



www.biovigpa.fr

**Master 2<sup>ème</sup> année, cohabilité :**

**U.F.R Sciences, Université d'Angers**

**Université d'Angers, de Bretagne Occidentale,  
de Nantes, de Poitiers, de Rennes 1, de Tours  
et Agrocampus Ouest**

**Centre INRA d'Angers-Nantes**

**RAPPORT de STAGE EFFECTUE**

**par**

**Thibault LEROY**

**au**

**Laboratoire UMR PaVé  
Centre INRA Angers - Nantes**

**en vue de l'obtention du**

**Master 2<sup>ème</sup> Année STS**

**Spécialité : Biologie Végétale Intégrative : Gène, Plante, Agrosystème (BioVIGPA)**

**(Responsables de stage : Bruno Le Cam, Chargé de recherche INRA,  
& Christophe Lemaire, Maître de Conférences, Université d'Angers)**

**Titre du Sujet :**

**HETEROGENEITE INTER-LOCUS DE STRUCTURE GENETIQUE ENTRE POPULATIONS DE  
*VENTURIA INAEQUALIS* PATHOGENES DE DIFFERENTES ESPECES DE *MALUS***

**Soutenance devant le jury d'examen le 15 juin 2009**

## **Remerciements**

Principalement,

à Charles Manceau, en sa qualité de directeur de l'UMR PaVé pour m'avoir permis d'effectuer mon stage au sein de son laboratoire.

à Bruno Le Cam et Christophe Lemaire, merci de m'avoir fait confiance à nouveau pour ce stage, d'avoir été à l'écoute et de bonne humeur dans tous les moments de surcroît ; c'est apaisant quand les expériences ne fonctionnent pas correctement et que l'échéance avance! Merci à vous de m'avoir appris à mieux maîtriser les concepts et moult logiciels en génétique des populations. Enfin, un grand merci à vous pour la relecture et les corrections apportées à ce mémoire.

Particulièrement,

aux autres membres du laboratoire, Valérie Caffier, Véronique Bignon, Martine Devaux, Pascale Expert, Amandine Le Van, Natalia Sapoukhina. Un grand merci pour votre accueil, et pour m'avoir donné des conseils qui se sont révélés bien pratiques par la suite. Je pense aussi à Frank Dunemann et ses collègues du Julius Kühn Institute (JKI) pour les informations précieuses concernant les hôtes du verger.

Egalement,

aux autres personnels et aux stagiaires qui peuplent (ou ont peuplé) le rez-de-jardin du bâtiment et ont absorbé mes moments d'euphorie et de stress. Plus particulièrement, un grand merci à Bérénice, Sophie, Perrine, Sabrina, Noelline et Céline.

Enfin,

je souhaite remercier l'ensemble du personnel de l'UMR PaVé pour l'accueil, l'écoute et l'aide qu'ils ont pu me donner au cours de ses 6 derniers mois.

## Table des abréviations

**2EPF** : Ecologie évolutive des pathosystèmes fongiques

**5-TAMRA** : 5-carboxytetramethylrhodamine

**6-FAM** : 5-carboxyfluorescéine

**6-HEX** : 6-carboxyhexafluorescéine

**AFC** : Analyses Factorielles des Correspondances

**BAC** : souches isolées de *M. baccata*

**COR** : souches isolées de *M. coronaria*

**dNTP** : désoxyribonucléotides tri-phosphates

**EDTA** : acide éthylène diamine tétraacétique

**fAFLP** : fluorescent Amplified Fragment Length Polymorphism  
(polymorphisme de taille de fragments amplifiés détecté par fluorescence)

**FLO** : souches isolées de *Malus floribunda*

**FLO1** : souches isolées de *Malus floribunda* clone 821

**FLO2** : souches isolées de *Malus floribunda* non 821

**GPS** : Global Positioning System  
(système de positionnement global)

**HCl** : acide chlorhydrique

**IBD** : Isolation By Distance  
(Isolement par la distance)

**INRA** : Institut National de Recherche Agronomique

**IOE** : souches isolées de *Malus ioensis*

**JKI** : Julius Kühn Institute

**MgCl<sub>2</sub>** : Chlorure de Magnésium

**NaCl** : Chlorure de Sodium

**ORI** : souches isolées de *Malus. orientalis*

**PaVé** : Pathologie Végétale

**pb** : paire de base

**PCR** : polymerase Chain Reaction  
(réaction de polymérisation cyclique)

**p/v** : poids sur volume

**SDS** : dodécylsulfate de sodium

**SIE** : souches isolées de *Malus sieversii*

**SSR** : Simple Sequence Repeats  
(répétitions de séquences simples en tandem)

**SYL** : souches isolées de *Malus sylvestris*

**UMR** : Unité Mixte de Recherche

**v.** : version

**v/v** : volume sur volume

**XPUR** : souches isolées de *Malus x purpurea*

**XZUM** : souches isolées de *Malus x zumi*

# Sommaire

<b>Partie 1 : Introduction bibliographique</b>	1
<b>Partie 2 : Matériel et Méthodes</b>	5
A. Matériel	5
B. Méthodes	5
1. Analyse sur marqueurs SSR	5
2. Analyse sur fluorescent AFLP ( <i>f</i> AFLP)	6
3. Analyse des données	7
<b>Partie 3 : Résultats</b>	9
A. Détection des loci sous sélection	9
B. Isolement par la distance	9
C. Structuration des populations selon les hôtes	11
1. Assignations des individus à différentes populations	11
2. Variabilité entre populations	11
<b>Partie 4 : Discussion</b>	13
A. Signature de sélection chez <i>V. inaequalis</i>	13
B. Dispersion de <i>V. inaequalis</i>	14
C. <i>SYL</i> , une population localement adaptée	15
D. Spécialisation des souches <i>FLO1</i> et <i>FLO2</i>	15
<b>Partie 5 : Conclusion</b>	17
<b>Partie 6 : Perspectives</b>	19
<b>Partie 7 : Références bibliographiques</b>	20
<b>Partie 8 : Annexes</b>	23

## **Partie 1 : Introduction**

Les populations de microorganismes et particulièrement d'organismes phytopathogènes semblent présenter des capacités exceptionnelles d'adaptation (Slev et Potts, 2002). Il est en effet attendu que la sélection naturelle joue un rôle important dans les mécanismes d'interactions avec l'hôte, de spécialisation et de spéciation. En outre, les bouleversements récents apportés par l'agriculture moderne (application de fongicides, cultivars résistants, homogénéité génétique des hôtes sur de grandes surfaces) imposent une sélection drastique sur les populations de phytopathogènes. Face à la pression sélective, déterministe, générée par l'environnement, d'autres facteurs, stochastiques, contribuent eux aussi à l'évolution des populations de phytopathogènes. Parmi ces facteurs on peut identifier la mutation, le taux de reproduction sexuée, le flux de gènes et la taille efficace de la population (McDonald, 1997).

En génétique des populations, l'estimation des paramètres stochastiques repose sur l'hypothèse de neutralité sélective des marqueurs (Lewontin et Krakauer, 1973). Depuis quelques années, les développements importants de la biologie moléculaire et la connaissance plus précise des génomes permettent l'analyse d'un grand nombre de marqueurs. Dans la grande majorité des organismes couramment étudiés (plantes, animaux) une large proportion du génome est supposée se comporter de manière neutre (*i.e.* en l'absence de contraintes sélectives liées à l'environnement). Cependant, les génomes des organismes phytopathogènes (champignons, bactéries, protozoaires, virus...) seraient plus susceptibles de montrer des déviations significatives aux modèles neutres de variation géographique du polymorphisme. Deux raisons peuvent expliquer ce comportement particulier : *i)* l'interaction avec l'hôte est couramment décrite dans le cadre des relations gène-pour-gène (Flor, 1971). Il est en effet admis qu'à un gène d'avorulence chez le pathogène correspond un gène de résistance chez l'hôte. De plus, l'interaction avec l'hôte est considérée le plus souvent comme multifactorielle. Ceci implique que plusieurs zones du génome seront directement influencées par la sélection naturelle. Subséquemment, *ii)* la multiplicité des gènes sous sélection associée à certains traits d'histoire de vie telles qu'un faible potentiel de dispersion induit un large déséquilibre de liaison sur l'ensemble du génome (Giraud *et al.*, 2006). Il est donc attendu qu'un grand nombre de marqueurs génétiques intrinsèquement neutres présentent une variation génétique corrélée à celle des gènes impliqués dans l'interaction avec l'hôte. Il apparaît donc nécessaire de distinguer clairement les comportements des gènes neutres des gènes sous sélection ou en déséquilibre de liaison avec ceux-ci, dans le cadre d'études impliquant un grand nombre de marqueurs génétiques chez les organismes pathogènes de plantes.

Chez les organismes phytopathogènes, il est attendu que la différenciation génétique soit le plus souvent liée à la spécialisation d'hôte. Fondamentalement, la différenciation génétique peut se produire selon deux modalités : (i) deux individus ne peuvent se trouver sur la même aire au moins lors de la reproduction (liée à la présence de barrières géographiques), correspondant à une différenciation de type allopatrique ou (ii) parce qu'elle a été introduite par un processus de spéciation adaptative. Cette dernière modalité est difficilement réalisée en sympatrie chez la plupart des organismes (Coyne et Orr, 2004) du fait de l'action homogénéisatrice du flux génique. Chez un organisme phytopathogène, la spéciation en sympatrie est corrélée à la spécificité d'hôte.

Chez les champignons pathogènes de plantes, les spores sont dispersées de façon passive par le vent et l'eau, par conséquent, la divergence sympatrique ne peut être expliquée que par le choix du partenaire. Pour les pathogènes ayant un cycle de vie où la dispersion du pathogène est restreinte entre le développement sur l'hôte et le croisement, comme *V. inaequalis*, une étude a montré que le flux de gènes entre population en adaptation sympatrique à deux hôtes peut être réduit à une conséquence directe de la spécialisation (Giraud *et al.*, 2006). En effet, puisque la dispersion des gamètes est limitée, la probabilité de croisement entre deux individus dépend directement de leur capacité à infecter le même hôte, créant une pléiotropie entre spécialisation écologique et choix du partenaire. D'ailleurs plusieurs études ont récemment mis en avant chez des champignons phytopathogènes l'existence d'un processus de spécialisation écologique à leurs hôtes comme chez *Mycosphaerella graminicola* (Stukenbrock *et al.*, 2007), *Claviceps purpurea* (Douhan *et al.* 2008) et *Rhynchosporium secalis* (Zaffarano *et al.*, 2008).

Une différenciation génétique neutre peut cependant être observée chez les organismes phytopathogènes. Le flux de gènes est alors uniquement liée au potentiel de dispersion, à l'existence de barrières (ou de corridors) ainsi qu'à la taille efficace des populations. La présence de barrière géographique constitue un des facteurs déterminants dans la diminution du flux génique entre populations. Au sein d'une même espèce, des populations sont présentes et peuvent être définies comme un ensemble d'individus d'une même espèce capables d'interagir au moment de la reproduction. Les migrations efficaces (c'est à dire le flux de gènes) par leur effet homogénéisateur vont réduire la différenciation entre populations. Cependant, sous certaines conditions de faible taux de dispersion et d'habitat quasi-continu, des patrons de différenciation spatiale peuvent émerger. Ainsi, l'hétérogénéité spatiale des populations peut souvent être corrélée aux distances limitant la possibilité de flux de gènes, on parle d'isolement par la distance (IBD pour 'Isolation By Distance'). Sous l'hypothèse d'IBD, les individus vivant à proximité partagent plus de traits que ceux vivant éloignés (Malécot 1948, Malécot 1955, Malécot 1959). Les analyses de structures des populations naturelles selon des modèles de diffusion continue sont généralement basées sur le modèle dit en

« pas japonais » présenté par Kimura, où les populations peuvent échanger à chaque génération des migrants écartés d'un seul pas selon un axe (Figure 1a) ou deux axes (Figure 1b) (Kimura et Weiss, 1964). Dans le cas d'un modèle en pas japonais, Rousset a montré que l'attendu sous le modèle en pas japonais en une dimension est une diminution de l'apparentement des individus suivant la distance et une diminution selon le logarithme népérien de la distance dans un modèle en pas japonais en deux dimensions.

Comme cela a été précisé plus haut, l'ensemble du génome n'est pas contraint aux mêmes forces évolutives. Les régions neutres ou sous sélections ne répondent aux mêmes contraintes. Les loci neutres sont soumis aux effets conjoints de la mutation, la migration et la dérive aléatoire. Dans le cadre d'un modèle spatial en pas japonais, il est alors aisé de tester la relation entre apparentement et distance géographique que doivent théoriquement suivre les loci neutres. Pour la détection des loci sous sélection, l'idée mise en place par Lewontin et Krakauer (1973) basée sur l'estimation de la variance de  $F_{st}$  entre loci est comparée aux attendus selon le modèle neutre a été très utilisée. Elle permet la détermination des loci soumis à une sélection soit positive (qualifiée d'adaptative ou de directionnelle) où la différenciation génétique est supérieure aux loci neutres, soit équilibrante visualisable par une plus petite différenciation génétique. Les loci qui diffèrent de l'attendu sous le modèle neutraliste sont qualifiées de « outlier ». Cette notion de  $F_{st}$  outlier est une théorie utilisée pour la recherche sur l'ensemble du génome de régions soumises à la sélection naturelle (*ie* la recherche de candidats liée à l'adaptation) (Lewontin et Krakauer, 1973). Désormais, de nouveaux tests de détection de loci sous sélections sont basés sur des algorithmes bayésiens (Beaumont et Balding, 2004 ; Foll et Gaggiotti, 2008).

Le champignon *Venturia inaequalis*, agent de la tavelure, est le pathogène du pommier (*Malus x domestica*) le plus préjudiciable économiquement. Présent également sur de nombreuses espèces sauvages du genre *Malus*, cet ascomycète haploïde a été capable de coloniser l'ensemble des zones cultivées du pommier. Il constitue un bon objet d'étude pour traiter des questions de différenciation génétique liée à l'hôte ou à la distance car ses traits d'histoire de vie sont précisément ceux utilisés dans le modèle de Giraud *et al.* (2006) lequel prédit une spécialisation rapide entraînant un déséquilibre de liaison sur l'ensemble du génome. Le verger, contrairement aux cultures annuelles, permet l'établissement d'une structure génétique à l'équilibre. La géométrie des vergers en rangs et colonnes équidistants correspond en outre à celle d'un modèle spatial en pas japonais à deux dimensions. Il existe enfin un grand nombre de marqueurs moléculaires polymorphes disponibles (Simple Sequence Repeats (SSR), et Amplified Fragment Length Polymorphism (AFLP)).

Bien que les spores de ce pathogène ne soit pas connues pour se disséminer sur de longues distances, l'interaction qu'il forme avec le pommier domestique lui a permis de franchir tous les continents. En effet, le pommier étant une plante d'importance économique majeure, le pathogène a ainsi pu se déplacer via l'expansion des zones de cultures de son hôte, lui-même étroitement guidé par l'homme (Gladieux *et al.*, 2008). L'extension de l'aire de répartition de la maladie s'est faite en parallèle de celle de son hôte le pommier. Elle s'est réalisée en plusieurs étapes, les marchands de la route de la Soie ont d'abord domestiqué le pommier puis l'ont introduit en Europe. Après des siècles de domestication et d'amélioration les pommiers furent introduits dans les nouveaux continents à partir principalement du continent européen dès le XVIème siècle. *V. inaequalis* a ensuite suivi les migrations de son hôte et ainsi atteint la phase ultime de son expansion. (Gladieux *et al.*, 2008).

Bien que *Venturia inaequalis* soit surtout étudié pour son interaction avec le pommier, il attaque de nombreux autres *Malus* dont certains présentent des gènes de résistances majeurs (*Vf* chez *M. floribunda*, *Vb* chez *M. baccata*). Le genre *Malus* compterait 41 espèces et de nombreux hybrides (Juniper et Mabblerley, 2006).

De nombreuses études rapportent que le centre d'origine du pommier se situe en Asie centrale (Juniper et Mabblerley, 2006). Le géniteur principal du pommier domestique serait *M. sieversii* espèce endémique d'Asie Centrale, le pommier sauvage européen *M. sylvestris* aurait également participé d'une façon significative au pool génique du pommier cultivé (Harris *et al.*, 2002 ; Robinson *et al.*, 2001, Coart *et al.*, 2006).

Pour analyser la différenciation de *V. inaequalis* liée aux différents *Malus*, il est nécessaire de s'affranchir de biais de structuration liés à des d'effets spatiaux. Le verger expérimental échantillonné situé à Dresde en Allemagne présente de nombreux intérêts pour notre étude puisque (i) de nombreuses espèces différentes de *Malus* y sont présentes dont certains portent des gènes de résistance majeure, *M. floribunda* (*Vf*) ou *M. baccata* (*Vb*), (ii) aucune méthode de lutte dirigée vers cette maladie n'y est développée, (iii) l'Europe est un continent où la diversité génétique de *V. inaequalis* est très importante (Gladieux *et al.*, 2008).

Dans cette étude nous testerons l'hypothèse d'hétérogénéité de réponse des loci dans les patrons observés de structure génétique spatiale. L'étude nous permettra de répondre à différentes questions : *i*) parmi les marqueurs utilisés est-il possible de détecter des loci dont la variation génétique spatiale se démarque des attendus neutres ? *ii*) existe-t-il une structure génétique spatiale ? *iii*) si oui, cette structure est elle homogène sur l'ensemble des loci ? *iv*) sinon, certains loci reflètent-ils un IBD ? *v*) si certains loci ne montrent pas d'IBD, la structure génétique observée est-elle liée à la nature des hôtes ?

## **Partie 2 : Matériel & Méthodes**

### **A| Matériel biologique**

Les souches de *V. inaequalis* proviennent d'un verger non traité chimiquement situé à Dresde, en Allemagne composé de plusieurs espèces de *Malus*. Les souches ont été collectées en août 2008 sauf les souches 06M06 et 06M07 collectées en 2006. La localisation et l'espèce d'hôte de chaque arbre du verger sont données dans la figure 2. La nomenclature de chaque souche, le nom d'espèce de chaque *Malus* échantillonné et la détermination des coordonnées spatiales de chaque arbre (lieu d'échantillonnage des 337 souches) sont donnés (annexe I).

### **B| Méthodes**

Deux méthodes d'extraction d'ADN ont été utilisées (annexe II). L'ADN extrait à partir de culture mycélienne de souches monosporées permet le génotypage par SSR et *f*AFLP, l'autre méthode consistant à extraire l'ADN à partir de lésions de feuilles infectées, ne permet que le génotypage par SSR en raison de la présence d'ADN d'origine végétale.

#### **1) Analyse sur marqueurs SSR**

##### *a) Amplifications*

Les échantillons ont été génotypés sur 11 loci microsatellites : *Itc1a*, *Itc1b*, *Itc1g*, *Iaac3b* (Tenzer *et al.*, 1999), *Vitcca7/P*, *Vitg11/70*, *Vicacg8/42*, *Viga7/116*, *Vica 9/X* (Guérin et Le CaM. 2004) et *M42* (Guérin *et al.*, 2007). Les amorces utilisées pour la polymérase chain reaction (PCR) sont présentées dans le tableau 1. La PCR a été réalisée dans un volume de 10µL avec 10 à 60 ng d'ADN de l'échantillon comme matrice et utilisant 2µL de tampon 5x, 0.5µM d'amorces marquées avec un fluorochrome (5-carboxytetramethylrhodamine (5-TAMRA), 6-carboxyhexafluorescéine (6-HEX) ou 6-carboxyfluorescéine (6-FAM)) sur l'amorce sens en 5', 500 µM de désoxyribonucléotides tri-phosphates (dNTP), 1.5mM de chlorure de magnésium (MgCl<sub>2</sub>), 1 unité de GoTaq (Promega, Charbonnières-les Bains, France). Après une première étape de dénaturation à 95°C (5 minutes), 40 cycles composés d'une dénaturation à 95°C (45 secondes), d'une hybridation de 30 secondes) à la température de l'amorce (tableau 1), et d'une élongation à 72°C (45 secondes). L'élongation finale est réalisée à 72°C pendant 5 minutes.

##### *b) Génotypage*

Une dilution 1:60 est réalisée pour les ADNs extraits de culture mycélienne et 1:10 pour les ADNs extraits à partir de feuilles présentant des lésions. Les ADN dilués sont alors associés à un

mélange de formamide (Sigma-Aldrich, Saint Louis, Missouri, USA), et un marqueur moléculaire de taille GS400HD (Applied Biosystems, Foster City, California, USA) selon un ratio 200 :800 :1 respectivement. Chaque séquence sens étant marquée par un fluorochrome, le génotypage est effectué sur un séquenceur ABI 3130 et les données analysées manuellement sur le programme Genemapper 4.0 (Applied Biosystems).

## 2) Analyse sur fluorescent AFLP (fAFLP)

### *a) digestion – ligation des adaptateurs*

Toutes les souches monosporées ont été analysées en fAFLP. La méthode employée est celle de Guérin *et al.*, (2004) basée sur la mise au point de deux solutions. La première contient 3.6 unités de *Pst*I et de *Msp*I (Eurogentec SA, Herstal, Belgique), 2 µL de NaCl 0,5M, 0,2µL d'albumine de sérum bovin (New England Biolabs, Beverly, Massachusetts, USA), 1,5 µL de tampon 10X de T4 ligase (New England Biolabs), 1µL à 2µM de l'adaptateur *Pst*I et 1µL à 20µM de l'adaptateur *Msp*I dans un volume final de 10µM. Les amorces utilisées sont présentées dans le tableau 2. La seconde solution contient 0,5µL de tampon de la ligase T4 10X et 0,2µL de ligase T4. Cent cinquante nanogrammes d'ADN génomique sont mélangés à 10 µL de la solution 1 et 5 µL de la solution 2 puis incubés 3 heures à 37°C.

### *b) amplification préselective*

Une première amplification est réalisée dans 20 µL contenant 4µL de produits ADN dilués (1/10) après l'étape de restriction-ligation, 2 µL de tampon Goldstar 10X (Eurogentec), 5mM MgCl<sub>2</sub>, 2 mM dNTPs (Eurogentec), 0,25 µM de chaque amorce « core » *Pst*I et *Msp*I (tableau 2) et 0,5 unités de Taq polymérase (Goldstar Red, Eurogentec). Les échantillons sont alors incubés 2 minutes à 72°C suivi de 20 cycles composés de 20 secondes de dénaturation à 94°C, 30 secondes d'hybridation à 56°C et 2 minutes d'élongation à 72°C.

### *c) amplification sélective*

Après dilution au dixième des produits de l'amplification préselective, la seconde amplification est réalisée dans 20µL contenant 4µL de produits dilués, 2µL de tampon Goldstar 10X (Eurogentec), 5mM MgCl<sub>2</sub>, 2mM dNTPs (Eurogentec), 50nM de l'amorce *Pst*I +A marquée au fluorochrome 6-FAM (tableau 2), 250nM de l'amorce *Msp*I +AA et 0,5 unités de Taq polymérase (Goldstar Red, Eurogentec). Après 2 minutes de dénaturation initiale, l'amplification est effectuée par 33 cycles composés d'une étape de dénaturation à 94°C pendant 20 secondes, suivie d'une hybridation de 30 secondes et d'une élongation à 72°C pendant 2 minutes, l'élongation finale est réalisée à 60°C pendant 30 minutes. L'hybridation est réalisée à 66°C pour la première étape et

ensuite décroît de 1°C par cycle pour les 9 cycles suivants, enfin pour les 23 derniers cycles les hybridations ont lieu à 56°C.

#### *d) Génotypage des loci fAFLP*

2,4µL de produit amplifié est alors ajouté à un mélange de formamide (Sigma-Aldrich), et du marqueur moléculaire GS500 (Applied Biosystems,) en ratio 30 :1 respectivement. Le génotypage est effectué sur un séquenceur ABI 3130 et les données analysées sur le programme Genemapper 4.0 (Applied Biosystems). Seules les bandes fAFLP comprises entre 70 et 395 pb ont été annotées. Les bandes présentes dans certains isolats et absentes dans les autres sont considérées polymorphes.

### 3) Analyse de données

La collecte des échantillons étant réalisée pendant le stade asexué du cycle de vie de *V. inaequalis*, les individus ayant les mêmes allèles à tous les individus ont été écartés (jeu de données clone-corrigé).

#### *a) Sélection des marqueurs non neutres*

Puisque les marqueurs AFLP sont issus d'une digestion non ciblée de l'ensemble du génome, aucune connaissance n'existe sur l'histoire évolutive des marqueurs générés. La recherche d'outliers a été effectuée sur BayeScan (Foll et Gaggiotti, 2008), cet algorithme bayésien détermine une probabilité postérieure que des loci soient sous sélection. La probabilité d'un loci à être sous sélection naturelle dépend du facteur bayésien déterminé par l'algorithme. Les analyses de détection de loci sous sélection ont été réalisées avec un intervalle de confiance de 99%. Les loci outliers s'ils existent ont été retirés dans les analyses subséquentes.

#### *b) Analyses des variances*

Les Analyses Factorielles des Correspondances (AFC) ont été réalisées sous Genetix version (v.) 4.05 (Belkhir *et al.*, 1996-2004). Guinand (1996) a montré que dans ce cas les valeurs d'inertie le long de chaque axe pouvaient être assimilées à des combinaisons linéaires de valeurs de Fst (Guinand, 1996). Dans les analyses toutes les souches provenant d'hôtes différents ont été considérées comme appartenant à des populations différentes. Pour plus de lisibilité, les populations sont représentées par le barycentre des individus appartenant à chaque population.

#### *c) Assignations des individus*

Les assignations ont été effectuées sous Structure v. 2.2 (Prichard *et al.*, 2000; Falush *et al.*, 2003, Falush *et al.*, 2007) et BAPS v. 4.14 (Corander *et al.*, 2004 ; Corander et Marttinen, 2006)

afin de déterminer l'appartenance relative des individus en groupes (k). Pour Structure, les valeurs de vraisemblance de chaque nombre de groupe (k) compte tenu des données ont été transformées selon la méthode proposée par Evanno (Evanno *et al.*, 2003). Cette méthode est basée sur le taux de changement de la vraisemblance entre k successifs et ce, en tenant compte des écart-types entre répétitions. Chaque analyse a été réalisée 5 fois avec une Monte Carlo Chaîne de Markov de 500 000 étapes après une période de préchauffage de 50 000 étapes.

#### d) Structure génétique spatiale

Le test pour l'IBD (isolement par la distance) a été effectué à l'aide du logiciel SPAGeDI v1.2 (Hardy et Vekemans, 2002). Ce programme associe les données générées par les statistiques par paires d'individus à la distance euclidienne les séparant. Le coefficient d'apparentement  $F_{ij}$  de Loiselle *et al.* (1995) a été calculé entre chaque pair d'individu. Pour tester la robustesse de l'analyse, les localisations et les individus ont été permutés 20000 fois.

Dans le cas de l'existence d'un IBD, l'apparentement entre paires d'individus décroît de manière linéaire avec la distance spatiale dans le cas d'un habitat en une dimension, et du logarithme népérien de la distance dans le cas d'un espace en deux dimensions (Rousset, 1997, Hardy et Vekemans, 1999). La valeur  $F_d$  représente la moyenne des apparentements entre paires d'individus  $F_{ij}$  par classe de distance. En effet, pour étudier la diminution de l'apparentement par rapport à la distance, des classes de distances contenant le même nombre d'individus sont définis. Or, Rousset (2000) a montré que dans le cas d'un modèle en pas japonais en deux dimensions les attendus sont:

$$F_{(d)} = \ln(R) * \hat{b} + \text{constante} \quad \text{où } \hat{b} = 1/(4D\pi\sigma^2)$$

D représente la taille efficace et  $\sigma^2$  la variance de la distance parent-enfant. Par conséquent, la pente de la régression ( $\hat{b}$ ) est fonction de la moyenne des apparentements des individus par classe de distance par rapport au logarithme népérien de la distance. La statistique  $Sp$  (Hardy et Vekemans, 2004) quant à elle représente le taux de décroissance de l'apparentement par paire avec la distance. La valeur de  $Sp$  est donnée par la relation :

$$Sp = -\hat{b} / (1 - F_{(d=1)})$$

Où  $F_{(d=1)}$  représente la moyenne des apparentements des individus pour la première classe de distance.

## **Partie 3 : Résultats**

Afin de bien caractériser l'impact de chaque hôte sur la diversité des souches, les individus qui sont issus du même hôte seront regroupés sous le terme de « population ».

### **A] Détection des loci sous sélection**

Les 337 échantillons disponibles ont été analysés sur les 11 loci SSR et les 127 souches monosporées sur 81 loci AFLP. Pour les microsatellites, seul les 248 individus dont le typage a été possible sur au moins 7 loci ont été conservés, 183 haplotypes différents ont alors été obtenus soit une fraction clonale de 0,27.

Pour les analyses *f*AFLP, 127 des 139 souches ont pu être correctement annotées sur l'ensemble des 81 loci générés aléatoirement dans le génome. 4 clones identifiés sur les 92 loci (81 AFLP + 11 SSR) ont été éliminés du jeu de données global (jeu de données clone-corrigé) soit une fraction clonale de 0,03. Cette fraction est faible, plusieurs individus différents ont pu être identifiés sur la même feuille.

Pour la détection des loci sous sélection, l'analyse a été effectuée sur l'ensemble des 92 loci. Chaque locus est représenté par un point dans le nuage (figure 3). Au seuil 1%, une détection d'outliers ne révèle que 1 seul locus AFLP sous sélection diversifiante, il s'agit du locus L5 tandis qu'aucun locus ne semble être sous sélection équilibrante. Le locus obtenu par cette analyse est un candidat pour tester l'effet d'une adaptation à l'hôte. Certaines populations présentent cet allèle pour l'ensemble de leurs individus (annexe I), c'est le cas des individus issus de *M. sieversii*, *a contrario* certaines ne possèdent jamais cet allèle comme *M. coronaria* et *M. x purpurea* ; des situations nuancées sont observées avec *M. sylvestris*, *M. baccata*, *M. ioensis*, *M. orientalis* et *M. x zumi*, où certains individus possèdent cet allèle et d'autres non. En conséquence, le locus L5 ne discrimine pas totalement deux groupes de populations selon la présence/absence à ce locus.

### **B] Isolement par la distance**

La caractérisation du flux génique neutre a été réalisée sur les 91 marqueurs non détectés sous sélection, c'est-à-dire, l'ensemble des loci excepté le locus sous sélection adaptative L5. Il est bien établi que les loci neutres sont les meilleurs marqueurs d'une proximité géographique entre individus (Lemaire *et al.*, 2000). La première analyse a montré qu'une régression linéaire entre le coefficient

d'apparement de Loiselle (Loiselle *et al.*, 1995) et le logarithme de la distance entre individus est possible. Sur l'ensemble des loci, la régression prend des valeurs fortes ( $\hat{b} = -0.028$ ,  $p < 0,001$ ). Les régressions linéaires entre le coefficient d'apparement de Loiselle et la distance ont été ensuite réalisées sur quatre classes de distances ayant les mêmes effectifs (0-31,5 ; 31,5-55,9 ; 55,9-88,5 ; 88,5-147). Un autocorrélogramme (trait plein, figure 4) a pu être réalisé basé sur l'obtention des valeurs du  $F_{(d)}$ , c'est-à-dire la moyenne de l'ensemble des valeurs d'apparements par paires d'individus ( $F_{ij}$ ) pour chaque classe de distance. Les valeurs de  $F_{(d)}$  diminuent lorsque le logarithme népérien de la distance augmente (figure 4). La valeur  $Sp$ , représentant le taux de décroissance de l'apparement avec la distance (Vekemans et Hardy, 2004) est de 0,028.

Les valeurs du  $F_{(d)}$  sont positives jusqu'au logarithme népérien de la distance égal à 3.78. Par conséquent, aucun apparement entre individu n'est détectable au delà d'une distance de 43,8 mètres. Même si cette donnée est imparfaite puisque dépendant fortement des classes définies (Vekemans et Hardy, 2004), il apparaît qu'à chaque génération, la propagation des descendants est limitée à quelques dizaines de mètres au sein de ce verger puisque au-delà aucun apparement entre individus ne peut être expliqué sur l'ensemble des souches échantillonnées.

Une analyse basée sur la détection des loci expliquant le profil d'IBD a ensuite été initiée. Soixante deux loci sur quatre-vingt onze montrent un coefficient directeur  $\hat{b}$  de régression négatif et montre ainsi une décroissance de l'apparement suivant la distance. Basée sur une transformation mathématique de la probabilité d'existence de l'IBD *via* une fonction logit, 24 loci (en rouge, figure 5) ne peuvent expliquer l'existence d'un IBD par le hasard à la probabilité de 5% soit  $\text{logit}(2 * |0.05 - 0.5|) = 2,19$  où  $\text{logit}(x) = \ln(x/(x-1))$  (Beaumont et Balding, 2004). Ces 24 loci présentent quasi exclusivement les coefficients directeurs les plus forts ( $\hat{b} = -0,065$   $p = 0$ ) expliquant la grande influence de ces loci sur l'analyse globale présentée précédemment, basée sur l'ensemble des loci excepté L5. L'autocorrélogramme basée sur le  $F_{(d)}$  pour les 24 loci répondant significativement à l'IBD (trait pointillé, figure 4) décroît plus rapidement que sur l'ensemble des loci (trait plein) avec un taux de décroissance suivant la distance ( $Sp$ ) de 0,07. De plus, une corrélation entre l'apparement par paire d'individus et la distance peut être perçue sur une distance plus importante sur ces loci sous IBD, un apparement est trouvé sur le deuxième intervalle alors que pour l'analyse sur les 92 loci l'apparement est nul. Ce résultat semble indiquer que les loci ne répondent pas de la même façon au flux de gènes, ce qui sous-entend qu'une force agit contre l'isolement par la distance. Or, par ailleurs, il est bien établi que la sélection naturelle agit contre le flux de gènes (Lenormand, 2002) d'où la mise en place d'une étude de structuration sur ces mêmes populations.

## C] Structuration des populations selon les hôtes

Les études de structuration des populations ont été effectuées sur l'ensemble des loci ne présentant pas statistiquement un profil d'IBD soit les 67 loci. En effet, dans le but de maximiser les possibilités d'observer une structure, les informations portées par les loci sous IBD ont été ignorées. L'objectif étant ainsi de déterminer si les populations au sein du verger sont structurées par un autre effet que la distance. Les populations définies par l'espèce de l'hôte d'origine ont été analysées. Seules les populations présentant au minimum 4 souches par type d'hôte ont été conservées soit 117 souches de 9 populations dont 7 espèces différentes : *M. sieversii* (SIE), *M. sylvestris* (SYL), *M. baccata* (BAC), *M. ioensis* (IOE), *M. coronaria* (COR), *M. orientalis* (ORI), *M. floribunda* (FLO) et 2 hybrides : *M. x zumi* (XZUM), *M. x purpurea* (XPUR).

### 1) Assignations des individus à différentes populations

Deux analyses d'assignation des individus ont été menées en parallèle. Le logiciel Structure utilise une méthode sans *a priori* sur les différentes populations présentes. Via la transformation mathématique d'Evanno (Evanno *et al.*, 2005), le meilleur nombre de groupes expliquant le jeu de données est de 4 (figure 1, annexe III). L'analyse présentée (figure 6) représente l'appartenance relative de chaque individu (en abscisse, figure 6) aux 4 groupes trouvés. Deux populations semblent bien expliquer l'appartenance à 3 groupes distincts. Les souches isolées (SYL) appartiennent majoritairement à un groupe. Au sein des souches isolées de *M. floribunda*, deux groupes semblent fortement différenciés. Les individus issus de *M. floribunda clone* 821 forment un groupe distinct des autres. Les autres populations issues de prélèvement sur d'autres types d'hôtes ne semblent pas expliquer d'appartenance majoritaire à un des 3 autres groupes et présentent une structure majoritairement expliquée dans un quatrième groupe.

Sous BAPS 4.14, un autre logiciel d'assignation, le même nombre de groupe est observé avec deux populations différenciées chez *M. floribunda* (figure 2, annexe III). *M. sylvestris* forme aussi une population différenciée des autres populations de souches isolées des autres *Malus* du verger.

Pour la suite de l'analyse, on définira deux populations (FLO1 pour les *M. floribunda clone* 821 et FLO2 pour les autres *M. floribunda*).

### 2) Variabilité entre populations

Afin de mieux caractériser les variations intrapopulations, des analyses factorielles des Correspondances (AFC) (Benzecri, 1973) ont été effectuées basée sur l'ensemble des 67 loci neutres

qui ne répondent pas significativement à l'IBD (en vert, figure 5). L'AFC présentée (figure 7) projette les barycentres des populations précédemment définies en fonction de deux axes. Sur le premier axe, les souches isolées de *FLO1* et de *SYL* sont isolées de l'ensemble des autres souches du verger selon un axe expliquant 23.2% de la variance. Le premier axe d'une AFC représente un bon estimateur du  $F_{st}$  global entre les populations (Guinand, 1996), ce qui indique que ces deux populations sont bien différenciées de l'ensemble des autres souches. Le second axe (17,6% de l'inertie) sépare les autres souches du verger sauf les souches *XZUM* et *XPUR* qui ne semblent pas différenciées. De façon intéressante, *FLO2* semble sur cet axe être assez différencié des autres souches du verger. L'analyse AFC semble ainsi corroborer l'hypothèse précédemment vérifiée par les sorties de logiciels d'assignation selon laquelle 4 populations fortement différenciées existent au sein du verger : *FLO1*, *FLO2*, *SYL* et enfin les individus issus des autres espèces.

## **Partie 4 : Discussion**

L'équipe Ecologie Evolutive des Pathosystèmes Fongiques (2EPF) de l'Unité mixte de recherche (UMR) Pathologie Végétale (PaVé) du Centre INRA d'Angers-Nantes s'intéresse à l'étude des meilleures stratégies de mise en place de vergers de pommiers à résistance durable contre la tavelure. Cette étude de génétique des populations a pour objectif de mieux comprendre la spécialisation de *Venturia inaequalis* aux différentes espèces de *Malus* sources de résistance dans les programmes de sélection.

En génétique des populations la détermination de deux dynamiques semblent fondamentales, l'évolution neutre du génome basée sur l'existence des processus de mutation, de migration et de dérive où le devenir des allèles est stochastique et où les allèles se fixent sur des temps longs ; et, *a contrario*, le processus sélectif darwinien, un processus adaptatif où les allèles se fixent sur des temps bien plus courts et permettent une adaptation rapide à l'environnement.

Afin de répondre à la présence éventuelle d'une spécialisation, question complexe en génétique des populations, nous avons développé plusieurs approches pour (i) rechercher des signatures de sélection liée aux différents *Malus*, (ii) évaluer les flux géniques neutres et (iii) rechercher une structure des populations en s'appuyant sur les loci neutres ne répondant pas significativement à l'IBD

### **A) Signature de sélection chez *V. inaequalis***

La première étape de l'étude s'est basée sur la détection de loci sous sélection, un locus sous sélection positive diversifiante a pu être mis en avant et ce sur un nombre limité d'une centaine de loci. La variation au locus L5 semble être la signature de l'existence d'une région du génome de *V. inaequalis* sous sélection positive et donc en adaptation à l'environnement. La nature biallélique du marqueur (AFLP) et le faible nombre de réplicats de chaque espèce ne nous permet cependant pas de conclure sur la structure génétique à ce locus. Il est donc pour l'instant impossible de déterminer si la variation chez L5 est véritablement liée à un type d'hôte. Toutefois, il est possible de poser quelques hypothèses quant à la nature de L5. En effet, chez les pathogènes, il est attendu que les protéines effectrices soient très souvent sous sélection diversifiante. En effet, le polymorphisme observé à ce type de protéine est censé évoluer en réponse aux changements d'hôte et donc refléter les interactions avec ce dernier. Il a été montré récemment que le champignon *Rhynchosporium* est un complexe d'espèces cryptiques qui est sous divergence écologique liée à différents hôtes (en l'occurrence le blé, le riz et le chiendent) et que cette spécialisation est extrêmement récente

(Zaffarano *et al.*, 2008). L5 peut ainsi se situer dans une région proche d'un effecteur et montrer une variation génétique sous sélection diversifiante *via* la liaison génétique existant entre L5 et un effecteur fongique. La découverte d'une signature de sélection chez *V. inaequalis* est une piste intéressante quant à la détection d'une spécialisation aux différents *Malus*. Le pathogène et son hôte sont dans une perpétuelle course à l'armement où le pathogène cherche à être virulent et l'hôte essaye de résister à l'évolution des effecteurs afin de contrecarrer le pathogène.

### **B) Dispersion de *V. inaequalis***

L'existence d'un IBD peut parfois être difficile à détecter pour certains pathogènes, de part la nécessité d'avoir une échelle géographique adaptée (Prospero *et al.*, 2008). Chez *Melampsora larici-populina*, agent de la rouille du peuplier, un IBD a été démontré indiquant des migrations régulières dans l'espace (Barrès *et al.*, 2008). Bien que l'échelle de leur étude soit tout autre (IBD sur toute l'Europe), nous avons montré, au sein d'un petit verger où les arbres les plus distants sont séparés de 150 mètres, la présence d'isolement par la distance. Pour la première fois chez *V. inaequalis*, ce type d'étude d'IBD est menée et donnent des informations importantes sur la dissémination d'un pathogène dans son environnement. Les paramètres de dispersion estimés à partir de l'autocorrélogramme semblent indiquer que *V. inaequalis* ne disperserait qu'à quelques dizaines de mètres dans le verger de Dresde. Ce résultat peut sembler faible comparé à ceux obtenus par simulations par Aylor (1998) et montrant que les ascospores de *V. inaequalis* sont capables de se disperser jusqu'à 5 kilomètres. Cette distance reflétait déjà pourtant une faible dispersion de ce pathogène par rapport à de nombreux autres organismes phytopathogènes, capables de se disperser sur des centaines à des milliers de kilomètres (Brown et Hovmoller, 2002). Cependant, la comparaison ne considère pas les mêmes processus. Les simulations d'Aylor (1998) reflètent une dispersion brute et n'implique pas une distance de diffusion des gènes. C'est justement ce dernier paramètre qui est estimé dans l'autocorrélogramme. La pente de l'IBD est en effet inversement proportionnelle à la variance de la distance parent-enfant. Ainsi, la capacité réelle de dispersion de *V. inaequalis*, c'est-à-dire le flux de gènes possible au sein du verger échantillonné est de quelques dizaines de mètres, ce qui tend à montrer que les événements de dispersion à moyenne distance (quelques kilomètres) sont des processus rares au sein des vergers.

Etonnamment, *V. inaequalis* est un pathogène à faible capacité intrinsèque de dispersion, mais il a malgré tout atteint le dernier stade de son processus invasif puisqu'il est présent dans toutes les zones de culture de *M. x domestica* (Gladieux *et al.*, 2008) Ces mêmes auteurs concluent leur

étude en estimant que ce sont les activités anthropiques via le commerce des plants contaminés qui ont permis à cette maladie d'être présente sur tous les continents.

### **C) SYL, une population localement adaptée**

Une forte différenciation des souches issues de *M. sylvestris* existe au sein du verger. Récemment, Gladieux (2007) a démontré que l'attaque des *M. sylvestris* était un processus lié à l'introduction fortuite et répétée en Europe de souches lors de la domestication du pommier à partir de l'Asie (Gladieux, 2007). *V. inaequalis* a pu contaminer et s'installer sur les populations endémiques allemandes de *M. sylvestris* depuis donc bien plus longtemps que sur les autres espèces de *Malus* non endémiques présentes dans le verger. Les populations SYL que nous avons échantillonnées révéleraient la diversité génétique des populations de *V. inaequalis* infectant les *M. sylvestris* présents dans les forêts environnantes. A l'inverse les populations sur les autres hôtes non endémiques du verger proviendraient de contaminations plus récentes dues à des souches présentes sur les pommiers domestiques environnants. En effet, il apparaît peu probable que ces souches soient représentatives de la variabilité de *V. inaequalis* des populations naturelles des différentes espèces de *Malus*. L'observation d'un IBD montre que même si les souches provenaient des sites naturels des différentes espèces de *Malus*, l'adaptation à leur hôte respectif n'est pas suffisamment forte pour empêcher significativement tout flux de gènes entre elles, comme cela semble être le cas chez SYL. Toutefois, nous pouvons supposer que les populations présentes sur les hôtes non endémiques sont elles-mêmes en cours de spécialisation car un nombre relativement important de loci ne répond pas à l'IBD, or la spécialisation induit un important déséquilibre de liaison (Giraud *et al.*, 2006).

### **D) spécialisation des souches FLO1 et FLO2**

Les souches attaquant *M. floribunda* (FLO) sont divisées en deux populations : les FLO1 isolées de *M. floribunda* clone 821 et les FLO2 isolées d'autres *M. floribunda* du verger. Or, Gladieux (2007) et Guérin *et al.* (2007) ont montré que des souches issues de cultivars de pommier, peuvent être isolées reproductivement, en raison de spécificités parasitaires alors que des souches issues de différents genres de *Maloïdeae* peuvent présenter des spectres de virulences chevauchants. Ils montrent également que les populations virulentes du gène de résistance *Vf* introgressé dans le pommier domestique *M. x domestica* à partir du clone *M. floribunda* 821 sont fortement différenciées des populations avirulentes.

Dans cette étude, nous avons recherché à caractériser si la capacité à infecter les différents *Malus* a pu entraîner une spécialisation des populations chez *V. inaequalis*. L'introgression du gène *Vf* chez un pommier sensible permet sa résistance à la tavelure (Silfverberg-Dilworth *et al.*, 2005) mais des résistances aux champs ont rapidement été observées après introgression (Parisi *et al.*, 1993). Les populations virulentes sont très différenciées génétiquement des populations issues d'hôtes ne portant pas le gène *Vf* que ce soit à l'échelle nationale ou européenne (Guérin et Le Cam, 2004 ; Guérin, 2004 ; Gladieux, 2007). La nette différenciation obtenue avec *FLO1* et *FLO2* semble également révéler l'effet structurant lié à la présence de gènes majeurs. Toutefois, l'ensemble des *M. floribunda* du verger semblent présenter le même allèle au gène *Vf* que le clone 821 (Dunemann, communication personnelle). Dans l'état de nos connaissances actuelles, cette différenciation génétique entre *FLO1* et *FLO2* pourrait s'expliquer par la présence du gène de résistance *Vfh* identifié chez *M. floribunda* clone 821 (Benaouf et Parisi, 2000) et non étudié chez les autres hôtes. Par ailleurs, il est connu que la probabilité de croisement des individus dispersant leurs gamètes à faible distance dépend de leur capacité à infecter le même hôte, ce qui induit une pléiotropie entre spécialisation à l'hôte et choix de croisement (Giraud *et al.*, 2006). La résultante de ce processus est un isolement reproducteur rapide des deux populations. Il est alors possible que la stricte spécialisation de ces deux populations aux deux types de *M. floribunda* induira un isolement reproducteur. Le clonage du gène *Vfh* est nécessaire pour vérifier la présence de deux allèles différents chez les différents *M. floribunda*, si tel était le cas, cela mettrait en avant l'effet structurant du gène majeur *Vfh* au sein du verger mais aussi la faible fitness des individus virulents-*Vfh* sur les autres hôtes.

## **Partie 5 : Conclusion**

La compréhension des processus d'adaptation constitue un enjeu majeur pour établir la relation entre un pathogène et son hôte. En effet, pour un organisme pathogène, l'adaptation à un hôte est indispensable à sa survie. Les connaissances actuelles sur la spécialisation des agents pathogènes aux différentes espèces végétales sont parcellaires. Afin de répondre à la présence éventuelle d'une spécialisation de *V. inaequalis* aux différents *Malus*, plusieurs approches ont été développées (i) la recherche de signature de sélection liée aux différents *Malus* a été initiée basée sur un jeu de 92 marqueurs, (ii) la détermination du flux génique neutre entre l'ensemble des individus du verger, c'est à dire l'isolement par la distance et le flux génique entre populations. En effet, en génétique des populations la détermination de deux dynamiques semble fondamentale, l'évolution neutre du génome basée sur l'existence des processus de mutation, de migration et de dérive où le devenir des allèles est stochastique et le processus sélectif déterministe.

A travers l'étude multilocus, une hétérogénéité de signal a été observée en fonction des loci, en effet, (i) un locus est sous sélection, (ii) 24 montrent un profil d'isolement par la distance et (iii) 67 permettent de déterminer des spécialisations du pathogène liées aux hôtes.

La présence du locus sous sélection au sein du génome de *V. inaequalis* semble être la signature d'un processus de spécialisation des populations aux différents *Malus*. Malheureusement, ce locus ne différencie pas l'ensemble des individus d'une population d'une autre et ne permet pas, par conséquent, de déterminer une spéciation écologique de certains hôtes via ce locus.

L'étude du flux génique neutre a mis en évidence l'existence d'un isolement des populations par la distance. De façon intéressante, la majorité des ascospores semblent se disséminer à quelques mètres ou quelques dizaines de mètres, même si quelques cas sporadiques de dissémination à plusieurs kilomètres peuvent avoir lieu (Aylor, 1998). Il en résulte que chez *V. inaequalis*, le flux de gènes ne serait possible qu'entre vergers proches. Pour les vergers distants, le flux de gènes ne serait possible que par la mise en contact accidentelle des populations par l'homme. De plus, basée sur l'analyse des 91 loci, seul *SYL* montre le même signal sur les 91 loci impliquant une adaptation locale.

Les souches isolées de *M. floribunda* clone 821 et des autres *M. floribunda* non 821 sont fortement différenciées des autres souches du verger mais aussi entre-elles, si et seulement si l'analyse est basée sur les 67 loci, impliquant une forte hétérogénéité de signal et ainsi des régions du génome en spécialisation à l'hôte et des régions sous IBD. Du point de vue de la spécialisation, la différence entre *FLO1* et *FLO2* peut être expliquée par l'existence d'un gène majeur supplémentaire chez le clone 821 structurant les populations en deux groupes. *Vf* le gène majeur le mieux documenté

ne peut expliquer cette différence, puisque le même allèle est présent chez l'ensemble des *M. floribunda* du verger. La présence du gène *Vfh* chez le clone 821 contraindrait les contournants à porter au moins deux virulences (*Vf* et *Vfh*). La différenciation observée entre les deux populations *FLO1* et *FLO2* traduirait une difficulté pour ces deux populations à infecter un même hôte et ensuite à se croiser et pose donc à nouveau le coût pour les souches contournantes à porter ces virulences.

## **Partie 6 : Perspectives**

Nos résultats permettent d'initier de nouvelles pistes concernant l'analyse de la spécialisation des agents pathogènes aux hôtes attaqués. Notre projet met en avant de nouvelles pistes de recherche à plus ou moins court terme.

A court terme, l'augmentation du nombre de marqueurs pour rechercher des loci sous sélection semble être une information importante pour détecter des combinaisons de loci expliquant l'infection préférentielle de telle espèce de *Malus*, la cartographie de ces loci sera ensuite effectuée. Un projet en cours au sein du laboratoire s'intéresse déjà au polymorphisme de 16 effecteurs potentiels identifiés chez *V. inaequalis* (Bowen *et al.*, 2009; Kucheryava *et al.*, 2008). Des corrélations sont recherchées pour étudier (i) comment les pathogènes se structurent par rapport à ces facteurs de virulence (McDermott et McDonald, 1993) (ii) l'évolution moléculaire de ces effecteurs, en nombre et en qualité, pour expliquer l'attaque des nouveaux hôtes notamment *M. sylvestris* depuis la domestication.

Enfin, l'obtention prochaine du génome de *V. inaequalis* et le développement d'outils bioinformatiques développés chez des champignons proches tels que *Leptosphaeria maculans* devraient nous permettre une identification plus rapide d'effecteurs que nous utiliserons, par exemple, pour cribler des gènes de résistances dans les ressources génétiques (Song *et al.*, 2009).

## Partie 7: Références bibliographiques

- Aylor, D.E. (1998) The aerobiology of apple scab. *Plant Disease*, **82**, 838-349.
- Barres, B., Halkett, F., Dutech, C., Andrieux, A., Pinon, J. and Frey, P. (2008) Genetic structure of the poplar rust fungus *Melampsora larici-populina*: evidence for isolation by distance in Europe and recent founder effects overseas. *Infect Genet Evol*, **8**, 577-587.
- Beaumont, M.A. and Balding, D.J. (2004) Identifying adaptive genetic divergence among populations from genome scans. *Mol Ecol*, **13**, 969-980.
- Belkhir, K., Borsa, P., Chikhi, L., Raufaste, N. and Bonhomme, F. (1996-2004) GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations. Montpellier: Laboratoire Génome, Populations, Interactions, CNRS UMR 5000, Université de Montpellier II.
- Benaouf, G. and Parisi, L. (2000) Genetics of Host-Pathogen Relationships Between *Venturia inaequalis* Races 6 and 7 and *Malus* Species. *Phytopathology*, **90**, 236-242.
- Bowen, J.K., Mesarich, C.H., Rees-George, J., Cui, W., Fitzgerald, A., Win, J., Plummer, K.M. and Templeton, M.D. (2009) Candidate effector gene identification in the ascomycete fungal phytopathogen *Venturia inaequalis* by expressed sequence tag analysis. *Mol Plant Pathol*, **10**, 431-448.
- Brown, J.K. and Hovmoller, M.S. (2002) Aerial dispersal of pathogens on the global and continental scales and its impact on plant disease. *Science*, **297**, 537-541.
- Coart, E., Van Glabeke, S., De Loose, M., Larsen, A.S. and Roldan-Ruiz, I. (2006) Chloroplast diversity in the genus *Malus*: new insights into the relationship between the European wild apple (*Malus sylvestris* (L.) Mill.) and the domesticated apple (*Malus domestica* Borkh.). *Mol Ecol*, **15**, 2171-2182.
- Corander, J. and Marttinen, P. (2006) Bayesian identification of admixture events using multilocus molecular markers. *Mol Ecol*, **15**, 2833-2843.
- Corander, J., Waldmann, P., Marttinen, P. and Sillanpaa, M.J. (2004) BAPS 2: enhanced possibilities for the analysis of genetic population structure. *Bioinformatics*, **20**, 2363-2369.
- Coyne, J.A. and Orr, H.A. (2004) *Speciation*. Sinauer Associates. Sunderland, Massachusetts.
- Douhan, G.W., Smith, M.E., Huynh, K.L., Westbrook, A., Beerli, P. and Fisher, A.J. (2008) Multigene analysis suggests ecological speciation in the fungal pathogen *Claviceps purpurea*. *Mol Ecol*, **17**, 2276-2286.
- Evanno, G., Regnaut, S. and Goudet, J. (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol*, **14**, 2611-2620.
- Falush, D., Stephens, M. and Pritchard, J.K. (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics*, **164**, 1567-1587.
- Falush, D., Stephens, M. and Pritchard, J.K. (2007) Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Mol Ecol Notes*, **7**, 574-578.
- Flor, H.H. (1971) Current status of the gene-for-gene concept. *Annu Rev Phytopathol*, **9**, 275-278.
- Foll, M. and Gaggiotti, O. (2008) A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. *Genetics*, **180**, 977-993.
- Giraud, T., Villareal, L.M., Austerlitz, F., Le Gac, M. and Lavigne, C. (2006) Importance of the life cycle in sympatric host race formation and speciation of pathogens. *Phytopathology*, **96**, 280-287.
- Gladieux, P. (2007) Biogéographie et biologie des populations de *Venturia inaequalis*, champignon invasif responsable de la tavelure du pommier. Thèse de doctorat. Université d'Angers, pp. 140.
- Gladieux, P., Zhang, X.G., Afoufa-Bastien, D., Valdebenito Sanhueza, R.M., Sbaghi, M. and Le Cam, B. (2008) On the origin and spread of the Scab disease of apple: out of central Asia. *PLoS ONE*, **3**, e1455.

- Goodwin, D.C. and Lee, S.B.** (1993) Microwave miniprep of total genomic DNA from fungi, plants, protists and animals for PCR. *Biotechniques*, **15**, 438, 441-432, 444.
- Guérin, F.** (2004) Mise en évidence d'une population génétiquement différenciée de *Venturia inaequalis*, agent de la tavelure du Pommier, associée au contournement du gène majeur de résistance *Vf*. Thèse de doctorat. Université d'Angers, pp. 120.
- Guérin, F., Gladieux, P. and Le Cam. B.** (2007) Origin and colonization history of newly virulent strains of the phytopathogenic fungus *Venturia inaequalis*. *Fungal Genet Biol*, **44**, 284-292.
- Guérin, F. and Le Cam. B.** (2004) Breakdown of the Scab Resistance Gene *Vf* in Apple Leads to a Founder Effect in Populations of the Fungal Pathogen *Venturia inaequalis*. *Phytopathology*, **94**, 364-369.
- Guinand, B.** (1996) Use of a multivariate model using allele frequency distributions to analyse patterns of genetic differentiation among populations. *Biol J Linn Soc Lond*, **58**, 173-195.
- Hardy, O.J. and Vekemans, X.** (1999) Isolation by distance in a continuous population: reconciliation between spatial autocorrelation analysis and population genetics models. *Heredity*, **83** ( Pt 2), 145-154.
- Harris, S.A., Robinson, J.P. and Juniper, B.E.** (2002) Genetic clues to the origin of the apple. *Trends Genet*, **18**, 426-430.
- Juniper, B.E. and Maberley, D.J.** (2006) *The story of the apple*. Timber Press. Grantham, United Kingdom.
- Kimura, M. and Weiss, G.H.** (1964) The Stepping Stone Model of Population Structure and the Decrease of Genetic Correlation with Distance. *Genetics*, **49**, 561-576.
- Kucheryava, N., Bowen, J.K., Sutherland, P.W., Conolly, J.J., Mesarich, C.H., Rikkerink, E.H., Kemen, E., Plummer, K.M., Hahn, M. and Templeton, M.D.** (2008) Two novel *Venturia inaequalis* genes induced upon morphogenetic differentiation during infection and in vitro growth on cellophane. *Fungal Genet Biol*, **45**, 1329-1339.
- Lemaire, C., Allegrucci, G., Naciri, M., Bahri-Sfar, L., Kara, H. and Bonhomme, F.** (2000) Do discrepancies between microsatellite and allozyme variation reveal differential selection between sea and lagoon in the sea bass (*Dicentrarchus labrax*)? *Mol Ecol*, **9**, 457-467.
- Lenormand, T.** (2002) Gene flow and the limits to natural selection. *Trends Ecol Evol*, **17**, 183-189.
- Lewontin, R.C. and Krakauer, J.** (1973) Distribution of gene frequency as a test of the theory of the selective neutrality of polymorphisms. *Genetics*, **74**, 175-195.
- Loiselle, B.A., Sork, V.L., Nason, J. and Graham. C.** (1995) Spatial genetic structure of a tropical understory shrub, *Psychotria officinalis* (Rubiaceae). *Am J Bot*, **82**, 1420-1425.
- Malécot, G.** (1948) *Les mathématiques de l'hérédité*. Masson et Cie. Paris, France.
- Malécot, G.** (1955) The decrease of relationship with distance. *Cold Spring Harbor Symp Quant Biol*, **20**, 52-53.
- Malécot, G.** (1959) Les modèles stochastiques en génétique de population. *Publications de L'Institut de Statistiques de l'Université de Paris*, **8**, 173-210.
- McDermott, J.M. and McDonald, B.A.** (1993) Gene flow in plant pathosystems., **31**, 353-373.
- McDonald, B.A.** (1997) The population genetics of fungi: tools and techniques. *Phytopathology*, **87**, 448-453.
- Parisi, L., Lespinasse, Y., Guillaumes, J. and Krüger, J.** (1993) A new race of *Venturia inaequalis* virulent to apples with resistance due to the *Vf* gene. *Phytopathology*, **83**, 533-537.
- Pritchard, J.K., Stephens, M. and Donnelly, P.** (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945-959.
- Prospero, S., Lung-Escarmant, B. and Dutech, C.** (2008) Genetic structure of an expanding *Armillaria* root rot fungus (*Armillaria ostoyae*) population in a managed pine forest in southwestern France. *Mol Ecol*, **17**, 3366-3378.
- Robinson, J.P., Harris, S.A. and Juniper, B.E.** (2001) Taxonomy of the genus *Malus* mill. (Rosaceae) with emphasis on the cultivated apple, *Malus domestica* Borkh. *Plant Syst Evol*, **226**, 35-58.

- Rousset, F.** (1997) Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics*, **145**, 1219-1228.
- Rousset, F.** (2000) Genetic differentiation between individuals. *J Evol Biol*, **13**, 58-62.
- Silfverberg-Dilworth, E., Besse, S., Paris, R., Belfanti, E., Tartarini, S., Sansavini, S., Patocchi, A. and Gessler, C.** (2005) Identification of functional apple scab resistance gene promoters. *Theor Appl Genet*, **110**, 1119-1126.
- Slev, P.R. and Potts, W.K.** (2002) Disease consequences of pathogen adaptation. *Curr Opin Immunol*, **14**, 609-614.
- Song, J., Win, J., Tian, M., Schornack, S., Kaschani, F., Ilyas, M., van der Hoorn, R.A. and Kamoun, S.** (2009) Apoplastic effectors secreted by two unrelated eukaryotic plant pathogens target the tomato defense protease Rcr3. *Proc Natl Acad Sci U S A*, **106**, 1654-1659.
- Stukenbrock, E.H., Banke, S., Javan-Nikkhah, M. and McDonald, B.A.** (2007) Origin and domestication of the fungal wheat pathogen *Mycosphaerella graminicola* via sympatric speciation. *Mol Biol Evol*, **24**, 398-411.
- Tenzer, I., Degli Ivanissevich, S., Morgante, M. and Gessler, C.** (1999) Identification of Microsatellite Markers and Their Application to Population Genetics of *Venturia inaequalis*. *Phytopathology*, **89**, 748-753.
- Vekemans, X. and Hardy, O.J.** (2004) New insights from fine-scale spatial genetic structure analyses in plant populations. *Mol Ecol*, **13**, 921-935.
- Zaffarano, P.L., McDonald, B.A. and Linde, C.C.** (2008) Rapid speciation following recent host shifts in the plant pathogenic fungus *Rhynchosporium*. *Evolution*, **62**, 1418-1436.

## Partie 8 : Annexes

### Annexe I : souches utilisées pour l'étude

Tableau 1 : souches utilisées à la fois pour les analyses sur les loci AFLP et SSR. Les colonnes X et Y représentent les coordonnées spatiales de l'arbre. La colonne L5 représente l'allèle (présence/absence) au locus AFLP L5.

\* individu ne provenant pas du verger étudié, mais de vergers avoisinants

<b>Souche</b>	<b>Hôte</b>	<b>Arbre</b>	<b>Rang</b>	<b>X</b>	<b>Y</b>	<b>locus L5</b>
08M91	<i>M. sieversii</i>	7	1	21	85,5	1
08M92	<i>M. sieversii</i>	9	1	28	85,5	1
08M94a	<i>M. sieversii</i>	9	1	28	85,5	1
08M94b	<i>M. sieversii</i>	9	1	28	85,5	1
08M95	<i>M. sieversii</i>	10	1	31,5	85,5	1
08M96	<i>M. sieversii</i>	10	1	31,5	85,5	1
08M97	<i>M. sieversii</i>	10	1	31,5	85,5	1
08M98a	<i>M. sieversii</i>	3	4	7	72	1
08M108	<i>M. sieversii</i>	5	16	14	18	1
08M110	<i>M. sieversii</i>	5	16	14	18	1
08M111	<i>M. sieversii</i>	6	16	17,5	18	1
08M112	<i>M. sieversii</i>	6	16	17,5	18	1
08M113	<i>M. sieversii</i>	6	16	17,5	18	1
08M114	<i>M. sieversii</i>	6	16	17,5	18	1
08M116	<i>M. sylvestris</i>	38	1	129,5	85,5	1
08M128	<i>M. sylvestris</i>	14	3	45,5	76,5	1
08M129	<i>M. sylvestris</i>	14	3	45,5	76,5	1
08M130b	<i>M. sylvestris</i>	14	3	45,5	76,5	1
08M133a	<i>M. sylvestris</i>	32	3	108,5	76,5	0
08M133c	<i>M. sylvestris</i>	32	3	108,5	76,5	0
08M134a	<i>M. sylvestris</i>	32	3	108,5	76,5	0
08M135	<i>M. sylvestris</i>	32	3	108,5	76,5	0
08M136	<i>M. baccata</i>	1	10	0	45	1
08M137	<i>M. baccata</i>	1	10	0	45	1
08M146	<i>M. baccata</i>	7	10	21	45	1
08M153	<i>M. baccata</i>	12	10	38,5	45	1
08M154	<i>M. baccata</i>	13	10	42	45	1
08M155	<i>M. baccata</i>	13	10	42	45	1
08M156	<i>M. baccata</i>	14	10	45,5	45	1
08M157	<i>M. baccata</i>	14	10	45,5	45	1
08M158	<i>M. baccata</i>	15	10	49	45	1
08M159	<i>M. baccata</i>	15	10	49	45	1
08M160	<i>M. baccata</i>	16	10	52,5	45	0
08M161	<i>M. baccata</i>	16	10	52,5	45	1
08M162a	<i>M. baccata</i>	20	10	66,5	45	1
08M162b	<i>M. baccata</i>	20	10	66,5	45	1
08M163	<i>M. baccata</i>	21	10	70	45	1
08M167	<i>M. baccata</i>	31	10	105	45	1
08M168	<i>M. baccata</i>	31	10	105	45	1

08M170	<i>M. baccata</i>	32	10	108,5	45	1
08M172	<i>M. baccata</i>	33	10	112	45	1
08M173	<i>M. baccata</i>	5	11	14	40,5	1
08M174	<i>M. baccata</i>	5	11	14	40,5	1
08M175	<i>M. baccata</i>	4	11	10,5	40,5	1
08M177	<i>M. baccata</i>	3	11	7	40,5	1
08M178	<i>M. baccata</i>	3	11	7	40,5	1
08M180b	<i>M. ioensis</i>	11	6	35	63	1
08M181a	<i>M. ioensis</i>	11	6	35	63	1
08M185	<i>M. ioensis</i>	25	7	84	58,5	1
08M192	<i>M. ioensis</i>	27	7	91	58,5	1
08M195	<i>M. ioensis</i>	28	7	94,5	58,5	1
08M199a	<i>M. ioensis</i>	31	7	105	58,5	0
08M199b	<i>M. ioensis</i>	31	7	105	58,5	1
08M200a	<i>M. ioensis</i>	31	7	105	58,5	0
08M200b	<i>M. ioensis</i>	31	7	105	58,5	0
08M201	<i>M. ioensis</i>	31	7	105	58,5	1
08M202b	<i>M. ioensis</i>	32	7	108,5	58,5	0
08M203	<i>M. ioensis</i>	32	7	108,5	58,5	0
08M204	<i>M. ioensis</i>	32	7	108,5	58,5	0
08M205	<i>M. ioensis</i>	32	7	108,5	58,5	1
08M206	<i>M. ioensis</i>	32	7	108,5	58,5	0
08M209	<i>M. floribunda</i> clone 821	5	20	14	0	1
08M211	<i>M. floribunda</i> clone 821	6	20	17,5	0	1
08M214	<i>M. floribunda</i> clone 821	6	20	17,5	0	0
08M395a	<i>M. floribunda</i> clone 821	5	20	14	0	0
08M395b	<i>M. floribunda</i> clone 821	5	20	14	0	0
08M396	<i>M. floribunda</i> clone 821	5	20	14	0	0
08M397	<i>M. floribunda</i> clone 821	6	20	17,5	0	0
08M399	<i>M. floribunda</i> clone 821	6	20	17,5	0	0
08M233	<i>M. floribunda</i> (non 821)	29	6	98	63	0
08M237	<i>M. floribunda</i> (non 821)	30	6	101,5	63	1
08M239	<i>M. floribunda</i> (non 821)	31	6	105	63	1
08M243	<i>M. floribunda</i> (non 821)	32	6	108,5	63	0
08M401b	<i>M. floribunda</i> (non 821)	25	6	84	63	0
08M406	<i>M. floribunda</i> (non 821)	27	6	91	63	0
08M407	<i>M. floribunda</i> (non 821)	27	6	91	63	0
08M408	<i>M. floribunda</i> (non 821)	27	6	91	63	0
08M409	<i>M. floribunda</i> (non 821)	28	6	94,5	63	0
08M410	<i>M. floribunda</i> (non 821)	28	6	94,5	63	0
08M411	<i>M. floribunda</i> (non 821)	30	6	101,5	63	0
08M412	<i>M. floribunda</i> (non 821)	30	6	101,5	63	0
08M413	<i>M. floribunda</i> (non 821)	*	*	*	*	0
08M414	<i>M. floribunda</i> (non 821)	*	*	*	*	0
08M415	<i>M. floribunda</i> (non 821)	*	*	*	*	1
06M06	<i>M. floribunda</i> (non 821)	*	*	*	*	0
06M07	<i>M. floribunda</i> (non 821)	*	*	*	*	0
08M248b	<i>M. coronaria</i>	2	6	3,5	63	0
08M249b	<i>M. coronaria</i>	2	6	3,5	63	0
08M251b	<i>M. coronaria</i>	7	7	21	58,5	0

08M251c	<i>M. coronaria</i>	7	7	21	58,5	0
08M252a	<i>M. coronaria</i>	7	7	21	58,5	0
08M256a	<i>M. coronaria</i>	8	7	24,5	58,5	0
08M257b	<i>M. coronaria</i>	8	7	24,5	58,5	0
08M258	<i>M. coronaria</i>	8	7	24,5	58,5	0
08M261	<i>M. x zumi</i>	11	12	35	36	0
08M267	<i>M. x zumi</i>	12	12	38,5	36	0
08M268	<i>M. x zumi</i>	12	12	38,5	36	0
08M269	<i>M. x zumi</i>	12	12	38,5	36	0
08M271	<i>M. x zumi</i>	3	19	7	4,5	0
08M273	<i>M. x zumi</i>	3	19	7	4,5	0
08M275	<i>M. x zumi</i>	4	19	10,5	4,5	0
08M278	<i>M. x zumi</i>	4	19	10,5	4,5	0
08M279	<i>M. x zumi</i>	4	19	10,5	4,5	1
08M294a	<i>M. lancifolia</i>	7	6	21	63	1
08M297b	<i>M. lancifolia</i>	8	6	24,5	63	0
08M301	<i>M. fusca</i>	21	6	70	63	0
08M304a	<i>M. fusca</i>	22	6	73,5	63	0
08M306	<i>M. x purpurea</i>	31	4	105	72	0
08M307	<i>M. x purpurea</i>	31	4	105	72	0
08M309a	<i>M. x purpurea</i>	31	4	105	72	0
08M309b	<i>M. x purpurea</i>	31	4	105	72	0
08M315	<i>M. x purpurea</i>	19	13	63	31,5	0
08M320	<i>M. x purpurea</i>	19	13	63	31,5	0
08M321	<i>M. x purpurea</i>	19	13	63	31,5	0
08M327	<i>M. x purpurea</i>	20	13	66,5	31,5	0
08M329	<i>M. x purpurea</i>	20	13	66,5	31,5	0
08M336	<i>M. x purpurea</i>	15	13	49	31,5	0
08M343	<i>M. bhutanica</i>	13,5	8	43,75	54	1
08M347	<i>M. orientalis</i>	1	1	0	85,5	1
08M348a	<i>M. orientalis</i>	1	1	0	85,5	0
08M348b	<i>M. orientalis</i>	1	1	0	85,5	0
08M352	<i>M. orientalis</i>	1	1	0	85,5	0

## Annexe II : Méthodes utilisées pour l'isolement et les extractions ADN.

### Isolement monospores

Les souches de *V. inaequalis* sont isolées à partir de lésions sporulantes sur feuilles. Les spores sont récupérées par dépôt et aspersion de 100µL d'eau stérile sur la lésion. Cette solution est ensuite étalée sur milieu Malt-Agar (1%, 1,5% p/v respectivement) supplémenté en streptomycine (25µg/L), penicilline G (12,5 µg/L) et chlortétracycline (5µg/L).Après 24H de croissance dans une chambre de culture à 17°C, les spores germées sont isolées sous loupe binoculaire et transférées sur milieu Malt-Agar sans antibiotiques. Après six semaines à 17°C, la souche monosporee a colonisé une partie de la boîte de Pétri, un étalement des souches est alors effectué sur un milieu Malt-Agar contenant une membrane de cellophane stérile (préalablement découpée en disque et déposée dans une boîte de Pétri).

### Extraction d'ADN selon le protocole de Goodwin et Lee (1993)

Le mycélium pousse à la surface de la membrane tout en se nourrissant du milieu nutritif et colonise en 2 à 3 semaines l'ensemble de la boîte. Une extraction d'ADN génomique est possible selon la méthode décrite par Goodwin et Lee en 1993. Le mycélium est prélevé en milieu stérile et déposé dans des tubes Phase Lock Gel Heavy 2 ml (5 Prime , Gaithersburg, Maryland, USA) et 100µL de tampon de lyse (Tris-HCl pH 7,2 50 mM. EDTA 50 mM. SDS 3% (p/v); 2-mercaptoéthanol 1% (v/v)) est ajouté. Les échantillons sont alors soumis aux microondes à 850 W pendant 15, 10 puis 5 secondes. Après ajout de 300µL de tampon de lyse et incubation 10 minutes dans un bain-marie à 80°C, 1 volume de Phénol-Chloroforme-Alcool Isoamylique (25:24:1) est ajouté et les tubes sont agités vigoureusement. Après centrifugation (15 minutes, 14000 g, 4°C), la phase aqueuse est récupérée et la précipitation de l'ADN génomique est réalisée par ajout de 10µL d'acétate de sodium 5 M ainsi que 0,5 volumes d'isopropanol. Après incubation 15 minutes à -20°C, une centrifugation (4 minutes, 14000 g, 4°C) et un lavage à l'éthanol 70% (v/v) sont effectués, et le culot est séché et repris dans de l'eau ultra pure.

### Extraction d'ADN pour les échantillons non monospores

Les échantillons pour lequel un isolement monospore n'a pu être réalisé ont été étudiés par extraction des ADN génomiques à partir de disques foliaires, permettant l'extraction de l'ADN de l'hôte et du pathogène. Cette extraction a été réalisée par l'utilisation du kit Nucleospin Plant (Macherey-Nagel, Düren, Allemagne), en suivant strictement les instructions données par le fournisseur.

### Annexe III : Assignations des 117 haplotypes

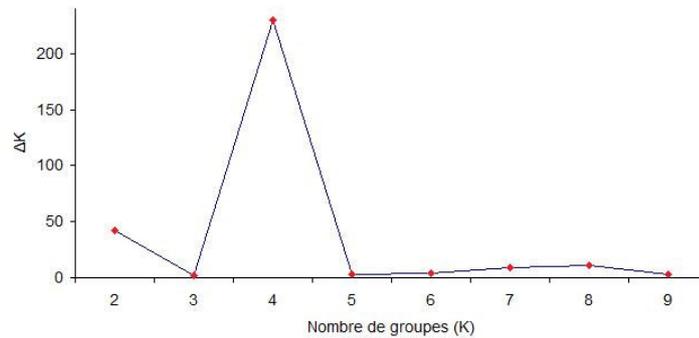


Figure 1 : Détermination du k optimal sous Structure 2.2 (Prichard *et al.*, 2000) à partir de l'analyse des 117 haplotypes après modification de la vraisemblance selon la méthode décrite par Evanno *et al.* (2005).

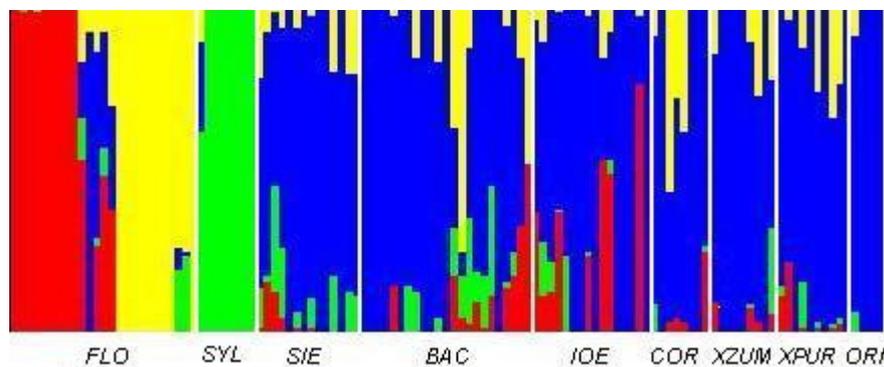


Figure 2 : Assignation de chacun des 117 haplotypes aux 4 groupes par BAPS 4.14 (Corander *et al.*, 2004 ; Corander et Marttinen, 2006).

Les individus sont regroupés en population séparées par les barres verticales où *FLO*, *SYL*, *SIE*, *BAC*, *IOE*, *COR*, *XZUM*, *XPUR*, *ORI* représentent les souches isolées de *M. floribunda*, *M. sylvestris*, *M. sieversii*, *M. baccata*, *M. ioensis*, *M. coronaria*, *M. x zumi*, *M. x purpurea* et *M. orientalis* respectivement.

## Figures et tableaux en opposition du texte dans la version papier

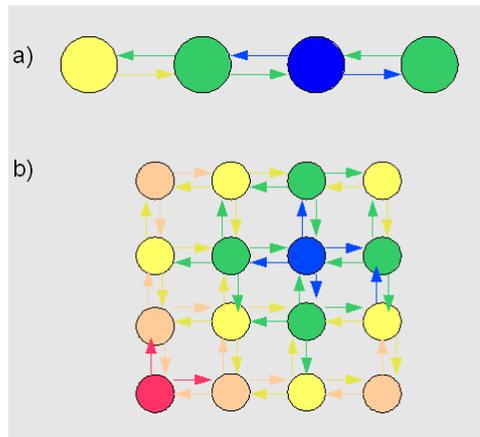


Figure 1 : Modèles en pas japonais de Kimura et Weiss (1964)

a) modèle en une dimension et b) modèle en deux dimensions

La propagation des migrants (bleu) est suivie au cours des générations. Plusieurs générations sont nécessaires pour migrer jusqu'au dernier clade (jaune en a, rouge en b). Les flèches représentent les échanges de migrants entre les clades.

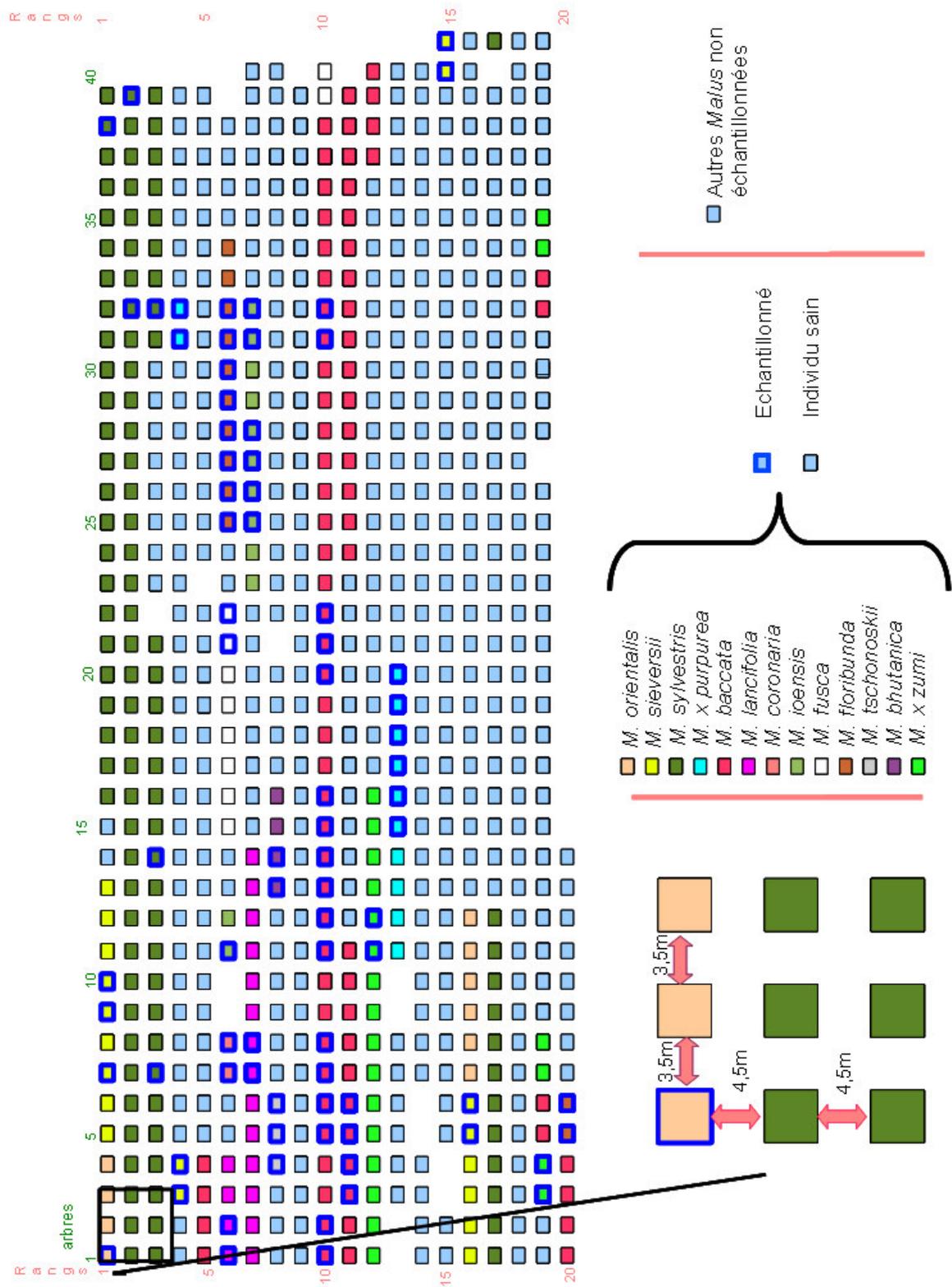


Figure 2 : Représentation simplifiée de la localisation des arbres échantillonnés au sein du verger selon le nom d'espèces de *Malus* en présence et des distances les séparant.

L'arbre 1 rang 1 est accessible aux coordonnées Global Positioning System (GPS) 50°59,915 Nord ; 13°53,275 Est.

Tableau 1 : Séquences des amorces et températures d'hybridation utilisées pour l'amplification des loci SSR.

loci	F	R	Temp. Hyb. (°C)
<i>1tc1a</i>	TCGAGATCCTCAAACCTTCCTT	TTTTAACTGTGCGGCCTG	58
<i>1tc1b</i>	CGATTGGGGATATGAAGACTT	TTAGTAATCAAATCGCACCCA	58
<i>1aac3b</i>	AGCGCTAGGTCGTGAAATC	TTTCTGAAGTGTGTGGGACAT	58
<i>1tc1g</i>	TCACTCAACAATACAGTTTCTTACG	TTTCACGGTAGCGATAGGAG	50
<i>Vitca7/P</i>	GAATACTTCCAAGTGCACAG	GTGACGCGGATATGGTAC	58
<i>Vica9/152</i>	GCACCTGCTCTGTCTATCTC	AAGGTTCAAGCACTGGAG	58
<i>Vitg11/70</i>	GAAGAGGTTGGAGTGGTTG	GAACCGAATCTGTACAGGAC	58
<i>Viga7/116</i>	GCCTGGTTGTGGATCTGTC	ATCCTGCTACATCGACCTTC	58
<i>Vica9/X</i>	TCGCGCATCACTATCTACAC	AGACAGGAATGTGGTGGAAAG	58
<i>Vicacg8/42</i>	TGTCAGCCACGCTAGAAG	CACCGGACGAATCATGC	58
<i>M42</i>	TAACTCCTGAAGACGGCATG	CCAGACCTCCTTATTCACG	58

Où F, R et Temp. Hyb. correspondent respectivement à la séquence sens, antisens, et à la température d'hybridation en degrés Celsius utilisée lors des PCR.

Tableau 2 : Séquences des amorces utilisées pour les analyses *f* AFLP

PstI Aaptateur	5'-TCGTAGACTGAGTACGTGCA-3'
	3'-CATCTGACTCATGC-phosphorylation-5'
Amorce PstI « core »	5'-GACTGAGTACGTGCAG-3'
Amorce sélective PstI « core » + A	5'-6-FAM-GACTGAGTACGTGCAGA-3'
MspI Adaptateur	5'-GACGATGAGTCCTGAG-3'
	3'-CTACTCACGACTCGC-phosphorylation-5'
Amorce MspI « core »	5'-CGATGAGTCCTGAGCGG-3'
Amorce sélective MspI « core » + AA	5'-CGATGAGTCCTGAGCGGAA-3'

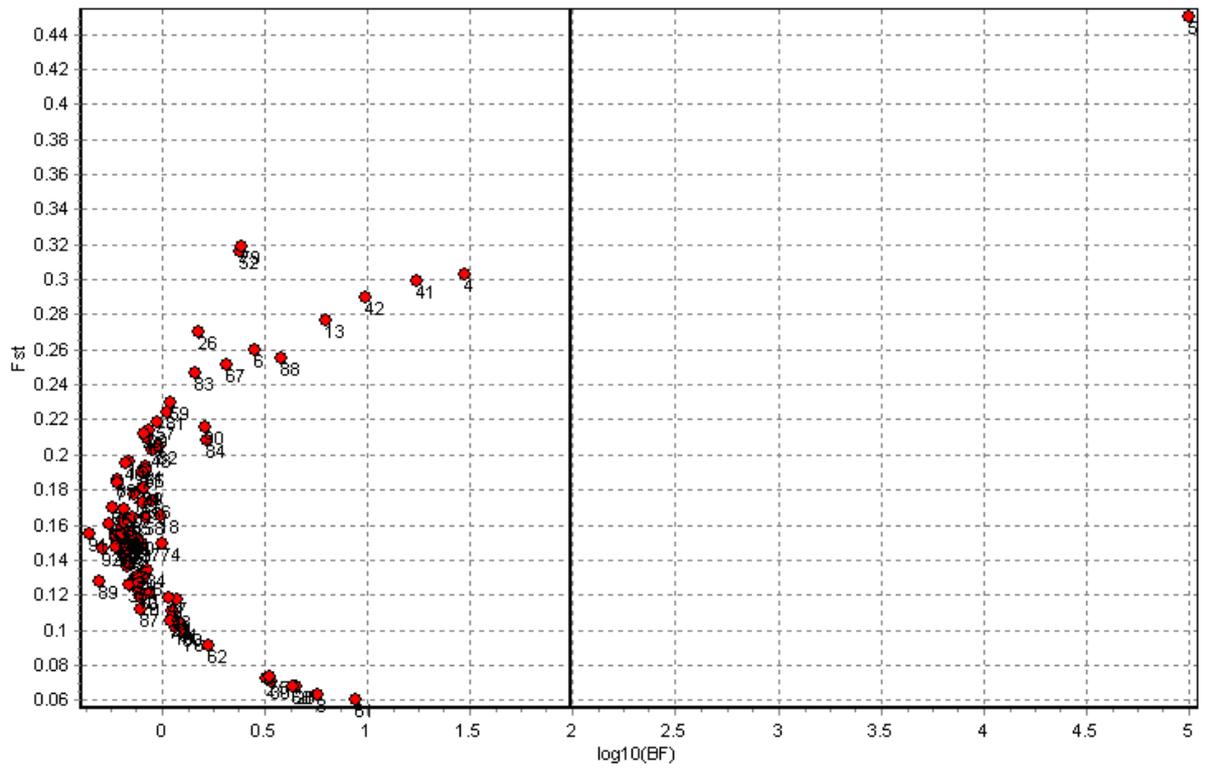


Figure 3 : Projection de chaque locus en fonction de la différenciation en sous populations et de la vraisemblance à être sous sélection. La limite fixée représente la zone d'acceptation d'un loci comme outlier au seuil 1%

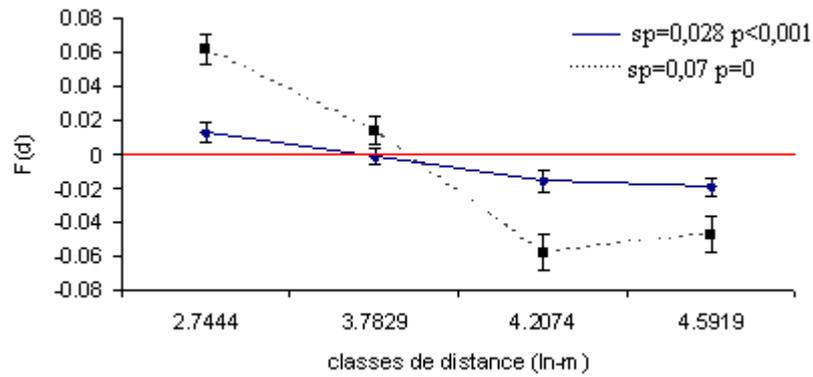


Figure 4 : Autocorrelogrammes montrant la moyenne des coefficients d'apparements par classe de distance [F(d)]. La courbe en trait plein représente l'analyse effectuée sur tous les loci (91) et la courbe en trait pointillé est réalisée sur les 24 loci significativement sous IBD. Les barres d'erreurs représentent la moyenne de l'écart-type, cette valeur étant déterminée par le jackknife à travers les 91 ou 24 loci respectivement. Le seuil indique la zone où F(d)=0.

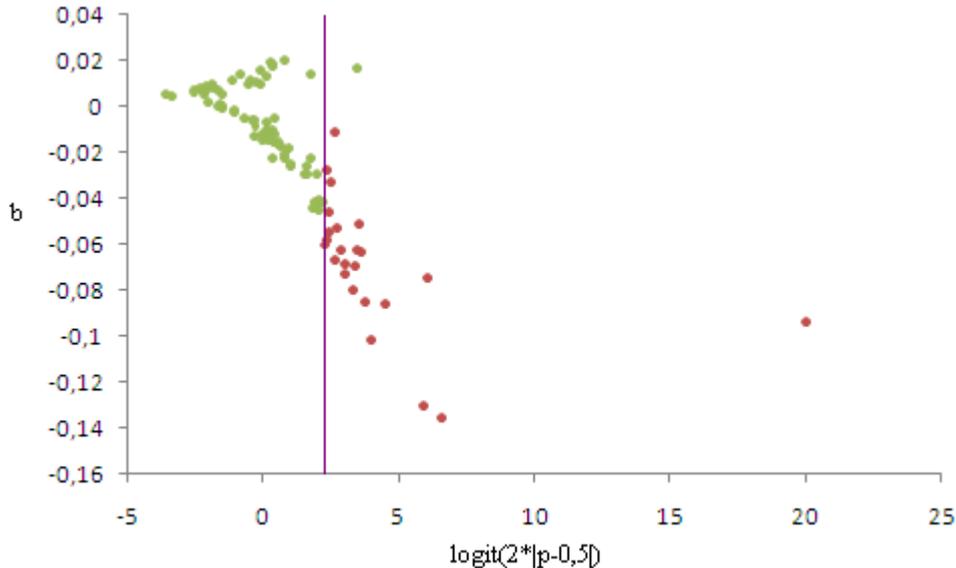


Figure 5 : Projection de chaque locus en fonction du coefficient directeur b de la régression linéaire et le  $\text{logit}(2*p-0,5)$  où  $\text{logit}(x)=\ln(x/(1-x))$ . La valeur p représente la probabilité d'existence d'un IBD sur les valeurs permutées 20 000 fois sur les individus et 20 000 fois sur leurs localisations. La barre verticale montre la valeur de  $p = 0,05$  critique utilisée pour détecter les loci sous IBD.

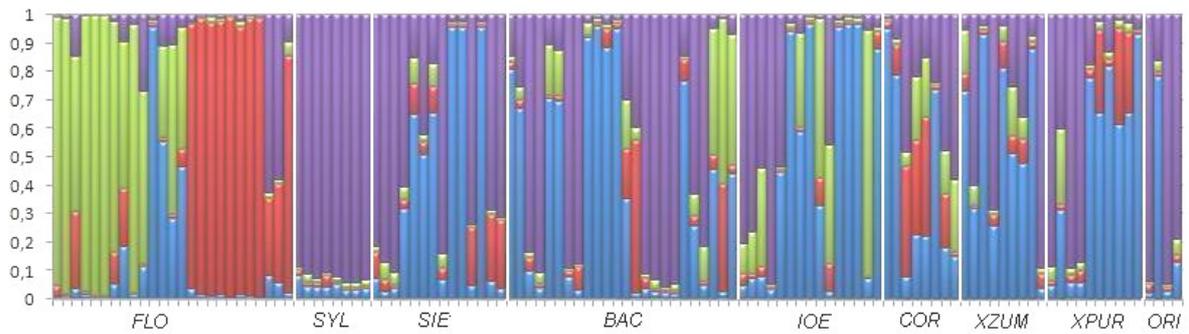


Figure 6 : Assignment de chacun des 117 haplotypes aux 4 clusters par Structure.

Les individus sont regroupés en population séparées par les barres verticales blanches où *FLO*, *SYL*, *SIE*, *BAC*, *IOE*, *COR*, *XZUM*, *XPUR*, *ORI* représentent *M. floribunda*, *M. sylvestris*, *M. sieversii*, *M. baccata*, *M. ioensis*, *M. coronaria*, *M. x zumi*, *M. x purpurea* et *M. orientalis* respectivement. Le nombre de clusters optimal défini par la méthode Evanno *et al* (2005) est de 4.

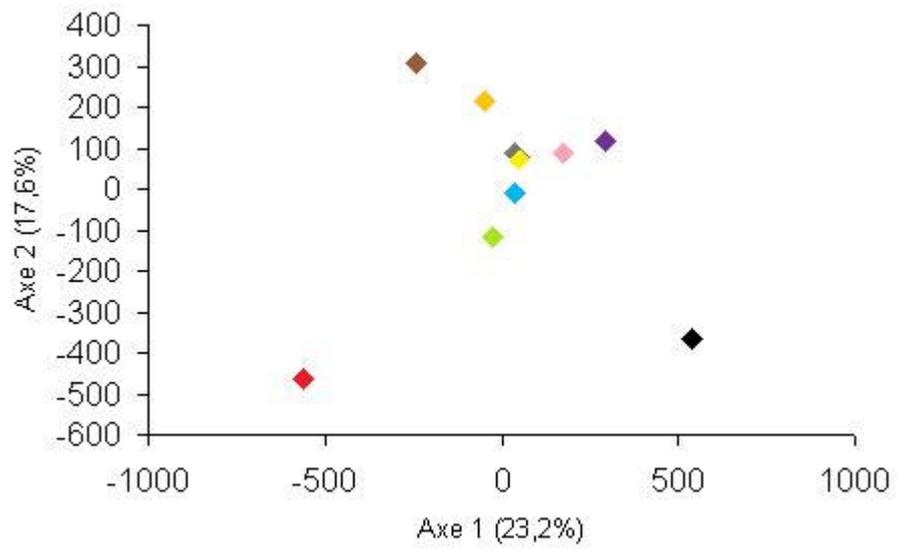


Figure 7 : Analyse Factorielle des Correspondances (AFC) basée sur les barycentres des populations issues de 10 espèces différentes d'hôtes. Les losanges représentent chaque population selon la légende suivante : marron (*FLO1*), rouge (*FLO2*), orange (*COR*), gris (*XZUM*), jaune (*XPUR*), bleu (*BAC*), vert (*IOE*), rose (*SIE*), violet (*ORI*) et noir (*SYL*).