé a été évalué à 33, et d'après la qualité de l'estimation, le nombre de pères est au minimum de 27 dans cet échantillon (figure 8). Dans le dispositif où cet échantillon a été récolté, environ 40 pères potentiels sont présents, ce qui est cohérent avec l'estimation.

Cette méthode permet donc une bonne estimation du nombre de pères intervenant dans la constitution d'un lot de graines issues d'une même mère, et peut être appliquée à toute espèce pour laquelle des locus polymorphes et liés sont disponibles. Néanmoins, du fait du polymorphisme qui ne conduit pas à un nombre illimité d'haplotypes

paternels, cette méthode s'applique à des lots pour lesquels le nombre de pères n'est pas trop élevé.

Un travail d'estimation du nombre de pères a également été mis en place avec une étudiante suédoise, suite à ma collaboration (soutenue par un fond INRA-Formas) avec Martin Lascoux, professeur à l'Université d'Uppsala. Dans le cadre de sa thèse, Susanne Gustafsson a étudié le régime de reproduction d'une orchidée, l'orchis mouche, *Gymnadenia conopsea*. Cette plante présente des capsules contenant des graines très petites qu'il est impossible d'individualiser. Une population suédoise isolée de 24 plantes fleuries a été génotypée pour 7 locus microsatellites nucléaires. Des capsules (357 au total) de chacune de ces plantes ont été génotypées également : le génotype moyen de l'ensemble des graines présentes dans la capsule est alors obtenu.

La question est de savoir combien de pères ont contribué à la fécondation des graines présentes dans une capsule, le génotype de la mère étant connu. Suite à des critiques sur la méthode utilisée par l'étudiante pour retrouver ces pères, nous avons réfléchi à une approche statistique pour étayer les résultats. J'ai proposé de calculer, pour chaque capsule, la probabilité d'un échantillon de pères (qui correspondent aux plantes fleuries, hermaphrodites) divisée par le produit des fréquences des allèles observés dans la capsule. En testant tous les échantillons de pères possibles, il suffit d'extraire la combinaison présentant la vraisemblance la plus forte.

On trouve pour chaque capsule entre 1 et 17 pères probables, 7,4 en moyenne sur l'ensemble, soit 30 % des plantes de la population (figure 9). Les 24 plantes participent toutes à la reproduction comme père pour 15 à 255 capsules (soit 4 à 71 % du total des capsules), avec 110 capsules (31 %) en moyenne par père (et un écart-type de 66 capsules). Le père possédant le succès reproducteur le plus élevé contribue pour 34 capsules de plus que le précédent (191 à 225 capsules fécondées respectivement soit 54 % à 71 %). Une grande gamme de variation des succès reproducteurs est observée, ce qui témoigne d'une taille de population effective vraisemblablement inférieure à la taille observée, déjà faible à la base (Gustafsson et Gerber, en préparation).

2.4. Populations améliorées et flux de gènes

Les marqueurs sont des outils importants dans la gestion des populations végétales cibles de la sélection. Les approches biométriques autorisent des tests puissants d'assignation d'échantillons à des populations de référence.

2.4.1. Test variétal et marqueurs chez le pin maritime

Dans la forêt des Landes, des pins maritimes d'origine ibérique (espagnole et portugaise) ont été plantés dans les années 1950, du fait d'incendies advenus pendant et après la seconde guerre mondiale, qui ont détruit de grandes surfaces boisées (en 1943 et 1949). Suite à un hiver exceptionnellement froid en 1985 (températures de -22°C), ces provenances se sont révélées très sensibles au froid, 30000 hectares ont alors été perdus. Depuis 1986, les provenances plantées sont soumises à des tests pour garantir leur origine landaise. Ce test est basé sur la composition terpénique de rameaux frais d'un mélange de 30 arbres du peuplement à tester, analysé par chromatographie en phase gazeuse. L'idée de ce travail était de proposer une alternative à ce test, en se basant sur des données moléculaires.

Pour ce faire, les compositions haplotypiques chloroplastiques des populations landaises et ibériques ont été comparées. L'information obtenue grâce à 6 amorces chloroplastiques a été combinée pour obtenir les haplotypes d'échantillons d'origine portugaise (235 arbres), espagnole (68) et landaise (450). Un total de 71 haplotypes différents a été identifié dont 21 % étaient communs aux deux provenances, 55 % et 24 % étant spécifiques aux arbres français et portugais. L'étude de différenciation entre les provenances a montré que l'on ne distinguait pas significativement les deux provenances ibériques, réunies par conséquent en une seule. En revanche, la mise au point d'un test statistique d'écart entre les provenances landaise et ibérique a permis de construire des courbes de densités très disjointes, ce qui suggère une bonne puissance dans la capacité à déterminer l'origine d'un lot inconnu. L'erreur de type II (à savoir accepter un lot comme étant d'origine landaise alors qu'il ne l'est pas), qui doit être la plus faible possible, est inférieure à 1 % dans ce test. Le test moléculaire s'avère plus performant que le test terpénique (Ribeiro *et al.* 2002).

Une approche similaire a été développée au même moment, mais en utilisant trois locus marqueurs microsatellites nucléaires (Derory *et al.* 2002). Dans 47 populations de pins maritimes (France, Corse, péninsule ibérique, Maroc) des échantillons de 30 individus ont été génotypés aux trois locus. Une analyse en composantes principales des diversités génétiques des populations a permis de distinguer trois groupes géographiques : France-ouest, France-sud-est (comprenant la Corse) et péninsule ibérique (comprenant le Maroc).

Le test statistique construit précédemment a été adapté à des données diploïdes. Le test a été appliqué à une population plantée dans les Landes, suspectée d'origine Corse. Cette population a été comparée aux populations de référence landaise et corse et des simulations réalisées pour déterminer un seuil de rejet de l'hypothèse que la population suspectée était bien d'origine landaise. À l'issue de l'analyse en composantes principales, cette population était déjà localisée dans le groupe France-sud-est, le test statistique a confirmé qu'elle ne pouvait être distinguée des provenances corses. Cette distinction pouvait même être établie à partir de l'information d'un seul des trois locus, discriminant efficacement les provenances landaise et corse. Les informations fournies par ces marqueurs génétiques se révèlent donc intéressantes pour mieux connaître la structuration de la variabilité chez le pin maritime et aider non seulement à la gestion des ressources génétiques mais également à l'identification de l'origine des populations (Derory *et al.* 2002).

2.4.2. Fonctionnement des vergers à graines

Dans la gestion des populations améliorées par l'homme, le contrôle des flux de gènes va permettre de garantir les gains génétiques prédits, en limitant les effets d'un pollen non amélioré. J'ai contribué à plusieurs études de système de reproduction au sein de verger à graines destinés à produire des variétés améliorées d'arbres forestiers. Les résultats de ces études moléculaires et statistiques montrent l'importance de la contribution du pollen extérieur dans les dispositifs (avec absence d'isolement des arbres améliorés et pollinisation libre), ce qui conduit inmanquablement à diminuer le progrès génétique attendu suite à l'amélioration génétique.

Dans le cadre du programme d'amélioration génétique du pin maritime en France, des vergers à graines ont été installés pour produire une variété plus performante,

constitués notamment par les plants issus des croisements entre les meilleurs arbres de la génération de sélection (verger de polycross). Si l'on souhaite atteindre le gain génétique espéré, il faut néanmoins s'assurer que les géniteurs des plants ont contribué de façon égale à cette descendance, et que le flux de pollen en provenance de l'extérieur du verger ne soit pas trop important. Une étude a été faite pour répondre à ces questions et pour élaborer un test moléculaire permettant de garantir l'origine d'un lot de graines du verger, en se basant sur la diversité des haplotypes chloroplastiques des arbres, obtenue à partir de 6 locus.

Le verger étudié a été créé avec les plants issus des croisements de 28 arbres aux performances maximales dans la population d'amélioration. Ces arbres, conservés sous forme de copies clonales, ainsi que 320 arbres constituant une tranche du verger, ont été échantillonnés. L'haplotype chloroplastique de l'ensemble de ces arbres a été déterminé, 17 haplotypes différents ont été détectés. La contribution égale des géniteurs aux arbres du verger a été testée en comparant les profils alléliques des deux groupes d'arbres. Pour savoir si la différence de fréquences observée entre ces groupes était significative ou non, j'ai construit un test de rééchantillonnage (tirage avec remise parmi les parents), avec un recalcul de la statistique utilisée à chaque fois (différence au carré entre les fréquences des parents et celles du tirage), la valeur observée étant comparée à la distribution des valeurs obtenues après rééchantillonnage, pour juger si elle est extrême (auquel cas on peut conclure à la non représentativité des parents dans les descendants, ceux-ci étant considérés dans la construction du test comme un tirage aléatoire des haplotypes parentaux). Le résultat obtenu conduit à rejeter l'hypothèse de contribution égale des parents aux plants du verger, néanmoins, seul un des 6 locus est responsable de cette observation, les autres locus ne permettant pas de rejeter l'hypothèse.

Un test similaire a été construit pour permettre de tester l'origine de lots de graines, afin d'évaluer notre capacité à garantir leur provenance du verger. Quatre lots de graines récoltées à différents endroits du verger (53 individus à chaque fois), un lot issu du Médoc (50 individus), un lot de référence d'origine Aquitaine (371 arbres de 13 populations) et un lot de référence du verger (93 individus) ont été analysés pour leurs haplotypes chloroplastiques. La même statistique que précédemment a été utilisée pour comparer deux hypothèses alternatives : origine Aquitaine du lot contre origine du verger. Deux locus se sont révélés les plus discriminants pour distinguer les deux populations et un effet de la taille de l'échantillon a pu être noté, un échantillon de 93 graines permet une meilleure distinction qu'un échantillon de 50 (figure 10). Le test appliqué aux lots de graines nous a conduits à rejeter l'origine "verger" du lot du Médoc mais également de 3 des 4 lots du verger. Ceci suggère une contamination pollinique importante en provenance de l'extérieur du verger qui vient modifier significativement la composition des lots de graines du verger. Un lot de 309 graines récoltées dans le verger a permis de donner une idée de l'importance de la pollution pollinique. D'après le nombre de nouveaux haplotypes détectés, absent parmi les parents des croisements, le taux détecté était de 36,5 %. Ce taux représente une estimation minimum, en effet, des haplotypes identiques à ceux des parents peuvent avoir été apportés par des arbres extérieurs.

Le travail de thèse de Gilles Chaix, réalisé au CIRAD, concernait l'étude du régime de reproduction à l'intérieur d'un verger de production de graines d'*Eucalyptus grandis* à

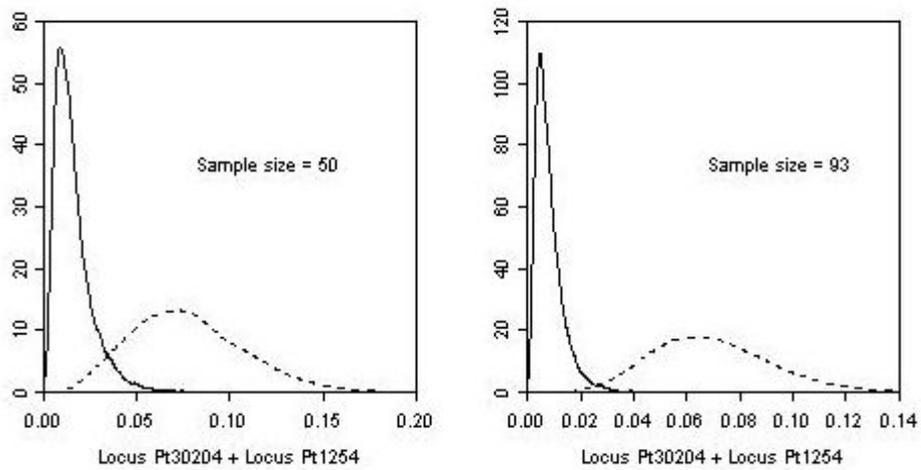


Figure 10 : distribution de la statistique de test pour le verger (ligne continue) et pour la population Aquitaine (pointillés) pour les deux locus les plus discriminants et avec deux tailles d'échantillon.

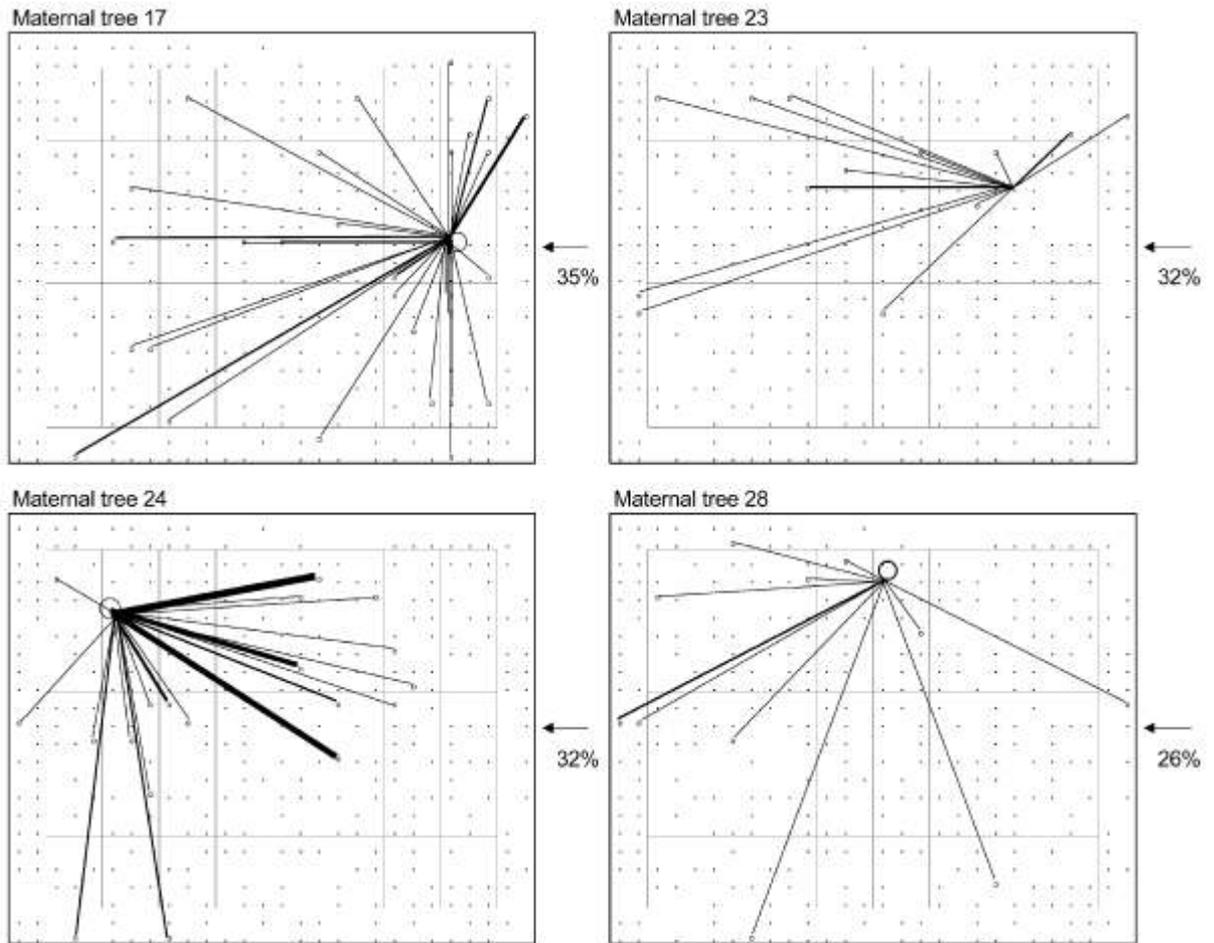


Figure 11 : résultat des recherches de paternité dans un verger à graines d'eucalyptus. Les traits relient 4 arbres mères et les pères retrouvés pour les graines analysées, l'épaisseur du trait est proportionnelle au nombre de graines fécondées par le père (pourcentages de flux extérieur au verger à droite des schémas).

Madagascar. J'ai participé à son comité de thèse et contribué à l'analyse de ses données. En génotypant à 6 locus microsatellite nucléaires 349 arbres adultes du verger (échantillonnage exhaustif) et 724 descendants récoltés sur 30 arbres mères parmi les adultes, une analyse de paternité a pu être conduite, en appliquant les méthodes que j'ai développées.

L'allofécondation observée sur les 30 mères varie de 73 à 100 % (96,7 % en moyenne), le pollen provient à 60,8 % de l'intérieur du verger (30,4 à 73,7 %) et donc près de 40 % provient de l'extérieur. Le succès reproducteur des pères suit une courbe en "L", classique, avec la moitié d'entre eux donnant un seul descendant et quelques-uns en donnant plusieurs (12 au maximum). Chaque père participant à la reproduction féconde 1 à 7 mères parmi les 30. Les différentes provenances constituant le verger ne sont pas également représentées dans les nuages polliniques, mais les différences restent faibles : sur 11 provenances présentes, seules deux sont sous-représentées (ratio de 0,5) et une seule est sur-représentée (ratio de 2). L'organisation spatiale des événements de pollinisation observés autour de chaque mère ne présente pas de structuration particulière (figure 11). Il n'y a pas de direction privilégiée chez cette espèce pollinisée par des insectes, pas d'excès de pollinisation à courte distance (seules 2 mères sur les 30 montrent un écart significatif) et pas d'effet notable de la position de la mère dans le verger. La distance moyenne de pollinisation est de 32 m avec un maximum de 85 m : des valeurs du même ordre de grandeur que la taille du verger (68 × 85m).

Le fonctionnement d'un verger à graines de pin maritime, constitué de 60 clones (728 copies au total sur 4 ha) a été étudié dans la forêt domaniale d'Escarpoum, au Portugal. Les 60 clones ont été décrits pour 3 locus microsatellites nucléaires, ainsi que 206 graines récoltées sur 7 arbres répartis selon les vents dominants. Parmi les 60 clones, seule une paire s'est révélée identique d'un point de vue génétique avec les trois marqueurs. Les pères des graines ont été recherchés avec le logiciel FaMoz, pour estimer le degré de contamination pollinique du verger : entre 46 et 56 % du pollen ayant fécondé les graines analysées provient de l'extérieur du verger. Seuls 20 % des clones ont contribué à ce lot de graines, soit 42 graines pour lesquelles un seul père a pu être assigné (11 pères différents retrouvés au total). Aucun lien n'a été détecté entre le nombre de copies d'un père donné présentes dans le verger et le nombre de graines pollinisées. Sur les 11 pères, seuls deux clones ont laissé une dizaine de graines, les autres en ont laissé entre 1 et 5 : non seulement peu de pères interviennent dans la reproduction, devancés par la masse de pollen provenant de l'extérieur, mais en plus, leurs contributions individuelles sont inégales.

Le verger n'est pas complètement isolé et les arbres les plus proches sont à 1-2 km, distance facilement parcourue par le pollen. Les estimations de flux de pollen en provenance de l'extérieur, de parcelles d'arbres d'autres espèces (conifères ou feuillus) sont généralement du même ordre de grandeur que le flux détecté dans cette étude. Il apparaît donc que dans le cas d'un verger peu isolé comme celui-ci, la contamination pollinique est importante et la contribution des arbres sélectionnés dans les vergers est amoindrie, baissant de ce fait le gain génétique attendu pour les graines produites (Fernandes *et al.*, soumis).

Chez le pin radiata de Nouvelle-Zélande, 15 parents de la population d'amélioration ont été utilisés en 1988 pour réaliser des croisements avec un mélange de pollens de ces

mêmes 15 arbres (croisement polycross). Parmi les 15 familles, 5 ont été retenues pour les analyses ultérieures. Un génotypage a été fait avec 4 marqueurs microsatellites nucléaires et 4 marqueurs microsatellites chloroplastiques pour 85 à 90 graines de 3 familles et une cinquantaine de plants de 5 familles (dont les 3 précédentes). L'assignation d'un seul parent a été possible pour 73 à 88 % des descendants (81 % en moyenne), 13 % apparaissaient comme fécondés par du pollen extérieur. Sur les 8 descendances analysées, une contribution équilibrée des 15 pères possibles parmi les pères assignés a pu être observée, sauf dans un cas. Cet équilibre entre contributions paternelles semble être lié, et ceci est confirmé dans la littérature, au nombre de pères mélangés dans le pollen, ce nombre doit être assez grand pour éviter que quelques génotypes prennent le dessus. En revanche, les descendances se sont révélées significativement différentes pour la composition des pères intervenant dans la reproduction. Aucun père en particulier ne s'est trouvé favorisé dans l'ensemble des familles, et leurs contributions relatives pouvaient être très variables d'une famille à l'autre: le génotype maternel intervient dans la sélection du génotype du gamète mâle. Cette étude permet de conclure que le croisement polycross est une bonne façon de représenter un ensemble d'arbres dans un croisement, à condition d'en utiliser au moins une dizaine. Aucun parent n'est sur-représenté et les contributions en terme de succès reproducteur sont assez équilibrées et non systématiquement biaisées. Ceci autorise une bonne estimation des aptitudes générales à la combinaison, et cette technique de croisement apparaît dès lors comme une bonne méthode à utiliser dans les programmes de sélection (Kumar *et al.* 2006).

2.4.3. Apparemment chez les vignes cultivées

Les vignes constituent une espèce domestiquée particulière : elles sont reproduites de façon clonale, mais de nombreux croisements sont à l'origine de leur diversification. L'analyse *a posteriori* de leur structure d'apparement va permettre de reconstituer une partie de cette histoire. Dans le cadre du travail de thèse de Manuel Di Vecchi, à l'Université de Montpellier en co-tutelle avec l'Université de Florence, un ensemble de 2919 cépages de vigne (comprenant quasiment l'ensemble des cépages cultivés en France et à peu près la moitié de l'ensemble des cépages connus dans le monde !) a été étudié à l'INRA de Montpellier pour sa structure d'apparement, afin de reconstituer l'histoire des croisements à l'origine des différents cépages. Ces cépages ont été typés pour 20 locus microsatellites, largement répartis dans le génome. L'association des logiciels FaMoz et Kinship (non limités en terme de nombre d'individus à analyser) a permis de reconstituer des liens parent-descendant et de calculer les degrés d'apparement entre individus sur la base des données génétiques, les familles de cépages étant reconstruites par ce moyen. Les apparentés au cépage Chardonnay, très largement diffusé et d'histoire ancienne, ont été plus particulièrement explorés. Les deux parents les plus vraisemblables de ce cépage ont été identifiés (Pinot noir et Gouais blanc), relation confirmée par des critères morphologiques. Environ 265 apparentés au cépage Chardonnay, parmi lesquels 109 dont les deux parents ont pu être identifiés, ont été détectés parmi les variétés incluses dans l'échantillon. Des pleins-frères (22), demi-frères (109), descendants (9) et neveux (26) ont pu être distingués : Pinot noir et Gouais blanc ont été utilisés dans un grand nombre de croisements dans l'histoire viticole, malgré la reproduction traditionnellement clonale de l'espèce depuis très longtemps. La variété Chardonnay se révèle ainsi impliquée dans un réseau de

relations denses parmi les cépages de vigne. La procédure développée dans ce travail peut être étendue à tous les cépages typés mais également appliquée à d'autres espèces (Di Vecchi *et al.* soumis).

2.5. Paternité et parenté chez les chênes européens : le projet OAKFLOW

La création du logiciel FaMoz avait été affichée comme "milestone" n°1 (première étape) du projet européen OAKFLOW. Dans ce projet, la mise en œuvre, à une très grande échelle, d'analyses de paternité et de parenté à des parcelles de chênes à travers l'Europe, va me permettre dans les prochains mois de collecter une grande série de données précises sur les flux de gènes contemporains, aussi bien intra-qu'interspécifiques, pour ces arbres forestiers. Un travail de synthèse sera réalisé avec les membres du projet européen OAKFLOW, maintenant officiellement achevé, en s'appuyant sur ces résultats et sur la mise en commun des ressources générées par le projet. Ces résultats orienteront vraisemblablement mes perspectives de recherche futures. L'analyse de l'ensemble du dispositif fera l'objet de la publication du numéro spécial d'une revue scientifique.

2.5.1. Principe

L'expérience de la Petite Charnie, parcelle de 5 ha de chênes sessiles et pédonculés de la Sarthe dans laquelle tous les arbres adultes avaient été génotypés ainsi qu'un millier de glands récoltés sur 13 arbres mères des deux espèces (Streiff *et al.* 1998, 1999), a été étendue à des parcelles de 9 autres pays européens. Une étude fine des flux de gènes contemporains (par pollen et graine) et de l'hybridation entre espèces de chênes à l'aide de marqueurs microsatellites hautement polymorphes a donc été conduite à cette échelle au cours de l'un des thèmes du projet européen OAKFLOW (2000-2005) : "Flux de gènes intra- et interspécifiques chez les chênes : mécanismes favorisant la diversité génétique et le potentiel adaptatif", qui associait 14 partenaires de 11 pays. Les analyses finales sont en cours, en appliquant notamment les outils que j'ai développés.

Les dix parcelles contenaient de 300 à 750 arbres, tous génotypés pour 5 à 8 locus microsatellites. Une moyenne de 22 arbres mères ont été échantillonnés dans chaque parcelle pour 48 descendants en moyenne soit 260 à 1540 descendants analysés au total (tableau 3). Tous les arbres ont été cartographiés dans les parcelles. En utilisant la méthode de recherche de paternité proposée par FaMoz, nous avons pu retrouver les pères d'un certain nombre de descendants de ces parcelles.

2.5.2. Premiers résultats

Le flux de gènes mesuré dans ces expériences, c'est-à-dire le pourcentage de grains de pollen fécondants provenant de l'extérieur de la parcelle est compris entre 42,78 et 89,87 % (69,98 % en moyenne avec un écart-type de 16,52, tableau 3). Ces flux sont donc élevés, ce qui n'est pas une surprise pour un arbre avec une dispersion pollinique anémophile, mais variable d'une parcelle à l'autre, et vraisemblablement aussi d'une année à l'autre (par exemple 4 % de différence entre 1998 et 2002 en Hollande). Un important brassage génétique est ainsi assuré. Les pères retrouvés présentent des succès reproducteurs variables avec en moyenne 1,2 à 4,3 descendants laissés par père

Tableau 3 : Assignation de paternité dans les expériences du projet OAKFLOW, résultats sans erreur de génotypage.

Pays	Parcelle	N	Nm / Nd	Ntd	Np	Fg
France	Petite Charnie	354	13 / 76	992	329	66,84
Italie	Tatti	295	30 / 28	841	467	44,47
Hollande	Meinweg 1998	369	5 / 41	700	188	73,14
Hollande	Meinweg 2002	369	3 / 147	440	136	69,09
Espagne	Islands	239	11/24	264	50	81,06
Ecosse	Dalkeith	754	44 / 26	1136	650	42,78
Danemark	Velling	365	33 / 17	553	56	89,87
Suède	Båtfors	631	47 / 23	1061	246	76,81
Suisse	Bueren	496	19 / 81	1543	656	57,49
Hongrie	Sopron	430	20 / 20	382	45	88,22

N : nombre de parents génotypés de la parcelle

Nm / Nd : nombre de mères échantillonnées / nombre de descendants moyen par mère

Ntd : nombre total de descendants analysés

Np : nombre de descendants ayant un père assigné

Fg : pourcentage de flux de gènes en provenance de l'extérieur de la parcelle

Tableau 4 : Analyse spatiale des résultats de recherche de paternité du projet OAKFLOW.

Pays	Parcelle	Np	Nd/p	Ne	Ne / Np	D	P	ss / pp / sp / ps	Aut
France	Petite Charnie	121	3,1	51,8	42,8	106	7,1	40 / 13 / 38 / 9	4,6
Italie	Tatti	152	5,5	49,5	32,6	0,002	0	78 / 1 / 18 / 2	2,8
Hollande	Meinweg 1998	86	2,2	35,6	41,4	67	0	5 / 65 / 27 / 3	2,7
Hollande	Meinweg 2002	61	2,2	24	39,3	72	0	0 / 80 / 16 / 4	0
Ecosse	Dalkeith	280	2,3	147	52,5	165	0	68 / 2 / 9 / 21	2,3
Danemark	Velling	41	1,4	30,6	74,6	49	0	0 / 100 / 0 / 0	5,4
Suède	Båtfors	162	1,5	126,6	78,1	182	0	0 / 100 / 0 / 0	0,8
Suisse	Bueren	154	4,3	40,8	26,4	70	0	78 / 20 / 2 / 0	3,4
Hongrie	Sopron	38	1,2	30	79,0	95	24	0 / 100 / 0 / 0	0

Ne : nombre efficace de pères

$(1/\sum(fi^2))$, où *fi* est le succès reproducteur relatif de chaque père)

Ne/Np : en %

D : distance moyenne (m) de pollinisation

P : probabilité d'avoir observé une distance inférieure en tirant des pères au hasard

ss / pp / sp / ps : pourcentage de croisement mère × père (sessile : s, pédonculé : p)

Aut : pourcentage d'autofécondation

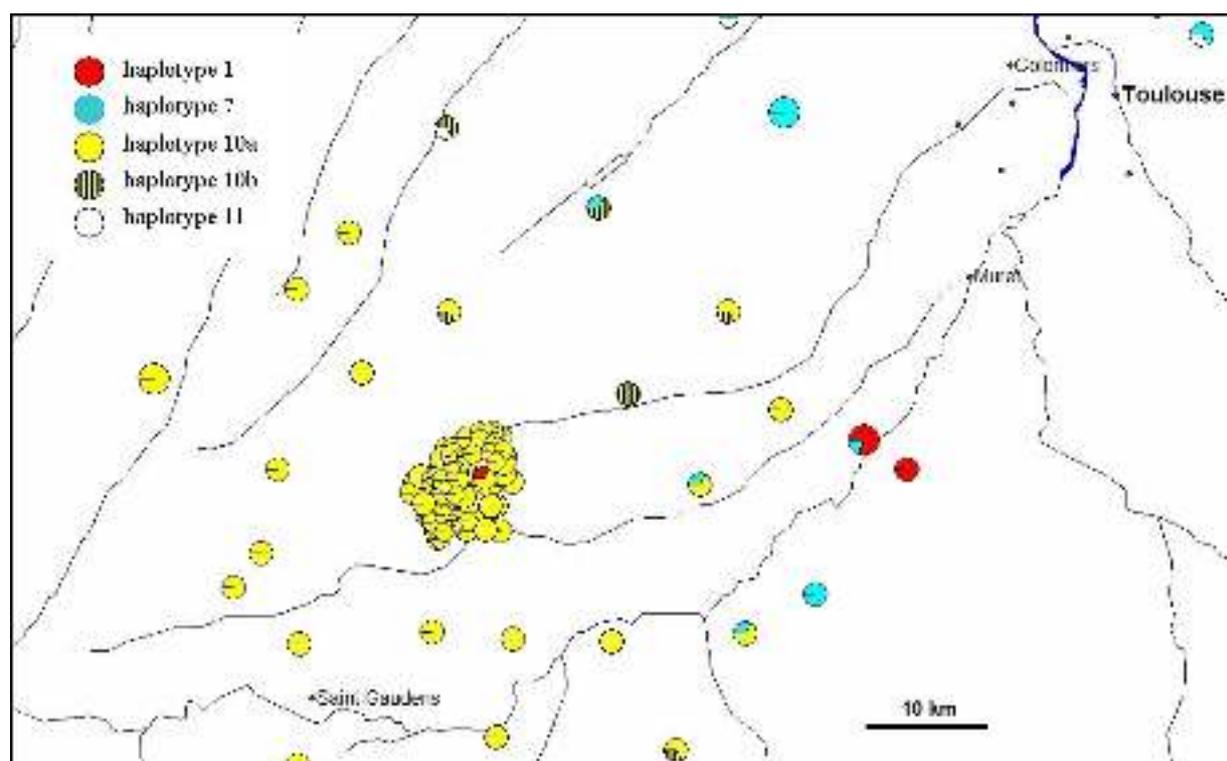
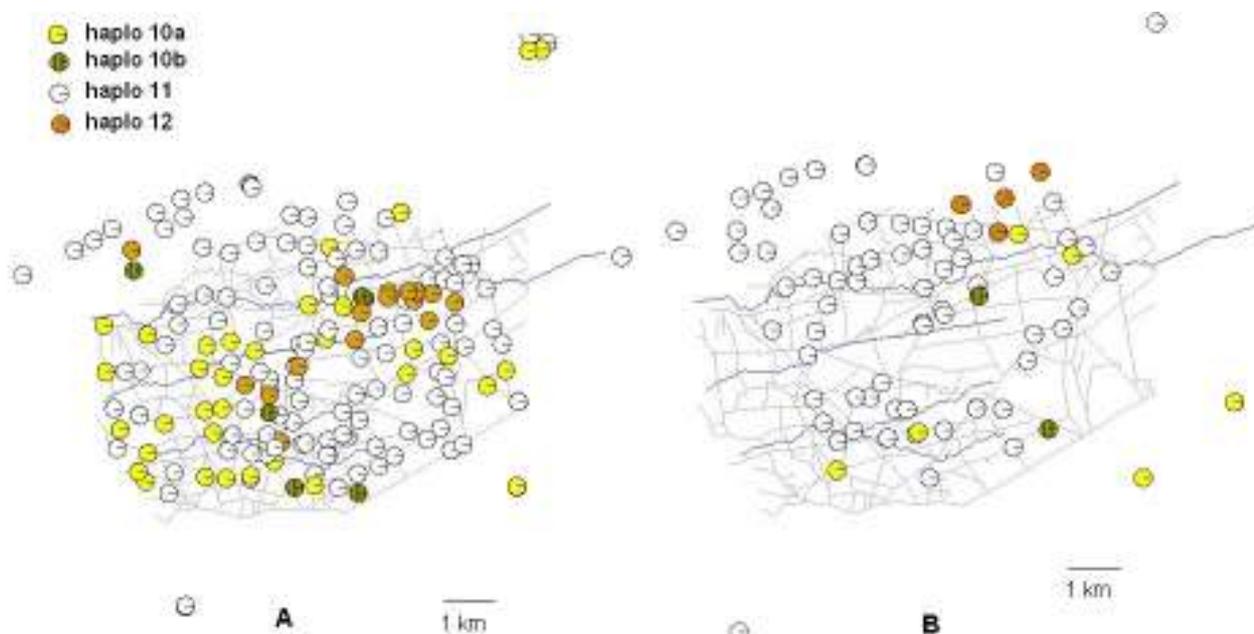


Figure 12 : structure génétique de marqueurs chloroplastiques des chênes de deux paysages du sud-ouest de la France, Tagon (haut, un arbre par cercle, A chênes pédonculés, B chênes tauzins) et Aurignac (bas, 3 à 5 arbres par cercle).

et un nombre efficace de pères (pondéré par le nombre de descendants par père) qui représente entre 26,4 % et 79 % du nombre de pères différents retrouvés (54,3 en moyenne et un écart-type de 20,3). Là encore, la variabilité entre parcelles est grande, et l'unique comparaison de deux années montre une stabilité. L'autofécondation, traditionnellement faible chez les arbres, hautement allogames, est néanmoins présente à faible taux, 2,4 % en moyenne avec un écart-type de 2, donc une gamme de variation là aussi significative (tableau 4). Les distances moyennes de pollinisation varient mais sont toutes significativement inférieures aux valeurs attendues avec un choix aléatoire de pères, en dehors des parcelles française et hongroise (tableau 4). Globalement, les flux de pollen intra-parcelle ont lieu sur des distances réduites, favorisant les arbres proches spatialement. Lorsque les espèces sessiles et pédonculés coexistent, les croisements intraspécifiques sont majoritaires, et l'hybridation se fait préférentiellement dans le sens mère sessile et père pédonculé, à l'exception de l'Écosse. Cette observation ne coïncide pas avec le sens de pollinisation préférentiel préconisé par l'hypothèse d'une recolonisation post-glaciaire des chênes sessiles *via* leur pollen, par fécondation des chênes pédonculés présents, plus pionniers (Petit *et al.* 2003).

2.5.3. Perspectives

La suite des analyses consistera notamment à reconstruire les courbes de dispersion pollinique, à partir des dispersions individuelles. Nous envisageons également d'étudier les descendants pour lesquels aucun père n'a pu être retrouvé grâce à des reconstitutions des apparentements sans pedigree, en collaboration avec Johan Fogelqvist, doctorant suédois (voir 3.2.2.). Une recherche de parenté sera également réalisée à partir de plantules récoltées dans certaines des parcelles. Elle nous renseignera sur les dispersions combinées du pollen et des graines.

2.6. Génétique, chênes et paysages

Parallèlement aux travaux présentés plus haut, spécifiquement consacrés à l'estimation des flux de gènes contemporains, je souhaitais présenter une étude à une autre échelle, qui concerne les chênes, mais au niveau du "paysage", une échelle, qui, associée à la génétique, est en vogue dans les publications scientifiques actuelles (Holderegger et Wagner (2006), préface à un numéro entier consacré à la question). Dans le cadre d'un projet financé par le Bureau des Ressources Génétiques (appel à propositions 2003), nous nous sommes intéressés à l'impact de l'écologie, de l'histoire et de la gestion sur la diversité et la structuration génétique des chênes dans deux paysages. Ce projet a fait l'objet de deux stages de master II "recherche" de l'Université de Bordeaux I successifs (Lepais 2004, Lavabre 2005).

Nous avons pour cela décrit la diversité génétique chloroplastique et nucléaire des chênes de deux paysages forestiers fragmentés du sud-ouest de la France. Le premier est situé en Gironde, dans le bassin du Tagon, à proximité d'Arcachon, et présente des chênes pédonculés et tauzins. L'échantillonnage comprend 250 chênes prélevés le long d'un maillage régulier de 500 m. À chaque point de prélèvement un chêne par espèce présente a été échantillonné. Le second paysage se trouve dans le canton d'Aurignac, en Haute-Garonne, au sud de Toulouse, et possède des chênes pédonculés, sessiles et pubescents. Au niveau régional, 66 arbres ont été échantillonnés à raison de 3 chênes

tous les 10 km dans un rayon de 30 km autour du canton d'Aurignac. A l'échelle du canton, 250 chênes ont été récoltés à raison de 3 arbres par population tous les 500 m. Enfin, dans un seul bois (la Paguères), 220 chênes ont été échantillonnés le long d'une trame régulière.

La diversité chloroplastique des deux paysages se révèle très contrastée (figure 12). Quatre haplotypes différents sont présents au Tagon, alors qu'un seul haplotype est exprimé à Aurignac, à l'exception d'un seul arbre sur les 250 du canton. Cet arbre est vraisemblablement le résultat d'un événement de dispersion à longue distance, puisque le même haplotype est retrouvé à une distance de 30 km du canton. La zone du Tagon a été décrite finement pour ses différents milieux, ce qui nous a permis de tester la liaison entre haplotype et milieu. Celle-ci est significative, certains haplotypes étant liés à certains milieux, sans que nous ayons pu trouver une explication convaincante à cette observation. Nous nous sommes heurtés à cet endroit à une question ancienne et récurrente en génétique écologique, sans être cependant allés très loin dans l'exploration des causes possibles de ce lien, à savoir "un problème avec trop de solutions" (Jones *et al.* 1977). Les bois de la région d'Aurignac sont transmis de génération en génération, cette stabilité s'accompagne d'une gestion très locale basée sur la régénération naturelle. L'exploitation du pin maritime au Tagon conduit à une dynamique paysagère qui se traduit par des cycles de colonisation - extinction pour les chênes, espèces non cibles de la gestion forestière. Leur structure haplotypique dans la zone est alors constamment remaniée, comme en témoignent les corrélations entre haplotypes et milieux. Contrairement à la région d'Aurignac, le Tagon est une zone instable où la dispersion des chênes par graine est plus intense, en relation avec les perturbations anthropiques du milieu. De plus, au Tagon, les deux espèces présentes diffèrent pour la composition spatiale en haplotypes, les chênes tauzins étant plus faiblement structurés que les chênes pédonculés, ce qui suggère des comportements de dispersion différents (Lepais *et al.* 2006a).

La diversité nucléaire (10 marqueurs microsatellites) du bois de la Paguères, géré en taillis sous futaie, a été comparée à celle d'une forêt gérée en futaie régulière en peuplement continu (Streiff *et al.* 1998). Les diversités sont équivalentes, le mode de gestion n'a pas ici de conséquence détectable sur la diversité intraspécifique. En revanche, les locus étudiés présentent un déficit en hétérozygotes. Si les chênes pédonculés de la Paguères de type franc-pied (issus de graine) et cépées (issus de rejet de souche) sont distingués, ce déficit devient non significatif pour chaque groupe : ces deux types d'arbres appartiennent à des générations différentes et leur mélange conduit à un déficit apparent en hétérozygotes (effet Wahlund).

Les chênes pubescents sont significativement plus diversifiés que les chênes pédonculés (respectivement, $H_e = 0,88 \pm 0,06$ et $H_e = 0,78 \pm 0,17$; test $F : p < 0,01$). Le contexte écologique semble jouer un rôle prédominant dans la diversité interspécifique ce qui pourrait influencer, par l'hybridation, la diversité intraspécifique. Le chêne pubescent semble en effet avoir recolonisé les milieux par introgression récurrente, ce qui expliquerait l'importante diversité nucléaire de cette espèce (Lepais *et al.*, soumis).

Alors que les marqueurs chloroplastiques décrits montrent que les chênes des deux paysages ont une origine maternelle ibérique commune, leur structuration géographique contrastée témoigne de dynamiques paysagères différentes. Nous n'avons pas pu déceler d'effet du traitement sylvicole sur les paramètres de diversité que nous avons examinés,

mais il existe des effets démographiques, et les espèces ont des comportements un peu différents.

Ce type d'approche en génétique du paysage sera certainement poursuivi dans la suite de mes travaux, sur d'autres espèces et dans la perspective de l'étude des dispersions à longue distance, élément clé de l'évolution des populations.

3. Perspectives : flux de gènes intra- et interspécifiques

Pour la suite de mes travaux de recherches, je souhaite conserver le thème des flux de gènes comme thématique centrale, tout en introduisant des considérations plus biologiques et écologiques dans les approches. L'estimation de l'impact des activités humaines, et notamment de la sylviculture, sur les populations d'arbres vues sous l'angle de la génétique sera également un aspect à approfondir. Mes recherches vont prendre une tournure plus interspécifique et élargir l'étude des arbres aux espèces associées. Dans notre unité de recherche, la volonté d'un travail interdisciplinaire a été affichée dès le début, entre les équipes de génétique, de pathologie, d'entomologie et d'écologie des communautés. Les deux projets structurants engagés vont dans ce sens (voir 3.3.). Dans le même esprit, le réseau d'excellence européen piloté par Antoine Kremer, qui regroupe 25 équipes de 15 pays différents, va orienter nos recherches dans les années à venir : EVOLTREE, "EVOLution of TREEs as drivers of terrestrial biodiversity" (évolution des arbres, moteurs de la biodiversité terrestre). Parallèlement aux chênes, espèce principale de mes recherches jusqu'à présent - et qui va rester centrale - nous nous intéresserons aux espèces liées, oiseaux, insectes, pathogènes, voire mycorhizes, et à leurs interactions. Les travaux vont petit à petit porter en grande partie sur un site intensivement étudié, sélectionné dans le cadre de ce réseau (ISS - Intensive Study Site). Il est situé dans les Landes sur une surface de 25000 hectares, et contient des sites Natura 2000. Il a été retenu dans la catégorie "gestion intensive", qui concerne le pin maritime, les chênes étant des espèces modèles dans le système.

Mes perspectives s'articulent autour de trois points majeurs. Le premier vise à étudier de façon approfondie la dynamique d'hybridation au sein du complexe d'espèce des chênes blancs, un travail en cours qui fait l'objet d'une thèse. En parallèle, il s'agit de continuer à développer les approches méthodologiques autour de ces recherches d'apparement, en affinant et enrichissant les méthodes disponibles, en complétant et adaptant le logiciel FaMoz, et en comparant les méthodes de reconstruction de fratrie, notamment sur les données "chêne" dont nous disposons. Le troisième point va dans le sens de la construction de mon unité de recherche d'un point de vue scientifique, en étendant les études d'abord aux espèces directement associées aux chênes (vertébrés, pathogènes), puis, ultérieurement, à des espèces indépendantes mais qui présentent des traits d'histoire de vie variés, en étudiant l'impact de ces caractéristiques sur les dispersions à longue distance.

3.1. Flux de gènes dans le complexe d'espèces des chênes blancs

Dans le cadre du travail de thèse d'Olivier Lepais (Université Bordeaux I, 2004-2007), nous avons choisi d'étudier de façon précise les flux de gènes contemporains parmi les chênes blancs, en forêt mais aussi grâce à des croisements contrôlés. Les quatre espèces de chênes considérées dans cette étude, chênes pédonculé (*Quercus robur*), sessile (*Q. petraea*), pubescent (*Q. pubescens*) et tauzin (*Q. pyrenaica*) font partie du même complexe d'espèces. Ils possèdent une large aire de distribution et les deux premières espèces constituent un tiers de la surface de la forêt française.

Ces quatre espèces sont parfaitement identifiables et apparaissent très nettement séparées d'un point de vue morphologique (feuilles et rameaux), les chênes pédonculés étant complètement séparés des pubescents, alors qu'un continuum est observé entre

ces derniers et les sessiles, qui présentent quelques formes intermédiaires avec les pédonculés. Les chênes pubescents et sessiles possèdent une grande variabilité morphologique, les chênes pédonculés étant plus homogènes (Badeau 1990, Dupouey et Badeau 1993). La différenciation morphologique des chênes sessiles et pédonculés se retrouve également de façon stable à travers toute l'Europe (Kremer *et al.* 2002). Ces espèces ont des préférences écologiques différentes. Les chênes pubescents sont cantonnés à des substrats calcaires secs. Les chênes pédonculés, présents sur des sols riches et humides, sont exigeants en eau et en éléments minéraux. Les chênes sessiles se trouvent sur des stations acides et neutres, mais sont très plastiques (Badeau 1990). De plus, ils occupent des places différentes dans la dynamique de succession des espèces d'un peuplement forestier : les chênes pédonculés sont plus pionniers que les chênes sessiles, pubescents et tauzin (Rameau *et al.* 1989).

Les chênes sessiles et pédonculés ont été étudiés conjointement dans l'équipe, notamment pour leur distinction moléculaire à l'aide d'isozymes (Zanetto *et al.* 1994, Bacilieri *et al.* 1996), de fragments d'ADN amplifiés aléatoirement (RAPD, Bodénès *et al.* 1997), de marqueurs microsatellites (P. Goikoetxea *et al.*, non publié). Récemment, les zones génomiques impliquées dans cette différenciation ont pu être localisées (Saintagne *et al.* 2004). Dans tous les cas, avec ces marqueurs neutres ou anonymes, ces espèces apparaissent peu différenciées, ce qui traduit un flux de gènes important au sein du complexe d'espèces.

L'étude des flux de gènes et de l'hybridation des chênes sessiles et pédonculés a été conduite dans le cadre du projet européen OAKFLOW (voir 2.5.). Le travail de thèse est orienté par les résultats de ce projet, sur les mouvements de graines et de pollen, le comportement des hybrides, et l'importance de la sylviculture sur les flux de gènes, et il est étendu en plus à deux autres espèces du complexe.

Notre but est d'étudier la dynamique d'hybridation entre les quatre espèces, *in-* et *ex-situ*. En milieu naturel, dans une forêt avec un mélange des espèces en proportions variables (forêt de Briouant / la Paguères de Gasque, commune du Fousseret, Haute-Garonne, au sud de Toulouse), un génotypage d'individus adultes a été réalisé (800 arbres des 4 espèces, principalement pédonculés et tauzins, pubescents en nombre moindre et sessiles très minoritaires), pour établir un profil de fréquences alléliques des quatre espèces. Les arbres adultes ont été localisés géographiquement. Une récolte de plantules (200 au total) a également été réalisée sur 5 carrés de 1 m de côté, localisés sur le terrain. Chaque plantule a été repérée spatialement, et sa hauteur mesurée. Elles sont toutes génotypées et suivies dans le temps, pour leur survie, leur hauteur, leur diamètre.

Une trentaine d'arbres ont été observés au printemps 2006 pour leurs date de débourrement, afin d'estimer en quoi les décalages phénologiques pourraient expliquer l'isolement des espèces. Les floraisons des quatre espèces sont en effet décalées mais chevauchantes, ce qui exclut cette seule explication dans l'éventuelle absence d'hybridation.

Dix locus microsatellites ont été sélectionnés parmi plus d'une centaine lors du projet européen OAKFLOW. Ils discriminent efficacement les chênes sessiles et pédonculés. La discrimination des chênes pubescents avec les mêmes locus a été testée au laboratoire, et est probante (Lepais 2004). Les génotypages sont donc réalisés pour 10 locus microsatellites, associés dans deux PCR et révélés sur séquenceur 96 capillaires, grâce à

la mise au point technique réalisée dans le cadre de la thèse, qui fait l'objet d'une première publication (Lepais *et al.* 2006b).

La description génétique fine de l'ensemble des individus est dès à présent acquise, et le degré d'introggression de chaque arbre peut être mesuré à l'aide des outils disponibles, tels le logiciel Structure (Pritchard *et al.* 2000). Une analyse de parenté réalisée sur les plantules a montré l'absence d'hybrides, alors que certains arbres adultes apparaissent nettement introgressés. En vérifiant si, parmi les adultes, des relations parents-enfants existaient, l'un des arbres s'est avéré résulter du croisement de deux autres adultes, de deux espèces différentes, et ce même arbre apparaît en effet comme un hybride F1 entre les deux espèces, avec l'approche du logiciel Structure, ce qui valide les deux analyses.

Une récolte de glands sera réalisée à l'automne 2006 sur une trentaine d'arbres mères des quatre espèces. Une récolte est prévue sur l'arbre même et au sol, afin d'apprécier un éventuel effet de la dispersion. Une trentaine de glands par arbre seront prélevés et génotypés. La contribution paternelle pourra alors être déduite, ainsi que son appartenance spécifique, d'après les profils de fréquences précédemment déterminés. Les techniques d'assignation à mettre en œuvre feront l'objet de développements méthodologiques, en comparant les performances des outils existants, et en développant de nouveaux, à l'aide de la modélisation et de simulations.

Parallèlement, nous avons procédé au laboratoire à une série de croisements contrôlés entre les quatre espèces, une technique désormais bien maîtrisée (Roussel 2005). Un mélange des pollens des mêmes espèces ou de différentes espèces en proportions variables sera fait, et les croisements seront conduits sur les quatre espèces, dans tous les sens possibles. Plus d'un millier de glands a été obtenu suite à ces croisements à l'automne 2005. Un morceau de chaque gland a été prélevé avant plantation, pour le génotypage ultérieur. La germination et la croissance des plants sont suivis. Les marqueurs génétiques (typage de 5 locus fin 2006, début 2007), nous permettront de vérifier que le croisement a bien eu lieu et de mesurer le succès reproducteur relatif de chacun des cinq arbres présents dans les mélanges polliniques. La qualité génétique des glands pourra être comparée à leur taux de survie et à leur croissance. Les résultats obtenus *ex-situ* seront comparés aux résultats *in-situ*.

Des hybrides entre ces espèces réalisés en 1989 et à présent florifères sont également disponibles. On vérifiera avec des marqueurs leur caractère hybride ou non. Les succès de fécondation seront comparés, et la nature des croisements réussis sera contrôlée *a posteriori* grâce aux marqueurs. Des croisements contrôlés ont été réalisés en 2006 sur ces arbres avec les mêmes pollens que précédemment. Une analyse des glands obtenus, de leur germination, de la viabilité et de la morphologie des jeunes plantules en pépinière sera également réalisée. La fréquence et le sens des croisements préférentiels seront précisés. Les résultats des croisements sur les hybrides apporteront des informations précieuses sur les flux suite à une première génération d'hybridation : le mécanisme de maintien des espèces pourra ainsi être mieux compris.

Les bénéfices attendus de cette étude sont une bonne estimation des taux d'hybridation moyens au sein des quatre espèces, aussi bien en milieu naturel qu'en conditions artificielles. La comparaison des deux types d'approches permettra d'évaluer un effet éventuel du milieu naturel sur le potentiel des espèces.

Cette étude profitera de l'approfondissement des approches méthodologiques que j'entends mener sur les questions de recherche d'apparement, directement applicables aux données disponibles.

3.2. Approches méthodologiques

3.2.1. Flux de gènes instantanés intra- et inter-population

Dans le prolongement des travaux développés au préalable et valorisés notamment dans la mise à disposition du logiciel FaMoz (voir 2.2.), je souhaite continuer à approfondir des questions méthodologiques.

Les méthodes d'assignation de paternité ou de parenté développées dans le logiciel FaMoz sont basées sur des calculs de rapport de vraisemblance, uniquement liés au génotype de chacun des individus considérés et aux fréquences alléliques moyennes de la population de référence. Grâce à cette approche, la probabilité qu'un individu donné soit le père ou le parent d'un descendant particulier est divisée par la probabilité que n'importe quel autre individu de la population soit le vrai parent. Plus cette probabilité est élevée, plus la décision penchera en faveur de l'individu pour l'attribution de paternité.

L'idée serait d'utiliser des données biologiques mesurées sur chaque individu de la population et d'affiner ainsi les recherches de parenté, en ajoutant dans le calcul des vraisemblances des variables autres que purement génétiques. Ainsi, les caractéristiques phénologiques, qui vont isoler les arbres les uns des autres en les rendant incompatibles suite à des décalages de floraison, les caractéristiques phénotypiques, classes de fertilités - et donc succès reproductif - liées à un diamètre par exemple. Il serait également utile d'intégrer les informations spatiales directement dans l'analyse. L'une des sorties directe des études serait alors, sans intermédiaire de calcul, les courbes de dispersions les plus adaptées aux observations de parenté. Une collaboration avec Etienne Klein, biométricien INRA à Avignon, pourrait porter sur ce point.

Du point de vue des marqueurs, il serait à prévoir d'inclure dans FaMoz les calculs concernant les marqueurs SNP (Single Nucleotide Polymorphism), qui touchent des bases uniques de l'ADN, et ne présentent que deux allèles. Leur analyse au laboratoire est largement automatisable et donne lieu à des génotypages de très nombreux locus, à très haut débit (par exemple 10000 locus chez l'homme dans une publication récente (Carothers *et al.* 2006)...). Pour parvenir à la même puissance d'analyse que les microsatellites, un nombre nettement plus grand de marqueurs SNP est nécessaire - de l'ordre de 150 - mais ils présentent l'avantage de donner accès à des résolutions analytiques d'équations, impossibles avec des marqueurs multialléliques : ainsi l'erreur de génotypage peut être considérée de manière efficace et rigoureuse, et des tests statistiques peuvent être construits sur des bases solides (Anderson et Garza 2006). Chez les chênes, nous ne disposons pas encore de tels marqueurs en nombre suffisant pour une utilisation en routine, mais cela est à prévoir pour l'avenir.

Si le nombre de marqueurs utilisés augmente, l'hypothèse de leur indépendance, nécessaire au calcul simple des vraisemblances (additivité des valeurs sur l'ensemble des locus, Meagher et Thompson 1986), devient caduque. Il faut alors inclure dans l'analyse la liaison entre locus. L'accès à l'information sur les phases de locus liés peut se révéler utile dans l'assignation génétique (Lexer *et al.* 2000, voir 2.3.2.). La liaison entre locus

induit néanmoins des modifications dans les calculs, et si des travaux existent pour des estimations d'apparentement (par exemple Thompson et Meagher 1998, Ayres et Balding 2004), l'adaptation des calculs de vraisemblance dans le cas parent-descendant reste à mettre au point, à ma connaissance.

Très récemment, des auteurs ont proposé une méthode d'assignation de parenté associée à des méthodes bayésiennes pour intégrer des données qualifiées de "comportementales" et des données spatiales (Hadfield *et al.* 2006). Il reste à tester ce type d'approche sur les données dont nous disposons chez les arbres.

La réflexion sur l'évolution du logiciel FaMoz doit être faite par rapport à ces développements. L'article qui décrit le logiciel (Gerber *et al.* 2003) a été cité 11 fois (6 fois en 2005, 5 fois en 2006), ce qui reste assez faible en 3 ans, alors que la description de la méthode mise en œuvre (Gerber *et al.* 2000b) dans le logiciel a été citée 53 fois en 6 ans (avec régularité, 10 citations par an environ depuis 2002) : elle a plus de succès... Il est difficile de prévoir quel avenir aura ce logiciel. La recherche de paternité/parenté va continuer à se faire à l'aide de marqueurs, et pour la paternité, le logiciel Cervus (Marshall *et al.* 1998) domine tous les autres largement en tête des classements, les études basées sur des marqueurs dominants restant peu nombreuses. L'avenir du logiciel n'est donc pas garanti, il pourrait néanmoins gagner des utilisateurs en proposant les méthodes décrites plus haut, et en étant d'utilisation plus aisée, ce qui nécessiterait l'aide d'un informaticien. Un interfaçage avec le logiciel libre R, très utilisé dans notre communauté, sous forme d'un "package" indépendant est à envisager. L'accès à des graphes de grande qualité par ce biais est un des avantages.

3.2.2. Méthodes de reconstitution des apparentements sans pedigree

A l'issue d'une analyse de paternité, et notamment chez les arbres, où des flux de gènes élevés sont généralement la règle, le pourcentage de "déchets" dans les données est important : chez les chênes par exemple, dans les données du projet OAKFLOW, 70 % des descendants analysés n'ont pas de père attribué parmi les parents génotypés, les nombreux génotypes correspondants peuvent donc paraître inutilisables en l'état. Pour remédier à cela, il serait intéressant de tenter une reconstitution des apparentements parmi les descendants sans père, afin d'estimer le nombre de pères différents intervenant dans le pollen en provenance de l'extérieur des dispositifs, et de comparer les données intra- et hors-parcelle.

Les méthodes à mettre en œuvre pour ce type d'analyse sont toutes basées sur des estimations de degré d'apparentement à l'aide de l'information fournie par les marqueurs (Capy et Brookfield 1991). Celles-ci s'appuient sur l'exclusion d'apparentement (Almudevar et Field 1999), sur des calculs bayésiens (Painter 1997, Emery *et al.* 2001), sur une approche probabiliste basée sur les fréquences alléliques (Fiumera *et al.* 2001), sur du maximum de vraisemblance (Letcher et King 2001). Cette dernière approche a été également combinée à la théorie des graphes (Beyer et May 2003). L'introduction des chaînes de Markov associées à des algorithmes de Monte-Carlo améliore la qualité des résultats (Smith *et al.* 2001). Quatre de ces méthodes appliquées à des données réelles et simulées conduisent à une moyenne de 70 à 98 % des individus, affectés par une erreur de génotypage ou une mutation, à être mal classés (Butler *et al.* 2004). Pour intégrer cette grande sensibilité aux erreurs, Wang (2004) ajoute à une méthode de maximum de vraisemblance un modèle d'erreur de génotypage qui limite les mauvaises classifications. L'intérêt et l'utilité de ce type d'analyse sont démontrés dans

une étude de populations menacées de saumons, mais 9 locus polymorphes sont utilisés (Herbinger *et al.* 2006) : le projet OAKFLOW n'en propose pas autant. Enfin une méthode, utilisant le "recuit simulé", mérite également d'être testée (Fernández et Toro 2006).

La bibliographie du sujet esquissée ici montre la richesse des approches. Nous disposons d'un jeu de données conséquent, et j'ai confié à Johan Fogelqvist, doctorant à l'Université d'Uppsala, le soin de comparer les différentes méthodes grâce aux programmes et logiciels disponibles auprès des auteurs, et d'en déduire une description fine des flux de gènes "hors parcelle" de nos dispositifs.

3.3. Interactions biotiques et biodiversité

Mes recherches vont évoluer de plus en plus vers la prise en compte d'espèces associées à notre espèce cible, les chênes. Les premières études viseront à comprendre comment les interactions entre espèces, positives ou négatives, influent les flux de gènes de l'espèce cible. Ensuite, nous nous intéresserons aux flux au sein d'espèces différentes, pour mesurer en quoi les traits d'histoire de vie vont les influencer. Pour stimuler une dynamique entre les équipes de l'UMR Biogeco, deux projets structurants interdisciplinaires ont été lancés. Je suis plus particulièrement concernée par le second, "Les interactions biotiques comme moteurs de la biodiversité lors de la régénération des chênes", piloté par Rémy Petit. Dans ce projet, nous voulons, pour commencer, étudier deux espèces ou groupe d'espèces et leur association avec les chênes, *via* les flux de gènes, pour un vertébré disséminateur de graines, le geai, d'une part et pour les pathogènes associés aux régénérations, d'autre part.

3.3.1. Dispersion des chênes sous l'effet des geais

Le geai est connu pour favoriser l'installation de nouveaux chênes, grâce à sa dispersion active de glands, suggérant une véritable symbiose avec l'arbre (Ducouso et Petit 1994). Nous souhaiterions mieux connaître l'impact d'un tel disperseur sur les chênes et sur les dispersions réalisées. Un projet dans ce sens a été proposé par Arndt Hampe et Rémy Petit dans le cadre d'un programme d'accueil de chercheur de l'Union européenne, le "Marie Curie Intra-European Fellowships", il s'intitule "ACORNDISP: Demo-genetic consequences of acorn dispersal and predation by vertebrates on oak populations". Arndt Hampe est ainsi actuellement en post-doctorat pour deux ans dans notre équipe. Une placette d'observation a été installée au sein d'une zone étudiée par nos collègues entomologistes. Sur 6 hectares, elle comprend des pins maritimes, qui font l'objet de la gestion, et des groupes de chênes, essentiellement pédonculés mais également tauzins (environ 10 %), qui ont 50 à 60 ans pour les plus gros arbres. 200 arbres adultes ont été repérés et cartographiés. Leur hauteur, leur degré d'isolement, et la quantité de cupules sous l'arbre, supposée corrélée à la quantité de glands produits, ont été estimés.

Environ 400 plantules ont été échantillonnées (prélèvement d'une feuille) et marquées de façon permanente sur l'ensemble de la placette. Ces plantules sont toutes issues de la glandée de l'année précédant l'échantillonnage, ceci vérifié par la présence du gland encore visible à la base, et dispersées probablement par des geais, les groupes d'arbres adultes les plus proches étant distants de 50 à 100 m. Une grande densité de plantules est observée le long du chemin qui coupe le dispositif et autour d'un plan d'eau : les

geais sont connus pour apprécier ces zones ouvertes. De plus, on observe des alignements sur des plants plus âgés, ce qui témoigne de l'installation des chênes entre les lignes de pins, coupés ensuite en conservant les chênes présents.

Un second échantillonnage de plantules a été réalisé à proximité des arbres adultes, en prenant un individu toutes les 5 plantules pour parvenir à un effectif total gérable, d'environ 500 individus. Cet échantillon, également issu des glands de l'automne précédant la récolte, est supposé non dispersé. Toutes les plantules ont été caractérisées pour leur hauteur, leur survie au cours du temps, leur degré d'infestation. L'infestation est d'ailleurs plus élevée à proximité des chênes adultes (voir 3.3.2.).

En utilisant les méthodes de génotypage à haut débit disponibles (Lepais *et al.* 2006b), nous caractériserons les structures génétiques des arbres adultes et des deux échantillons de plantules, et les comparerons, notamment d'un point de vue spatial. Les plantules seront suivies pendant plusieurs années, pour tracer l'évolution de la composition génétique de la population. Une analyse de parenté sera réalisée par mes soins pour retrouver le ou les parents de chacune des plantules et estimer les dispersions. Il sera alors possible de vérifier *a posteriori* si les deux échantillons de plantules correspondent bien à des dispersions moyennes différentes. L'observation parallèle du comportement des geais, qui sera faite à l'automne, permettra de voir si l'on retrouve au niveau génétique des traces de ce comportement (trajets répétés d'un endroit de récolte à un lieu de stockage des glands par exemple). Une reconstitution de parenté sans pedigree (voir 3.2.2.) sera appliquée aux plantules pour lesquelles aucun parent n'aura pu être retrouvé. Leur structuration familiale pourra ainsi être reconstituée, et notamment le nombre de mères c'est-à-dire le nombre d'arbres sources différents échantillonnés par les geais. La différence entre graines dispersées ou non sera également étudiée de ce point de vue, mesurant ainsi l'impact évolutif du geai sur le chêne.

3.3.2. Dispersion des graines : génétique et pathologie

Si les glands sont dispersés efficacement et à de grandes distances par des vertébrés comme les geais, il est logique de penser que, d'après leur morphologie et sous l'effet de la barochorie (dissémination due au poids), la dispersion de ces graines a lieu à courte distance et essentiellement à proximité de l'arbre mère. La recherche des parents de jeunes plants dans une parcelle où tous les adultes alentour avaient été génotypés (expérience répétée dans les parcelles du projet OAKFLOW, en cours d'analyse), a montré que de nombreux plants n'avaient aucun parent à proximité, pas même leur mère, suggérant par là une dispersion de glands plus importante que prévue. Un apport extérieur d'au moins 20 % des glands avait pu être détecté (Gerber *et al.* 2004). Une des explications à cette observation est dérivée de l'hypothèse de Janzen-Connell (Janzen 1970, Connell 1971), élaborée pour rendre compte de la grande diversité des forêts tropicales. Le cortège d'espèces pathogènes associées à un arbre, du fait d'une certaine coévolution, serait mieux adapté aux génotypes apparentés à cet arbre. Ainsi, les graines dispersées à proximité de l'arbre mère seraient plus à même d'être détruites par ces pathogènes que les autres. Le résultat serait alors une augmentation des distances de dispersion, les graines de provenance plus lointaine étant favorisées du fait de l'inadaptation des pathogènes.

Nous souhaitons tester cette hypothèse chez le chêne en collaboration avec Cyril Dutech, chercheur de l'UMR Biogeco dans l'équipe de pathologie. Pour ce faire, nous allons étudier d'un point de vue génétique l'expérience mise en place par ce chercheur. Le dispositif a été installé en forêt domaniale de Campet, dans le Lot-et-Garonne, grâce à la coopération de l'Office National des Forêts et du Département de la Santé des Forêts, dans une parcelle de chênes pédonculés d'environ 80 ans. Des récoltes de descendances ont été effectuées sur 21 arbres (suffisamment isolés pour que les glands récoltés à leur pied puissent leur être attribués) à l'automne 2003. Ces 21 arbres ont été choisis dans 3 zones, de façon à échantillonner une gamme de distances inter-arbres de 5 à 500 m. Une éclaircie a été réalisée au cours de l'hiver 2003-2004 et une placette expérimentale a été définie (nettoyage superficiel, barrière de grillage) à proximité de chacun des 21 arbres. En mars 2004, le dispositif expérimental a été installé : les glands (vernalisés au laboratoire) des 21 descendances ont été semés dans chaque placette, à raison de 10 blocs (lignes) par placette, avec une répétition par descendance et par bloc.

Des notations sur la survie et la croissance des chênes, l'infection par l'oïdium et les autres dégâts éventuels ont été réalisées de façon hebdomadaire d'avril à juillet, puis mensuellement, en 2004, 2005 et 2006. Une feuille a été prélevée par plantule et conservée pour une analyse ultérieure. Le génotypage des arbres mères et de leurs descendants sera réalisé en utilisant les techniques mises au point dans l'équipe de génétique (Lepais *et al.* 2006b), lors d'un stage de master II en 2007. Nous confronterons alors la constitution génétique des plantules, leur distance à l'arbre mère, leur survie, leur croissance et leur sensibilité aux parasites afin de mieux comprendre la dynamique de mise en place d'une régénération aux tous premiers stades. Nous suivrons le dispositif sur une période plus longue pour étudier son évolution dans le temps.

Les études futures se concentreront sur les ISS (Intensive Study Site) du réseau Evoltree, et sur l'ISS des Landes plus particulièrement. Dans ce contexte, au-delà de l'analyse des chênes et de leurs espèces associées, l'idée serait de comparer les flux de gènes observés dans des espèces différentes, aux caractéristiques biologiques et démographiques contrastées, en mesurant par exemple la vitesse de dispersion d'un gène d'un endroit à l'autre du paysage, sur de longues distances, en fonction de ces critères. L'approche multidisciplinaire possible au sein même de notre unité sera un plus pour proposer des interprétations aux phénomènes observés. Les dispersions à longue distance sont un moteur important dans la dynamique des espèces, posent des problèmes méthodologiques, et sont un défi motivant pour l'étude de l'évolution actuellement (Cain *et al.* 2000, Nathan *et al.* 2003, Nathan 2006).

3.4. Une école-chercheur "philosophie et histoire des sciences biologiques"

Pour dépasser les questions purement disciplinaires rencontrées par chaque chercheur dans son travail au quotidien, il me semble important de trouver un moyen d'examiner les questions de la science de façon plus générale dans le cadre professionnel. Ceci est possible au sein de l'INRA, où de nombreuses disciplines sont représentées, et alentour, où des éléments de philosophie et d'histoire des sciences peuvent être abordés.

J'ai ainsi participé en janvier 2002 à une école-chercheur sur le thème "science et éthique : une raison commune ?", organisée par le groupe Ethos, "Éthique Économique et Sociale", de l'INRA. En mai 2005, j'ai assisté à un séminaire international au Muséum

National d'Histoire Naturelle intitulé "Entre la philosophie de la biologie et la philosophie de l'écologie : évolutionnismes, écologies et éthiques". Je suis régulièrement les séminaires organisés à l'Université Bordeaux I par le laboratoire Epistémé, "Épistémologie - Histoire des sciences et des techniques".

A travers ma propre expérience, il me semble que la culture des chercheurs en matière de philosophie et d'histoire des sciences est faible. Elle me paraît pourtant fondamentale dans l'approche de ce métier et permet d'ajouter une dimension et une profondeur à notre réflexion. Une bonne introduction à la philosophie des sciences est proposée de façon abordable par A.F. Chalmers dans son livre "Qu'est-ce que la Science" (Chalmers 1987). N'ayant pas vu jusqu'à présent de proposition de ce type, j'aimerais organiser une école-chercheur au sein de l'INRA consacrée à la philosophie et à l'histoire des sciences, et plus particulièrement des sciences biologiques, adaptée à un public de non spécialistes, et donc soucieuse de faciliter une appropriation du domaine. Dans ma discipline plus spécifique, la génétique, un petit livre introductif "génétique et éthique" peut servir de repère (Allais 1995). De manière plus approfondie, Jean Gayon (professeur de philosophie à l'Université Paris I) offre des analyses extrêmement éclairantes dans le domaine de la philosophie et de l'histoire des sciences en génétique (Gayon 1992).

J'imagine cette école-chercheur organisée sur une semaine, avec une première partie consacrée à des éléments généraux d'épistémologie, et une seconde partie plus thématique, sur des questions ciblées, à partir de l'étude de textes, et de l'accompagnement de philosophes du domaine. Je voudrais élaborer cette école-chercheur en collaboration avec Anne Léon-Miehe, ancienne élève de l'ENS Ulm, agrégée et docteur en philosophie, qui a enseigné la philosophie des sciences aux élèves scientifiques de l'ENS-Lyon, et contribué à la création de l'enseignement de philosophie en première année de médecine à la Faculté Lyon-Sud, Université Claude Bernard-Lyon I (Miehe 1994, Léon-Miehe 2005). Nous aimerions proposer d'une part des interventions sur le développement des sciences du vivant au XVIII^{ème} à travers le renouvellement des modèles animaux et végétaux dans l'histoire naturelle et les théories de la reproduction. Nous souhaiterions aborder d'autre part les questions sur la théorie de l'évolution et les débats qu'elles suscitent actuellement : l'ouvrage majeur de Stephen Jay Gould, La structure de l'évolution, tout récemment traduit en français (Jay Gould 2006), pourrait être un bon point de départ. La mise en relation de ces deux champs d'investigation et de réflexion devrait permettre une bonne approche des notions de modèle et de théorie ainsi qu'un examen de leur fonction dans la structuration des sciences et de leur domaine d'application. Parmi les orateurs, nous souhaiterions inviter notamment François Dagognet, Jean Gayon, Jean-Didier Vincent. Le premier sondage entrepris il y a quelques temps auprès des partenaires INRA m'a permis de constater un intérêt certain de mon département de recherches, et de plusieurs directions scientifiques, et leur soutien assuré pour le projet. Les services de la formation permanente de l'INRA offrent de plus une aide à l'organisation de ce type d'événement.

Si cette expérience s'avère concluante, pourquoi ne pas à l'avenir créer un groupe d'animation plus durable, dont les modalités d'existence restent à définir ?

4. En guise de conclusion

En arrivant dans mon équipe actuelle, mes thématiques de recherche ont évolué, en conservant les aspects liant marqueurs génétiques et biométrie, présents dès mon DEA et toujours entretenus par la suite, mais en se tournant vers la génétique des populations. Cette dernière discipline a été l'un des moteurs de mes choix de spécialisation à l'Institut National Agronomique, j'ai donc été heureuse de la voir réapparaître dans mon travail. Des considérations esthétiques, qui jouent parfois dans certains décisions, m'incitent à être également satisfaite de travailler sur un arbre comme le chêne, rejoignant doublement ainsi, par les hasards des pérégrinations professionnelles, un sujet que j'aurais aimé travailler dès la thèse. Je réalise sans peine la chance de travailler dans un lieu comme Pierroton, où les éléments moteurs ne manquent pas, à commencer par un prix "Nobel" forestier (prix Marcus Wallenberg 2006) à qui nous devons beaucoup. Participer de sa petite contribution à l'édifice est un privilège. La collaboration constante avec des étudiants, de Bordeaux, d'ailleurs en France ou de l'étranger est toujours motivante et enrichissante.

L'ouverture multispécifique et multidisciplinaire qui se profile à l'horizon de notre unité est également stimulante, le travail du chercheur étant malgré tout très solitaire à la base, quoi que l'on en dise. Le souci de ne pas s'isoler doit être constant. Nous faisons partie d'une communauté internationale structurée par des moyens informatiques extraordinaires, les réseaux tout particulièrement : l'illusion de la communication universelle peut aussi avoir tendance à nous confiner devant notre écran, en discutant seul face au monde entier.

J'ai été très régulièrement amenée à manipuler les outils de biométrie lors de mes travaux et ce en collaboration avec des biométriciens de l'INRA (Jean-Baptiste Denis (dans ma préhistoire), François Rodolphe, Bruno Goffinet, Magali San Cristobal, Joël Chadoeuf, Etienne Klein), et je compte bien poursuivre dans cette voie. Une formation complémentaire en biométrie serait bienvenue, mais n'existe pas en tant que telle dans ma région. Je trouverais intéressant de pouvoir suivre des cours de master en biométrie (à Paris ou Lyon) mais le problème pratique me paraît pour le moment insurmontable...

Enfin, j'ai à cœur de mener à bien le projet de cette école-chercheur "philosophie et histoire des sciences" et espère lancer la procédure dès le début de l'année 2007. C'est une chance du métier de chercheur que de pouvoir élargir son horizon de la sorte.

5. Références

- Allais C. (1995) Génétique et éthique. Éditions Hachette.
- Almudevar A., Field C. (1999) Estimation of single generation sibling relationships based on DNA markers. *Journal of Agricultural, Biological and Environmental Statistics* 4: 136-165.
- Anderson E.C., Garza J.C. (2006) The power of single-nucleotide polymorphisms for large-scale parentage inference. *Genetics* 172: 2567-2582.
- Arcade A., Labourdette A., Falque M., Mangin B., Chardon F., Charcosset A., Joets J. (2004) BioMercator: integrating genetic maps and QTL towards discovery of candidate genes. *Bioinformatics* 20 (14): 2324-2326.
- Ayres K.L., Balding D.J. (2005) Paternity index calculations when some individuals share common ancestry. *Forensic Science International* 151: 101-103.
- Bacilieri R., Ducouso A., Kremer A. (1996) Comparison of morphological characters and molecular markers for the analysis of hybridization in sessile and pedunculate oak. *Ann Sci For* 53: 79-91.
- Bacles C.F.E., Lowe A.J., Ennos R.A. (2006) Effective seed dispersal across a fragmented landscape. *Science* 311 (5761): 628.
- Badeau V. (1990) Étude de la variabilité morphologique des chênes de Lorraine. DEA de biologie forestière, Université Nancy I, 35 p.
- Bailey N.T.J. (1961) Introduction to the mathematical theory of genetic linkage. Oxford at the Clarendon Press.
- Beyer J., May B. (2003) A graph-theoretic approach to the partition of individuals into full-sib families. *Molecular Ecology* 12: 2243-2250.
- Bodénès C., Joandet S., Laigret F., Kremer A. (1997) Detection of genomic regions differentiating two closely related oak species *Quercus petraea* (Matt.) Liebl. and *Quercus robur* L. *Heredity* 78: 433-444.
- Bossart J.L., Prowell D.P. (1998) Genetic estimates of population structure and gene flow: limitations, lessons, and new directions. *Trends Ecol Evol* (13): 202-206
- Boyd W.C. (1949) Use of blood groups in cases of disputed paternity. *N Engl J Med.* 241 (19): 759.
- Butler K., Field C., Herbinger C.M., Smith B.R. (2004) Accuracy, efficiency and robustness of four algorithms allowing full sibship reconstruction from DNA marker data. *Molecular Ecology* 13: 1589-1600.
- Cain M.L., Milligan B.G., Strand A.E. (2000) Long-distance seed dispersal in plant populations. *Amer J Bot* 87: 1217-1227
- Campbell I., McDonald K., Flannigan M., Kringayark J. (1999) Long-distance transport of pollen into the Arctic. *Nature* 399: 29-30.
- Capy P., Brookfield J.F.Y. (1991) Estimation of relatedness in natural populations using highly polymorphic genetic markers. *Genet Sel Evol* 23: 391-406.
- Carothers A. D., Rudan I., Kolcic I., Polasek O., Hayward C., Wright A.F., Campbell H., Teague P., Hastie N.D., Weber J.L. (2006) Estimating human inbreeding coefficients: comparison of genealogical and marker heterozygosity approaches *Annals of Human Genetics* 70(5): 666-676
- Chalmers A.F. (1987) *Qu'est-ce que la Science*. Éditions la Découverte.
- Chardon F., Hourcade D., Combes V., Charcosset A (2005) Mapping of a spontaneous mutation for early flowering time in maize highlights contrasting allelic series at two-linked QTL on chromosome 8. *Theoretical and Applied Genetics* 112 (1): 1-11.
- Chardon F., Virlon B., Moreau L., Falque F., Joets J., Decousset L., Murigneux A., Charcosset A. (2004) Genetic architecture of flowering time in maize as inferred from quantitative trait loci meta-analysis and synteny conservation with the rice genome. *Genetics* 168 (4): 2169-2185.
- Connell J.H. (1971) On the role of natural enemies in preventing competitive exclusion in some marine animals and in forest trees. in P.J. den Boer and G.R. Gradwell, editors. *Dynamics of populations*. Centre for Agricultural Publishing and Documentation, Wageningen, The Netherlands: 298-312.
- Derory J., Mariette S., González-Martínez S.C., Chagné D., Madur D., Gerber S., Brach J., Persyn F., Ribeiro M.M., Plomion C. (2002) What can nuclear microsatellites tell us about maritime pine genetic resources conservation and provenance certification strategies ? *Ann. For. Sci.* 59, Iss 5-6, Sp. Iss.: 699-708.

- Di Vecchi M., Laucou V., Lacombe T., Varès D., Gerber S., Boselli M., This P. (soumis) FaMoz: a software for large scale parentage analysis in *vitis vinifera* L. species.
- Ducouso A., Petit R. (1994) Le geai des chênes - premier reboiseur européen. *Forêts-entreprise* 97: 60-64.
- Dupouey J.L., Badeau V. (1993) Morphological variability of oaks (*Quercus robur* L, *Q. petraea* (Matt) Liebl, *Q. pubescens* Willd) in northeastern France: preliminary results. *Ann Sci For* 50 (Suppl 1): 35s-40s.
- Ellstrand N.C. (1984) Multiple paternity within the fruits of the wild radish, *Raphanus sativus*. *American Naturalist* 123 (6): 819-828.
- Ellstrand N.C., Marshall D.L. (1985) Variation in extent of multiple paternity among plants and populations of wild radish. *American Journal of Botany* 72 (6): 876.
- Emery A.M., Wilson I.J., Craig S., Boyle P.R., Noble L.R. (2001) Assignment of paternity groups without access to parental genotypes: multiple mating and developmental plasticity in squid. *Molecular Ecology* 10: 1265-1278.
- Fabre F. (1998) Amélioration génétique de la qualité de la graine et fixation symbiotique de l'azote chez le soja (*Glycine max* L.), Thèse de l'INP - ENSA Toulouse.
- Fernandes L., Rocheta M., Cordeiro J., Pereira S., Gerber S., Oliveira M.M., Ribeiro M.M. (soumis) Paternity analysis and mating system in a *Pinus pinaster* Aiton clonal seed orchard using microsatellites.
- Fernández, J., Toro M.A. (2006) A new method to estimate relatedness from molecular markers. *Molecular Ecology* 15: 1657-1667.
- Fiumera A.C., DeWoody Y.D., Dewoody J.A., Asmussen M.A., Avise J.C. (2001) Accuracy and precision of methods to estimate the number of parents contributing to a half-sib progeny array. *J Hered* 92: 120-126.
- Gallais A. (1990) Théorie de la sélection en amélioration des plantes. Masson, Paris.
- Gayon J. (1992) Darwin et l'après Darwin : une histoire de l'hypothèse de sélection naturelle. Éditions Kimé.
- Gerber S., Chabrier P., Kremer A. (2003) FaMoz: a software for parentage analysis using dominant, codominant and uniparentally inherited markers. *Molecular Ecology Notes* 3: 479-481.
- Gerber S., Fabre F., Planchon C. (2000a) Genetic of seed quality in soybean analysed by capillary gel electrophoresis. *Plant Science* 152 (2): 181-189.
- Gerber S., Latouche-Hallé C., Lourmas M., Morand-Prieur M.E., Oddou-Muratorio S., Schibler L., Bandou E., Caron H., Degen B., Frascaria-Lacoste N., Kremer A., Lefèvre F., Musch B. (2004). Mesure directe des flux de gènes en forêt. Actes du 4ème Colloque National "Le patrimoine génétique : la diversité et la ressource", La Châtre, 14-16 octobre 2002 : 349-368.
- Gerber S., Streiff R., Bodénès C., Mariette S., Kremer A. (2000b) Comparison of microsatellites and AFLP markers for parentage analysis. *Molecular Ecology* 9: 1037-1048.
- Goffinet B., Gerber S. (2000) Quantitative trait loci: a meta-analysis. *Genetics* 155: 463-473.
- González-Martínez S.C., Gerber S., Cervera M.T., Martínez-Zapater J.M., Alía R., Gil L. (2003) Selfing and sibship structure in a two-cohort stand of maritime pine (*Pinus pinaster* Ait.) using nuclear SSR markers. *Ann. For. Sci.* 60: 115-121.
- González-Martínez S.C., Gerber S., Cervera M.T., Martínez-Zapater J.M., Gil L., Alía R. (2002) Seed gene flow and fine-scale structure in a Mediterranean pine (*Pinus pinaster* Ait.) using nuclear microsatellite markers. *Theor. Appl. Genet.* 104: 1290-1297.
- Gustafsson S., Gerber S. (en préparation) Mating pattern in the orchid *Gymnadenia conopsea*.
- 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.
- Herbinger C.M., O'Reilly P.T.J., Verspoor E. (2006) Unravelling first-generation pedigrees in wild endangered salmon populations using molecular genetic markers. *Molecular Ecology* 8(15): 2261-2275.
- Holderegger R., Wagner H.H. (2006) Preface - A brief guide to Landscape Genetics. *Landscape Ecology* 21: 793-796.
- Janzen D.H. (1970) Herbivores and the number of tree species in tropical forests. *American Naturalist* 104: 501-528.
- Jay Gould S. (2006) La structure de la théorie de l'évolution. Éditions Gallimard.
- Jones A.G., Ardren W.R. (2003) Methods of parentage analysis in natural populations. *Molecular Ecology* 12: 2511-2523

- Jones G.P., Planes S., Thorrold S.R. (2005) Coral reef fish larvae settle close to home. *Current Biology* 15 (14): 1314-1318.
- Jones J.S., Leith B.H., Rawlings P. (1977) Polymorphism in *Cepaea*: a problem with too many solutions? *Annual Review of Ecology and Systematics* (8): 109-143.
- Kimura M. (1953) "Stepping-stone" model of population. *Annual Report of the National Institute of Genetics* 3: 62-63.
- Kremer A., Dupouey J.L., Deans J.D., Cottrell J., Csaikl U., Finkeldey R., Espinel S., Jensen J., Kleinschmit J., Van Dam B., Ducouso A., Forrest I., de Heredia U.L., Lowe A.J., Tutkova M., Munro R.C., Steinhoff S., Badeau V. (2002) Leaf morphological differentiation between *Quercus robur* and *Quercus petraea* is stable across western European mixed oak stands. *Annals of Forest Science* 59: 777-787.
- Kumar S., Gerber S., Richardson T.E. (2006) Testing for non-random contribution of pollen parents using microsatellites and chloroplast markers in polycross families of *Pinus radiata* D. Don. *Tree Genetics & Genomes*, sous presse.
- Lavabre J. (2005) Conséquences de la gestion forestière sur la structure génétique d'une chênaie. Master 2 Recherche "Sciences et Technologies", mention "Systèmes écologiques", Université Bordeaux I.
- Lanier L. (1994) Précis de sylviculture. 2^{ème} édition, ENGREF, Nancy.
- Léon-Miehe A. (2005) Le paysage entre description et fiction, ETA Hoffmann : l'image du dépaysement. Dans "Paysage et ornement", sous la direction de Baldine Saint Girons et Didier Laroque, Paris, Éditions Verdier.
- Lepais O. (2004) Étude comparée de la structuration génétique des chênes à l'échelle d'un paysage. Master Recherche "Génétique et Développement des Plantes", Université Bordeaux I et II.
- Lepais O., Gonzalez M., Cabanettes A., Gerber S. (soumis) Spatial structure of chloroplast DNA polymorphism of European white oaks in two landscapes: historical and management output.
- Lepais O., Lavabre J., Gonzalez M., Willm J., Cabanettes A., Gerber S. (2006a). Diversité et structuration génétique des chênes à l'échelle de deux paysages : impact de l'écologie, de l'histoire et de la gestion. *Les Actes du BRG n°6 (6ème Colloque National, La Rochelle, 2-4 Octobre 2006)*, 543-557.
- Lepais O., Léger V., Gerber S. (2006b) High throughput microsatellite genotyping in oak species. *Silvae Genetica* 55: 4-5.
- Letcher B.H., King T.L. (2001) Parentage and grandparentage assignment with known and unknown matings: application to Connecticut River Atlantic salmon restoration. *Canadian Journal of Fisheries and Aquatic Sciences* 58: 1812-1821.
- Lexer C., Heinze B., Gerber S., Kampfner-Macalka S., Steinkellner H., Kremer A., Glössl J. (2000) Microsatellite analysis of maternal half-sib families of *Quercus robur*, pedunculate oak (II): Inferring the number of pollen donors from the offspring. *Theor. Appl. Genet.* 100 (6): 858-865.
- Malécot G. (1950) Quelques schémas probabilistes sur la variabilité des populations naturelles. *Annales de l'Université de Lyon*, 3e série, A, fasc. 13.
- Mariette S. (2001) Mesure de la diversité génétique intra et inter-populations : échantillonnage intra-génome et choix des marqueurs. Applications expérimentales sur *Quercus petraea*, *Quercus robur* et *Pinus pinaster*. Thèse de l'ENGREF.
- Marshall T.C., Slate J., Kruuk L.E.B., Pemberton J.M. (1998) Statistical confidence for likelihood-based paternity inference in natural populations. *Molecular Ecology* 7: 639-655.
- Meagher T.R., Thompson E.A. (1986) The relationship between single parent and parent pair genetic likelihoods in genealogy reconstruction. *Theoretical Population Biology* 29: 87-106.
- Miehe A. (1994) La raison par alphabet : un instrument à la mesure de la démesure à l'époque de l'Encyclopédie. Dans "La mesure, instruments et philosophie", sous la direction de Jean-Claude Beaune, Seyssel, Éditions Champ Vallon.
- Mookerjee S., Guerin J., Collins G., Ford C., Sedgley M. (2005) Paternity analysis using microsatellite markers to identify pollen donors in an olive grove. *Theoretical and Applied Genetics* 111 (6): 1174-1182.
- Morris R. W., Critchfield W.B., Fowler D.P. (1980) The putative Austrian × red pine hybrid: a test of paternity based on allelic variation at enzyme-specifying loci. *Silvae Genetica* 29 (3/4): 93-100.
- Nathan R., Perry G., Cronin J.T., Strand A.E., Cain M.L. (2003) Methods for estimating long-distance dispersal. *Oikos* 103: 261-273.
- Nathan R. (2006) Long-distance dispersal of plants. *Science* 313: 786-788.

- Neigel J.E. (1997) A comparison of alternative strategies for estimating gene flow from genetic markers. *Annu.Rev.Ecol.Syst* 28: 105-128
- Neigel J.E. (2002) Is F_{ST} obsolete? *Conservation Genetics* 3 (2): 167-173.
- Ouborg N., Piquot Y. Vangroenendael J. (1999) Population genetics, molecular markers and the study of dispersal in plants. *Journal of Ecology* 87 (4): 551-568.
- Painter I. (1997) Sibship reconstruction without parental information. *J. Agricultural, Biol. Environ. Statistics* 2: 212-229.
- Paterson A.H., Lander E.S., Hewitt J.D., Peterson S., Lincoln S.E., Tanksley D.T. (1988) Resolution of quantitative traits into mendelian factors by using a complete linkage map of restriction fragment length polymorphisms. *Nature* 335: 721-726.
- Petit R.J., Bodénès C., Ducouso A., Roussel G., Kremer A. (2003) Hybridization as a mechanism of invasion in oaks. *New Phytologist* 161: 151-164.
- Plomion C., LeProvost G., Pot D., Vendramin G.G., Gerber S., Decroocq S., Brach J., Raffin A., Pastuszka P. (2001) Pollen contamination in a maritime pine (*Pinus pinaster* Ait.) polycross seed orchard and certification of improved seeds using cpSSR. *Can. J. For. Res.* 31: 1816-1825.
- Pospíšková M., Šálková I. (2006) Population structure and parentage analysis of black poplar along the Morava River. *Can J For Res* 36 (5): 1067-1076.
- Pritchard J.K., Stephens M., Donnelly P.J. (2000) Inference of population structure using multilocus genotype data. *Genetics* 155: 945-959.
- Rameau J.C., Mansion D., Dumé G. (1989) Flore forestière française. 1. Plaines et collines. Institut pour le Développement Forestier, 565-571.
- Renshaw M.A., Saillant E., Broughton R.E., Gold J.R. (2006) Application of hypervariable genetic markers to forensic identification of 'wild' from hatchery-raised red drum, *Sciaenops ocellatus*. *Forensic Science International* 156: 9-15.
- Ribeiro M.M., LeProvost G., Gerber S., Vendramin G.G., Anzidei M., Decroocq S., Marpeau A., Mariette S., Plomion C. (2002) Origin identification of maritime pine stands in France using chloroplast simple-sequence repeats. *Ann. For. Sci.* 59: 53-62.
- Rodzen J.A., Famula T.R., May B. (2004) Estimation of parentage and relatedness in the polyploid white sturgeon (*Acipenser transmontanus*) using a dominant marker approach for duplicated microsatellite loci. *Aquaculture* 232 (1-4): 165-182.
- Roussel G. (2005) Techniques de pollinisation sur les chênes européens au laboratoire de génétique et d'amélioration des arbres forestiers de Pierroton. Cahier technique, INRA, UMR Biogeco, 49 p.
- Saintagne C., Bodénès C., Barreneche T., Pot D., Plomion C., Kremer A. (2004) Distribution of genomic regions differentiating oak species assessed by QTL detection. *Heredity* 92: 20-30.
- Sezen U.U., Chazdon R.L., Holsinger K.E. (2005) Genetic consequences of tropical second-growth forest regeneration. *Science* 307: 891.
- Smith B.R., Herbinger C.M., Merry H.R. (2001) Accurate partition of individuals into full-sib families from genetic data without parental information. *Genetics* 158: 1329-1338.
- Streiff R., Ducouso A., Lexer C., Steinkellner H., Gloessl J., Kremer A. (1999) Pollen dispersal inferred from paternity analysis in a mixed stand of *Quercus robur* L. and *Quercus petraea* (Matt.) Liebl. *Molecular Ecology* 8: 831-841.
- Streiff R., Labbe T., Bacilieri R., Steinkellner H., Gloessl J., Kremer A. (1998) Within-population genetic structure in *Quercus robur* L. and *Quercus petraea* (Matt.) Liebl. assessed with isozymes and microsatellites. *Molecular Ecology* 7: 317-328.
- Tero N., Aspi J., Siikamäki P., Jäkäläniemi A. (2005) Local genetic population structure in an endangered plant species, *Silene tatarica* (Caryophyllaceae). *Heredity* 94: 478-487.
- Thoday J.M. (1961) Location of polygenes. *Nature* 192: 368-370.
- Thompson E.A., Meagher T.R. (1998) Genetic linkage in the estimation of pairwise relationship. *Theor Appl Genet* 97: 857-864.
- Wang J.L. (2004) Sibship reconstruction from genetic data with typing errors. *Genetics* 166: 1963-1979.
- Wiener A.S. (1950) Heredity of the Rh blood types. IX. Observations in a series of 526 cases of disputed parentage. *Am J Hum Genet.* 2(2): 177-197.
- Zanetto A., Roussel G., Kremer A. (1994) Geographic variation of inter-specific differentiation between *Quercus robur* L. and *Quercus petraea* (Matt.) Liebl. *Forest Genetic* 1: 111-123.

II. LISTE DES PUBLICATIONS

1. Publications scientifiques

1.1. Articles primaires dans périodique à comité de lecture

1.1.1. Imbert A., Gerber S., Tran V., Perez S., 1990. Data bank of Three-dimensional structures of disaccharides, a tool to build 3-D structures of oligosaccharides. *Glycoconjugate Journal* 7: 27-54.

1.1.2. Gerber S., Rodolphe F., Bahrman N., Baradat P., 1993. Seed protein variations of a pine (*Pinus pinaster* Ait) revealed by two-dimensional electrophoresis: genetic determinism and construction of a linkage map. *Theor. Appl. Genet.* 85: 521-528.

1.1.3. Gerber S., Rodolphe F., 1994. An estimation of the genome length of maritime pine (*Pinus pinaster* Ait.). *Theor. Appl. Genet.* 88: 289-292.

1.1.4. Gerber S., Rodolphe F., 1994. Estimation and test for linkage between markers: a comparison of lod score and X^2 test in a linkage study of maritime pine (*Pinus pinaster* Ait.). *Theor. Appl. Genet.* 88: 293-297.

1.1.5. Gerber S., Baradat Ph, Marpeau A., Arbez M., 1995. Geographic variation in terpene composition of *Pinus nigra* ARN. *Forest Genetics* 2(1): 1-10.

1.1.6. de Vienne D., Burstin J., Gerber S., Leonardi A., Le Guilloux M., Murigneux A., Beckert M., Bahrman N., Damerval C., Zivy M., 1996. Two-dimensional electrophoresis of proteins as a source of monogenic, codominant markers for population genetics and mapping the expressed genome. *Heredity* (76): 166-177.

1.1.7. Gerber S., Lascoux M., Kremer A., 1997. Relation between protein markers and quantitative traits in maritime pine (*Pinus pinaster* Ait.). *Silvae Genetica* 46(5): 286-291

1.1.8. Lexer C., Heinze B., Gerber S., Kampfer-Macalka S., Steinkellner H., Kremer A., Glössl J. 2000. Microsatellite analysis of maternal half-sib families of *Quercus robur*, pedunculate oak (II): Inferring the number of pollen donors from the offspring. *Theor. Appl. Genet.* 100 (6): 858-865.

1.1.9. Gerber S., Fabre F., Planchon C. 2000. Genetic of seed quality in soybean analysed by capillary gel electrophoresis. *Plant Science* 152 (2): 181-189.

1.1.10. Goffinet B., Gerber S. 2000. Quantitative trait loci: a meta-analysis. *Genetics* 155: 463-473.

1.1.11. Gerber S., Streiff R., Bodénès C., Mariette S., Kremer, A. 2000. Comparison of microsatellites and AFLP markers for parentage analysis. *Molecular Ecology* 9: 1037-1048.

1.1.12. Plomion C., LeProvost G., Pot D., Vendramin G.G., Gerber S., Decroocq S., Brach J., Raffin A., Pastuszka P., 2001. Pollen contamination in a maritime pine (*Pinus pinaster* Ait.) polycross seed orchard and certification of improved seeds using cpSSR. *Can. J. For. Res.* 31: 1816-1825.

1.1.13. Lexer C., Heinze, B., Gerber S., MacalkaKampfer S., Steinkellner H., Ziegenhagen B., Kremer A., Glosl J., 2001. Microsatellite analysis of small anonymous seedlot samples from pedunculate oak (*Quercus robur*): A promising approach to monitor the number of different seed parents and pollen donors. *Genetic Response of Forest Systems to Changing Environmental Conditions*. G. MullerStarck and R. Schubert. PO Box 17/3300 AA Dordrecht/Netherlands, Kluwer Academic Publ: 239-250.

1.1.14. Ribeiro M.M., LeProvost G., Gerber S., Vendramin G.G., Anzidei M., Decroocq S., Marpeau A., Mariette S., Plomion C 2002. Origin identification of maritime pine stands in France using chloroplast simple-sequence repeats. *Ann. For. Sci.* 59: 53-62.

1.1.15. Derory J., Mariette S., González-Martínez S.C., Chagné D., Madur D., Gerber S., Brach J., Persyn F., Ribeiro M.M., Plomion C., 2002. What can nuclear microsatellites tell us about maritime pine genetic resources conservation and provenance certification strategies ? *Ann. For. Sci.* 59, Iss 5-6, Sp. Iss.: 699-708

1.1.16. González-Martínez S.C., Gerber S., Cervera M.T., Martínez-Zapater J.M., Gil L., Alía R. 2002. Seed gene flow and fine-scale structure in a Mediterranean pine (*Pinus pinaster* Ait.) using nuclear microsatellite markers. *Theor. Appl. Genet.* 104 : 1290-1297.

1.1.17. Gerber S., Chabrier P., Kremer A. 2003. FaMoz: a software for parentage analysis using dominant, codominant and uniparentally inherited markers. *Molecular Ecology Notes* 3: 479-481.

1.1.18. González-Martínez S.C., Gerber S., Cervera M.T., Martínez-Zapater J.M., Alía R., Gil L. 2003. Selfing and sibship structure in a two-cohort stand of maritime pine (*Pinus pinaster* Ait.) using nuclear SSR markers. *Ann. For. Sci.* 60: 115-121

1.1.19. Chaix G., Gerber S., Razafimaharo V., Vigneron P., Verhaegen D., Hamon S. 2003. Gene flow estimation with microsatellites in a Malagasy seed orchard of *Eucalyptus grandis*. *Theor. Appl. Genet.* 107: 705-712

1.1.20. Gerber S., Latouche-Hallé C., Lourmas M., Morand-Prieur M.E., Oddou-Muratorio S., Schibler L., Bandou E., Caron H., Degen B., Frascaria-Lacoste N., Kremer A., Lefèvre F., Musch B. 2004. Flux de gènes par pollen et par graines chez quelques espèces forestières, exemples des chênes, de l'angélique, de l'alisier, du cèdre et du frêne. *Rendez-vous technique de l'ONF, hors-série n° 1* : 16-23.

1.1.21. Kumar S., Gerber S., Richardson T.E. 2006. Testing for non-random contribution of pollen parents using microsatellites and chloroplast markers in polycross families of *Pinus radiata* D. Don. *Tree Genetics & Genomes*, sous presse.

1.1.22. Lepais O., Léger V., Gerber S. 2006 High throughput microsatellite genotyping in oak species. *Silvae Genetica* 55 (4-5): 238-240.

1.1.23. Lowe A., Unsworth C., Gerber S., Davies S. White A., Munro R., Kelleher C., King A., Brewer S., Cottrell J. 2006. Route, speed and mode of oak postglacial colonisation across the British Isles : integrating molecular ecology, palaeoecology and modelling approaches. *Botanical Journal of Scotland, Special Issue*, sous presse.

1.1. Articles soumis

1.1.1. Fernandes L., Rocheta M., Cordeiro J., Pereira S., Gerber S., Oliveira M.M., Ribeiro M.M. Paternity analysis and mating system in a *Pinus pinaster* Aiton clonal seed orchard using microsatellites. (*Tree Genetics & Genomes*)

1.1.2. Di Vecchi M., Laucou V., Lacombe T., Varès D., Gerber S., Boselli M., This P. FaMoz: a software for large scale parentage analysis in *Vitis vinifera* L. species. (*Molecular Ecology*)

1.1.3. Lepais O., Gonzalez M., Cabanettes A., Gerber S. Spatial structure of chloroplast DNA polymorphism of European white oaks in two landscapes: historical and management output (*Forest Ecology and Management*)

1.1. Articles en préparation

1.1.1. Gustafsson S., Gerber S. Mating pattern in the orchid *Gymnadenia conopsea*.

1.2. Actes de colloques

1.2.1. Gonzalez Martinez S.C., Gerber S., Cervera M.T., Martinez-Zapater J.M., Gil L., Alia R. 2002. Detecting reliable parent-offspring matches in parentage analysis: a case study. In: Degen B., Loveless M.D., Kremer A., (Eds.). Modelling and experimental research on genetic processes in tropical and temperate forests. Embrapa Amazonia Oriental, Belem, PA: 159-164.

1.2.2. Gerber S., Latouche-Hallé C., Lourmas M., Morand-Prieur M.E., Oddou-Muratorio S., Schibler L., Bandou E., Caron H., Degen B., Frascaria-Lacoste N., Kremer A., Lefèvre F., Musch B. 2004. Mesure directe des flux de gènes en forêt. Actes du 4^{ème} Colloque National "Le patrimoine génétique : la diversité et la ressource", La Châtre, 14-15-16 octobre 2002 : 349-368.

1.2.3. Gerber S., Mariette S., Streiff R., Bodénès C., Kremer, A. 2002. Using oaks to compare microsatellites and AFLP markers for parentage analysis In: Degen B., Loveless M.D., Kremer A., (Eds.). Modelling and experimental research on genetic processes in tropical and temperate forests. Embrapa Amazonia Oriental, Belem, PA: 69-76.

1.2.4. Di Vecchi M., Laucou V., Lacombe T., Varès D., Gerber S., Boselli M., This P., FaMoz : a software for large scale parentage analysis in *Vitis vinifera* L. species. International Workshop on Advances in Grapevine and Wine Research, Venosa (Italy), September 2005 (sous presse).

1.2.5. Dreyfus P., Pichot C., de Coligny F., Gourlet-Fleury S., Cornu G., Jéssel S., Dessard H., Oddou-Muratorio S., Gerber S., Caron H., Latouche-Hallé C., Lefèvre F., Courbet F., Seynave I. 2005. Couplage de modèles de flux de gènes et de modèles de dynamique forestière. Un dialogue pour la diversité génétique - Actes du 5^{ème} colloque national BRG, Lyon, 3-5 novembre 2004 - Les Actes du BRG n°5, ISBN 2-908447-33-9, 231-250.

1.2.6. Lepais O., Lavabre J., Gonzalez M., Willm J., Cabanettes A., Gerber S. 2006. Diversité et structuration génétique des chênes à l'échelle de deux paysages : impact de l'écologie, de l'histoire et de la gestion. Les Actes du BRG n°6 (6^{ème} Colloque National, La Rochelle, 2-4 Octobre 2006), 543-557.

1.2.7. Pichot C., Bastien C., Courbet F., Demesure-Musch B., Dreyfus P., Fady B., Frascaria-Lacoste N., Gerber S., Lefèvre F., Morand-Prieur M.E., 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 (6^{ème} Colloque National, La Rochelle, 2-4 Octobre 2006), 277-297.

1.3. Communications courtes dans congrès, symposiums (posters)

1.3.1. Bahrman N., Gerber S., Puyo C., Baradat P. 1992. Genetic polymorphism in maritime pine (*Pinus pinaster* Ait.) assessed by two-dimensional gel electrophoresis of needles proteins. "Molecular biology of forest trees", Carcans-Maubuisson, 10-15 Juin.

1.3.2. Bahrman N., Gerber S., Baradat P. 1992. Genetic polymorphism in maritime pine (*Pinus pinaster* Ait.) assessed by two-dimensional gel electrophoresis of needles

and bud proteins. IUFRO meeting "International symposium on population genetics and gene conservation of forest trees", Carcans-Maubuisson, 24-28 Août.

1.3.3. Gerber S., Durel C.E., Costa P., Baradat P. 1992. Geographic variation of European black pines (*Pinus nigra* ARN) for their terpenic composition. IUFRO meeting "International symposium on population genetics and gene conservation of forest trees", Carcans-Maubuisson, 24-28 Août.

1.3.4. Le Provost, G., Gerber, S., Decroocq, S., Ribeiro, M., Pot, D., Mariette, S., Plomion, C. 2000. Use of chloroplast microsatellites for stand and seed certification in maritime pine. Plant & Animal Genome VIII Conference, San Diego (CA), 9-12 Janvier. Abstract 340.

1.3.5. Gerber S., Schibler L., Streiff R., Ducouso A., Louvet J.M., Kremer A. 2000. Pollen and seed flow inferred from parentage analysis in a mixed oak stand comprising *Quercus petraea* and *Q. robur*. OAK 2000 "Improvement of Wood Quality and Genetic Diversity of Oaks" IUFRO International Conference, Zagreb, 20-25 May.

1.3.6. Lepais O., Gonzalez M., Gerber S. 2004. Trois espèces de chênes et le paysage. 27^{ème} réunion du groupe de biologie et génétique des populations, Paris, Museum National d'Histoire Naturelle, 24-27 août.

1.3.7. Di Vecchi M., Lacombe T., Laucou V., Gerber S., Varès D., This P., Boselli M. 2005 FaMoz : a software for a large scale pedigree relationship reconstruction in *Vitis vinifera* L. International Workshop on Advances in Grapevine and Wine Research, Venosa (Italy), September 15-17.

1.3.8. Lepais O., Gerber S. 2006 Hybridisation dynamics in the white oaks species complex (*Quercus robur*, *Q. petraea*, *Q. pubescens* and *Q. pyrenaica*). Population Genetics and Genomics of Forest Trees: from gene function to evolutionary dynamics and conservation, IUFRO conference, 1-4 octobre, Madrid.

1.4. Autres supports

1.4.1. Gerber, S., Mariette, S., Streiff, R., Bodénès, C., Kremer, A., 1999. Comparison of microsatellites and AFLP markers for parentage analysis. Chapter 4 in: Which DNA Marker for Which Purpose? Final Compendium of the Research Project Development, optimisation and validation of molecular tools for assessment of biodiversity in forest trees in the European Union DGXII Biotechnology FW IV Research Programme Molecular Tools for Biodiversity. Gillet, E.M. (ed.). 1999. URL: <http://webdoc.sub.gwdg.de/ebook/y/1999/whichmarker/index.htm>

1.4.2. Gerber S., Mariette S., Streiff R., Bodenes C., Kremer A. 2000. Comparison of microsatellites and AFLP markers for parentage analysis. Mitteilungen der Bundesforschungsanstalt für Forst- und Holzwirtschaft, Hamburg 198: 19-25.

1.4.3. Lexer C., Heinze B., Gerber S., Steinkellner H., Ziegenhagen B., Kremer A., Glossl J. 2000. Microsatellite analysis of anonymous seedlot samples from oak: a promising approach to monitor the number of different seed parents and pollen donors. Mitteilungen der Bundesforschungsanstalt für Forst- und Holzwirtschaft, Hamburg 198: 39-47.

2. Synthèses scientifiques

2.1 Chapitre d'ouvrage

2.1.1. Plomion C., Bahrman N., Costa P., Dubos C., Frigerio J.M., Gion J.M., Lalanne C., Madur D., Pionneau C., **Gerber S.** 2004. Proteomics for genetics and physiological studies in forest trees: application in maritime pine. in *Molecular Genetics and Breeding of Forest Trees*. Kumar S., Fladung M. editors. Food Products Press; Binghamton; USA 53-79.

2.2. Conférences dans congrès ou symposium (orateur souligné)

2.2.1. **Gerber S.** 1991. Marquage du génome du Pin maritime et sélection précoce. In "Meribel 91, marquage moléculaire et sélection" Séminaire du département de génétique et d'amélioration des plantes, INRA : 172-176.

2.2.2. **Gerber S.**, Rodolphe F., Bahrman N., Baradat P. 1992. A linkage map of maritime pine based on seed protein polymorphism. Communication au congrès "Molecular biology of forest trees", Carcans-Maubuisson, 10-15 Juin.

2.2.3. **Gerber S.**, 1993. Variabilité des protéines de l'endosperme du Pin maritime révélée par électrophorèse bidimensionnelle : interprétation génétique, cartographie et relation avec des caractères quantitatifs. Communication au Groupe de travail Marqueurs Moléculaires chez les Végétaux, Méribel, 22-26 Mars.

2.2.4. **Gerber S.**, Streiff R., Bodénès C., Mariette S., Kremer, A. 2000. Using oaks to compare microsatellites and AFLP markers for parentage analysis. Communication au Symposium Sylvolab : "Modelling and experimental research on genetic processes in tropical and temperate Forests", 18-22 Septembre, Kourou, Guyane.

2.2.5. **Gerber S.** 2001. Flux de gènes, recherche de paternité et arbres forestiers. Commission du Génie Biomoléculaire, séminaire "Étude des flux de gènes chez les plantes", 23-24 Janvier, Paris.

2.2.6. **Gerber S.**, Latouche-Hallé C., Lourmas M., Morand-Prieur M.E., Oddou-Muratorio S., Schibler L., Bandou E., Caron H., Degen B., Frascaria-Lacoste N., Kremer A., Lefèvre F., Musch B. 2002. Mesure directe des flux de gènes en forêt. Bureau des Ressources Génétiques, 4^{ème} colloque national "Le patrimoine génétique : la diversité et la ressource", 14-16 octobre, La Châtre.

2.2.7. **Gerber S.**, Austerlitz F., Klein E., Latouche-Hallé C., Oddou-Muratorio S., Bandou E., Caron H., Degen B., Kremer A., Musch B. 2002. Estimating pollen flow in tree species with various life-history traits. International conference DYGEN "Dynamics and conservation of genetic diversity in forest ecosystems" 2-5 décembre, Strasbourg.

2.2.8. Lepais O., Gonzalez M., Lavabre J., Cabanettes A., **Gerber S.** 2006 Impact de l'écologie, de l'histoire et du mode de gestion des populations sur la structure de la diversité en forêts fragmentées. Bureau des Ressources Génétiques, 6^{ème} colloque national, "Des ressources partagées", 2-4 octobre, La Rochelle.

3. Documents à vocation de transfert

3.1. Créations informatiques

3.1.1. Logiciel **FaMoz**, écrit dans le langage de script TclTk avec appel de programmes de calculs en langage C.

<http://www.pierroton.inra.fr/genetics/labo/Software/Famoz/index.html>

3.1.2. Conception de programmes répondant à une question spécifique d'étudiants, thésards ou chercheurs.

3.2 Travaux encadrés ou coordonnés par l'auteur

3.2.1 Mémoires de stages

3.2.1.1. Baurès S., 1998. Mise au point d'un protocole d'électrophorèse de protéines de soja. IUT Génie Biologique, Université Paul Sabatier, Auch.

3.2.1.2. Schibler L., 1999. Analyse de parenté dans une forêt de chênes. Mémoire pour l'obtention du titre d'Ingénieur des Techniques Agricoles. Ecole Nationale d'Ingénieurs des Travaux Agricoles de Bordeaux, 77 pp.

3.2.1.3. Johan Fogelqvist, 2001-2006. Study of mating system and gene flow in *Quercus Robur*, in the northern limit of its distribution range. PhD, Uppsala University, Sweden, en co-encadrement avec Martin Lascoux, Professeur dans cette université. Coopération franco-suédoise (INRA/FORMAS) soutenue financièrement.

3.2.1.4. Hakim Medour, 2001-2002. Génotypages des individus de l'extension d'une parcelle d'étude intensive (ISP) de chênes et conséquence sur l'étude des flux de gènes. Diplôme d'Étude Universitaire, Université Bordeaux I.

3.2.1.5. Catherine Schehr 2002. Étude des flux de gènes dans une chenaie. Mémoire de deuxième année, École Nationale d'Ingénieurs des Travaux Agricoles de Bordeaux.

3.2.1.6. Olivier Lepais 2004. Étude comparée de la structuration génétique des chênes à l'échelle d'un paysage, Master Recherche "Génétique et Développement des Plantes", Université Bordeaux I et II.

3.2.1.7. Jessica Lavabre 2005. Conséquences de la gestion forestière sur la structure génétique d'une chenaie. Master 2 Recherche "Sciences et Technologies", mention "Systèmes écologiques", Université Bordeaux I.

3.2.1.8. Philippe Matter 2005. Estimation de l'héritabilité in-situ des formations épïcormiques chez le chêne sessile (*Quercus petraea*). Stage de césure 2^{ème}/3^{ème} année de cycle ingénieur, École Nationale Supérieure Agronomique de Toulouse.

4. Réponses à appels d'offre

4.1. Coordinatrice du projet "Mesure directe des flux de gènes en forêt : approches théorique et expérimentale" (Projet n°47 de l'Appel à Propositions 1999/2000 du Bureau des Ressources Génétiques, 6 partenaires, budget total 600 000 FF (91 470 euros), 2000-2002).

4.2. Collaboration franco-suédoise (INRA-FORMAS 2002-2003 et 2004-2005, 10 k€ sur deux ans), avec Martin Lascoux, Université d'Uppsala.

4.3. Échange franco-polonais (soutien (voyages) EGIDE 2004-2005), avec Jarosław Burczyk, Université de Bydgoszcz.

4.4. Coordinatrice du projet "Impact de l'écologie, de l'histoire et du mode de gestion des populations sur la structure de la diversité en forêts fragmentées." (Projet de l'Appel à Propositions 2003-2004 du Bureau des Ressources Génétiques, 2 partenaires, budget total 45 400 euros, 2004-2006).

4.5. Coordinatrice du projet "HYPER: Hybridization in plants: ecological impacts and evolutionary outcomes", soumis (octobre 2004) à l'appel "Challenges of Biodiversity Science EUODIVERSITY", European Science Foundation, 18 partenaires européens, budget total 3 M€ (non retenu).

4.6. Coordinatrice d'un projet pour l'appel à proposition du Bureau de Ressources Génétiques 2005-2006 : "A la croisée des chênes : flux de gènes et hybridation dans le complexe d'espèces des chênes blancs, analyses génétiques, écologiques et écophysiological, et projection évolutive". 3 partenaires, budget total 80 k€ (non retenu).

4.7. Coordinatrice d'une lettre d'intention au programme Agriculture et Développement Durable (2006) : "Projet Querman : A la croisée des chênes : interaction entre gestion forestière, histoire du paysage et dynamique dans le complexe d'espèces des chênes blancs". 2 partenaires, budget total 150 k€ (non retenu).

4.8. Co-coordinatrice d'un projet de collaboration franco-suédoise, INRA-Formas, (2006) : "Hybridization in *Quercus* and *Betula*: dynamic processes and evolutionary outcomes", avec Martin Lascoux, Professeur à l'Université d'Uppsala. Budget 80 k€ (non retenu).

III. CURRICULUM VITAE

Née le 19 janvier 1967 à Strasbourg, mariée, deux enfants

FORMATION ET DIPLÔMES

Juin 1984 :	Baccalauréat série C. Lycée Fustel de Coulanges, Strasbourg.
1984/1986 :	Mathématiques supérieures et spéciales biologie. Lycée Jean Rostand, Strasbourg.
1986/1989 :	Institut National Agronomique Paris-Grignon.
Septembre 1989 :	Diplôme d'Ingénieur de l'INA P.G.
Septembre 1989 :	DEA "Ressources génétiques, amélioration des plantes et création de variétés" Paris XI, INA P.G.
Octobre 1992 :	Thèse de Docteur de l'INA P.G.

STAGES ET ACTIVITÉS PROFESSIONNELLES

- Avril-Mai 1988, Station d'Amélioration des Plantes, INRA, Versailles (H Bannerot, J-B Denis) : Comparaison, sur la base d'une étude microscopique du pollen, de différentes stérilités mâles-cytoplasmiques du haricot (*Phaseolus vulgaris*).
- Juillet-Août 1988, Laboratoire de Physico-chimie des Macromolécules, INRA, Nantes (S Perez) : Étude conformationnelle de polysaccharides. Recherche des conformations de faible énergie, et représentation des structures correspondantes grâce au logiciel SYBYL de graphisme moléculaire.
- Mars-Septembre 1989, stage de DEA. Laboratoire d'Amélioration des Arbres Forestiers (INRA), Cestas (P Baradat) :
Chimiotaxinomie et hybridation interraciale chez le Pin noir (*Pinus nigra* Arn.).
- Octobre 1989 - Octobre 1992, Thèse de doctorat. Laboratoire d'Amélioration des Arbres Forestiers (INRA), Cestas (P Baradat, A Kremer), en collaboration avec le Laboratoire de Génétique des Systèmes Végétaux (INRA-CNRS-Paris XI), Gif sur Yvette (A Gallais, D de Vienne), et le Laboratoire de Biométrie (INRA), Jouy en Josas (F Rodolphe) :
Variabilité des protéines de l'endosperme du Pin maritime (*Pinus pinaster* Ait.) révélée par électrophorèse bidimensionnelle : interprétation génétique, cartographie et relation avec des caractères quantitatifs.
- 1992-1994 : Assistante Temporaire d'Enseignement et de Recherche à l'Université Paris XI. Laboratoire d'Écologie, Systématique et Évolution (P-H Gouyon) :
Travaux dirigés et enseignement intégré en génétique des populations et génétique évolutive (niveaux DEUG, licence, maîtrise, DEA).
- 1994-1998 : **Maître de Conférences** en génétique végétale à l'École Nationale Supérieure Agronomique de Toulouse. Laboratoire de Biotechnologie et Amélioration des Plantes, Équipe Physiologie (C Planchon) :
Enseignement - Travaux dirigés de mathématiques appliquées et de biométrie (1^{ère} et 2^{ème} années de formation ingénieur), cours de génétique des populations et génétique évolutive (2^{ème} années), cours "marqueurs et amélioration des plantes" (3^{ème} année).
Recherche - Utilisation de l'électrophorèse capillaire pour l'étude de la génétique de la qualité de la graine du soja (*Glycine max*).
En collaboration avec B Goffinet (INRA biométrie, Toulouse) : méta-analyse de locus à effet quantitatif (QTL).

1998-2000 : **Détachée** dans l'Équipe de Génétique et d'Amélioration des Arbres

Forestiers (INRA) Cestas (A Kremer) :

Étude des flux de gènes par recherche de paternité et de maternité dans les populations de chênes à l'aide de marqueurs microsatellites. Approches théoriques probabilistes, simulations et application à des données expérimentales.

Août 2000 : Recrutée (**CR1 INRA**) au Laboratoire de Génétique et d'Amélioration des Arbres Forestiers, Cestas :

Impact de la sylviculture sur la diversité génétique des arbres forestiers.

PROJETS

Coordinatrice du projet "Mesure directe des flux de gènes en forêt : approches théorique et expérimentale" (Projet n°47 de l'Appel à Propositions 1999/2000 du Bureau des Ressources Génétiques, 6 partenaires, budget total 600 000 FF (91 470 euros), 2000-2002).

Coordinatrice du projet "Impact de l'écologie, de l'histoire et du mode de gestion des populations sur la structure de la diversité en forêts fragmentées." (Projet de l'Appel à Propositions 2003-2004 du Bureau des Ressources Génétiques, 2 partenaires, budget total 45 400 euros, 2004-2006)

Collaboration franco-suédoise (INRA-FORMAS 2002-2003 et 2004-2005), avec Martin Lascoux, Université d'Uppsala

Échange franco-polonais (soutien EGIDE 2004-2005), avec Jarosław Burczyk, Université de Bydgoszcz.

Participations aux projets

- "OAKFLOW: intra & interspecific gene flow in oaks as mechanisms promoting genetic diversity and adaptive potential" (projet européen 2001-2004)

- "Couplage de modèles de Flux de Gènes et de modèles de Dynamique Forestière. Simulation de l'impact de la sylviculture sur la diversité infra-spécifique des espèces forestières exploitées. Mise au point d'indicateurs et de méthodes de gestion." (Appel à Propositions 2001-2002 du Bureau des Ressources Génétiques).

- "Evolutions démographiques et génétiques durant la phase initiale de régénération des peuplements forestiers: approches expérimentales et simulations, proposition d'indicateurs pour la gestion *in situ* des ressources génétiques" (Appel à Propositions 2003-2004 du Bureau des Ressources Génétiques).

- "EvoTREE : evolution of trees as drivers of terrestrial biodiversity" (réseau d'excellence européen, 2006-2010).

DIVERS

Communication scientifique : Rédaction d'articles concernant l'actualité de la recherche pour un journal médical (*Médecine et enfance* 1993-1994).

Représentante des personnels du collège B au sein du «Conseil de l'Enseignement et de la Vie Universitaire» de l'ENSAT (1995-1998).

Directrice du département d'enseignement «Sciences de l'Ingénieur» de l'ENSAT (1997-1998)

Secrétaire de l'association des usagers de la cantine de Pierroton (1998-2002).

Relectrice d'articles soumis à des journaux internationaux à comité de lecture : American Journal of Botany, Annals of Botany, Australian Journal of Botany (2 articles), Canadian Journal of Forest Research, Conservation Genetics (4 articles), Forest Ecology and Management, Forest Genetics, Genetics, Heredity (2 articles), Journal of Ecology (2 articles), Molecular Ecology (19 articles), New Forest, Plant Systematics and Evolution, TREES Structure and Function.

Membre du "review board" de la revue Molecular Ecology (2006-2009).

Évaluation d'un projet de recherche soumis au conseil néerlandais pour la recherche (NWO), 2006.

Thèses :

- Johan Fogelqvist (2001-2006), Université d'Uppsala, Suède, co-encadrement avec Martin Lascoux
- Olivier Lepais (2004-2007), Université Bordeaux I, co-encadrement avec Antoine Kremer

Comités de thèse :

- Gilles Chaix, CIRAD, Montpellier (1999-2002)
- Maya Gonzalez, INRA, Toulouse (2003-2005)
- Manuel Di Vecchi (2003-2005), INRA, Montpellier / Université de Florence, Italie
- Jessica Lavabre (2006-2009), Université de Séville, Espagne

Jurys de thèse :

- Gilles Chaix, Décembre 2002 "Flux de gènes dans un verger à graines d'*Eucalyptus grandis*". Cirad, École Nationale Supérieure Agronomique de Montpellier.
- Cristina Garcia, Novembre 2006 "Studying pollen and seed dispersal patterns assisted by animals in heterogenous populations" Université de Séville, Espagne.

Membre :

- élue du *conseil scientifique* du département INRA EFPA "Écologie des Forêts, Prairies et Milieux Aquatiques" (2002-2010).
- de la *commission scientifique spécialisée INRA* "génétique moléculaire quantitative des populations végétales" (2003-2007, évaluation des chercheurs).
- de *jurys de recrutement* INRA (Ingénieurs d'études, 3 postes (2004), Chargés de Recherche 1^{ère} et 2^{ème} classe, 10 postes (2004), 5 postes (2006)).
- de la commission d'évaluation de l'UMR Biologie et gestion des adventices, Dijon (2006)
- élue (représentante des chercheurs) au *conseil de service* de l'UMR Biogeco.

Responsable, pour l'équipe de génétique, de l'animation scientifique de L'UMR Biogeco et membre du groupe d'animation de l'IFR "Biologie végétale intégrative" (INRA, Université Bordeaux I et II, CNRS).

Organisatrice, avec Pauline Garnier-Géré, de la 28^{ème} réunion du groupe de biologie et génétique des populations "le petit pois déridé", à Bordeaux du 29 août - 1^{er} septembre 2005, 160 participants.

Organisatrice de journée de réunion des doctorants de l'UMR Biogeco, première rencontre : novembre 2006.

