



**HAL**  
open science

# Caractérisation des déterminants génétiques et moléculaires liés à la résistance au dépérissement bactérien chez l'abricotier et analyse des risques associés

Mariem Omrani

## ► To cite this version:

Mariem Omrani. Caractérisation des déterminants génétiques et moléculaires liés à la résistance au dépérissement bactérien chez l'abricotier et analyse des risques associés. Sciences agricoles. Université d'Avignon, 2018. Français. NNT : 2018AVIG0698 . tel-02354903

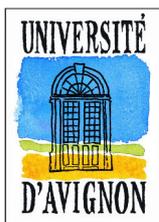
**HAL Id: tel-02354903**

**<https://theses.hal.science/tel-02354903>**

Submitted on 8 Nov 2019

**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 d'Avignon et des Pays de Vaucluse  
Ecole Doctorale 536 « Agrosiences et Sciences »

## Thèse

Pour obtenir le grade de  
Docteur de l'Université d'Avignon et des Pays de Vaucluse

présentée et soutenue publiquement par  
Mariem OMRANI  
le 6 novembre 2018

---

# Caractérisation des déterminants génétiques et moléculaires de la résistance au déperissement bactérien de l'abricotier et analyse de risques associés

---

Thèse préparée au sein des unités de recherche Génétique et Amélioration des Fruits et Légumes  
& Pathologie Végétale du centre INRA PACA d'Avignon

Mme Laurence MOREAU	Directrice de Recherches, INRA UMR GQE, Le Moulon	<i>Rapporteur</i>
M. Didier MERDINOGLU	Directeur de Recherches, INRA UMR SVQV, Colmar	<i>Rapporteur</i>
Mme Marie-Anne BARNY	Directrice de Recherches, INRA UMR IEES, Paris	<i>Examinatrice</i>
M. Patrice THIS	Directeur de Recherches, INRA UMR AGAP, Montpellier	<i>Examineur</i>
Mme Cindy MORRIS	Directrice de Recherches, INRA UR Pathologie Végétale, Avignon	<i>Directrice de thèse</i>
M. Jean-Marc AUDERGON	Ingénieur de Recherches, INRA UR GAFL, Avignon	<i>Co-encadrant de thèse</i>



*“La Science n’a jamais tout à fait raison, mais elle a rarement tout à fait tort, et en général, elle a plus de chance d’avoir raison que les théories non scientifiques. Il est donc rationnel de l’accepter à titre d’hypothèse.”*

Bertrand Russell, Histoire de mes idées philosophiques, 1961



## Remerciements

Et voilà, j'y suis... Après trois années de dur travail, d'heures innombrables à réessayer, ajuster, avancer, reculer, douter puis espérer, il est temps pour moi de remercier un certain nombre de personnes. Ce travail de thèse ne serait pas ce qu'il est sans la contribution de celles et ceux qui m'ont accompagnée dans cette expérience riche d'enseignements scientifiques mais aussi (et je dirais surtout) personnels et humains.

Je tiens tout d'abord à remercier les membres du jury d'avoir accepté d'évaluer mon travail, Laurence Moreau et Didier Merdinoglu en tant que rapporteurs ainsi que Patrice This et Marie-Anne Barny pour leur travail d'examineur.

Un très grand merci à Charles-Eric Durel, Michel Dron et Luciana Parisi pour leurs nombreux conseils et remarques constructives, donnés à l'occasion des comités de thèse, qui m'ont été particulièrement utiles afin d'avancer.

Mes remerciements s'adressent ensuite naturellement à mes deux formidables encadrants de thèse ! Jean-Marc, Cindy, un énorme merci pour la confiance et l'autonomie de travail que vous m'avez accordée, ainsi que pour la disponibilité dont vous avez fait preuve jusqu'aux dernières minutes de rédaction de ce manuscrit (et ce, malgré des emplois du temps bien chargés). La passion, l'enthousiasme communicatif et les nombreuses qualités humaines dont vous faites preuve au quotidien sont des éléments forts que je retiendrai particulièrement de mon séjour parmi vous.

Quoi de plus agréable que de se sentir épaulée par les collègues dans les grands chantiers expérimentaux ? L'inoculation et le décorticage de rameaux n'a vraiment plus de secrets pour nous ! Mille mercis à mes supers collègues Guillaume (le rossignol), Jean (la bricole), Eric (M. Rouge de FourneSSe), Sabrina, Anne-Marie au GAFL, Charlotte et Caroline en Patho, ainsi que Guy (ou l'art de jouer avec les mots), Laurent, Simon, Christophe, Pedro, Freddy de Gotheron et enfin Alain de l'Amarine pour leur participation enthousiaste dans la partie expérimentale de cette thèse, riche en sueur, fatigue, parafilm, pipettes, copeaux et neige en bonus (je garderai un souvenir de cette froide et surprenante journée de novembre qui faute d'inoculations nous aura tout de même permis d'admirer la beauté des vergers drômois enneigés) !

Un énorme merci à Carole (alias Carole 'les bons plans'), pour son aide, sa patience et son accompagnement dans la manipulation des données de marquage moléculaire, mon baptême avec Linux ainsi que pour les nombreux bons moments passés ensemble lors des diverses randos, restos, cinés et soirées jeux. Je remercie également Naïma (la relève !) avec qui j'ai pu travailler lors de cette fin de thèse.

Un grand merci à Benoît Moury pour son intérêt dans mon travail, son implication dans les analyses du plan factoriel d'interaction, et pour l'animation scientifique autour des résistances durables. J'ai beaucoup appris sur ce monde des interactions plante-pathogène qui m'était très peu familier au début de ma thèse.

Je tiens également à remercier Brigitte Mangin, Joël Chadoeuf et Christopher Sauvage qui m'ont apporté une aide précieuse du côté modélisation et traitement statistique des données. Un grand merci pour votre disponibilité et réactivité face à mes nombreuses sollicitations.

Plusieurs générations de stagiaires se sont succédées avant moi afin d'apporter une pierre supplémentaire à l'édifice de la connaissance : toute ma gratitude à Morgane, que j'ai pu solliciter par la suite, une pensée pour Mouna et Carlos, que je n'ai pas eu l'opportunité de rencontrer mais dont le travail dans le recueil de données phénotypiques, a été utile pour mon travail. Merci aux stagiaires et CDD, Audrey, Hanifé, Amandine pour leur implication et leur aide dans la partie expérimentale et avec qui j'ai eu la chance d'avoir une première expérience d'encadrement.

Merci à l'ensemble de l'équipe administrative des deux unités, Claudine, Pascale, Evelyne JM., Evelyne J., Astrid, Sébastien, Annick et Philippe pour la gestion des aspects administratifs.

Merci aux collègues du CPER, Marie-Noëlle, Rémy, Bénédicte, Véronique, Mathilde, Renaud, Caroline, Laure (la mommy), Frédérique, Justine, Esther, Nasradine, Elodie, Rebecca, Gisèle, Thierry, Henri, pour les moments conviviaux autour d'un gâteau à la pause-café, lors du déjeuner ou en dehors du bureau.

Je remercie aussi les collègues de Patho, Odile, Jean-François, Christelle et Magali pour leur accueil et leur sympathie lors de mes quelques visites de l'autre côté de la route.

Cette thèse s'inscrit dans le cadre d'une Formation Complémentaire par la Recherche (FCPR), financée par le Ministère de l'Agriculture et de l'Alimentation. Je suis donc reconnaissante envers les membres du Conseil Général de l'Alimentation, de l'Agriculture et des Espaces Ruraux (CGAAER), pour avoir retenu ma candidature et m'avoir permis de conduire mon projet de thèse. Merci par ailleurs à Emmanuel K., qui m'a soutenue et été d'une aide précieuse dans la constitution du dossier FCPR.

Un grand merci à Robert Tessier, membre du CGAAER et Alexandre Péry d'AgroParisTech pour leur suivi régulier au cours de ces 3 ans, et leurs conseils dans la préparation de mon insertion d'après-thèse.

Je tiens également à remercier les membres de l'Académie d'Agriculture, de la Société Française de Phytopathologie ainsi que l'Université d'Avignon pour leur aide

financière dans ma participation au PAG XXVII, marquant un temps fort de ma thèse.

Merci à l'ensemble des organisateurs, participants et coachs du concours régional 'ma thèse en 180 secondes', cette étape m'a été vraiment très enrichissante et m'a permis d'entrevoir la Science sous un autre angle ; j'en garde un souvenir très amusant.

Parce qu'il est aussi important de prendre du temps pour se ressourcer lors de ce marathon doctoral, merci à tous les amis thésards, stagiaires, techniciens, anciens ou actuels avec qui j'ai pu passer de très bons moments en dehors du bureau.

Ma toute première 'copine de thèse' Steph, (papy) Isidore (au moment où tu liras ces lignes, je serai déjà loin, je n'ai pas peur des représailles...), Lucie, Elise, ma co-pilote badiste Marine, Max (tomate), Pierre, Zoé, Gaëtan, Anna, Jiantao, Rafael, Lorraine, Coralie, Margaux, mes co-bureaux Léandro, Caterina, (bientôt ex-'mini') Mariem et Hussein et enfin la bande de Saint-Paul: mes copines twin-peaks et gastronomes Héléne et Johanna, Cathleen, Max (forêt), Violette, Claire, Stéphanie C., Anne-Sophie, Coffi, Camille, Enrico : les soirées explo, les randos, restos (et autres événements en 'o'), me laissent d'excellents souvenirs de mon passage à Avignon.

Une petite dédicace musicale aux ratatouilles, nos sessions gratte vont me manquer ! Merci également à mes coéquipiers de bad de la D3, qui ont contribué au maintien de ma forme physique (et mentale !).

Merci aux amis de longue date, Coco, Juliette, Mathilde, Laurence, Anne-Laure, Marion pour vos visites estivales, les bons moments passés dans la ferveur du festival, et pour votre amitié depuis près de 15 ans maintenant ! Merci à mes trois copines de l'agro, Marina, Chloé et Caroline pour le tour de France de la saint-sylvestre. Merci à vous tous, de m'avoir permis de couper court à l'univers de la recherche quand c'était nécessaire. Vous êtes des amis merveilleux !

Enfin, à ma famille, merci pour tout, merci pour votre soutien sans faille en particulier lors de ces deux derniers mois de thèse. Papa, Maman, vous m'avez toujours accompagnée et je ne vous en remercierai jamais assez. Un nouveau docteur rejoint bientôt les rangs de la famille, je l'espère ! Merci à mon frérot pour son aide informatique à distance et à ma sœur pour être tout simplement elle-même.

Une pensée pour mes grands-parents, sans doute les premiers à avoir repéré ma fibre agronome, quand pas plus haute que trois pommes, je passais mon temps en Tunisie à explorer le jardin, à faire un inventaire des arbres fruitiers. Même à distance, je sais que vous avez toujours été là pour moi, merci.



## **Thèse de Doctorat**

Caractérisation des déterminants génétiques  
et moléculaires liés à la résistance au  
déperissement bactérien chez l'abricotier et  
analyse de risques associés

## **PhD Thesis**

Characterization of genetic and molecular  
determinants of apricot resistance to bacterial  
canker and analysis of associated risks

Mariem OMRANI, 2018



## Résumé

Parmi les *Prunus*, genre botanique d'intérêt économique important, l'abricotier (*Prunus armeniaca* L.) est une culture emblématique du Bassin Méditerranéen. Il y est soumis à des contraintes biotiques importantes, parmi lesquelles le dépérissement bactérien, causé par *Pseudomonas syringae* (*Psy*), peut mener à des phénomènes de mortalité en verger au niveau des régions à hivers froids et humides. La mise en évidence de différences variétales en verger offre potentiellement des perspectives de contrôle de la maladie à travers le levier génétique. Aussi, ce travail de thèse avait pour principaux objectifs (i) d'identifier chez la plante des régions génomiques liées à la résistance partielle à la bactérie et (ii) d'étudier un plan factoriel d'interaction entre les diversités de la plante et de la bactérie (GxG) afin d'apprécier la généricité de la résistance et sa durabilité. Afin de répondre au premier objectif, deux approches complémentaires ont été mobilisées : une cartographie de QRLs (Quantitative Resistance Loci) sur quatre populations biparentales dont trois sont issues du croisement avec un géniteur commun ainsi qu'une analyse d'association sur une core-collection. Les données phénotypiques mobilisées correspondent à des symptômes issus d'inoculations contrôlées ainsi que des notes de mortalité obtenues suite à infection naturelle en verger. Ces deux approches (analyse de liaison et d'association) ont permis de mettre en évidence 22 QRLs de résistance, parmi lesquels seuls 2 QRLs sur les chromosomes 6 et 7 colocalisent entre les deux approches. Deux régions majeures détectées en étude d'association sur les chromosomes 5 et 6 se sont révélées être en déséquilibre de liaison et contrôlent près de 26 et 43% de la variation des symptômes. Deux mécanismes complémentaires reposant sur le blocage de l'infection de *Psy* et sur la limitation de la progression locale de la bactérie dans les tissus ont été mis en évidence à travers la détection de QRLs sur les chromosomes 3, 6, 8 d'une part et 1,4 et 6 d'autre part. Le second objectif a été abordé grâce à une étude d'un plan factoriel d'interaction entre 20 accessions d'abricotier et 9 souches de *Psy*, échantillonnées d'après la connaissance de l'épidémiologie de la maladie en verger. L'analyse statistique de ce dispositif mis en œuvre à la fois en verger et en laboratoire a démontré la prédominance de l'effet du facteur souche dans la variabilité des symptômes étudiés et la très faible importance du facteur d'interaction GxG, indiquant une potentielle généricité des facteurs de résistance et des perspectives favorables à leur durabilité en verger.

Les résultats issus de cette thèse contribuent à offrir une meilleure compréhension des mécanismes de résistance partielle au dépérissement bactérien de l'abricotier et fournissent des marqueurs et haplotypes, potentiellement mobilisables dans le cadre de programmes d'innovation variétale.

**Mots-clés :** Abricotier, *Pseudomonas syringae*, Quantitative Resistance Loci, durabilité des résistances, analyse de liaison, analyse d'association

## Abstract

Within the genus *Prunus*, that contains highly valuable species, apricot (*Prunus armeniaca* L.) is an emblematic Mediterranean crop. But apricot cultivation is constrained by many biotic stresses, among which bacterial canker caused by *Pseudomonas syringae* (*Psy*) is particularly severe and can lead to the death of the trees in regions with humid and cold winters. Differences of susceptibilities have been observed between cultivars in orchards and create opportunities for disease management through genetic improvement. This thesis aimed to (i) identify genetic determinants linked to partial resistance to the bacterium and to (ii) study a factorial interaction design between both diversities of the plant and the pathogen (GxG interaction) in order to assess resistance genericity and durability. With regard to the first objective, two complementary approaches were used : QRL (Quantitative Resistance Loci) mapping over four biparental progenies, among which three were obtained with a cross involving a common genitor, and a genome-wide association study on a core-collection. The phenotypic data mobilized in this work rely on symptoms issued from controlled inoculations and on mortality notations following natural infections in the orchard. These approaches led to the detection of 22 QRLs among which only 2 QRLs, located on chromosomes 6 and 7, co-localized between the two methods. Two main regions detected in the association study, over the chromosomes 5 and 6, appeared to be in linkage disequilibrium and controlled 26 and 43% of the variation of the symptoms. A complementarity between two mechanisms, one that involves blocking the infection of *Psy* and the other that limits bacterial mobility in the tissues has been revealed through the detection of QRLs over chromosomes 3, 6, 8 for one mechanism and 1,4, 6 for the other, respectively. The second objective was fulfilled with a study of a factorial interaction design between 20 apricot accessions and 9 *Psy* strains, which were sampled according to the previous knowledge of the disease epidemiology in the orchard. Statistical analyses of phenotypic data obtained both from the orchard and a laboratory test showed a clear predominance of the strain effect on symptom variability and a weak importance of the GxG interaction factor. This last result highlighted a potential genericity of the resistance factors and favorable perspectives of durability in the orchard. The results issued from this thesis contribute to a better understanding of the mechanisms underlying partial resistance of apricot to bacterial canker. Moreover, it provide markers and haplotypes of interest which could be mobilized in breeding programs.

**Keywords :** apricot, *Pseudomonas syringae*, Quantitative Resistance Loci, resistance durability, linkage mapping, genome-wide association study

## Liste des publications et communications scientifiques

Omrani, M., Roth, M., Roch, G., Blanc, A., Morris, C.E., Audergon, J.M. (submitted publication), 'Genome-wide association multi-locus and multi-variate linear mixed models reveal two linked loci with major effects on partial resistance of apricot to bacterial canker', BMC Plant Biology

Omrani, M. (Présentateur, Oral+Poster), Roch, G., Blanc, A., Morris, C.E., Audergon, J.M. (2018). 'Genome-wide association study of partial resistance to bacterial canker of apricot'. Presented at PAG XXVI - Plant and Animal Genome Conference, San Diego, USA (2018-01-13 - 2018-01-17)

Omrani, M. (Présentateur, Oral), Roch, G., Blanc, A., Morris, C.E., Audergon, J.M. (2017). 'Genome-wide association study of partial resistance to bacterial canker of apricot'. Presented at Journées Jeunes Chercheurs du département BAP 2017, Bordeaux, France (2017-04-20 - 2017-04-21)

Omrani, M. (Présentateur, Poster), Confolent, C., Clauzel, G., Brun, L., Parisi, L., Chandeysson, C., Martin, E., Ferreol, A.M., Leonetti, J., Lambert, P., Roch, G., Morris, C.E., Audergon J.M., (2016). 'Characterization of genetic determinants of resistance to bacterial canker in Apricot' Presented at Journées Jeunes Chercheurs du département BAP 2016, Lyon, France (2016-04-11 - 2016-04-12)

N'sibi, M. (Présentateur, Poster), Confolent, C., Lambert, P., Bureau, S., Gouble, B., Blanc, A., Roch, G., Omrani, M., Flutre, T., Sauvage, C., Audergon, J. M. (2017). Sélection génomique - Quelles perspectives chez les Prunus ? Presented at Séminaire SelGen 2017 : "La sélection génomique, bilan et perspectives", Paris, FRA (2017-09-28 - 2017-09-29)

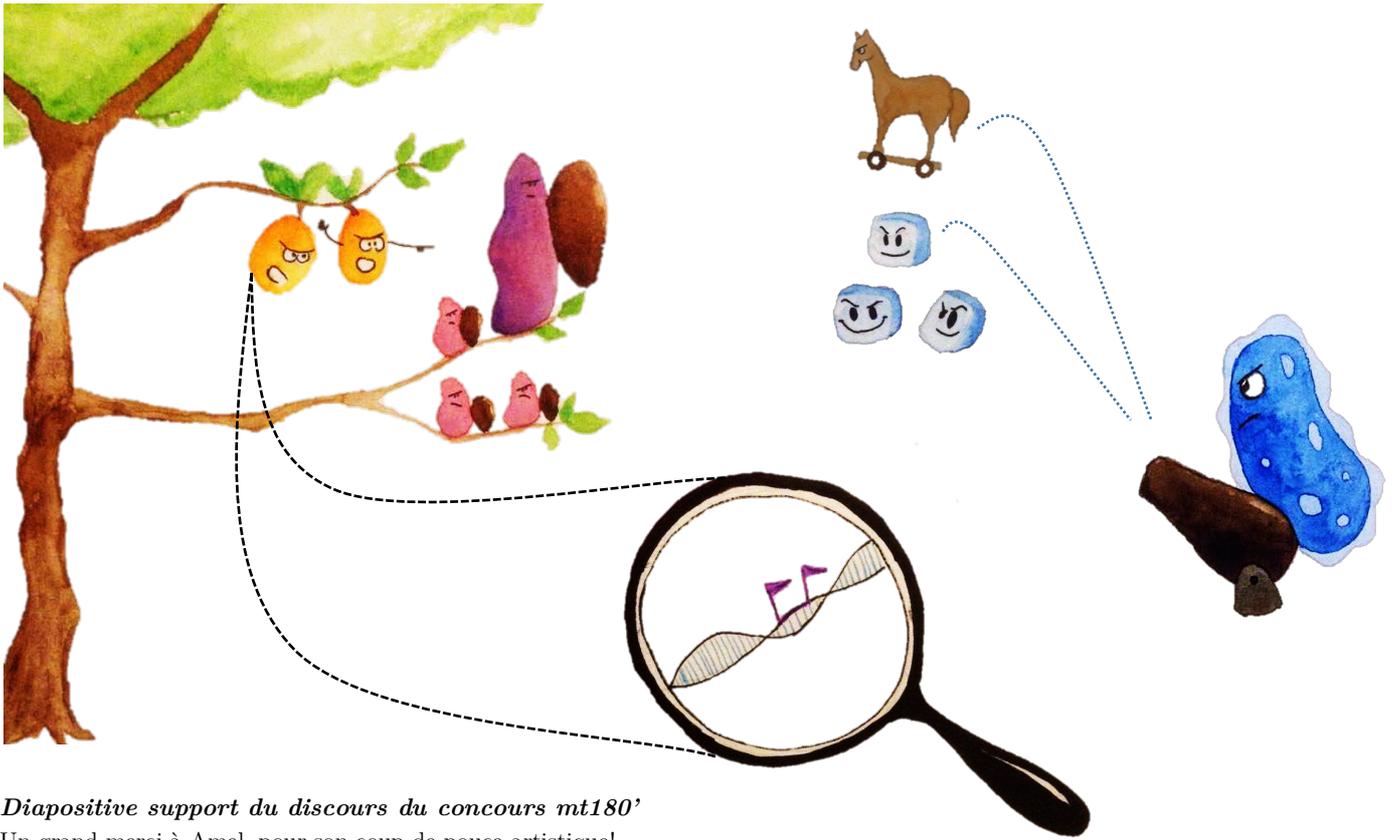
## Autres communications orales

Séminaire CTPS de restitution des projets CASDAR Semences 2015 - Paris, Octobre 2017 : Synthèse du projet ResiBac 'Prédiction de l'apparition du chancre bactérien (*Pseudomonas syringae*) chez les arbres fruitiers – preuve de concept'

Participation à la formation et au concours 'ma thèse en 180 secondes' - Marseille, avril 2017 : Finaliste de l'université d'Avignon et des Pays du Vaucluse à l'échelle de la finale du regroupement régional Aix-Marseille-Toulon-Avignon.

## Rapport

Audergon, J.M., Omrani, M., Morris, C.E., 2018 - Rapport final projet CASDAR : 'RésiBac - Prédiction de l'apparition du chancre bactérien (*Pseudomonas syringae*) chez les arbres fruitiers – Preuve de concept' – 35p + 6 annexes (synthèse vidéo : <https://www.youtube.com/watch?v=8p7rWDBj5fE>).



*Diapositive support du discours du concours mt180'*  
Un grand merci à Amel, pour son coup de pouce artistique!

**Texte déclamé lors de finale régionale du regroupement Aix-Marseille-Avignon-Toulon du concours « Ma thèse en 180 secondes » organisé par la Conférence des Présidents d'Université et le CNRS**

Le 28/04/2017 à Marseille

*« Si je vous dis, 'course aux armements', vous allez me répondre guerre froide. Et vous auriez tout à fait raison ! Pourtant mes recherches n'ont absolument rien à voir avec la géopolitique du XXe siècle mais portent sur la biologie végétale.*

*En fait le concept est également valable quand on étudie les couples formés entre un agent pathogène et une plante. C'est le principe de coévolution autrement dit de 'je ne suis pas assez armé'. Moi, mon couple préféré c'est l'abricotier et une bactérie du nom de *Pseudomonas syringae*, responsable d'une maladie mortelle !... Le chancre bactérien. Et au cours de leur histoire de couple vieille de milliers d'années, ces deux-là ont pris le temps de se connaître, de se titiller, voire de prendre le dessus sur l'autre ! Chacun a développé des armes pour se défendre et riposter aux attaques de son adversaire.*

*Pseudomonas que vous voyez en bleu, a un panel d'armes très diversifié. Entre autres, un système de sécrétion hyper élaboré qui agit comme un cheval de Troie en catapultant des protéines parasites (appelées effecteurs) à l'intérieur des cellules de la plante. Et le plus intéressant est à venir : c'est le pouvoir glacogène ! Cela permet à la bactérie de faciliter la prise en glace des tissus de la plante à des températures de 4°C supérieures à la normale. Avec l'alternance des gels/dégels en hiver, BIM !!! les cellules explosent et hop ! la bactérie en profite pour se faufiler et proliférer chez son hôte. De son côté, l'abricotier a acquis des gènes de résistance...et oui c'est une histoire d'amour unilatérale vous l'aurez compris... Dans certains cas, ils sont nombreux à petit effet, dans d'autres, ils sont plus rares mais pèsent plus lourd.*

*Et là, vous vous demandez peut-être où j'interviens au beau milieu de cette histoire rocambolesque !*

*Et bien, mon travail c'est justement d'identifier ces gènes de résistance et notamment leur localisation sur le matériel porteur de l'information génétique : l'ADN. Je fais ce qu'on appelle de la cartographie génétique. En connaissant la place exacte de ces gènes sur la carte que constitue l'ADN, je peux les suivre grâce à des espèces de balises auxquelles ils sont liés que l'on nomme des marqueurs génétiques. Ces balises peuvent ensuite servir dans des programmes d'amélioration des variétés.*

*Et là je vous vois venir... non je ne travaille pas sur le développement d'OGM ! On peut également faire de l'amélioration génétique en effectuant des croisements entre abricotiers de qualité mais sensibles et des très mauvais mais résistants. Parmi les descendants, on fait ensuite un tri à l'aide des balises pour ne garder que les abricotiers qui auront cumulé les deux caractères !*

*Oui aujourd'hui je peux le dire : les abricots français sont menacés : mais qui sait bientôt, j'espère bien devenir un vrai docteur...des abricotiers ! »*



## Abréviations

ABA : Acide abscissique  
ACP : Analyse en composantes principales  
ADN : Acide désoxyribonucléique  
ANOVA : Analyse de variance  
*avr*/*Avr* : gène/protéine d'avirulence  
BC : 'Backcross' = rétro-croisement  
bi : Brunissement interne  
BLAST : 'Basic local alignment search tool'  
BLUP: 'Best linear unbiased predictor'  
Bp ou pb : Paire de bases  
bs : Brunissement superficiel  
CIM : Cartographie d'intervalle composite  
cM : centiMorgan  
DL : Déséquilibre de liaison  
ETI : 'Effector triggered immunity'  
ETS : 'Effector triggered susceptibility'  
FDR : 'False-discovery rate'  
GBS : 'Genotyping by sequencing'  
GLM : Modèle linéaire généralisé  
GWAS : 'Genome-wide association study'  
GxE : Interaction génotype x environnement  
GxG : Interaction génotype (plante) x génotype (pathogène)  
GxY : Interaction génotype (plante) x année  
 $H^2$  : Héritabilité au sens large  
 $h^2$  : Héritabilité au sens strict  
HR : Réaction d'hypersensibilité  
Hrc : 'Hypersensitive reaction and pathogenicity protein conserved'  
Hrp : 'Hypersensitive reaction and pathogenicity protein'  
*ina*: 'ice nucleation activity'  
JA: Acide jasmonique  
Kb ou kbp : kilobase ou kilo paires de base  
lgc: Longueur de chancre  
lge: Longueur de la lésion sub-corticale  
LOD : 'Logarithm of the odds'  
MAF : Fréquence de l'allèle mineur  
Mb ou Mbp : megabase ou mega paires de base  
ML : Maximum de vraisemblance  
MLMM : 'Multi-locus linear mixed model'  
MQM : Modèle à QTL multiples  
mvLMM : 'Multi-variate linear mixed model'  
NBS-LRR : 'Nucleotide-binding site of leucine-rich repeats'  
NGS : 'Next generation sequencing'

MAMP : 'Microbe-associated molecular pattern'  
MAPK: 'Mitogen-activated protein kinase'  
miRNA: Micro-ARN  
ORF: 'Open-reading frame'  
PG : Phylogroupe  
*P.* : *Pseudomonas*  
*Psa* : *Pseudomonas syringae* pv. *actinidiae*  
*Psm* : *Pseudomonas syringae* pv. *morsprunorum*  
*Pss* : *Pseudomonas syringae* pv. *syringae*  
*Psy* : *Pseudomonas syringae*  
*Pto*: *Pseudomonas syringae* pv. *tomato*  
PTI : 'Pathogen triggered immunity'  
pv. : pathovar  
PVE : Pourcentage de variance expliquée  
QQ-plot : 'Quantile-quantile plot'  
QTL/QRL : 'Quantitative trait/resistance loci'  
Gène *R* : Gène de résistance  
REML : Estimation du maximum de vraisemblance restreint  
RIL : 'Recombinant inbred line'  
RLK/RLP : 'Receptor-like-kinase/protein'  
T3SS : Système de sécrétion de type III  
SA : Acide salicylique  
SAM : Sélection assistée par marqueurs  
SDN : Stimulateur de défenses naturelles  
SIM : Cartographie d'intervalle simple  
SNP : 'Single nucleotide polymorphism'  
SSR : 'Single sequence sepeat' = microsatellite

# TABLE DES MATIÈRES

<i>Préambule</i>	<b>1</b>
<i>Partie I - Bibliographie</i>	<b>3</b>
Chapitre 1. Cartographie des caractères quantitatifs : concepts et méthodologies	<b>5</b>
1.1. Analyse de liaison: cartographie génétique au sein des descendance	7
1.2. Analyse d'association : étude de la diversité naturelle de l'espèce	10
1.3. Dispositifs multi-parentaux	15
Chapitre 2. Résistances : déterminisme génétique, moléculaire et durabilité	<b>18</b>
2.1. Résistance non-spécifique ou non-hôte	19
2.2. Résistance hôte-spécifique	23
2.3. Modèles évolutifs des interactions plante x pathogène et contournement des résistances	27
2.4. Résistances quantitatives : déterminisme génétique	30
2.5. Durabilité des résistances : mécanismes et stratégies de déploiement	33
Chapitre 3. Pathosystème abricotier x <i>Pseudomonas syringae</i> ( <i>Psy</i> )	<b>38</b>
3.1. Abricotier	38
3.1.1. Origine, diversité naturelle et cultivée	38
3.1.2. Caractéristiques taxonomiques, botaniques et biologiques	42
3.1.3. Importance économique	44
3.1.4. Ressources génomiques et moléculaires	45
3.1.5. Objectifs et outils de l'amélioration variétale	46
3.1.5.1. Sélection Assistée par Marqueurs (SAM)	46
3.1.5.2. Cibles des programmes de sélection chez l'abricotier	48
3.1.5.2.1. Auto-fertilité	48
3.1.5.2.2. Phénologie et adaptation aux conditions climatiques	48
3.1.5.2.3. Qualité	49
3.1.5.2.4. Résistance aux maladies	49
3.2. Complexe d'espèces <i>Pseudomonas syringae</i>	51
3.2.1. Biologie, diversité et écologie	51
3.2.2. Composantes génétiques du pouvoir pathogène	56
3.2.3. Evolution et patho-adaptation	58
3.2.4. Bilan : modèle du mécanisme d'infection de <i>Psy</i>	61
3.3. Dépérissement bactérien de l'abricotier	65
3.3.1. Symptômes	65
3.3.2. Epidémiologie	67
3.3.3. Impact économique	69
3.3.4. Facteurs pédoclimatiques influençant le développement de la maladie	70
3.3.5. Lutte existante	70
<i>Contexte et objectifs de la thèse</i>	<b>73</b>
<i>Partie II - Matériel &amp; Méthodes</i>	<b>75</b>
1. Matériel végétal	77

2. Matériel bactérien	79
3. Données génotypiques et cartes génétiques	79
4. Design expérimental : protocole d'inoculations contrôlées	80
4.1. En verger : conditions in-vivo	80
4.1. En laboratoire : conditions in-vitro	82
5. Données phénotypiques	82
<i>Partie III- Analyse de la diversité naturelle de la sensibilité variétale à Psy à travers une analyse d'association</i>	<b>85</b>
<i>Partie IV- Cartographie de la résistance à Psy par analyse de liaison au sein d'une descendance pseudo-F1 et de trois descendances interconnectées de 2<sup>ème</sup> génération</i>	<b>131</b>
<i>Partie V - Caractérisation de l'interaction abricotier x Psy</i>	<b>153</b>
Introduction	<b>154</b>
Matériel & Méthodes	<b>155</b>
1. Plan factoriel d'interaction GxG	155
1.1. Conditions in-vivo en verger	156
1.2. Conditions in-vitro en laboratoire	156
2. Analyses statistiques	157
3. Étude de la migration bactérienne dans les tissus	158
Résultats et Discussion	<b>159</b>
<i>Partie VI - Discussion et perspectives</i>	<b>167</b>
1. Architecture génétique de la résistance partielle de l'abricotier au dépérissement bactérien	<b>168</b>
2. Comparaison des approches de cartographie : convergences et particularités	<b>170</b>
3. Regard critique et perspectives de recherches futures	<b>171</b>
4. Applications pour la sélection variétale	<b>173</b>
5. Vers une résistance durable des vergers à <i>Psy</i> ?	<b>176</b>
<i>Références bibliographiques</i>	<b>178</b>
<i>Table des annexes</i>	<b>209</b>

## TABLE DES FIGURES ET TABLEAUX

FIGURE 1 : COMPARAISON SCHEMATIQUE DE L'ANALYSE DE LIAISON ET L'ANALYSE D'ASSOCIATION.-----	8
FIGURE 2 : DIAGRAMME SCHEMATIQUE DU CONTRASTE ENTRE GWAS (DEMARCHE SANS A PRIORI SUR LES GENES CAUSAUX) ET LA DEMARCHE GENE-CANDIDATS (SELECTION PREALABLE DE GENES CANDIDATS SUITE A ANALYSE DE LIAISON, ANALYSES BIOCHIMIQUES OU MUTATIONNELLES).-----	10
FIGURE 3 : ILLUSTRATION D'UN EFFET D'ARTEFACT EN GWAS LIE A LA STRUCTURE ENTRE DEUX SOUS-POPULATIONS.-----	12
FIGURE 4 : EXEMPLE DE PEDIGREE D'UN DISPOSITIF DE 5 FAMILLES MULTI-PARENTALES (EN NOIR) CHEZ LE POMMIER.-----	16
FIGURE 5 : REPRESENTATION SCHEMATIQUE DES DIFFERENTS TYPES D'INTERACTION PLANTE X PATHOGENE.-----	18
FIGURE 6 : REPRESENTATION SCHEMATIQUE DES REACTIONS DE DEFENSE DES PLANTES.--	20
FIGURE 7 : REPRESENTATION SCHEMATIQUE DES INTERCONNEXIONS ENTRE DIFFERENTES VOIES DE SIGNALISATION DES PHYTOHORMONES EN REPONSE A <i>PSEUDOMONAS SYRINGAE</i> ( <i>Psy</i> ).-----	22
FIGURE 8 : ILLUSTRATION DU MODELE DE FLOR DE L'INTERACTION GENE-POUR-GENE.---	24
FIGURE 9 : REPRESENTATION SCHEMATIQUE DU MODELE DE GARDE SIMPLE (A) ET AVEC PLUSIEURS PROTEINES CIBLES (B) ET DU MODELE DU LEURRE (C).-----	25
FIGURE 10 : REPRESENTATION SCHEMATIQUE DE LA MODELISATION DES REACTIONS DE DEFENSE DES PLANTES PAR DES RESEAUX DE RECEPTEURS.-----	26
FIGURE 11 : REPRESENTATION SCHEMATIQUE DU MODELE EVOLUTIF EN 'ZIG-ZAG' DES INTERACTIONS PLANTE X PATHOGENE.-----	27
FIGURE 12 : REPRESENTATION SCHEMATIQUE DE LA COMPARAISON ENTRE LES MODELES EVOLUTIFS EN ZIG-ZAG (A) ET D'INVASION (B).-----	29
FIGURE 13 : ORIGINE ET DISSEMINATION DE L'ABRICOTIER A PARTIR DES PREMIERS CENTRES DE DOMESTICATION.-----	39
FIGURE 14 : REPRESENTATION D'UN ARBRE BAYESIEN ILLUSTRANT LES RELATIONS PHYLOGENETIQUES ENTRE <i>PRUNUS</i> .-----	41
FIGURE 15 : PHOTOGRAPHIES DE VERGER EN PERIODE DE FLORAISON (A), DE RAMEAUX A FLEURS (B ET C) ET DE FRUITS D'ABRICOTIER (D ET E).-----	43
FIGURE 16 : CARTE GENETIQUE DE REFERENCE (TxE) DES <i>PRUNUS</i> AVEC LA LOCALISATION APPROXIMATIVE DE 28 GENES MAJEURS SUR LES 8 CHROMOSOMES.-----	47
FIGURE 17 : CLASSIFICATION DU COMPLEXE D'ESPECES <i>Psy</i> .-----	53
FIGURE 18 : REPRESENTATION SCHEMATIQUE DE L'ECOLOGIE DE <i>Psy</i> A L'ECHELLE DU PAYSAGE.-----	55
FIGURE 19 : REPRESENTATION SCHEMATIQUE DE L'ORGANISATION ET DE LA VARIABILITE GENETIQUE DE L'ILOT DE PATHOGENICITE <i>HRP</i> CHEZ <i>Psy</i> .-----	57
FIGURE 20 : REPRESENTATION SCHEMATIQUE DES MECANISMES D'INFECTION DE <i>Psy</i> A L'ENTREE DE LA FEUILLE VIA LES STOMATES (A) ET AU NIVEAU DES CELLULES DU MESOPHYLLE (B ET C).-----	62
FIGURE 21 : PHOTOGRAPHIES DE SYMPTOMES DE DEPERISSEMENT BACTERIEN SUR ABRICOTIER.-----	66
FIGURE 22 : CYCLE EPIDEMIOLOGIQUE DU DEPERISSEMENT BACTERIEN DES ARBRES FRUITIERS CAUSE PAR <i>Psy</i> .-----	68

FIGURE 23 : PHOTOGRAPHIES DES DISPOSITIFS VEGETAUX EXPLOITES SUITE A INOCULATION DE <i>PSY</i> . -----	77
FIGURE 24 : CARTOGRAPHIE ET REPARTITION DE L'ORIGINE GEOGRAPHIQUE DES 73 ACCESSIONS DE LA CORE-COLLECTION. -----	78
FIGURE 25 : PHOTOGRAPHIES DU MODE OPERATOIRE DES INOCULATIONS ARTIFICIELLES EN VERGER. -----	81
FIGURE 26 : PHOTOGRAPHIES DES PHENOTYPES ETUDIES ISSUS DES EXPERIMENTATIONS D'INOCULATIONS CONTROLEES SUR RAMEAUX EN VERGER (A) SUR ET SUR BOUTURES EN LABORATOIRE (B). -----	83
FIGURE 27 : BOXPLOTS REPRESENTANT LA DISTRIBUTION DE L'ETENDUE DES SYMPTOMES (BLUP LGC ET BS) ISSUS DU PHENOTYPAGE DE VALIDATION EN 2018 EN FONCTION DES GENOTYPES AUX MARQUEURS CANDIDATS LG5_5394803, LG3_10897844, LG5_4842835. -----	129
FIGURE 28: PHOTOGRAPHIES DU PROTOCOLE GxG REALISE EN CONDITIONS CONTROLEES IN-VITRO. -----	156
FIGURE 29 : SCHEMA DE LA SEQUENCE TEMPORELLE D'INCUBATION DES BOUTURES EN CONDITIONS IN-VITRO DANS LE CADRE DU DISPOSITIF D'ETUDE GxG. -----	157
FIGURE 30: DISTRIBUTION ET CORRELATIONS ENTRE LES VARIABLES ETUDIEES DANS LE DISPOSITIF FACTORIEL GxG - LA NECROSE SUPERFICIELLE MESUREE EN CONDITIONS IN-VITRO ET LES VARIABLES LGC ET BS OBTENUES EN VERGER EN 2017 ET 2018. ----	160
FIGURE 31: GRAPHIQUES BIDIMENSIONNELS REPRESENTANT LA REPARTITION DE L'EFFET G SOUCHE (A) ET DE L'EFFET G ABRICOTIER (B) AU NIVEAU DE LA NECROSE SUPERFICIELLE, MESUREE EN CONDITIONS IN-VITRO, ET LGC ET BS, MESURES OBTENUES EN VERGER EN 2017 ET 2018. -----	161
FIGURE 32 : STRUCTURATION DE L'EFFET DE MODULARITE EN 2 OU 4 MODULES DETECTES SUR LA VARIATION DE LGC 2018 AU NIVEAU DU PLAN FACTORIEL D'INTERACTION EN VERGER 20 ACCESSIONS X 9 SOUCHES. -----	164
TABLEAU 1 : LISTE DES 23 GENES DE RESISTANCE <i>R</i> A <i>PSY</i> CLONES. ....	64
TABLEAU 2 : SYNTHESE DU MATERIEL VEGETAL, BACTERIEN, DES DONNEES PHENOTYPIQUES ET GENOTYPIQUES EXPLOITEES AU COURS DU TRAVAIL DE THESE. ....	76
TABLEAU 3 : CARACTERISTIQUES DES 9 SOUCHES EXPLOITEES LORS DU DISPOSITIF D'INOCULATIONS GxG .....	155
TABLEAU 4: RESULTATS DES ANOVA TESTANT LES EFFETS GENOTYPE PLANTE (G), SOUCHE (S), OPERATEUR, ANNEE (Y), ET LES DIVERS EFFETS D'INTERACTION G :Y, G :S, Y :S, G :Y :S AU NIVEAU DES DISPOSITIFS FACTORIELS GxG DEPLOYES EN VERGER ET EN LABORATOIRE.....	159
TABLEAU 5 : RELATION ENTRE LA PRESENCE ET LA MIGRATION DE <i>PSY</i> SOUCHE 41A ET L'APPARITION DE LA NECROSE SUPERFICIELLE AU NIVEAU DE TISSUS PRELEVES CHEZ 3 GENOTYPES – BAKOUR ET COLOMER (RESISTANTS EN VERT) ET BERGERON (SENSIBLE EN ROUGE). ....	162

# Préambule

Cette thèse vise à caractériser la variabilité génétique de l'abricotier quant à la résistance au dépérissement bactérien à travers l'intégration de données phénotypiques de sensibilité et de données génotypiques via des analyses de liaison et d'association. L'interaction de la diversité de l'hôte avec celle présentée par l'agent pathogène *Pseudomonas syringae* (*Psy*) est également abordée en lien avec la question de la durabilité des résistances identifiées.

Le manuscrit est composé de six parties :

La **partie I** présente une revue des connaissances liées au contexte de l'étude. Un premier chapitre décrit les méthodologies et concepts de la cartographie des caractères quantitatifs qui ont pu être mobilisées au cours de cette thèse. Puis, un deuxième chapitre explore l'état de l'art concernant le déterminisme génétique de l'interaction plante x pathogène (GxG); la question de la durabilité des résistances sera également abordée à travers l'évolution des interactions GxG. Enfin, le pathosystème abricotier x *Psy*, modèle d'étude de cette thèse est présenté.

La **partie II** récapitule l'ensemble du matériel végétal, bactérien exploité, et les méthodes employées. Ces points seront ensuite abordés en détail au cours des chapitres suivants.

La **partie III** est sous la forme d'un article soumis à *BMC Plant Biology*. L'article présente une étude d'association (GWAS) visant à identifier des QTL et gènes candidats sous-jacents à la résistance partielle à *Psy* au sein d'une collection de 73 accessions. Les données de sensibilité ont été recueillies au cours de 4 campagnes de mesure liées à des inoculations contrôlées sur rameaux en verger.

La **partie IV** est sous la forme d'un article en cours de préparation (projet de soumission à *Molecular Breeding*). Les travaux consistent en une analyse de liaison au sein d'un dispositif multiparental constitué d'une population biparentale et de 3 descendances interconnectées de 307 individus (3 hybrides F1 x géniteur commun). Les données de sensibilité sont issues respectivement, de 3 et 2 répétitions expérimentales d'inoculations contrôlées sur rameaux et de notations de mortalité en verger.

La **partie V** est sous forme manuscrite et présente l'étude des interactions GxG abordées dans le cadre de dispositifs à la fois d'inoculations sur (a) organes détachés en laboratoire et (b) rameaux en verger. La progression de la migration de la bactérie dans les tissus a été suivie sur plusieurs variétés (dispositif a). Puis, l'analyse statistique d'un plan factoriel d'interaction 20 génotypes abricotier x 9 souches *Psy* (dispositif b) nous permet de quantifier l'importance du facteur d'interaction GxG face aux facteurs génétiques individuels.

La **partie VI** propose une synthèse et discussion des principaux résultats puis dresse une conclusion et les perspectives issues de cette thèse.



# Partie I - Bibliographie



# Chapitre 1. Cartographie des caractères quantitatifs : concepts et méthodologies

Le champ d'étude de la génétique quantitative correspond à l'identification des régions génomiques responsables de la variation continue d'un caractère d'intérêt autrement appelées Quantitative trait loci (QTLs). Les caractères à variation continue tels que le rendement, le poids du fruit et les résistances à certaines maladies résultent souvent de l'association de plusieurs gènes à faible effet et de leur interaction avec l'environnement. Ces caractères sont alors souvent qualifiés de « complexes » à l'opposé des caractères « simples » (compatibilité pollinique, pilosité de l'épiderme des fruits), à ségrégation mendélienne dont l'architecture génétique repose très souvent sur un nombre limité de gènes à effet majeur.

A la fin des années 80, le développement des marqueurs génétiques a permis de visualiser les polymorphismes entre individus à l'échelle de l'ADN, et ainsi mettre en évidence la source de variation à l'échelle du phénotype. Un marqueur génétique est défini comme une zone du génome (locus) se présentant sous différentes formes (allèles) au sein d'une population. Les marqueurs ne représentent généralement pas directement les gènes étudiés mais étant localisés à proximité, ils sont très souvent hérités ensemble et fonctionnent comme des 'balises' de ces gènes. Le génotypage de ces marqueurs donne l'information du génotype pour chaque individu à un endroit bien précis du génome.

Les premiers marqueurs génétiques correspondaient à des marqueurs morphologiques et biochimiques (isozymes), qui en raison de leur faible fréquence sur le génome et de leur forte interaction avec l'environnement, ont été rapidement délaissés au profit des marqueurs moléculaires plus répétables : Restriction Fragment Length Polymorphisms (RFLPs), Random Amplified Polymorphic DNA (RAPDs), microsatellites (Single Sequence Repeat - SSRs), Amplified Fragment Length Polymorphisms (AFLPs) et enfin Single Nucleotide Polymorphisms (SNPs) (Collard et al., 2005).

Les marqueurs SNPs correspondent à des polymorphismes liés à la substitution, l'insertion ou la délétion d'un nucléotide. Ils présentent l'avantage d'être très nombreux sur le génome. Du fait de leur coût raisonnable et de la facilité de détection avec les moyens technologiques actuels, ils ont remplacé les SSRs (répétition de séquences très courtes), onéreux et moins nombreux. Les SSRs présentent néanmoins l'avantage d'être très polymorphes et transférables entre espèces, justifiant leur recours dans les études de caractérisation de la diversité génétique au niveau intra (Gasic et al., 2009) et inter-spécifique (Dirlewanger et al., 2004).

Avec l'avènement des nouvelles technologies de séquençage (**NGS**) et le faible coût associé au génotypage de populations de plusieurs centaines ou milliers d'individus, l'identification de milliers de SNPs sur génome complet a ouvert des perspectives nouvelles pour l'identification de l'architecture génétique des caractères quantitatifs (Davey *et al.*, 2011).

La cartographie de QTLs consiste en la recherche de corrélations sur l'ensemble du génome entre la variabilité du phénotype et le polymorphisme au marqueur au sein d'une population d'individus apparentés (cartographie par **analyse de liaison**) ou génétiquement plus éloignés (cartographie par analyse d'association ou **Genome Wide Association Study** - GWAS).

Bien que le principe soit identique dans ces deux types d'approches, les populations d'étude se distinguent sur leur structure génétique et leur niveau de diversité allélique (détaillés en *parties 1.1 et 1.2*). La conjonction de ces deux approches (que l'on abordera en *partie 1.3*) à travers les dispositifs des **familles multi-parentales** permet de tirer les avantages de chacune et d'en atténuer les limites à travers une validation croisée.

En termes de modélisation, l'approche quantitative se base sur des modèles considérant l'action d'un grand nombre de gènes, expliquant individuellement une faible partie de la variance du caractère. Ainsi le **modèle infinitésimal** de Fisher en 1918 partitionne la variance d'un caractère en une contribution génétique liée à un nombre important de loci ayant des effets additifs, de dominance (interaction entre allèles d'un locus), d'épistasie (interaction entre allèles de loci différents), et d'une contribution non-génétique (Fisher, 1919).

Ce modèle a pu notamment introduire la notion d'**héritabilité** (au sens-large)  $H^2$  comme la part de variance phénotypique attribuable à la variation génétique. Dans le cas d'une descendance, la valeur de l'héritabilité détermine alors la relation entre la valeur d'un caractère chez un parent et la distribution de ce même caractère au niveau de sa descendance. Plus l'héritabilité d'un caractère est élevée et proche de 1, plus l'influence de l'environnement est négligeable et plus la réponse à la sélection est efficace.

## 1.1. Analyse de liaison: cartographie génétique au sein des descendance

L'analyse de liaison exploite des populations de ségrégation issues du croisement entre individus choisis en fonction de leurs aptitudes opposées par rapport au phénotype étudié. Il s'agit d'établir un lien entre origine parentale de l'allèle et variabilité du phénotype grâce à l'information donnée par les marqueurs génétiques au niveau des descendants.

La cartographie de QTLs se base sur le principe que les gènes et les marqueurs ségrégent au cours de la **recombinaison** des chromosomes ('crossing-over') durant la méiose ce qui permet d'analyser la part des génotypes recombinants par rapport aux parentaux. Les gènes ou marqueurs proches (liés) présentent une plus forte probabilité d'être transmis ensemble à la descendance que ceux qui sont plus éloignés entre eux. Le calcul d'une fréquence de recombinaison permet ainsi d'inférer une **distance génétique** entre marqueurs estimée en centimorgans (cM).

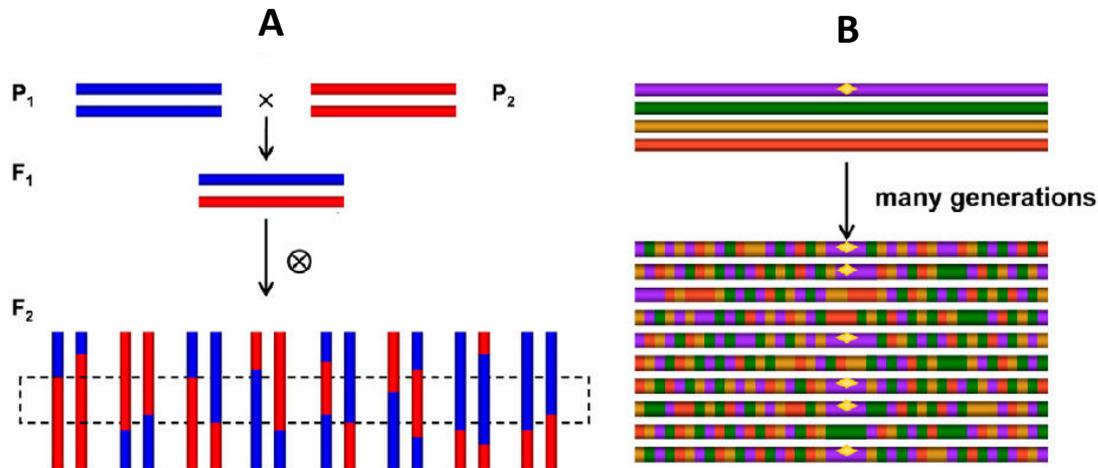
Plusieurs étapes sont nécessaires à l'analyse de liaison :

- (1) Obtention d'une population biparentale issue du croisement entre individus aux aptitudes contrastées par rapport au phénotype étudié
- (2) Recherche de marqueurs polymorphes entre les parents et génotypage de la population
- (3) Etablissement d'une **carte génétique** : agencement des marqueurs selon des groupes de liaison et calcul des distances génétiques entre marqueurs
- (4) Phénotypage de la population par rapport au caractère étudié
- (5) Cartographie de QTL : analyses statistiques de corrélation marqueur – phénotype.

Les populations de cartographie chez les espèces autogames sont très souvent issues du croisement entre deux lignées parentales homozygotes ('inbred') puis des descendants F1 (*Figure 1A*): back-cross BC (rétro-croisement d'un hybride F1 avec un des parents), Recombinant Inbred Lines (RIL) (autofécondations successives sur  $n$  générations d'individus  $F_n$  jusqu'à obtenir des lignées recombinantes homozygotes fixes) (Collard *et al.*, 2005). L'homozygotie peut également résulter de techniques d'haplo-diploïdisation afin d'obtenir des lignées haploïdes doublées. En considérant les espèces allogames, notamment les espèces pérennes très fortement hétérozygotes et pour lesquelles les temps de génération sont plus longs, les populations de cartographie peuvent être établies dès la 1<sup>ère</sup> génération de croisement (dit '**outbred**'). Les cartes génétiques sont alors souvent issues d'une stratégie '**pseudo-testcross**' (Grattapaglia and Sederoff, 1994): si un locus est hétérozygote chez un parent P1 et homozygote

chez l'autre parent P<sub>2</sub>, le locus ségrège au niveau de la descendance avec une proportion 1:1.

Les cartes génétiques sont alors établies de façon indépendante pour chaque parent avec les marqueurs hétérozygotes du parent correspondant.



**Figure 1 : Comparaison schématique de l'analyse de liaison et l'analyse d'association.**

(Zhu *et al.*, 2008)

- (A) l'analyse de liaison : exploitation de populations biparentales aux recombinaisons limitées, et
- (B) l'analyse d'association : collections diverses ayant accumulé de multiples évènements de recombinaison. Le locus fonctionnel en jaune est en faible déséquilibre de liaison (DL) avec les marqueurs moléculaires sauf ceux à proximité immédiate.

Diverses méthodes statistiques ont été développées afin d'évaluer la corrélation entre marqueur et phénotype. La méthode la plus simple correspond à une régression simple ou ANOVA (ANalysis Of VAriance) indiquant une potentielle détection par la comparaison des moyennes phénotypiques entre les différentes classes génotypiques au marqueur. Cette méthode bien que simple à utiliser présente un défaut majeur : son incapacité à estimer simultanément la position et l'effet du QTL en considérant les marqueurs pris individuellement (Doerge, 2002).

Ce problème a été levé grâce à la connaissance de l'ordre des marqueurs à travers la constitution d'une carte génétique et au développement de la méthode de cartographie par intervalle simple (SIM – '**Single Interval Mapping**') (Lander and Botstein, 1989). Il s'agit d'une procédure testant la présence d'un QTL dans un intervalle incrémenté par un pas (par exemple 2 cM) entre deux marqueurs consécutifs. A chaque pas, le rapport entre la vraisemblance du modèle avec QTL et celui d'hypothèse nulle (pas de QTL) est calculé et un **LOD** (logarithm of the odds) score est établi comme fonction de ce rapport. La position la plus probable du QTL est alors celle qui maximise la vraisemblance. Un autre avantage de cette méthode réside également dans

la possibilité d'inférer les génotypes manquants connaissant les marqueurs flanquants et la fréquence de recombinaison. En revanche, ce modèle ne considère que l'existence d'un unique QTL et ne permet pas de tester les phénomènes d'interaction entre QTLs qui sont pourtant vraisemblables à l'échelle du génome. Ces inconvénients ont été affranchis avec l'implémentation liée à la méthode de cartographie d'intervalle composite (CIM – '**Composite Interval Mapping**') (Zeng, 1994) ainsi que des **modèles à QTLs multiples** (MQM) (Jansen and Stam, 1994). Ces procédures s'effectuent en deux étapes : une phase de régression multiple sur l'ensemble des marqueurs pour le choix des covariables puis une procédure de SIM sur un intervalle amovible excluant les covariables. Ces algorithmes permettent ainsi de réduire la variance résiduelle du modèle, liée à la présence de multiples QTLs sur le génome ainsi que d'avoir une meilleure précision des paramètres (effets, position, intervalle de confiance). Toutefois, ces approches demeurent limitantes dans le cas d'effets d'épistasie ou de liaison entre QTLs, en raison de leur nature unidimensionnelle (Zeng *et al.*, 1999). Une extension à des scans multi-dimensionnels a été proposée via la cartographie d'intervalles multiples (MIM – '**Multiple Interval Mapping**') (Kao *et al.*, 1999) permettant une meilleure estimation des effets individuels de QTLs ainsi que de leurs effets d'interaction.

Un des principaux avantages de l'analyse de liaison tient à la faible densité de marqueurs nécessaires pour capturer les QTLs à l'échelle du génome, à condition d'avoir un niveau de recombinaison suffisant au sein de la population et donc un nombre d'individus approprié (Lynch and Walsh, 1998). En revanche, cette méthode de cartographie présente des inconvénients majeurs liés à :

- (1) sa faible résolution (précision sur la localisation), insuffisante afin d'aller vers le clonage positionnel<sup>1</sup> à moins d'apporter une haute densification du marquage permettant de réduire l'intervalle de confiance du QTL (Ronin *et al.*, 2003),
- (2) la nécessité de développer des populations spécifiques de grande taille afin de détecter des QTLs d'effets moyens à faible,
- (3) la capture d'une faible part de la diversité génétique disponible. Les effets des QTLs détectés se limitent au contraste entre allèles parentaux.

Ce dernier point pose la question de la validité des QTLs détectés au sein d'une population donnée sur tout autre fond génétique (Kenis *et al.*, 2008). Une méthode développée en 2000 sous le nom de '**Méta-analyse**' permet d'exploiter conjointement des résultats de cartographies individuelles issus de dispositifs différents afin d'en effectuer la synthèse à travers l'exploration d'une diversité plus large, tout en améliorant la résolution de détection avec la projection de **méta-QTLs** sur une **carte consensus** (Goffinet and Gerber, 2000; Veyrieras *et al.*, 2007).

---

<sup>1</sup> Identification du gène à partir de la seule connaissance de sa localisation chromosomique.

Plusieurs études de méta-analyses ont été effectuées chez les espèces fruitières en lien avec des traits portés sur la résistance à la sharka (Marandel *et al.*, 2009) chez l'abricotier et la qualité chez le pommier (Costa, 2015).

Une autre alternative à l'analyse de liaison, afin d'explorer une diversité génétique plus large, correspond à l'analyse d'association, dont l'essor est lié à l'avènement du marquage à haut-débit.

## 1.2. Analyse d'association : étude de la diversité naturelle de l'espèce

La cartographie d'association (GWAS) repose sur la détection de corrélations entre variants génotypiques et phénotypiques au niveau de populations d'individus (a priori non apparentés) caractérisant la diversité de l'espèce (*Figure 1B, Figure 2*) (Ogura and Busch, 2015). A l'origine, cette méthode issue de l'épidémiologie génétique humaine a d'abord été appliquée aux caractères qualitatifs avec les études cas/témoins (Corder *et al.*, 1994). Elle a ensuite été adaptée aux caractères quantitatifs, en particulier chez les plantes (Buckler and Thornsberry, 2002; Ingvarsson and Street, 2011).

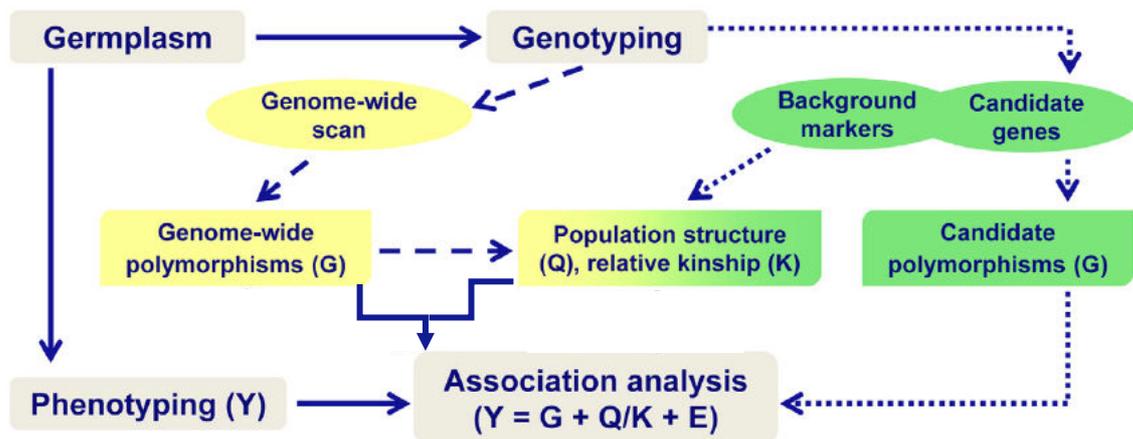


Figure 2 : Diagramme schématisé du contraste entre GWAS (démarche sans a priori sur les gènes causaux) et la démarche gène-candidats (sélection préalable de gènes candidats suite à analyse de liaisons, analyses biochimiques ou mutationnelles).

(Zhu *et al.*, 2008)

La prise en compte de la structure (Q) et/ou de l'apparentement (K) de la population dans l'analyse dépend de la relation génétique entre le panel variétal et la divergence du phénotype étudié.

Un des principaux avantages de cette approche correspond à l'exploitation de populations regroupant à la fois des cultivars issus de collections de ressources génétiques ainsi que de fonds génétiques issus de programme d'amélioration, ne

nécessitant donc pas le design de descendance de ségrégation établies a posteriori comme pour l'analyse de liaison (Korte and Farlow, 2013). Cet avantage est d'autant plus significatif en considérant les espèces pérennes fruitières pour lesquelles l'espace large occupé, l'importance du temps de génération et de la phase juvénile contraignent les schémas de sélection (Rikkerink *et al.*, 2007; Khan and Korban, 2012). La diversité allélique exploitée en GWAS est donc bien plus importante qu'au sein des populations biparentales, pour lesquelles elle est restreinte à la contribution des deux parents (Korte and Farlow, 2013).

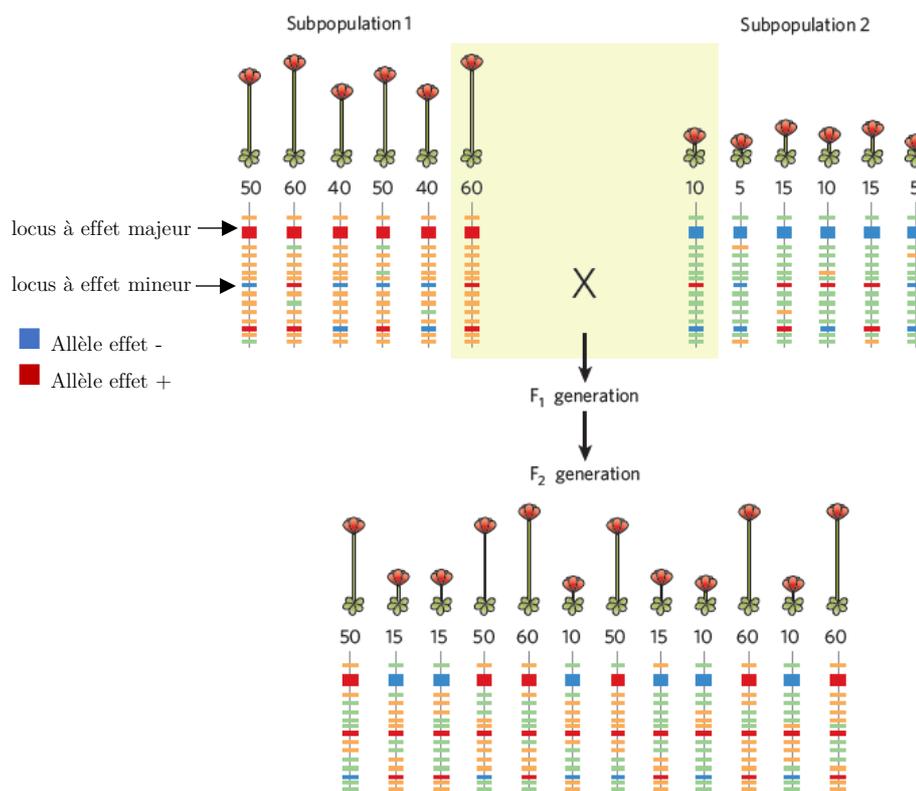
Les études d'association reposent sur l'exploitation du **déséquilibre de liaison** (DL) défini comme une association non aléatoire d'allèles situés à des loci différents au sein de l'ensemble des chromosomes (ou gamètes) présents dans la population. La détection d'une association en GWAS résulte donc du DL entre un marqueur et un QTL dont le polymorphisme est responsable d'une partie de la variabilité du caractère étudié (Nordborg and Weigel, 2008; Huang and Han, 2014). A l'échelle d'un chromosome, la résolution de cartographie dépend essentiellement de la décroissance du DL dans une région donnée : plus le DL décroît rapidement avec la distance physique, plus la résolution est fine et l'intervalle de confiance entourant le QTL est court (Brachi *et al.*, 2011). La décroissance du DL est intimement liée à l'importance des recombinaisons au sein d'une population (Flint-Garcia *et al.*, 2003). Les fonds génétiques exploités en GWAS sont issus de multiples événements de recombinaisons accumulés au cours de l'histoire évolutive de la population (*Figure 1B*), ce qui permet de rompre ces associations à l'échelle du génome et de réduire l'étendue du DL par rapport aux descendance de cartographie (limitées par le nombre de générations et donc de recombinaisons) (Bazakos *et al.*, 2017). L'étendue du DL dépend également de l'espèce, en particulier de son **système de reproduction** et de la **démographie** de la population étudiée.

Dans le cas des espèces autogames pour lesquelles les recombinaisons sont moins efficaces que chez les espèces à fécondation croisée (Nordborg and Donnelly, 1997), le DL est relativement important. Cela a pour avantage de nécessiter une couverture du génome moyenne mais a le revers de présenter une résolution de détection équivalente à celle de l'analyse de liaison. A l'inverse, chez les espèces allogames, en particulier la plupart des espèces pérennes fruitières, la décroissance du DL s'effectue en général sur des courtes distances permettant *in fine* une résolution de l'ordre du gène sous réserve de disposer d'une suffisante couverture de marquage suffisante sur le génome (Rikkerink *et al.*, 2007).

Les **puces SNPs** développées initialement chez la vigne (Myles *et al.*, 2010), le pêcher (Verde *et al.*, 2012) et le pommier (Chagné *et al.*, 2012) ont permis de lever les limitations technologiques liées à la densification du marquage et de rendre l'approche accessible aux espèces pérennes. La GWAS a eu un engouement encore plus marqué

chez les espèces fruitières avec l'avènement du **Genotyping-by-Sequencing** (GBS) permettant la détection simultanée de milliers de SNPs sur l'ensemble du génome sans avoir la nécessité d'utiliser un **génom de référence**. Ceci a rendu l'approche particulièrement utile chez de nombreuses espèces 'mineures' limitées par la disponibilité des ressources génomiques (Catchen *et al.*, 2011).

Les analyses d'association déployées chez les espèces fruitières ont ainsi permis de cartographier des QTLs avec une résolution suffisante pour mettre en évidence un ou plusieurs **gènes candidats** affectant des caractères agronomiques : qualité du fruit (Cevik *et al.*, 2010; Iwata *et al.*, 2013; Kumar *et al.*, 2013; Micheletti *et al.*, 2015; Cao *et al.*, 2016; Minamikawa *et al.*, 2017), résistance aux maladies (Iwata *et al.*, 2013; Mariette *et al.*, 2016; McClure *et al.*, 2018) et phénologie (Iwata *et al.*, 2013; Cao *et al.*, 2016; Urrestarazu *et al.*, 2017).



**Figure 3 : Illustration d'un effet d'artefact en GWAS lié à la structure entre deux sous-populations.**

(Nordborg *et Weigel*, 2008)

Les marqueurs orange (sous-population 1) et verts (sous-population 2) liés au fond polygénique sont spécifiques à chaque sous-population et fortement corrélés à la taille, bien qu'ils ne soient pas causaux. Le croisement de deux lignées parentales prises dans chaque sous-population et l'obtention d'une F2 ou d'une population de RIL permettent de rompre le DL entre fond polygénique et loci causaux.

L'erreur de type II (faux-positif) est alors évitée.

Face aux multiples avantages, les analyses d'association sont contraintes par des obstacles liés aux effets de **stratification (structure et apparentement)** de la population, pouvant mener à la détection de **faux-positifs** dans le cas où ils ne sont pas pris en compte (Ingvarsson and Street, 2011). Une différence de fréquence allélique entre groupes génétiques au niveau de loci non-causaux et la non-validité de la panmixie<sup>2</sup> au sein d'une population (à l'origine d'apparements hétérogènes entre individus) peuvent être des éléments menant à la détection de marqueurs artificiellement confondus avec le phénotype d'intérêt (*Figure 3*) (Astle and Balding, 2009). Des modèles statistiques ont néanmoins été développés afin de contrôler ces artefacts : (1) le **contrôle génomique** qui se base sur l'estimation d'un facteur d'inflation lié à la stratification de la population (Bacanu *et al.*, 2002), la prise en compte en effets fixes des **coefficients de structure ( $Q$ )** via (2) une matrice d'appartenance aux groupes ancestraux (Pritchard *et al.*, 2000) ou (3) les coordonnées des individus suite à une analyse en composantes principales (ACP) (Price *et al.*, 2006) et enfin (4) l'usage de **modèles mixtes**, une méthode qui s'est démontrée particulièrement efficace lors d'études de cartographie génétique dans les domaines animaux, végétaux et humains (Yu *et al.*, 2005; Kang *et al.*, 2008, 2010; Atwell *et al.*, 2010). Ces modèles mixtes testent l'association entre un phénotype et un marqueur  $l$  (les marqueurs étant pris un à un successivement) en intégrant la structure de la population ( $Q$ ) et l'apparement ( $K$ ) entre individus comme effet fixe et aléatoire, respectivement.

La matrice  $K$  d'apparement ('**kinship**') peut exploiter les données de parenté des individus ou bien être inférée sur la base des génotypes au marqueur quand le pedigree est inconnu (cas très fréquent de '**cryptic relatedness**') (Sillanpää, 2011). Cette matrice permet de structurer les dépendances des individus entre eux connaissant leur **covariance génétique**.

Considérant une étude d'association, les modèles mixtes développés initialement en génétique animale à partir de l'**équation d'Henderson** (Henderson, 1984), se déclinent sous la forme vectorielle suivante:

$$y = Wv + Q\beta + Zu + e$$

où  $y$  est le vecteur phénotypique de l'ensemble des individus ;  $W$  et  $Q$  les matrices d'incidence des effets fixes liant dosage de l'allèle mineur au SNP testé et fractions ancestrales (ainsi que d'autres covariables éventuelles), à chaque individu, respectivement ;  $v$  et  $\beta$  les vecteurs des effets fixes du marqueur et des covariables, respectivement ;  $u$  le vecteur des effets polygéniques aléatoires avec  $u \sim N(0, 2K\sigma_a^2)$  ;  $Z$  la matrice d'incidence des effets polygéniques liant chaque individu à son

---

<sup>2</sup> Croisements aléatoires entre individus d'une population.

phénotype ; et  $e$  le vecteur des effets résiduels aléatoires avec  $e \sim N(0, I\sigma_e^2)$  et  $\sigma_e^2$  la variance résiduelle.  $\sigma_a^2$  et  $\sigma_e^2$  correspondent aux variances génétique et résiduelle, respectivement ;  $I$  et  $K$  aux matrices identité et de covariance génétique entre individus.

Dans le cas d'études se focalisant sur des phénotypes fortement corrélés, l'adaptation de ces modèles mixtes sur des **données multivariées** (Korte *et al.*, 2012; Zhou and Stephens, 2014) s'est révélée être performante afin de caractériser à la fois la **plasticité** de phénotypes dans des essais multi-environnement (Albert *et al.*, 2016; Thoen Manus *et al.*, 2016) mais aussi des gènes **pléiotropes**, affectant plusieurs traits (Korte *et al.*, 2012). Les analyses multivariées d'association présentent également l'avantage d'être plus puissantes que des analyses uni-variées afin de détecter des variants génétiques n'affectant qu'un des phénotypes corrélés (Stephens, 2013). Enfin, ces modèles permettent aussi de décomposer la covariance de phénotypes en composantes génétique et environnementale, évaluant ainsi de façon précise une corrélation d'ordre génétique (Korte *et al.*, 2012).

L'ensemble des modèles mixtes décrits précédemment correspondent à des tests de comparaison de moyennes effectués variant par variant, la variance génétique étant évaluée uniquement au début de la procédure. Une implémentation a par la suite été développée afin de gérer les effets confondant du **fond polygénique**, liés à la présence de DL à l'échelle du génome, qui peuvent mener à la détection de nombreux signaux correspondant à des variants corrélés avec le SNP le plus significativement associé. Le **modèle mixte multi-locus** (Segura *et al.*, 2012), consiste ainsi en une régression multiple par étapes, réévaluant la variance génétique résiduelle à chaque mise en cofacteur du SNP le plus significatif. Ce modèle permet ainsi notamment la réduction du nombre de loci détectés par bloc de DL en intra-chromosome. Jusqu'à peu, toutes ces méthodologies étaient focalisées sur les **effets alléliques additifs** or une part importante de la variation des caractères quantitatifs est également liée à des effets de **dominance**, qui interviennent en particulier dans les phénomènes de vigueur hybride chez certaines cultures telles que le maïs (Yang, 2004), le riz (Zhou *et al.*, 2012) ou le tournesol (Dimitrijevic and Horn, 2018). Récemment, ces effets de dominance ont pu être implémentés en plus des effets parentaux mâle-femelle à travers un modèle mixte multi-locus, appliqué sur des traits de floraison au sein d'un panel de 452 hybrides de tournesol (Bonnaïfous *et al.*, 2018). L'étude a ainsi permis de détecter des associations supplémentaires (par rapport au modèle additif) et ainsi dans le même temps d'augmenter l'héritabilité au sens-large du trait.

Néanmoins, malgré le progrès des méthodologies statistiques récentes, il est reconnu que la GWAS s'avère limitante dans le cas des SNPs à effet faible, des variants sous-représentés au sein de la population, d'une forte corrélation du phénotype étudié avec la stratification de la population (Myles *et al.*, 2009) ainsi que pour la détection de

QTLs en épistasie (Bazakos *et al.*, 2017). Afin de contrebalancer ces difficultés, l’usage conjoint de la cartographie de liaison et d’association s’est révélé être une approche fiable afin de limiter à la fois les faux-positifs et faux-négatifs (Brachi *et al.*, 2010; El-Soda *et al.*, 2014). L’usage de la GWAS peut ainsi être conduit initialement afin d’identifier au sein d’un large panel, des cultivars intéressants dont le croisement générerait une population de ségrégation, utile pour une analyse de liaison à but de validation (Bazakos *et al.*, 2017). La détection des **allèles rares** (dans le cas de la GWAS) est ainsi envisageable par l’exploitation de ces populations de ségrégation qui permettent de rétablir un rééquilibrage des fréquences alléliques et donc une meilleure puissance statistique de détection (Korte and Farlow, 2013). A l’inverse, il est également possible de conduire un scan initial du génome à travers une analyse de liaison et d’effectuer une validation des régions détectées en exploitant la meilleure résolution de la GWAS, afin d’identifier les gènes candidats associés (Mackay, 2001).

Chez les espèces fruitières, plusieurs études ont mis en évidence la **colocalisation** de QTLs issus d’études de GWAS et d’analyse de liaison sur des traits portés sur la résistance aux maladies et la qualité du fruit chez le pommier (Cevik *et al.*, 2010; Kumar *et al.*, 2013), le pêcher (Cao *et al.*, 2012; Picañol *et al.*, 2013) et le poirier (Iwata *et al.*, 2013; Yamamoto *et al.*, 2014).

Des approches récentes de cartographie ont été développées afin de combiner la puissance de détection de l’analyse de liaison à la résolution précise de la GWAS à travers l’usage de populations multi-parentales.

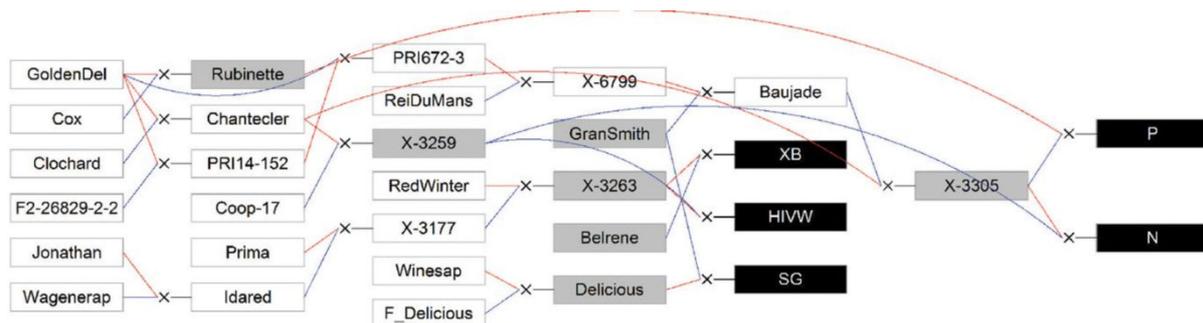
### 1.3. Dispositifs multi-parentaux

Les méthodes connues sous le nom de ‘LD-LA mapping’ (Wu *et al.*, 2002) ont été développées afin d’obtenir une convergence des approches de GWAS et d’analyse de liaison.

- (1) Par rapport aux croisements biparentaux, l’usage de dispositifs multi-parentaux permet à la fois d’exploiter la ségrégation d’une diversité allélique plus riche (liés à la contribution d’un plus grand nombre de géniteurs) et d’augmenter le nombre de recombinants et donc la précision de cartographie (Ogut *et al.*, 2015).
- (2) Par rapport aux panels de GWAS, les populations multi-parentales présentent des fréquences alléliques plus équilibrées permettant une meilleure détection des variants à faible effet. Ces populations ont également l’avantage de pouvoir exploiter les mêmes méthodologies statistiques qu’en GWAS, sans être pénalisées par un effet de structure de population.

Plusieurs dispositifs ont été développés chez des espèces modèles (*Arabidopsis thaliana*, maïs, riz, blé, tomate), suite à plusieurs cycles de reproduction afin d’augmenter efficacement le nombre de recombinants :

- (i) ‘**Advanced Intercross RIL (AI-RIL) populations**’: des géotypes F1 ou F2 sont croisés entre eux sur plusieurs générations (‘inter-mating’) avant fixation de lignées homozygotes suite à des auto-fécondations successives (Balasubramanian *et al.*, 2009),
- (ii) ‘**Multiple Advanced Generation Inter-Cross**’ (MAGIC) issue de croisements initiaux entre lignées parentales suivie par plusieurs générations de croisements avant fixation. Les lignées MAGIC correspondent à des mosaïques alléliques résultant de la contribution de l’ensemble des lignées parentales offrant une résolution de cartographie très précise (Kover *et al.*, 2009; Pascual *et al.*, 2015),
- (iii) ‘**Nested Population Mapping (NAM) design**’ correspondant à un ensemble de lignées RIL connectées par un parent commun (Yu *et al.*, 2008).



**Figure 4 : Exemple de pedigree d’un dispositif de 5 familles multi-parentales (en noir) chez le pommier.**

(Allard *et al.*, 2016)

Les géotypes parentaux sont représentés en blanc et les variétés génitrices fondatrices ainsi que les autres géotypes du pedigree en blanc. Les lignes bleue et rouge relient respectivement le père et la mère à la descendance.

Chez les espèces pérennes, des approches récentes de cartographie de QTLs initialement développées chez le pommier se basent sur l’exploitation des données de pedigree de populations interconnectées issues d’un set de géniteurs communs (‘**pedigree-based analysis**’) (Figure 4) (van de Weg *et al.*, 2004). Certains programmes spécifiques tels que FlexQTL (Bink *et al.*, 2014) et MCQTL (Jourjon *et al.*, 2005) permettent de suivre la transmission des allèles le long des générations au sein de ces populations multi-parentales et d’en évaluer la corrélation avec le trait étudié. Ces approches ont très souvent permis d’améliorer la puissance statistique et

la précision d'estimation des positions des QTLs et de leurs effets lors d'études effectuées chez le pommier (Bink *et al.*, 2014; Allard *et al.*, 2016; van de Weg *et al.*, 2018), le palmier à huile (Billotte *et al.*, 2010; Tisné *et al.*, 2017), le pêcher (Fresnedo-Ramírez *et al.*, 2016) et le cerisier (Rosyara *et al.*, 2013). Elles présentent également l'avantage d'évaluer les **interactions QTL x fond génétique** et d'adresser la question de la généricité des marqueurs en sélection, de façon plus précise qu'au travers de la méta-analyse chez les populations biparentales (Bink *et al.*, 2014; Tisné *et al.*, 2017).

## Chapitre 2. Résistances : déterminisme génétique, moléculaire et durabilité

L'interaction plante x pathogène peut être perçue comme un mode de communication bi-latéral dans lequel d'un côté, la plante a la capacité de percevoir un organisme étranger ('non-soi') et de s'en défendre et de l'autre, l'agent pathogène manipule et détourne les défenses de l'hôte afin de créer un environnement optimal pour sa propre croissance et reproduction.

Les agents pathogènes sont souvent catégorisés selon leur mode d'interaction avec la plante dont les mécanismes différents vont influencer sur le mode d'infection de la plante (Pieterse et al., 2009). Les agents biotrophes colonisent les tissus vivants de l'hôte alors que les nécrotrophes engendrent la mort cellulaire des tissus à l'aide de toxines avant d'envahir l'hôte. Une catégorie intermédiaire, les hémibiotrophes combinent les deux modes de vie au cours de leur cycle épidémiologique : une phase initiale biotrophe au cours de laquelle ils évitent la reconnaissance de l'hôte puis une phase nécrotrophe, correspondant souvent au début du développement des symptômes.

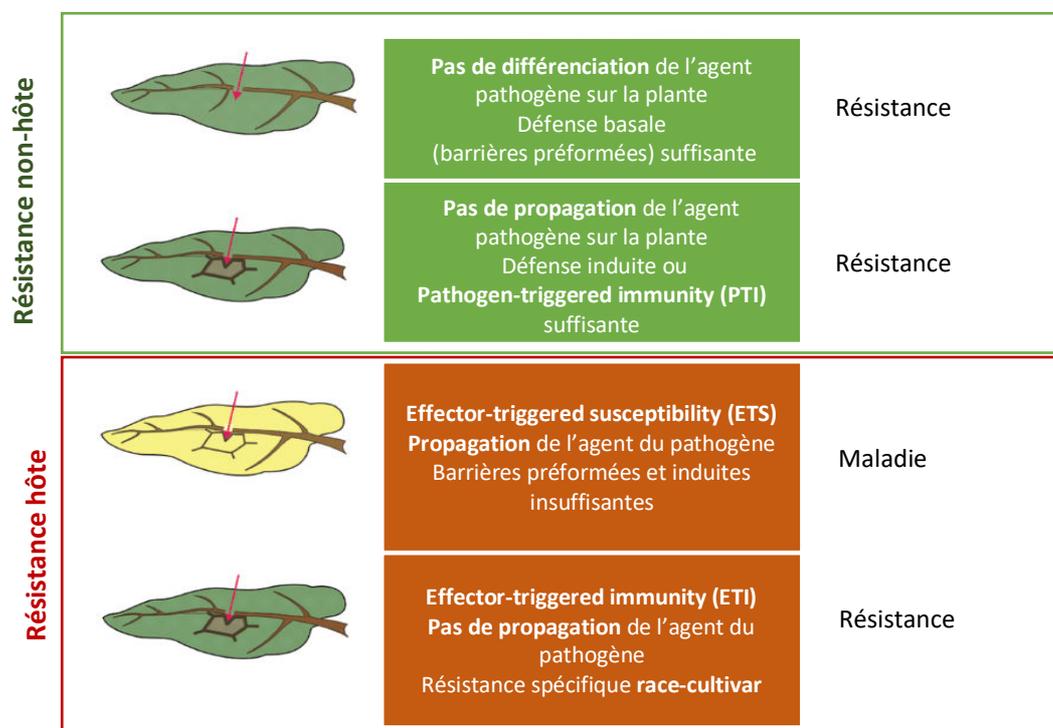


Figure 5 : Représentation schématique des différents types d'interaction plante x pathogène.

*Adapté de Nürnberger, 2004.*

Au cours de l'évolution, les plantes ont acquis un certain nombre de mécanismes les protégeant des invasions des agents pathogènes nuisibles. La conjonction de ces mécanismes de défenses fait qu'en réalité la maladie est davantage une exception qu'une règle générale.

Les deux grands modes de résistances connus (*Figure 5*) correspondent à la résistance non-spécifique (ou non-hôte) (partie 2.1) et la résistance hôte-spécifique (partie 2.2). La co-évolution entre plante et agent pathogène a également mené l'agent pathogène à acquérir des facteurs de virulence lui permettant dans certains cas de franchir l'ensemble des défenses de la plante (partie 2.3). Dans la 'course aux armements', l'agent pathogène est donné gagnant, la résistance est contournée. La colonisation de la plante est alors inévitable et la maladie se déclare. Néanmoins, nous verrons qu'il est possible d'agir grâce au levier de l'amélioration variétale et à l'exploitation des résistances quantitatives (partie 2.4) et également grâce au déploiement raisonné de ces résistances en champ afin d'en préserver la durabilité (partie 2.5).

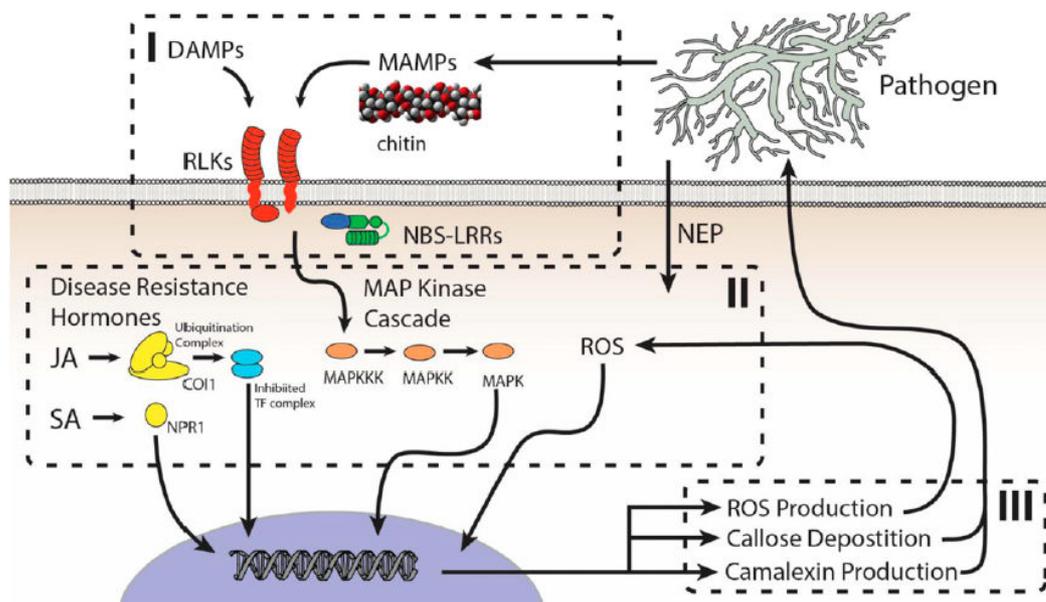
## 2.1. Résistance non-spécifique ou non-hôte

Une observation très souvent commune est le fait que la plupart des espèces végétales soient complètement résistantes à la majorité des agents pathogènes. Cela s'explique par la spécialisation assez étroite que présente une grande partie des micro-organismes pathogènes. Les espèces végétales pouvant être infectées par un agent pathogène donné sont considérées comme des espèces 'hôtes' et constituent la **gamme d'hôte** de cet agent pathogène. Toutes les autres sont alors qualifiées de 'non-hôtes'.

Parmi les premières lignes de défense de la plante, figure la **défense basale ou passive**. Les micro-organismes vont d'abord se heurter à un ensemble de barrières physiques dont la cuticule recouvrant l'épiderme végétal ainsi que les parois cellulaires. A ces barrières structurales s'ajoutent l'action de barrières chimiques : enzymes et métabolites secondaires antimicrobiennes (**phytoanticipines et phytoalexines**), afin de prévenir la progression ultérieure du pathogène (Nürnberger *et al.*, 2004; Senthil-Kumar and Mysore, 2013).

Dans le cas où, l'agent pathogène contourne toutes les barrières constitutives de défense, il est sujet à reconnaissance au niveau de la membrane plasmique des cellules (*Figure 6*). Des récepteurs extracellulaires '**pathogen recognition receptors**' (PRRs) ancrés au niveau des membranes cellulaires de la plante ont effectivement la capacité de reconnaître des motifs propres à l'agent pathogène, appelés **éliciteurs** ou '**pathogen/microbe – associated molecular patterns**' (PAMPs/MAMPs) (Jones and Dangl, 2006). Il s'agit de motifs assurant des fonctionnalités de base chez les micro-organismes. Ils sont de ce fait, en majorité faiblement soumis aux forces

évolutives<sup>3</sup> et fortement conservés entre agents pathogènes. La **flagelline**, protéine associée à la constitution du flagelle de nombreuses bactéries, le facteur d'élongation Tu (EF-Tu), ainsi que les lipopolysaccharides, composants principaux des membranes externes des bactéries Gram négatif sont des exemples de MAMPs bactériens reconnus par les plantes (Newman *et al.*, 2007; Zipfel, 2008). Les structures des PRRs correspondent souvent à des 'receptor-like kinases' (RLKs) ou 'receptor-like proteins' (RLPs) présentant des domaines fonctionnels modulaires (Monaghan and Zipfel, 2012).



**Figure 6 : Représentation schématique des réactions de défense des plantes.**

Le modèle peut être décrit en 3 phases: (I) perception, (II) transmission du signal et (III) réponse de défense.

(Corwin et Kliebenstein, 2017).

Abréviations: MAMP/DAMP, 'microbe/damage associated molecular pattern'; RLK, 'receptor like kinase'; NBS-LRR, 'nucleotide-binding site of leucine-rich repeats', JA et SA: acides jasmonique et salicylique; COI1, 'coronatine-insensitive 1'; NPR1, 'nonexpressor of pathogenesis-related genes 1'; MAPK, 'mitogen-activated protein kinase'; ROS: espèce réactive de l'oxygène; TF, 'transcription-factor'

D'autres types de motifs appelés '**damage – associated molecular patterns**' (DAMPs) sont également reconnus par la plante comme des éliciteurs endogènes ('soi infectieux'), synonymes de l'arrivée de l'agent pathogène. Il s'agit par exemple de

<sup>3</sup> En réalité, il peut exister des cas de pression de sélection sur les PAMPs afin d'abolir leur reconnaissance par les PRRs associés. Une mutation au niveau du peptide flg22, ancré dans le domaine d'interaction flagelline-récepteur FLS2 ('*receptor-like kinase flagellin sensing 2*') a par exemple permis d'abolir l'interaction et a fait l'objet d'une sélection positive chez les espèces bactériennes *Ralstonia solanacearum* et *Xanthomonas campestris* (Boller and Felix, 2009).

molécules libérées suite à la dégradation de la membrane plasmique (cutine, oligogalacturonides) ou de polypeptides signaux (systèmeine) agissant sur les cellules voisines (Newman *et al.*, 2013; Choi and Klessig, 2016).

La perception d'un MAMP par son PRR associé déclenche un premier niveau de défense active appelé '**PAMP-triggered immunity**' (PTI) engendrant une cascade de réactions. Parmi elles, on note la production de formes réactives de l'oxygène (ROS), de **phytohormones**, un dépôt de callose renforçant la paroi au niveau du site de l'infection et l'activation de la transcription de '**pathogenesis-related genes**' (*PR*) ou de **facteurs de transcription** de type WRKY, ERF et MYB (Pieterse *et al.*, 2009; Senthil-Kumar and Mysore, 2013). Cette cascade de signalisation est associée à l'intervention de '**Mitogen-Activated Protein Kinases**' (MAPKs), molécules centrales de la régulation de la transduction du signal de défense via des réactions de phosphorylation, chez les eucaryotes (Nürnberg *et al.*, 2004). En parallèle, des mécanismes complémentaires de défense sont mobilisés par la plante via le '**RNA silencing**' et les micro-ARN ('miRNA') (Vance, 2001). Des séquences spécifiques de miRNA permettent de réguler l'expression de gènes afin de prioriser la défense par rapport à d'autres voies métaboliques. Le premier miRNA, miR393, a été identifié chez *A. thaliana* en réponse à *Pseudomonas syringae* (*Psy*) (Huang *et al.*, 2016). miR393 est induit par la reconnaissance de flg22 et enclenche une série de réactions de défense tout en exerçant un contrôle négatif sur la voie de signalisation de l'auxine, priorisant ainsi la défense sur la croissance de la plante.

Le rôle des phytohormones, en particulier des **acides salicylique** (SA) et **jasmonique** (JA) ainsi que l'**éthylène** (ET) fait l'objet d'un pan important de la littérature portant sur les interactions plante x pathogène (pour revues, consulter : Pieterse *et al.*, 2009; Robert-Seilaniantz *et al.*, 2011; Gozzo and Faoro, 2013; Checker *et al.*, 2018).

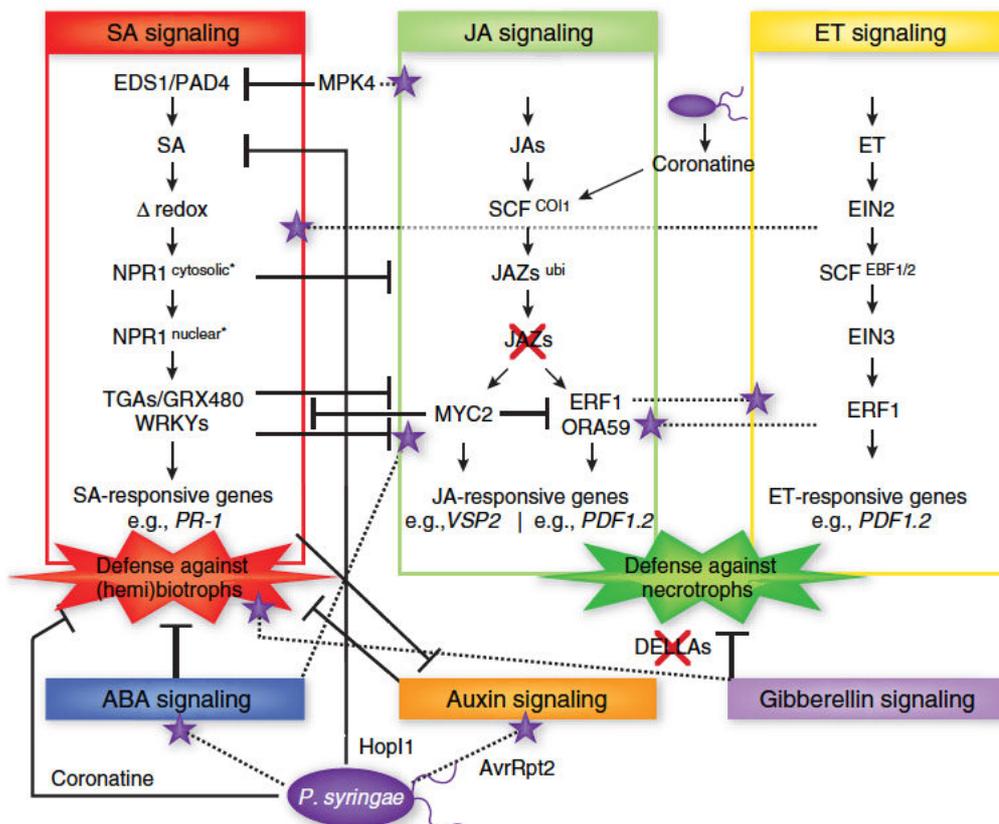
Le SA est un composé dérivé de phénols s'accumulant au niveau des tissus nouvellement infectés mais également au niveau des cellules adjacentes. Dans le cadre de la PTI, il induit la transcription des gènes *PR* associés à la '**résistance systémique acquise**' (SAR) par le biais d'une protéine de régulation NPR1 (Pieterse *et al.*, 2009). La SAR est par la suite responsable de la transmission du signal de défense à longue portée dans la plante. Ces gènes *PR* codent notamment des glucanases, chitinases et peroxydases dont l'action systémique au niveau des tissus leur confèrent une résistance vis-à-vis d'un large spectre d'agents pathogènes (Ryals *et al.*, 1996; Loake and Grant, 2007).

Le JA est une hormone dont la forme volatile, le méthyl-jasmonate (MeJA) présente un rôle clé dans les communications intra et inter-plantes (Checker *et al.*, 2018). La synthèse du JA requiert l'action de plusieurs enzymes dont les lipoxigénase et phospholipase, rapidement induites après la reconnaissance des éliciteurs et entraînant également la production de phytoalexines et d'inhibiteurs de protéases (Porta, 2002;

Canonne *et al.*, 2011). La plupart des signaux liés au JA sont modulés par une protéine F-box coronatine insensible 1 (COI1).

L'ET est une hormone volatile impliquée dans divers processus physiologiques, comme la germination ou la maturation des fruits, mais également en tant que signal de défense

chez les plantes. Son action principale est d'induire la production de phytoalexines et la synthèse de protéines PR (Tornero *et al.*, 1997; Fan *et al.*, 2000).



**Figure 7 : Représentation schématique des interconnexions entre différentes voies de signalisation des phytohormones en réponse à *Pseudomonas syringae* (*Psy*).**

(Pieterse *et al.*, 2009)

Les effecteurs (coronatine, HopI et AvrRpt2) interagissent avec le réseau afin de supprimer les défenses de l'hôte et assurer la virulence de *Psy*. ⊥, effet d'inhibition; étoile violette, effet positif.

L'implication de chacune de ces hormones dans les mécanismes de résistance dépend du mode d'interaction de l'agent pathogène avec la plante. Ainsi, la voie de signalisation du SA cible la résistance contre les agents biotrophes et hémibiotrophes alors que la combinaison des voies de signalisation du JA et de l'ET active la résistance contre les nécrotrophes (Glazebrook, 2005). Ces deux modes de signalisation sont très souvent **antagonistes** et **interdépendants** (Checker *et al.*, 2018). Les agents pathogènes ont la capacité de synthétiser des hormones 'mimiques' interagissant avec la modulation de la balance hormonale de la plante dont le déséquilibre peut fortement

compromettre la résistance de la plante (*Figure 7*) (Robert-Seilaniantz *et al.*, 2011). A l'action du JA, SA et ET, s'ajoutent d'autres hormones telles que l'acide abscissique (ABA) (Asselbergh *et al.*, 2008), les auxines (Wang *et al.*, 2007b), cytokinines (Walters and McRoberts, 2006), gibbérelines (DELLAs) (Navarro *et al.*, 2008) et brassinostéroïdes (Nakashita *et al.*, 2003) (*Figure 7*).

Compte-tenu de la grande diversité des mécanismes engendrés par la PTI, la résistance non-spécifique confère une protection à large spectre et généraliste, face à plusieurs agents pathogènes.

Afin de contrer la PTI, les agents pathogènes ont développé un répertoire de protéines **effectrices** pouvant agir dans l'espace intracellulaire afin de réprimer voire supprimer les défenses induites au cours de la PTI. Certaines protéines effectrices sont capables d'inhiber les dépôts de callose comme AvrPto chez *Psy* (Hauck *et al.*, 2003) ; d'autres, d'interférer avec les cascades régulées par les MAPKs afin d'empêcher l'activation de la PTI, c'est le cas de HopAI1 chez *Psy* (Zhang *et al.*, 2007). Dans le cas où l'agent pathogène inhibe la PTI, l'infection est alors déclarée, on parle alors d' '**effector triggered susceptibility**' (ETS).

Le terme 'effecteur', à l'origine employé pour qualifier des protéines, s'est peu à peu étendu à l'ensemble des molécules (toxines, polysaccharides, ARN) de l'agent pathogène ayant la capacité d'interférer avec le fonctionnement des cellules hôtes (Boller and Felix, 2009; Weiberg *et al.*, 2013).

## 2.2. Résistance hôte-spécifique

Face au déploiement des effecteurs par l'agent pathogène, la plante a développé un ensemble de mécanismes de détection de ces effecteurs ou de leur activité via des protéines de résistance R. Ces protéines R, correspondent souvent à des protéines 'nucleotide binding site leucine rich repeat' (NBS-LRR), une des familles les plus importantes et variables identifiées chez les plantes (Yue *et al.*, 2012). Ce second mode de défense correspond à l' 'effector-triggered immunity' (ETI) également appelée, 'résistance hôte spécifique'.

Le mode d'interaction mis à l'œuvre lors des réactions d'ETI correspond à une reconnaissance spécifique d'un effecteur chez le pathogène par un gène *R* chez la plante (*Figure 8*). Ce modèle d'interaction 'gène-pour-gène' a été établi par Flor (1955) lors de l'étude de l'interaction lin x *Melampsora lini*, champignon responsable de la rouille du lin (Flor, 1955, 1971). La reconnaissance mène alors à un arrêt précoce de l'infection, très souvent par l'intermédiaire d'une réaction de mort cellulaire programmée des tissus infectés, connue sous le nom d'hypersensibilité (HR).

La plante est totalement résistante et l'agent pathogène avirulent, l'interaction est alors dite 'incompatible'. L'effecteur prend alors la qualification de facteur d'avirulence

Avr. A l'inverse, dans le cas d'une interaction dite 'compatible', un des deux gènes est absent ou non-fonctionnel, la reconnaissance n'a pas lieu, la maladie se développe alors et la plante est sensible. Les caractères d'avirulence et de résistance dans le cadre du modèle gène-pour-gène sont tous deux de type monogénique et dominant.

		GÉNOTYPE DE LA PLANTE	
		R	r
GÉNOTYPE DE L'AGENT PATHOGENE	Avr	Interaction incompatible: résistance	Interaction compatible: sensibilité
	avr	Interaction compatible: sensibilité	Interaction compatible: sensibilité

Avr: gène d'avirulence présent et fonctionnel  
avr: gène d'avirulence absent ou non fonctionnel  
R: gène de résistance présent et fonctionnel  
r: gène de résistance absent ou non fonctionnel

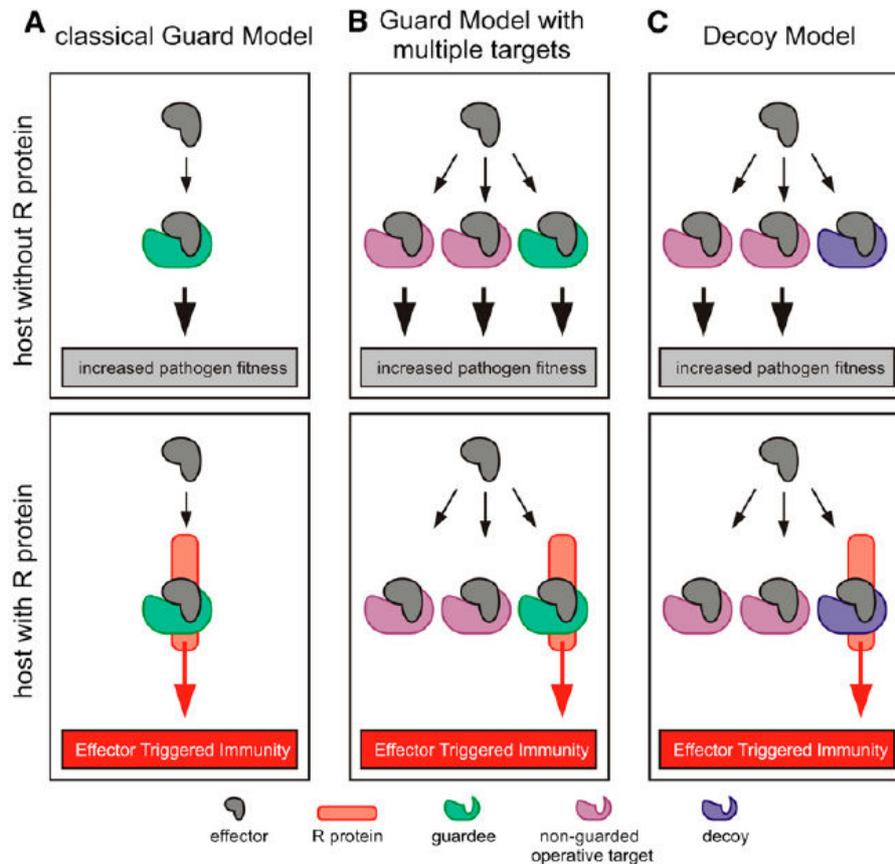
Figure 8 : Illustration du modèle de Flor de l'interaction gène-pour-gène.

Différents modèles biochimiques ont été proposés afin de caractériser l'interaction effecteur- protéine R.

Le modèle le plus simple correspond à une interaction physique directe de type ligand-récepteur entre la protéine R et le facteur Avr, initiant le signal de défense chez la plante (Keen, 1990). Bien que ce modèle ait été vérifié dans le cas de plusieurs couples effecteur-protéine R de divers pathosystèmes : RRS1-R/Pop2 chez *A.thaliana* x *Ralstonia solanacearum* (Deslandes et al., 2003), Protéine L/AvrL567 chez lin x *Melampsora lini* (Dodds et al., 2006), Pi-ta/Avr-Pita chez riz x *Magnaporthe grisea* (Jia et al., 2000), la majorité des interactions étudiées à ce jour semble suivre un mode de reconnaissance indirect.

Le modèle de garde, initialement proposé afin d'expliquer le mécanisme de perception indirecte de l'effecteur AvrPto de *P. syringae* pv. tomato (Pto) DC3000 par les protéines Pto et Prf chez la tomate (Van Der Biezen and Jones, 1998), s'est ensuite généralisé à d'autres interactions (Dangl and Jones, 2001). L'hypothèse de garde suggère que les protéines R (Prf) peuvent détecter les altérations ou modifications provoquées par l'effecteur (AvrPto) sur la protéine 'gardée' (Pto), qui est la cible de l'effecteur (Figure 9A et 9B). Un autre exemple particulièrement bien décrit de ce modèle correspond à l'association RIN4-RPM1-RPS2 chez *A.thaliana* (Jones and Dangl, 2006). RIN4, un régulateur négatif de la défense basale, est la cible des effecteurs AvrRmp1 et AvrRpt2 de *Psy* qui vont induire le clivage de la protéine cible en raison de leur activité phosphorylase et protéase, respectivement. Ce clivage est

par la suite détecté par les protéines RPM1 et RPS2 qui vont initier le signal de défense. Ce modèle est donc lié à la reconnaissance d'un signal impliquant plusieurs protéines et permet d'expliquer le fait que plusieurs effecteurs puissent être perçus par la même protéine R (Dangl and Jones, 2001). Ainsi un petit répertoire de gènes *R* a la capacité de cibler la grande diversité de pathogènes. Aussi, en l'absence de la protéine R, la présence de la protéine cible est indispensable pour la fonction de virulence de l'effecteur.



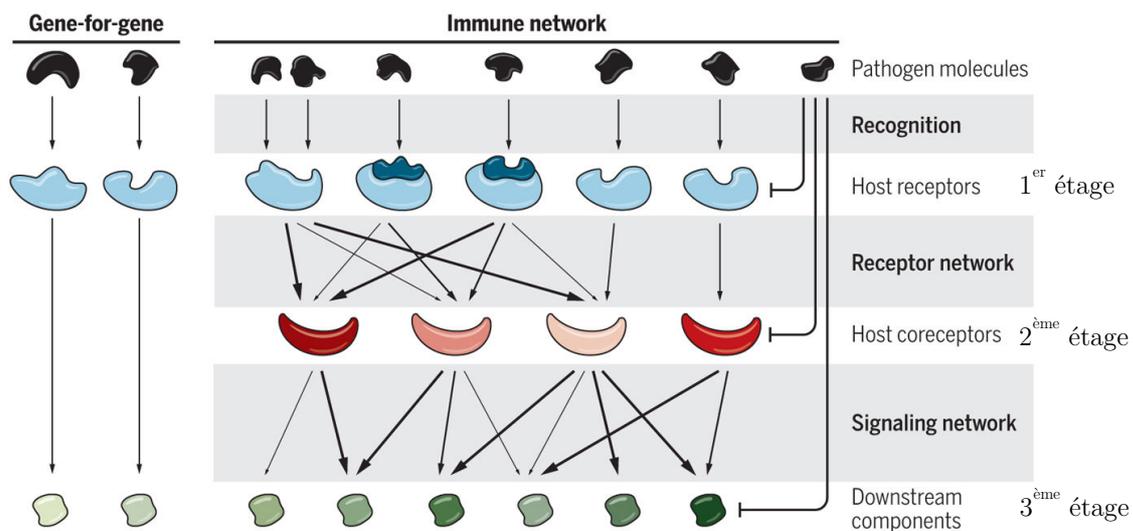
**Figure 9 : Représentation schématique du modèle de garde simple (A) et avec plusieurs protéines cibles (B) et du modèle du leurre (C).**

*(van der Hoorn et Kamoun, 2008)*

Ce modèle mène alors à une instabilité évolutive de par l'opposition des forces de sélection qui s'exercent sur les protéines 'gardées' en fonction de la présence ou l'absence de la protéine R. En effet, en cas d'absence de la protéine R, les protéines cibles sont sujettes à une pression de sélection afin de réduire l'interaction avec l'effecteur. En présence de la protéine R, elles sont au contraire sous pression de sélection afin d'améliorer l'interaction et augmenter la perception de l'effecteur. Ainsi une duplication de la protéine cible ou l'évolution indépendante d'une protéine cible 'imitatrice' pourrait consister en un 'leurre', sujet aux contrastes évolutifs, qui serait uniquement impliqué dans la reconnaissance de l'effecteur (*Figure 9C*) (van der Hoorn

and Kamoun, 2008). L'altération de cette protéine 'leurre' par l'effecteur n'augmente pas la virulence du pathogène en cas d'absence de la protéine R et pourrait à l'inverse initier un signal de défense chez les plantes présentant la protéine R. Ce modèle du leurre ('decoy model') (van der Hoorn and Kamoun, 2008) se distingue alors du modèle de garde, pour lequel, l'absence de la protéine R associée à une altération de la protéine cible de l'effecteur implique forcément une hausse de fitness<sup>4</sup> chez le pathogène.

Ce modèle permettrait d'expliquer plusieurs cas d'interactions n'entrant pas dans le cadre du modèle de garde. A titre d'exemple, chez *A. thaliana*, l'effecteur AvrPphB de *Psy* est une protéase capable de cliver la kinase PBS1. Cette dénaturation de forme de PBS1 est détectée par la protéine R, RPS5 (Ade et al., 2007). Deux autres kinases, BIK1 et PBL1 ont été également identifiées comme étant cibles de AvrPphB. Les mutants *pbs1* (absence de sécrétion de la protéine PBS1) d'*A. thaliana* ne sont pas plus sensibles que chez les plantes témoins (Warren et al., 1999), indiquant qu'en absence de RPS5, PBS1 n'influe pas sur la virulence et semble alors agir comme un 'leurre' de AvrPphB (Sun et al., 2017). En revanche, les mutants *bik1* et *pbl1*, sont affectés dans leur réponse de défense contre différents agents pathogènes, indiquant le fait que BIK1 et PBL1 soient de véritables protéines cibles de AvrPphB (Lu et al., 2010; Zhang et al., 2010a).



**Figure 10 : Représentation schématique de la modélisation des réactions de défense des plantes par des réseaux de récepteurs.**

(Wu et al., 2018)

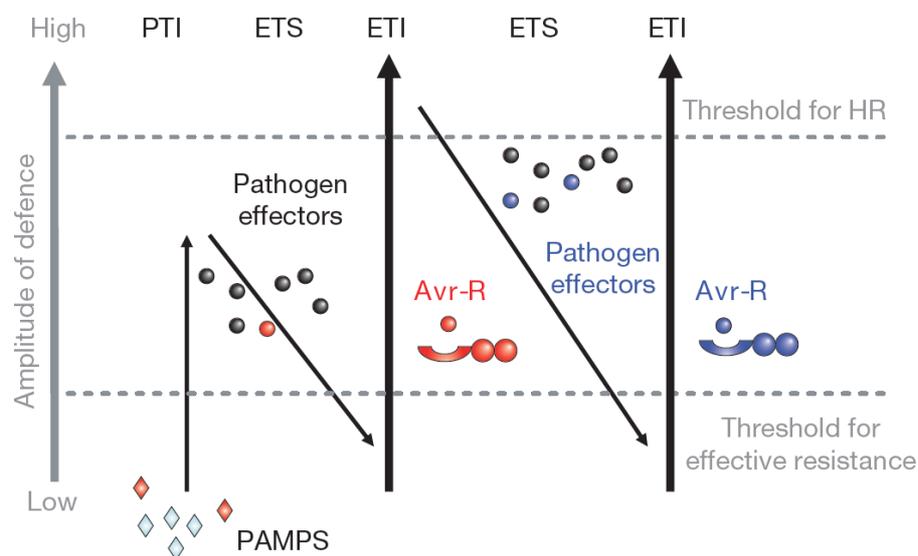
La base des interactions effecteurs-protéines R semble alors impliquer davantage de composantes que la vision binaire du modèle gène-pour-gène semblait l'indiquer. Aujourd'hui, les dernières recherches avancent des modèles de réactions de défense se basant sur des réseaux à trois étages mécanistiques et structurant des interconnexions

<sup>4</sup> Fitness = Valeur sélective, capacité d'un individu ayant un certain génotype de se reproduire.

entre récepteurs (*Figure 10*) (Wu et al., 2018). D’abord, une reconnaissance directe ou indirecte des MAMPs ou effecteurs est établie grâce à des récepteurs capteurs, PRR ou protéines R (1er étage). Ensuite, la majorité de ces récepteurs nécessitent la présence de co-récepteurs (2e étage) avant de traduire la reconnaissance des MAMPs/effecteurs en signaux ultérieurs de défense systémique via des composantes de signal en aval (3e étage). Ces différents étages constituent des nœuds présentant des redondances fonctionnelles, des dépendances différentielles (contrôle négatif/positif) et des spécificités. La redondance de ces récepteurs et co-récepteurs permet ainsi d’assurer l’opérationnalité du système en cas de non-fonctionnalité de l’un d’entre eux par mutation, délétion, ou immunosuppression provoquée par le pathogène.

### 2.3. Modèles évolutifs des interactions plante x pathogène et contournement des résistances

De par les différences de nature des molécules impliquées (MAMP-PRR *vs.* effecteur-protéine R) et d’étendue des résistances induites (générale *vs.* spécifique), la PTI et l’ETI ont longtemps été étudiées avec un regard dichotomique. Depuis une dizaine d’années, une vision intégratrice des réactions de défense et des interactions plante x pathogène a été proposée à travers un **modèle évolutif en ‘zig-zag’** (Jones and Dangl, 2006).



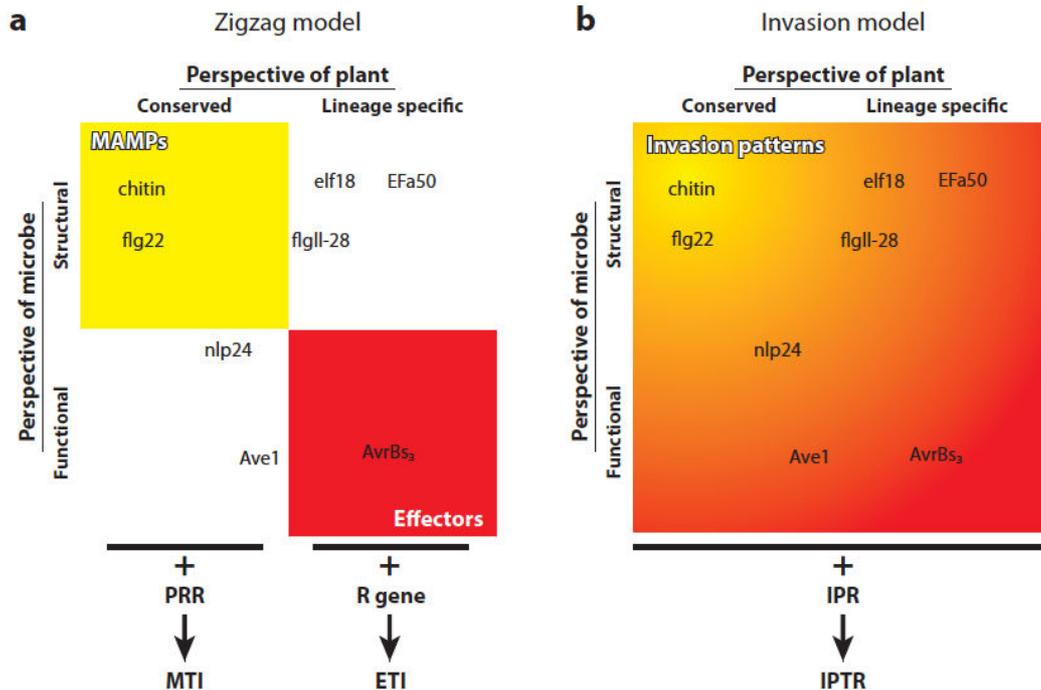
**Figure 11 :** Représentation schématique du modèle évolutif en ‘zig-zag’ des interactions plante x pathogène.  
(Jones et Dangl, 2006)

Ce modèle se déroule en 4 phases distinctes (*Figure 11*):

- (1) Reconnaissance des MAMPs par les PRR de la plante : initiation de la PTI.
- (2) Contournement de la PTI par plusieurs agents pathogènes grâce à leurs effecteurs : phase d'ETS.
- (3) Reconnaissance d'un ou plusieurs effecteurs par les protéines R de la plante : initiation de l'ETI via la HR.
- (4) Chez l'agent pathogène : sélection positive sur la perte d'avirulence et gain de nouveaux effecteurs. Suppression de l'ETI et inhibition de la PTI.  
Chez la plante : sélection positive sur de nouveaux allèles R pouvant reconnaître les nouveaux effecteurs du pathogène. Initiation de l'ETI à nouveau.

Ce modèle pose ainsi comme base de défense chez les plantes, une co-évolution continue au sein de laquelle la plante restaure l'ETI via l'acquisition de nouveaux gènes *R* et le pathogène restaure l'ETS en abolissant la reconnaissance de ses effecteurs. Ces processus évolutifs répondent à la **théorie de la reine rouge** (van Valen, 1973) : l'évolution permanente d'une espèce est nécessaire pour maintenir son aptitude (et sa survie) face aux évolutions des espèces avec lesquelles elle co-évolue. La principale avancée conceptuelle du modèle en 'zig-zag' fut également de présenter la dualité fonctionnelle des effecteurs qui ont le paradoxe d'avoir à la fois un rôle dans le pouvoir pathogène et dans l'élicitation des réactions de défense chez la plante hôte.

Aussi, de plus en plus d'études avancent la perception de la PTI et de l'ETI comme un continuum dans lequel la distinction entre PAMPs et effecteurs, PRR et protéines R n'est pas si évidente (Thomma *et al.*, 2011). La défense des plantes résulte alors d'une expression conjointe ou successive de la résistance basale à large spectre et d'une résistance spécifique ciblée. Ainsi, le modèle en 'zig-zag' semble présenter des limites concernant la vision réductionniste qu'il avance afin de caractériser les interactions moléculaires entre agent pathogène et plante. De plus, il ne prend pas compte de l'influence des DAMPs dans la PTI, limitant alors son cadre de lecture aux interactions entre plantes hôtes et agents pathogènes biotrophes (Pritchard and Birch, 2014). La prise en compte des limitations soulevées pour le modèle en 'zig-zag' a mené à la formulation du **modèle d'invasion** qui apparaît plus approprié dans l'explication du continuum de réponses entre le PTI et l'ETI (Cook *et al.*, 2015). En considérant ce modèle (*Figure 12*), la plante reconnaît des 'motifs d'invasion' (= 'invasion patterns' - IP) dérivés des agents pathogènes pouvant correspondre à la fois à des effecteurs, des DAMPs ou MAMPs. Ces IP sont par la suite reconnus par des 'IP-triggered receptors' (IPTRs) représentant des PRRs comme des protéines R.



**Figure 12 :** Représentation schématique de la comparaison entre les modèles évolutifs en zig-zag (a) et d'invasion (b).

- (a) Le modèle en 'zig-zag' est défini sur des critères évolutifs cloisonnés opposant soit MAMPs et PRR (domaines structuraux de l'agent pathogène et conservés chez la plante, respectivement : boîte jaune), soit effecteurs et gènes *R* (domaines fonctionnels de l'agent pathogène et cultivar-spécifique chez la plante, respectivement: boîte rouge). Certaines molécules peuvent néanmoins échapper à cette classification et se retrouver hors du cadre défini.
- (b) Le modèle d'invasion définit un continuum comprenant tous les types d'interaction possibles entre IPR et motifs d'invasion (= 'invasion pattern').

(Cook *et al.*, 2015)

Abréviations: PAMPs/MAMPs, 'pathogen/microbe associated molecular pattern'; PRR, 'pathogen recognition receptor'; MTI/ETI, 'MAMP/effector -triggered immunity'; IPR, 'invasion pattern receptor'; IPTR, 'invasion pattern triggered response'

Comme explicité par le modèle en 'zig-zag', l'ETI a la particularité d'exercer une pression de sélection sur les agents pathogènes, conduisant à l'adaptation relativement rapide de souches contournantes face à une résistance mono-génique. Ceci mène alors à une baisse d'efficacité voire une érosion totale du gène de résistance *R*. De plus, la défense basale fait intervenir une multitude de mécanismes et de gènes sous-jacents dont la diversité allélique peut mener à des niveaux intermédiaires de défense, qui varient de façon continue.

La conjonction de ces deux mécanismes mène à un niveau de résistance incomplète, qualifié de résistance quantitative, dont le déterminisme génétique sous-jacent est gouverné par de multiples loci d'effet faible à modéré appelés '**Quantitative Resistance Loci**' (QRLs), cas particuliers de QTLs liés à la résistance aux maladies (Poland *et al.*, 2009).

## 2.4. Résistances quantitatives : déterminisme génétique

La distinction entre résistance qualitative et résistance quantitative se fait aisément sur le plan structurel : déterminisme monogénique ou oligogénique<sup>5</sup> *vs* fortement polygénique. La nuance est en revanche complexe à établir sur le plan fonctionnel, les deux types de résistances pouvant être conditionnés par les mêmes mécanismes recouvrant à la fois la reconnaissance spécifique entre effecteurs de l'agent pathogène et protéines cibles chez la plante et la résistance basale à large spectre (Kamoun *et al.*, 1999).

Les résistances qualitatives se basant sur un faible nombre de gènes *R* à large effet, l'identification précise de ces gènes majeurs est facilitée par la force du signal en cartographie, offrant de bonnes perspectives de clonage et de validation fonctionnelle. A l'inverse, de par la multiplicité des gènes à effet faible ou modéré sous-jacents, la cartographie de QRLs offre des signaux de détection plus hétérogènes, avec des intervalles de confiance souvent plus étendus. Ainsi, de nombreux QRLs ont été cartographiés à ce jour sans que le nombre et la nature des gènes impliqués n'ait pu être précisée, notamment dans le cas d'études à puissance statistique et résolution faible, considérant des populations de faible effectif (Corwin and Kliebenstein, 2017).

Néanmoins, plusieurs hypothèses ont été émises afin de caractériser les mécanismes moléculaires sous-jacents aux QRLs (Poland *et al.*, 2009). Il est assez peu probable que l'une puisse expliquer en totalité ces mécanismes. Ainsi, la résistance quantitative résulterait plus vraisemblablement d'une conjonction de plusieurs de ces hypothèses.

### 1) Les QRLs sont conditionnés par des gènes impliqués dans la morphogénèse et le développement des plantes.

De nombreuses études font état d'une influence du stade de développement dans l'expression de la résistance, la plante étant souvent plus sensible aux maladies durant la phase juvénile (pour revue, Develey-Rivière and Galiana, 2007). Ce phénomène appelé 'résistance ontogénique' est bien souvent ignoré, ce qui peut induire des effets confondants avec les autres mécanismes de résistance lors d'analyses de cartographie, du fait du caractère pléiotrope des gènes de développement. Ce mode de résistance peut être restreint spécifiquement à un type de tissu ou organe. Chez la vigne, l'infection des grappes par *Uncinula necator*, l'ascomycète responsable de l'oïdium, est impossible durant le processus de maturation, au cours duquel une résistance s'exprime suite à l'augmentation au-delà d'un certain seuil, de la concentration en sucres solubles (Salzman *et al.*, 1998). Chez le pommier, la maturité foliaire est corrélée positivement avec la résistance à *Venturia inaequalis* (Li and Xu, 2002). D'autres traits

---

<sup>5</sup> Lié à un faible nombre de gènes.

morphologiques et liés à l'architecture de la plante tels que la densité stomatique, la hauteur de la plante, la surface et l'angle de la feuille peuvent être fortement corrélés avec la résistance quantitative (Albar *et al.*, 1998; Zhu *et al.*, 1999; Bradley *et al.*, 2003; Melotto *et al.*, 2006).

## **2) Les QRLs consistent en différentes mutations ou formes alléliques de gènes impliqués dans la défense basale.**

La défense basale liée à la PTI se base sur la reconnaissance de MAMPs chez l'agent pathogène, permettant à la plante d'initier les premières réactions généralistes de défense. Parmi les récepteurs PRRs de la plante assurant l'interaction avec le MAMP, le rôle de FLS2, récepteur NBS-LRR de type RLK de la flagelline a été particulièrement étudié (Zipfel *et al.*, 2004). Des mutations au niveau de *FLS2* sembleraient être responsables de la variation quantitative de phénotypes liés à la sévérité de l'infection et de la charge bactérienne, lors de tests en conditions naturelles sur *A. thaliana* (Zipfel *et al.*, 2004). De plus, une diversité allélique naturelle de *FLS2* a été observée au sein de plusieurs *Brassicacées* (Dunning *et al.*, 2007). La variation allélique de *FLS2* pourrait donc être considérée comme un QRL au sein d'une descendance de ségrégation.

## **3) Les QRLs sont des composantes impliquées dans des voies biochimiques**

Les diverses phytotoxines (ex : syringomycine, coronatine) synthétisées par les agents pathogènes (ex : *Psy*) sont reconnues pour leur rôle dans le pouvoir pathogène. Chez la plante, les enzymes détoxifiantes ainsi que les composés antibiotiques (phytoalexines) sont déployés en réponse afin d'assurer la défense. Suite à la cartographie de QRLs dans le cadre du patho-système *A.thaliana* x *Botrytis cinerea* (Denby *et al.*, 2004), des études biochimiques ont permis de démontrer la corrélation positive entre synthèse de camalexine (une phytoalexine) et réaction quantitative de défense (Kliebenstein *et al.*, 2005). Bien que la relation de cause à effet n'ait pas été démontrée, le QRL identifié semble être gouverné par des gènes responsables de la synthèse de phytoalexines.

## **4) Les QRLs sont impliqués dans la transmission des signaux de défense systémique**

Comme abordé précédemment, les phytohormones JA, SA et ET jouent un rôle particulièrement important dans la transduction systémique du signal de défense

(SAR) du site infecté à l'ensemble des organes de la plante (Pieterse *et al.*, 2009; Gozzo and Faoro, 2013). Différents allèles de gènes impliqués dans ces voies de signalisation pourraient correspondre à des QRLs. Ainsi, des mutations sur le facteur de transcription WRKY33 intervenant dans la voie de synthèse du JA, ont mené à l'augmentation de la sensibilité d'*A. thaliana* à *B. cinerea* (Zheng *et al.*, 2006). A l'inverse, des mutations au niveau d'une MAPK chez *A. thaliana*, MPK4 a permis l'augmentation de la résistance à *Pto* (Petersen *et al.*, 2000). Ces mutations ont souvent un rôle antagoniste au niveau de la résistance vis-à-vis des nécrotrophes ou biotrophes.

### 5) Les QRLs sont issus de l'érosion partielle de gènes *R*

Différentes études semblent indiquer que les QRLs correspondraient à des formes affaiblies de gènes NBS-LRR (Li *et al.*, 1999; Fukuoka *et al.*, 2014). Cette hypothèse est en partie étayée par la colocalisation de QRLs et de gènes *R* vérifiée à travers de nombreuses espèces : riz (Li *et al.*, 1999; Wisser *et al.*, 2005; Fukuoka *et al.*, 2014), *A. thaliana* (Debieu *et al.*, 2016), maïs (Xiao *et al.*, 2007) et pomme de terre (Gebhardt and Valkonen, 2001). Bien que la sélection ait tendance à favoriser le maintien de gènes *R* chez la plante, jugés efficaces de par l'importance de leur effet, les agents pathogènes ont la possibilité de contourner ces gènes *R*, les convertissant alors en QRLs. Il a été observé lors de plusieurs études, une baisse de la virulence du pathogène en présence d'un gène *R* érodé par rapport au cas où il est absent. Ces phénomènes de résistance résiduelle ont été observés chez plusieurs pathosystèmes : riz x *Xanthomonas oryzae* pv. *oryzae* (Li *et al.*, 1999), blé x *Puccinia graminis* (Brodny *et al.*, 1986) et blé x *Erysiphe graminis* (Nass *et al.*, 1981).

De nombreuses études ont avancé le caractère non spécifique des résistances quantitatives offrant une protection partielle très souvent contre un groupe d'agents pathogènes (biotrophes ou nécrotrophes) (Poland *et al.*, 2009). Des études de cartographie chez le maïs et le riz ont également pu mettre en évidence la colocalisation de QRLs ciblant plusieurs agents pathogènes indiquant un caractère potentiellement pléiotrope de certains clusters de gènes (bien qu'en raison de la faible résolution de détection, il est difficile de distinguer pléiotropie et liaison) (Wisser *et al.*, 2005, 2006). Toutefois, de plus en plus de cas d'identification de QRLs race-spécifique sont à noter indiquant un déterminisme proche de l'interaction gène-pour-gène, longtemps attribuée exclusivement aux gènes *R*. De tels QRLs ont été identifiés dans plusieurs pathosystèmes : *Brassica oleracea* x *Plasmiodiophora brassicae* (Rocherieux *et al.*, 2004), *Malus domestica* x *Venturia inaequalis* (Calenge *et al.*, 2004), *Capsicum annuum* x potato virus Y (Caranta *et al.*, 1997) entre autres. Aussi, une étude portée sur l'interaction riz x *Xanthomonas* évaluant la résistance de deux populations de cartographie à 10 souches pathogènes a pu mettre en évidence la

validité de nombreux QRLs à un set restreint de souches, les deux parents contribuant dans certains cas à la résistance à des souches différentes (Li *et al.*, 2006). D'un point de vue statistique, la variance phénotypique se décompose partiellement en une composante d'interaction **cultivar x race**, pouvant être expliquée d'un point de vue moléculaire par des interactions de type gène mineur-pour-gène mineur (Marcel *et al.*, 2008).

## 6) Les QRLs comportent des gènes de fonction encore non identifiée

Une étude sur le riz a permis de mettre en évidence un QRL recouvrant des gènes dont la fonction n'avait encore jamais été attribuée à la résistance aux maladies. Le QRL, *pi34*, relié à la résistance à la pyriculariose, a été réduit à une région de 65 kb contenant dix ORFs. Aucun des gènes candidats n'a pu présenter d'homologie de séquence avec des gènes de défense précédemment identifiés (Zenbayashi-Sawata *et al.*, 2007).

## 2.5. Durabilité des résistances : mécanismes et stratégies de déploiement

Le concept de résistance durable a été introduit par Johnson et Law en 1975 avec l'étude du comportement d'hybrides de blé face à la rouille jaune. Une résistance est dite durable lorsque son **efficacité** perdure pendant de **nombreuses années dans un environnement spatial étendu, à forte pression de l'agent pathogène** (Johnson and Law, 1975), favorisant *a priori* la sélection de variants virulents. La durabilité apparaît donc comme une mesure *a posteriori* correspondant à la durée nécessaire pour qu'une résistance soit contournée après son déploiement (Johnson, 1979). De ce fait, la notion de durabilité des résistances reste particulièrement complexe à caractériser. De nombreuses études ont alors cherché à identifier des critères prédictifs permettant d'évaluer les risques au contournement de ces résistances afin d'offrir des outils d'aide à la décision aux programmes d'amélioration génétique.

Un des premiers facteurs d'influence de la durabilité des résistances correspond au **déterminisme de l'interaction plante x pathogène**.

Les résistances quantitatives déployées en champ à travers le pyramidage de QRLs se sont en général avérées être plus durables que les résistances monogéniques (pour revues consulter St Clair, 2010; Niks *et al.*, 2015; French *et al.*, 2016; Pilet-Nayel *et al.*, 2017). Diverses stratégies de pyramidage chez de nombreuses espèces, combinant QRLs avec un gène majeur dans un fond génétique sensible (Miklas *et al.*, 2006; Palloix *et al.*, 2009; Barbary *et al.*, 2016), cumul de plusieurs gènes majeurs (Collard and Mackill, 2008; Ellis *et al.*, 2014), ou associant plusieurs QRLs (Thabuis *et al.*, 2004;

Wilde *et al.*, 2008) semblent être prometteuses en terme de durabilité. Bien que la résistance quantitative soit connue pour procurer une résistance partielle, dans certains cas le cumul de QRLs à fort effet peut mener à une résistance totale (Niks *et al.*, 2015) : exemple des QRLs *rx1*, *rx2*, *rx3* pour la résistance de la tomate à *Xanthomonas campestris* (Stall *et al.*, 2009).

L'augmentation de la durabilité liée aux résistances quantitatives est liée à plusieurs facteurs interagissant avec le **potentiel évolutif de l'agent pathogène**, deuxième facteur clé.

Les résistances quantitatives ont tendance à n'exercer une **pression de sélection limitée** sur l'agent pathogène contrairement aux résistances monogéniques (Pilet-Nayel *et al.*, 2017). Ainsi, l'adaptation d'un agent pathogène à un QRL donné mène à un bénéfice marginal dans le cas d'un pyramidage de QRLs et le variant n'aura pas tendance à augmenter en fréquence au sein de la population (Niks *et al.*, 2015). Ce bénéfice est également à relativiser par rapport à la fréquence du gène contourné au sein de la culture ; si ce dernier n'est présent qu'en faible proportion, l'avantage sélectif du variant pathogène est à relativiser (Niks *et al.*, 2015).

Aussi, la probabilité pour l'agent pathogène de présenter toutes les **mutations nécessaires** afin de contourner l'ensemble des QRLs est fonction du fond génétique de l'hôte ; plus le nombre de QRLs est important, plus cette probabilité est faible (Pilet-Nayel *et al.*, 2017).

De plus, l'adaptation évolutive nécessaire chez le pathogène est à balancer avec le **coût associé en terme de fitness** (Vera Cruz *et al.*, 2000; Leach *et al.*, 2001; McDonald and Linde, 2002). Si ce coût est trop important pour le pathogène, la probabilité de contournement de la résistance est amoindrie. Cet élément reste cependant à nuancer du fait de la redondance des effecteurs au niveau du génome, permettant de compenser le coût sur la valeur sélective de l'agent pathogène.

La durabilité des résistances quantitatives provient également vraisemblablement de l'importante **diversité des voies de défense** mobilisées par les QRLs dont la conjonction est difficile à contourner pour l'agent pathogène (Pilet-Nayel *et al.*, 2017).

Enfin, la combinaison de QRLs, induisant des mécanismes de résistances qui agissent **le long du cycle épidémiologique** de l'agent pathogène contribue à l'amélioration de la durabilité des résistances quantitatives (Palloix *et al.*, 2009; Mundt, 2014; Laloï *et al.*, 2017). La revue de Pilet-Nayel *et al.* 2017, propose un inventaire intéressant d'études ciblant l'impact, sur l'évolution de populations, de QRLs associés à des mécanismes de résistance affectant diverses étapes du cycle de vie de l'agent pathogène (latence, efficacité d'infection, colonisation, multiplication). A titre d'exemple, le pyramidage de 3 QRLs de résistances contre *V. inaequalis* chez le pommier, agissant à différents stades du cycle d'infection allant de la pénétration à des stades plus tardifs

comme la sporulation, confère une protection plus durable que les QRLs pris individuellement (Laloi *et al.*, 2017).

Le potentiel évolutif des agents pathogènes dépend grandement de leur cycle de vie et de la structure génétique des populations. L'étude de la **dynamique de ces populations** est déterminante afin de prédire leur aptitude au contournement de résistance (McDonald and Linde, 2002). Plusieurs forces évolutives gouvernent l'adaptation des populations pathogènes :

- 1) La **mutation**, qui permet l'apparition de nouveaux mutants dans une population, augmentant alors la probabilité d'avoir un variant virulent.
- 2) La **dérive génétique**, qui conduit à la disparition de certains allèles de façon aléatoire dans la population. La dérive est d'autant plus importante dans les populations à faible taille efficace de population  $Ne^6$ . Ainsi le maintien de populations de petite taille pourrait permet de réduire la vitesse de propagation des allèles de virulence de l'agent pathogène
- 3) La **migration**, qui consiste en des flux de gènes à des échelles plus ou moins larges en fonction de la capacité de dispersion de l'agent pathogène. La migration a pour conséquence : d'une part, d'augmenter la variabilité génétique dans les populations à partir de laquelle la sélection effectue son tri, et d'autre part, de permettre la dispersion de combinaisons alléliques favorables.
- 4) La **sélection**, qui favorise la fixation dans une population d'un nouveau variant (obtenu souvent par mutation) présentant un avantage sélectif.
- 5) La **recombinaison**, mise à l'œuvre lors de la reproduction sexuée, permettant l'association de nouvelles paires alléliques au sein de la descendance.

Le risque de contournement est davantage avéré en considérant des agents pathogènes présentant un fort potentiel de migration de gènes, caractérisés par des populations à  $Ne$  important, et un taux de mutation élevé (McDonald and Linde, 2002; Niks *et al.*, 2015). Les agents pathogènes disposant d'un mode de reproduction mixte (alternance de la reproduction sexuée et asexuée), tels que la plupart des champignons ont un potentiel de patho-adaptation encore plus élevé: les combinaisons génétiques favorables sont générées par la recombinaison en phase sexuée puis amplifiées en phase asexuée.

L'influence de ces forces évolutives sur les phénomènes de patho-adaptation est relativement bien étudiée chez les virus (García-Arenal and McDonald, 2003; Moury and Simon, 2011; Rousseau *et al.*, 2017). Des modèles prédictifs ont ainsi été élaborés afin d'estimer la durabilité des résistances sur la base de paramètres tels que le taux

---

<sup>6</sup> Nombre d'individus reproducteurs qui transmettent leurs gènes à la génération suivante dans une population.

de mutation, le nombre et la nature des mutations nécessaires au contournement ainsi que le coût de fitness occasionné (Fabre *et al.*, 2012*a,b*).

A l'inverse, les études portées sur la caractérisation du potentiel évolutif des champignons et bactéries sont bien moins courantes et le lien avec la durabilité des résistances reste difficile à établir. Chez les bactéries, il n'est pas irraisonnable de considérer la **mutation** comme une force évolutive particulièrement importante du fait des fortes tailles de population présentes à l'échelle de la plante (McDonald and Linde, 2002). Une étude portée sur le complexe *Psy* a permis de mettre en évidence le phénomène dit de 'lignées hyper-mutantes' à partir d'isolats de *P. viridiflava* prélevés sur plants de kiwi (Bartoli *et al.*, 2015*c*). Suite à croissance in-vitro, une proportion non négligeable des variants issus de ces isolats, a développé une résistance aux antibiotiques et perdu l'aptitude à la virulence sur haricot. Ce taux de mutation important, bien que pouvant potentiellement mener à une baisse globale de fitness de ces souches hyper-mutantes, pourrait dans certains cas contribuer à l'apparition de mutations bénéfiques à la patho-adaptation à des plantes résistantes (Baltrus *et al.*, 2017). La **recombinaison bactérienne par conjugaison** constitue également un facteur de risque. Enfin, le **transfert horizontal de gènes** semble également être un levier déterminant dans la structuration et l'évolution du génome de *Psy*, pouvant faciliter la patho-adaptation, comme cela a pu être suggéré chez les plantes ligneuses (Nowell *et al.*, 2016).

La compréhension de l'influence du potentiel évolutif des agents pathogènes dans le risque de contournement des résistances a mené certaines stratégies d'amélioration génétique à viser l'exploitation de **QRLs contrôlant l'évolution du pathogène** : soit en induisant des pressions de sélection contradictoire entre variants (Van *et al.*, 2013), soit en agissant sur la dérive génétique (Tamisier *et al.*, 2017), soit en contrôlant la fréquence de contournement (Quenouille *et al.*, 2014).

Au-delà des programmes d'amélioration ciblant des résistances quantitatives, la durabilité des résistances peut être également assurée par une stratégie parcimonieuse du déploiement de ces résistances en champ ou en verger (Mundt, 2014). La **gestion temporelle et spatiale du déploiement des résistances** (en particulier des gènes *R*) sur le terrain est une notion clé du fait de son impact direct sur le potentiel évolutif des agents pathogènes (McDonald and Linde, 2002; Sapoukhina *et al.*, 2009). L'usage d'associations variétales comprenant des gènes *R* ou QRLs différents (Sapoukhina *et al.*, 2013) ou la rotation spatiale et temporelle de gènes *R* (Papaïx *et al.*, 2011) sont des stratégies dont la modélisation semble indiquer une influence positive sur la durabilité des résistances. La gestion temporelle du déploiement des gènes de résistance, bien que concevable pour les cultures annuelles, demeure une limite pour les espèces pérennes en verger, à cycle de génération plus long.

La durabilité des résistances **en verger** est alors cruciale au vu de la longueur des schémas de sélection des espèces pérennes et de la durée de vie moyenne des arbres en verger de production. Dans les années 90, l'apparition d'une race de *V. inaequalis* (race 6) a permis de contourner la résistance monogénique à la tavelure, liée au gène majeur *Vf* et déployée pendant près de 50 ans chez le pommier (Parisi *et al.*, 1993). Cet épisode a largement contribué à rediriger les programmes de sélection des espèces fruitières vers la diversification des sources de résistance en exploitant les QRLs à résistance partielle en appui des gènes majeurs (Lespinasse *et al.*, 2002). Afin de construire des résistances *a priori* durables, la prise en compte des **interactions QRL x diversité de l'agent pathogène** semble alors primordiale. De plus en plus de dispositifs d'évaluation de la sensibilité face à plusieurs races (ou souches) de l'agent pathogène sont mis en place afin d'évaluer le caractère race-spécifique de QRL(s). Ces études chez les espèces fruitières sont à noter dans le cadre des pathosystèmes : *V. inaequalis* x pommier (Calenge *et al.*, 2004; Soufflet-Freslon *et al.*, 2008; Van *et al.*, 2013), *Xanthomonas arboricola* x abricotier (Socquet-Juglard *et al.*, 2013b) et *Erwinia amylovora* x poirier (Montanari *et al.*, 2016). Dans cette perspective, les QRL(s) non spécifiques à large spectre d'action semblent être à privilégier en sélection. Il est également intéressant de noter que dans le contexte épidémiologique en verger, différents groupes génétiques d'une même espèce pathogène peuvent être en **mélange**, comme cela a pu être mis en évidence pour *Psy* dans le cas d'abricotiers infectés (Parisi *et al.*, 2016). Aussi, le cumul des seuls effets additifs de QRLs race-spécifiques pourrait ne pas apporter une protection suffisante face à de potentiels **effets synergiques** résultant de la présence simultanée de différentes souches au contact d'une plante hôte. Les études de cartographie de QRLs basées sur des **co-inoculations** de plusieurs souches permettent ainsi de s'approcher des conditions réelles d'épidémiologie en verger. Ces études permettent également de suivre la dynamique de populations et de caractériser les pressions sélectives différentielles induites par les QRLs sur les mélanges de souches (Van *et al.*, 2013).

# Chapitre 3. Pathosystème abricotier x *Pseudomonas syringae* (*Psy*)

## 3.1. Abricotier

### 3.1.1. Origine, diversité naturelle et cultivée

De par son nom, *Prunus armeniaca*, on a longtemps attribué l'origine de l'abricotier à l'Arménie, zone de jonction géographique entre l'Europe de l'Est et l'Asie Centrale. Bien qu'il soit tout à fait probable que l'Arménie représenta une porte d'entrée de la culture en Europe, l'origine de l'abricotier est davantage à chercher au niveau de l'Asie Centrale, il y a près de 5000 ans (Faust *et al.*, 1998). A partir de ce foyer, l'expansion de la culture s'est alors faite à la fois vers l'Ouest et l'Est (Hormaza *et al.*, 2007).

Vavilov a suggéré l'existence de **trois centres d'origine** de l'espèce correspondant à la **Chine Centrale** (Chine et Tibet), au cœur de l'**Asie Centrale** (du Tien-Shan au Cachemire) et au **Proche-Orient** (Iran, Caucase, Turquie) (Vavilov, 1992). L'étude de la diversité de l'abricotier sur des critères géographiques et morphologiques a mené à l'identification de 4 à 6 groupes éco-géographiques : Asie Centrale, région Irano-Caucasienne, Europe, Dzunghar-Zailig, Chine du Nord et Chine de l'Est (Kostina, 1969; Bailey and Hough, 1975).

Le groupe le plus ancien et le plus **riche en diversité** correspond au groupe d'**Asie Centrale** ; la région regroupe plus de 500 cultivars issus de fonds génétiques sauvages (Zaurov *et al.*, 2013). L'importante diversité observable dans ce groupe résulte vraisemblablement des hybridations naturelles entre les fonds génétiques sauvages et partiellement domestiqués et leurs croisements avec les fonds génétiques cultivés de Chine et de la vallée de Ferghana<sup>7</sup> au cours d'une domestication vieille de milliers d'années (Zhebentyayeva *et al.*, 2003). Des études de diversité basées sur l'usage de marqueurs SSRs ont d'ailleurs permis de mettre en évidence une variabilité allélique significativement plus importante dans ce groupe d'Asie Centrale par rapport aux autres régions (Maghuly *et al.*, 2006). Aussi, l'isolement géographique liée au contexte géographique montagneux de la région, l'auto-incompatibilité des fonds génétiques présents, la nature des routes commerciales traversant l'Asie Centrale ainsi que le mode de propagation des vergers par semis sont autant d'éléments qui justifient la richesse allélique de ce groupe (Zaurov *et al.*, 2013).

---

<sup>7</sup> Vallée à la jonction entre l'est de l'Ouzbékistan, l'ouest du Kirgistan et le nord du Tadjikistan, regroupant 65% des vergers d'abricotiers d'Asie Centrale.

Le **second centre de diversification** génétique correspond à la région **Caucasienne** qui s'étend par la suite à l'Afrique du Nord, l'Espagne et l'Italie (Yilmaz and Gurcan, 2012). Il a été suggéré que la culture se soit disséminée du centre asiatique vers l'Ouest en traversant l'Iran et le Caucase au IV<sup>ème</sup> siècle avant J.C. à l'occasion des campagnes militaires d'Alexandre le Grand (Bailey and Hough, 1975). Ceci mène à un second phylum de diversification caractérisé par la mixture des fonds génétiques sauvages et domestiqués (Zhebentyayeva *et al.*, 2003).

Le **groupe européen**, considéré comme le plus jeune, présente la **variabilité phénotypique et génétique la moins importante** concernant un certain nombre de traits relatifs à l'architecture de l'arbre et la phénologie (Hormaza *et al.*, 2007; Zhebentyayeva *et al.*, 2012). Il se structure en trois sous-groupes : Europe Continentale, Bassin Méditerranéen et sous-groupe d'adaptation géographique réunissant des variétés présentant une base génétique plus large (Bourguiba *et al.*, 2012). La réduction de diversité génétique s'est notamment illustrée en comparant les groupes Irano-Caucasien et la partie sud du sous-groupe Bassin Méditerranéen (Hagen *et al.*, 2002; Bourguiba *et al.*, 2012). Cette perte s'élevant de 37 à 49% illustre une trace de domestication significative de la culture à travers le Bassin Méditerranéen, région symbolique d'expansion de la culture (Bourguiba *et al.*, 2012). L'adaptation au contexte climatique des zones continentales et méditerranéennes fut le facteur majeur impactant l'évolution de la culture. L'usage de la sélection clonale sur un faible nombre de cultivars spécifiques à une région donnée et de la multiplication par semis sur ces variétés adaptées a mené au développement de variétés locales à base génétique étroite et ayant un spectre d'adaptation géographique très limité (Zhebentyayeva *et al.*, 2012).

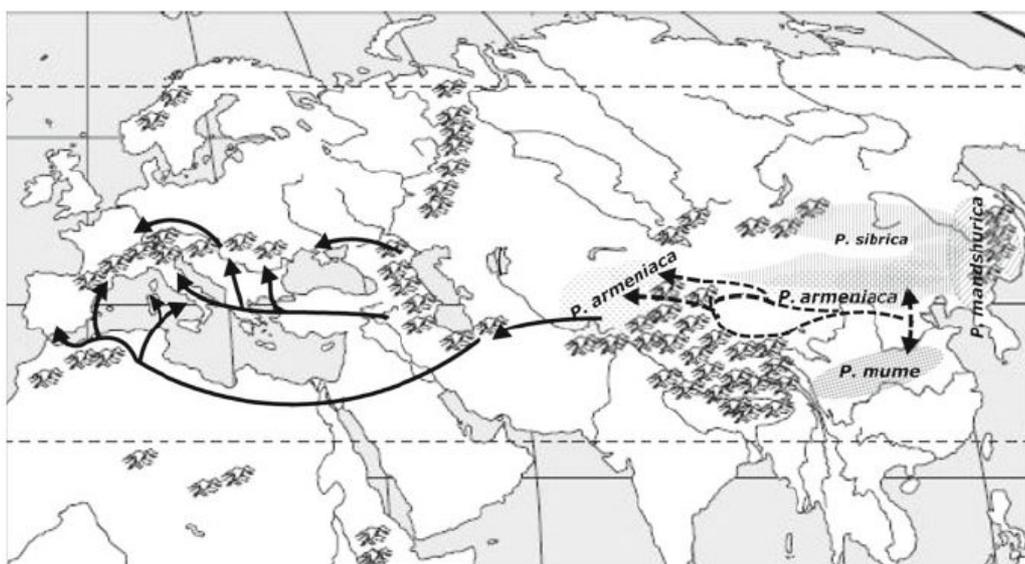


Figure 13 : Origine et dissémination de l'abricotier à partir des premiers centres de domestication.

(Zhebentyayeva *et al.*, 2012)

Deux différentes routes de propagation ont été proposées à partir de l'Asie Centrale (*Figure 13*) (Crossa-Raynaud, 1960). La **route africaine** a mené à la dispersion de la culture à travers le Moyen-Orient, l'Égypte et l'Afrique du Nord avec l'implantation de cultivars caractérisés par des besoins en froid limités pour la levée de dormance. Une deuxième route passant au **Nord de la Méditerranée** (Grèce, Europe Centrale puis Italie et France) a introduit des fonds génétiques nécessitant des besoins en froid plus variables mais globalement plus importants que les fonds génétiques du Bassin sud Méditerranée. Cette implantation de l'abricot au niveau des Balkans et en Europe Centrale est à relier d'abord à la présence romaine dans l'Antiquité puis a connu un essor considérable au XVI-XVIIe siècle avec la domination de l'Empire Ottoman dans la région (Faust *et al.*, 1998). L'apport de cultivars provenant de Turquie, plus particulièrement en Hongrie a été l'origine d'une sélection menant à l'obtention des fonds génétiques d'**Europe continentale** (Faust *et al.*, 1998).

Ces fonds génétiques constituent également la base génétique des cultivars développés en Amérique du Nord, Afrique du Sud, et Océanie avec l'apport des colons anglais, néerlandais et espagnols à partir du XVIIIe siècle (Bailey and Hough, 1975).

L'introduction de l'abricot en France résulte ainsi de deux courants historiques d'importation liés aux contraintes d'adaptation climatique de la culture sur le territoire (*Figure 13*). Le premier, situé autour de l'an 1000 est lié à l'apport, dans la région du Roussillon, de variétés du sous-groupe Bassin Méditerranéen par les Arabes, passant par l'Arménie ou l'Afrique du Nord et l'Espagne (Hagen *et al.*, 2002). Le second a lieu dès le XVIIIe siècle, introduisant des variétés des fonds génétiques turcs d'Europe Centrale et adaptés aux froids intenses (Faust *et al.*, 1998). Parmi ces fonds génétiques européens figurent des cultivars qui feront la renommée de l'abricot français à l'international: Pêche de Nancy (1755), Royal (nommé en 1815 en l'honneur du roi Louis XVIII), Rouge du Roussillon (1830), Luizet (1838), Paviot (fin du XIXe siècle) et Bergeron (1920) (Faust *et al.*, 1998). Hormis quelques cultivars en nombre limité dont la diffusion a été assurée par greffage, la plupart des fonds génétiques locaux sont issus de multiplications par semis, notamment du fait de la grande difficulté à assurer le greffage chez l'abricotier. Ils résultent de ce fait d'une sélection très fortement influencée par les interactions génotype x environnement justifiant les adaptations locales exacerbées rencontrées. Ce n'est qu'à partir des 30 dernières années que se sont développés des programmes qui ont conduit aux recombinaisons entre groupes géographiques.

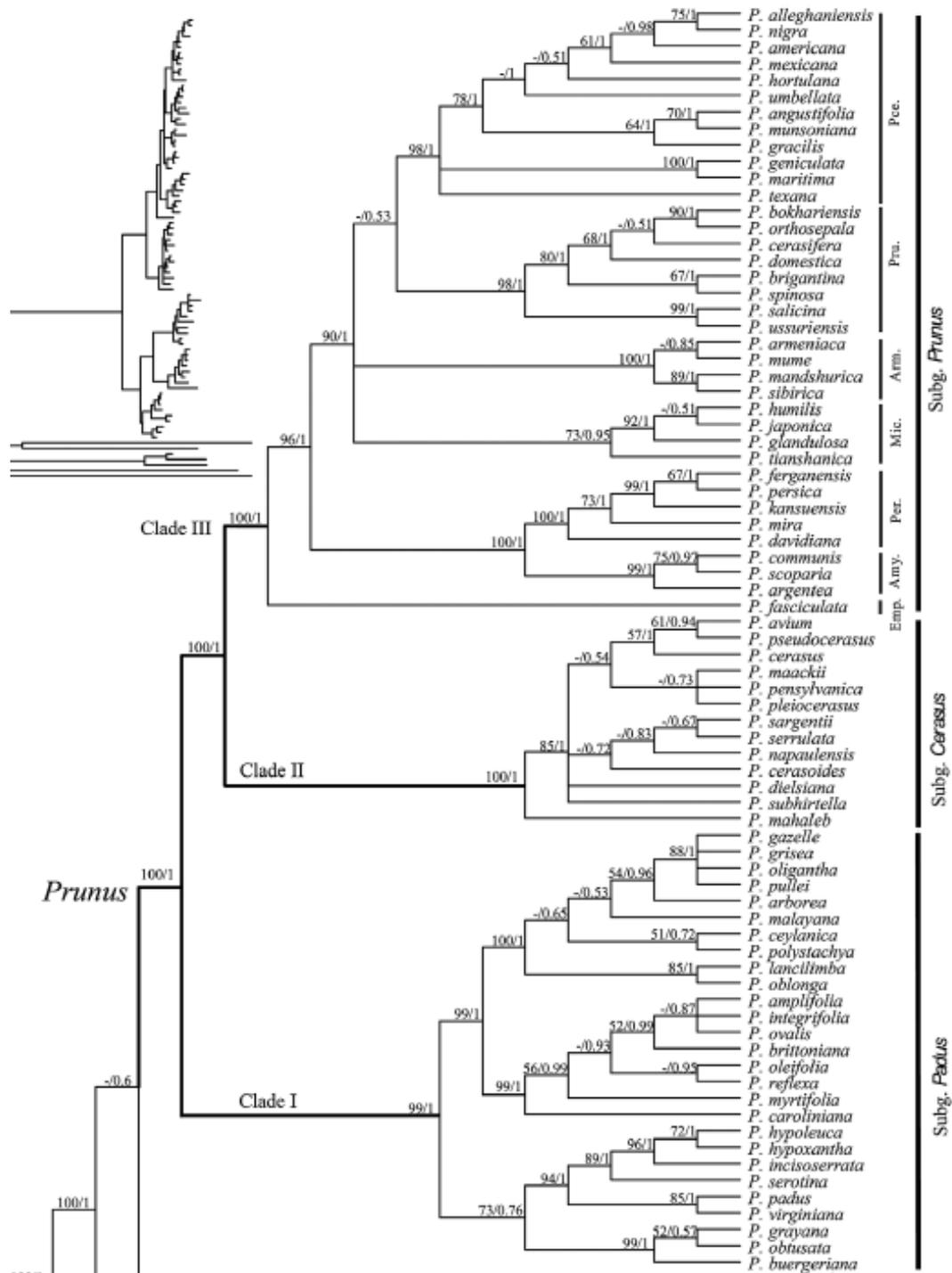


Figure 14 : Représentation d'un arbre bayésien illustrant les relations phylogénétiques entre *Prunus*.

L'abricotier (*Prunus armeniaca* L.) est représenté en orange.

(Shi et al, 2013)

### 3.1.2. Caractéristiques taxonomiques, botaniques et biologiques

L'abricotier commun *Prunus (P.) armeniaca* Linné (L.) appartient à la famille **Rosaceae** et la sous-famille **Prunoidea**. Selon la classification de Rehder, cette espèce est classée dans le genre **Prunus**, sous-genre **Prunophora** et dans la section **Armeniaca** (Rehder, 1949). Le genre *Prunus*, correspondant aux fruits à noyaux, comprend également les espèces proches cerisier (*P. avium* L.), prunier (*P. domestica* L.), amandier (*P. dulcis*) ainsi que pêcher (*P. persica* L. Batsch)(*Figure 14*). La section *Armeniaca*, au sens de Rehder (1949), comprend également quatre autres espèces apparentées à l'abricotier commun que sont l'abricotier Japonais (*P. mume*), l'abricotier noir (*P. dasycarpa*: hybridation naturelle *P. armeniaca* L. x *P. cerasifera*), l'abricotier de Briançon (*P. brigantina*) et l'abricotier du Tibet (*P. holosericea*). D'autres espèces ont été par la suite regroupées sous la section *Armeniaca*: l'abricotier de Sibérie (*P. armeniaca* var. *sibirica*), l'abricot de Manchourie (*P. armeniaca* var. *mandsurica*) et l'abricotier d'Ansu (*P. armeniaca* var. *ansu*).

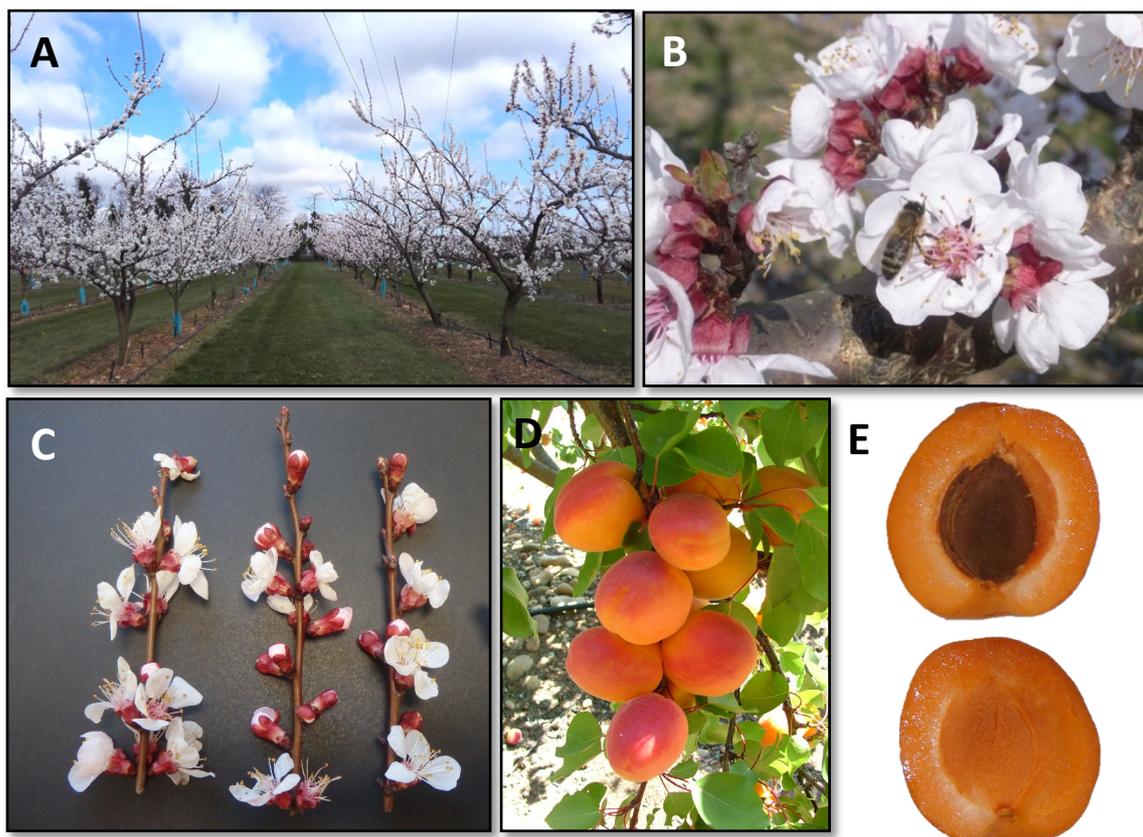
A l'origine, l'abricotier sauvage présente des conditions optimales de croissance dans les régions montagneuses présentant un climat continental caractérisé par des étés chauds et secs, et des hivers froids, justifiant son origine en Asie Centrale (Bailey and Hough, 1975). L'espèce est particulièrement résistante au froid et supporte des températures atteignant -30°C en hiver (Folta and Gardiner, 2009). Sa domestication a permis une plus large **extension climatique et géographique**, puisqu'on le retrouve désormais aussi bien sous les latitudes continentales du Canada qu'au niveau des oasis du Maghreb. Néanmoins, cette migration de l'espèce s'est accompagnée d'une **diminution du spectre d'adaptabilité pédoclimatique** des cultivars de sorte que, la plupart des fonds génétiques sont dans l'incapacité de produire régulièrement hors de leur zone optimale d'implantation (Bailey and Hough, 1975; Mehlenbacher *et al.*, 1991).

L'arbre présente une hauteur comprise entre 4 à 25 m (*Figure 15A*). La fleur est périgyne<sup>8</sup>, de couleur blanche ou rosâtre et comporte 5 sépales, 5 pétales et un unique pistil (*Figures 15B et 15C*). La levée de dormance et l'induction florale font suite à l'exposition à une certaine quantité de froid puis de chaleur correspondant à des paramètres dépendants du fond génétique (Ruiz *et al.*, 2007; Campoy *et al.*, 2012). Les besoins en froid selon les variétés peuvent varier de 300 à 1200 CU ('chilling units': nombre d'heures en dessous de 7,2°C selon le modèle de Weinberger (Weinberger, 1950)) et sont un facteur déterminant dans l'évolution phénologique de l'arbre (Campoy *et al.*, 2012). Ce paramètre conditionne l'implantation des cultivars

---

<sup>8</sup> Fleur ayant les étamines et les ovaires au même niveau que les carpelles.

en fonction du contexte climatique de la région afin d'éviter les pertes liées à une levée de dormance précoce suivie de gelées tardives (Hormaza *et al.*, 2007). La floraison de l'abricotier survient relativement tôt au printemps, de fin février à début avril ; elle est suivie par un développement du fruit s'étalant sur 3 à 6 mois selon la variété. Les cultivars commerciaux sont multipliés par greffage permettant notamment de conserver un bon niveau d'homogénéité et de protéger le scion des attaques de pathogènes telluriques.



**Figure 15 : Photographies de verger en période de floraison (A), de rameaux à fleurs (B et C) et de fruits d'abricotier (D et E).**

*Crédits photographiques: Laurent Brun, Eric Martin, Alain Blanc et Doriane Dam*

Le fruit est une drupe, pesant au niveau des fonds génétiques cultivés, de 40 à 100 g (*Figures 15D et 15E*), dans laquelle le noyau est constitué d'une amande protégée par un endocarpe lignifié, lui-même entouré du mésocarpe comestible et de l'exocarpe à aspect velouté, glabre ou pubescent selon les variétés. L'abricot est un fruit climactérique, nécessitant une récolte, une fois que le processus de biosynthèse de l'éthylène est engagé afin de permettre sa maturation après récolte. Riche en composés phénoliques aux propriétés antioxydantes (anthocyanes, procyanidines, flavonols), caroténoïdes, pro-vitamine A, et en composés minéraux (K, Mg en majorité), l'abricot

est également reconnu pour sa valeur nutritionnelle et ses bénéfices sur la santé humaine (Sochor *et al.*, 2010; Roussos *et al.*, 2016).

### 3.1.3. Importance économique

L'abricot correspond à la **septième production mondiale** parmi les fruits tempérés. La production mondiale est estimée à près de 3,9 millions de tonnes (Mt) en 2016 (FAO, 2018). Située loin derrière la production de pommes (près de 89,3 Mt) ou de pêches/nectarines (30,0 Mt), la production d'abricots est légèrement supérieure à celle des amandes (3,2 Mt) (FAO, 2018). Il s'agit d'une production majeure dans le **Bassin Méditerranéen**, zone géographique fournissant plus de la moitié de la production mondiale (FAO, 2018). Les principaux pays producteurs mondiaux d'abricots sont la **Turquie, l'Ouzbékistan, l'Iran et l'Algérie** assurant à eux quatre près de 62% de la production mondiale (FAO, 2018). La France, troisième producteur européen (157 000 tonnes récoltées en 2017) après l'Italie et l'Espagne et premier en abricots frais (non transformés), est également un opérateur majeur sur le marché mondial de l'abricot (FAO, 2018).

Son volume de production est en progression en réponse à la demande accrue du marché local français mais également des marchés à l'export. La très grande majorité est à destination du marché européen, faisant de la France, le deuxième exportateur mondial (après l'Espagne) reconnu pour son offre diversifiée et de qualité (FAO, 2018). Le verger occupe près de 8% de la surface de production fruitière nationale et les marchés exports rapportent près de 85 millions de dollars (FAO, 2018). De plus, d'un point de vue social, la filière abricot française représente en 2013 près de 5 700 emplois à temps plein (Interfel, 2013). La pérennisation de la qualité et du niveau des volumes de production proposés sur le marché international se pose donc comme un enjeu majeur pour la filière française.

La production française repose sur une trentaine de variétés dominantes parmi lesquelles le cultivar Bergeron figure de loin comme la variété de base représentant en 2009 près de 29% de la structure variétale du verger national (Lichou and Jay, 2012). La production sur le territoire national se concentre dans le quart Sud-Est au niveau des régions Occitanie (région du Roussillon et Gard), Provence-Alpes Côte d'Azur (Bouches-du-Rhône et Vaucluse) et Rhône-Alpes-Auvergne (Vallée du Rhône et Baronnies), assurant respectivement 29%, 14% et 55% de la production nationale (Ministère de l'Agriculture et de l'Alimentation, 2018). La récolte s'effectue successivement sur ces trois régions permettant ainsi d'assurer une campagne de commercialisation allant de mi-mai pour les variétés les plus précoces à mi-septembre pour les plus tardives.

### 3.1.4. Ressources génomiques et moléculaires

L'abricotier est une espèce diploïde, d'origine allogame, à 8 paires de chromosomes homologues ( $2n=16$ ) présentant un génome de petite taille (294 Mb/n) intermédiaire entre celui du pêcher (290 Mb/n) et du cerisier (338 Mb/n) (Arumuganathan and Earle, 1991; Folta and Gardiner, 2009). Il est également à peine plus long que le double de celui d'*A. thaliana* (119 Mb/n) (Haas *et al.*, 2005). L'abricotier est une espèce présentant un niveau d'hétérozygotie modéré, supérieur au pêcher (Hurtado *et al.*, 2002) et inférieur au cerisier, prunier ou amandier (Byrne and Littleton, 1989).

Chez les *Rosacées*, les premières études génomiques ont débuté avec l'avènement des marqueurs moléculaires et la construction de cartes génétiques notamment en considérant des populations de type pseudo-F1<sup>9</sup>, F2 ou Back-Cross (BC).

Le genre *Prunus* a pour spécificité de regrouper des espèces présentant une structure génétique commune et des génomes fortement colinéaires et synténiques comme le démontre l'ordre relativement conservé des marqueurs sur plusieurs cartes génétiques (Dirlewanger *et al.*, 2004; Illa *et al.*, 2009). Cette proximité génétique facilite les croisements interspécifiques, menant à l'obtention de nouveaux matériels fertiles tels que Plumcot (prune x abricot) ou Nectarcot (nectarine x abricot) (Folta and Gardiner, 2009).

La proximité des génomes a par ailleurs permis de construire une carte référence, la carte F2 pêcher 'Earliglow' x amandier 'Texas' (TxE) constituée de 826 marqueurs SSRs répartis sur une distance totale de 524 cM et transposables en majorité sur l'ensemble des *Prunus* (Dirlewanger *et al.*, 2004; Howad *et al.*, 2005). Cette carte sert de base pour la comparaison des positions de marqueurs ou d'intervalles contenant des gènes d'intérêt entre espèces du genre.

Le génome du pêcher est à ce jour, le plus caractérisé et le mieux connu parmi les *Prunus*. Son génome a été entièrement séquencé et publié en 2010 (version 1.0) (Verde *et al.*, 2013) puis reséquencé en 2015 (version 2.0) (Verde *et al.*, 2017), et se pose comme référence pour l'étude génomique des autres espèces de *Prunus*.

Concernant l'abricotier, un projet de séquençage, d'assemblage et d'annotation a été initié en 2018 et est actuellement en cours de réalisation. L'issue de ce travail devrait permettre de dépasser les limites et les approximations liées à l'adossement des séquences abricotier sur le génome du pêcher. Le génotype de référence choisi correspond à 'Marouch 14' (profondeur de séquençage 14X), un cultivar issu d'Afrique du Nord (Maroc) présentant un fort degré d'homozygotie globale. Les résultats préliminaires ont révélé une taille de génome plus réduite qu'attendue (208 Mb/n)

---

<sup>9</sup> Population issue du croisement entre deux parents hétérozygotes (n'étant pas des lignées fixées).

laissant supposer l'occurrence d'évènements d'insertion/délétion entre accessions provenant de différents groupes génétiques. A l'heure actuelle, l'assemblage a mené à la constitution des 8 chromosomes sur 61 scaffolds représentant 91% de la séquence *P. persica* v2.0 (Audergon, communication personnelle).

### 3.1.5. Objectifs et outils de l'amélioration variétale

#### 3.1.5.1. Sélection Assistée par Marqueurs (SAM)

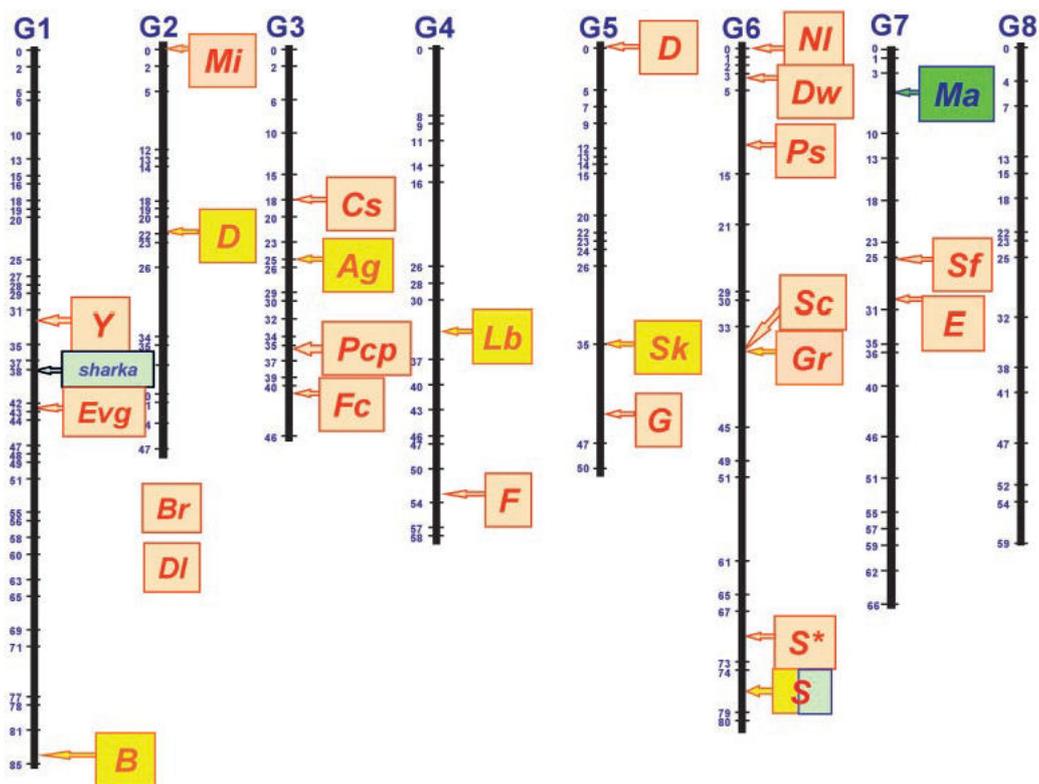
L'amélioration variétale chez l'abricotier, comme chez toute espèce pérenne, est contrainte par l'importance du temps entre générations et l'espace occupé pour le maintien en verger d'importantes populations parmi lesquelles seuls quelques génotypes présentent des aptitudes intéressantes au plan agronomique. La période juvénile de l'espèce en verger s'étale sur 3 à 5 ans avant que s'installe la production fruitière engendrant des frais de maintien en verger conséquents alors que toute évaluation phénotypique sur des critères de qualité et de phénologie est impossible.

Dans ce contexte, l'application des marqueurs moléculaires à travers la **sélection assistée par marqueurs** (SAM) s'avère particulièrement intéressante afin de réduire les coûts et accélérer le processus de sélection. La SAM se base sur l'identification précoce des plants portant les allèles favorables identifiés sur des marqueurs mis en évidence suite à une détection de QTLs ciblant un caractère d'intérêt (Dekkers and Hospital, 2002).

La SAM a été développée avec succès chez les espèces fruitières dans le cadre du projet RosBREED, lancé en 2009. Chez les *Prunus*, **28 gènes majeurs** affectant des caractères morphologiques et agronomiques initialement cartographiés sur différentes populations ont pu être localisés sur la carte de référence TxE grâce à la synténie observée entre les différents génomes (*Figure 16*) (Dirlewanger *et al.*, 2004). Les marqueurs identifiés servent comme base pour de multiples programmes de sélection de l'espèce.

La SAM, reconnue pour son efficacité afin de cibler des traits contrôlés par un faible nombre de gènes ou QTLs à effet majeur présente des limites dans le cas des traits fortement quantitatifs, gouvernés par un nombre plus important de gènes à effet mineur. Une méthode alternative, la '**sélection génomique**', a été développée récemment et présente l'avantage de s'affranchir des limites liées à l'importance des effets (Meuwissen *et al.*, 2001; Crossa *et al.*, 2017). Elle se base sur l'utilisation des données de marquage moléculaire pour établir un modèle de prédiction des traits d'intérêt. La sélection génomique a déjà démontré son efficacité afin de diminuer la durée des schémas de sélections chez de nombreuses espèces de grande culture (maïs, blé). Elle présente un intérêt tout particulier chez les espèces fruitières pour lesquelles quelques études préliminaires sont à relever chez le pommier et le pêcher (Kumar *et*

al., 2013; Muranty *et al.*, 2015; Biscarini *et al.*, 2017). Très récemment, le développement des outils moléculaires à haut débit (puces SNP 20K et 487K chez le pommier, 18K chez le pêcher notamment) a ouvert des perspectives favorables à l'approfondissement des approches de sélection génomique chez les espèces fruitières (Laurens *et al.*, 2018).



**Figure 16 :** Carte génétique de référence (TxE) des *Prunus* avec la localisation approximative de 28 gènes majeurs sur les 8 chromosomes.

Gènes cartographiés au sein de populations d'abricotiers (fond bleu), pêchers (fond orange), amandier ou hybride pêcher x amandier (fond jaune) et pruniers Myrobolan (fond vert).

(Dirlewanger *et al.*, 2004)

Abréviations: *Y*, chair sanguine; *B*, couleur des pétales et fleurs; *Mi*, résistance aux nématodes (pêcher); *D*, dureté de la coque; *Br*, caractère horizontal des branches; *DI*, fleur double; *Cs*, couleur chair autour du noyau; *Ag*, couleur de l'anthère; *Pcp*, polycarpelle; *Fc*, couleur de la fleur; *Lb*, date de floraison; *F*, adhérence de la chair au noyau; *D*, non-acidité du fruit; *Sk*, amertume de l'amande; *G*, pubescence de l'exocarpe; *Nl*, forme de la feuille; *Dw*, plante naine; *Ps*, stérilité mâle; *Sc*, couleur de l'exocarpe; *Gr*, couleur de la feuille; *S\**, forme du fruit; *S*, auto-incompatibilité; *Ma*, résistance aux nématodes (Myrobolan); *E*, forme du gland de la feuille; *Sf*, résistance à l'oïdium.

Les gènes *DI* et *Br* sont localisés au niveau d'une position inconnue sur le chromosome 2.

### 3.1.5.2. Cibles des programmes de sélection chez l'abricotier

Chez l'abricotier, jusqu'à peu, la majorité des programmes de sélection se basaient sur l'exploitation des ressources génétiques issues du groupe génétique Européen. Or l'importante diversité génétique que présente l'espèce au niveau des autres groupes, notamment d'Asie Centrale et du groupe Irano-Causasien peut être une source exploitable en amélioration (Bailey and Hough, 1975). Une grande partie des cartes génétiques développées ont été initialement destinées à caractériser le déterminisme génétique de la résistance à la sharka et de l'autofertilité, les deux cibles principales des programmes d'amélioration de l'espèce. Les objectifs de l'amélioration se sont ensuite étendus à la qualité, la phénologie et l'adaptation aux conditions climatiques, la résistance à d'autres bio-agresseurs.

#### 3.1.5.2.1. Auto-fertilité

L'abricotier est une espèce allogame et entomophile présentant un mécanisme d'incompatibilité gamétophytique, présent chez toutes les *Rosacées*. Cependant, les cultivars du groupe européen sont en majorité auto-fertiles (Kostina, 1969), ceci étant dû à l'occurrence d'une mutation génétique au niveau du locus *S* contrôlant l'auto-incompatibilité, qui par la suite a été sélectionnée (Vilanova *et al.*, 2006). Le locus *S* a été cartographié au niveau du chromosome 6 (*Figure 16*), en exploitant une population F2 d'abricotiers Lito x Lito (Vilanova *et al.*, 2003). Le locus *S* présente un polymorphisme allélique très important: un allèle unique *S<sub>c</sub>* code pour l'auto-compatibilité pollinique (Albuquerque *et al.*, 2002) et sept allèles *S<sub>1</sub>-S<sub>7</sub>* codent pour l'auto-incompatibilité (Burgos *et al.*, 1998). Le caractère d'auto-fertilité figure parmi les traits majeurs en sélection afin d'assurer une production régulière, non dépendante de la pollinisation entomophile qui peut être sérieusement impactée par de mauvaises conditions climatiques lors de la période de floraison.

#### 3.1.5.2.2. Phénologie et adaptation aux conditions climatiques

L'adaptation climatique devient un des critères majeurs d'amélioration des variétés en vue des bouleversements climatiques actuels et à venir. L'abricotier est une espèce caractérisée par une floraison précoce, limitant alors sa culture à des régions où les risques de gelées tardives sont négligeables. Ainsi, de nombreux programmes d'amélioration se consacrent à l'obtention de variétés à floraison tardive, afin de réduire le risque de perte de rendement lié aux gels dans les régions à froid printanier. L'allongement du calendrier de maturité est également un des objectifs de sélection à but d'approvisionnement continu du marché.

Comme déjà évoqué précédemment, le déterminisme de la floraison dépend de la satisfaction séquentielle des besoins en froid et en chaleur de la variété étudiée, qui sont deux facteurs dépendant du génotype (Ruiz *et al.*, 2007).

Plusieurs études QTLs ont été effectuées chez l'abricotier en ciblant plusieurs traits relatifs à la phénologie. Un QTL majeur contrôlant le besoin en froid, avec un pourcentage de variance expliquée (PVE) allant de 59 à 66% selon le parent, a été identifié dans une population F1 'Perfection' x 'A1740', au niveau du chromosome 1 (Olukolu *et al.*, 2009). Ce QTL colocalise avec une région identifiée par une étude analogue sur pêcher (Fan *et al.*, 2010), indiquant une conservation des gènes ('dormancy-associated MADS-box') contrôlant la dormance chez les deux espèces. Une autre étude effectuée sur une population BC1 'Z506-07' x 'Currot' a permis de localiser un QTL majeur sur le chromosome 5 de 'Z506-07' stable sur 3 ans et expliquant 60% de la variance totale du trait (Campoy *et al.*, 2011). Puis une étude de cartographie comparative de 3 *Prunus* (abricotier, cerisier et pêcher) a permis de mettre en évidence plusieurs QTLs contrôlant la date de floraison sur les chromosomes 2,3, 4 et 7 dont un majeur et stable sur le chromosome 4 chez l'abricotier (Dirlewanger *et al.*, 2012). Ce QTL semble être à proximité de régions contrôlant la date de maturité et la période de développement du fruit également identifiées sur le chromosome 4 au niveau d'une descendance 'Goldrich' x 'Currot' (Salazar *et al.*, 2016).

### 3.1.5.2.3. Qualité

De nombreux programmes se sont consacrés à la sélection de génotypes sur des critères de qualité, répondant aux attentes du marché du fruit frais: grosseur du fruit, chair ferme, résistance aux craquelures, couleur rouge attrayante (Hormaza *et al.*, 2007) mais également du marché de transformation : haute concentration de sucres solubles, couleur orangée, taille moyenne du fruit (Audergon, 1995). Malgré l'importance de ces activités de sélection et la multitude de traits ciblés, assez peu d'études sont à noter sur l'identification des QTLs contrôlant la variabilité de ces traits pour la plupart quantitatifs (Ruiz *et al.*, 2010; Cervellati *et al.*, 2012; Socquet-Juglard *et al.*, 2013a).

### 3.1.5.2.4. Résistance aux maladies

#### La sharka – Plum pox virus

Le programme de recherche et d'amélioration variétale le plus avancé chez l'abricotier en terme de compréhension des mécanismes de résistance aux maladies et d'identification des marqueurs associés est celui qui porte sur la lutte contre le **virus du plum pox** responsable de la sharka. Cette maladie, transmissible par puceron, et généraliste à l'ensemble des *Prunus* provoque essentiellement des lésions chlorotiques foliaires, une déformation des fruits et peut aboutir à une réduction significative du rendement. La maladie fait l'objet de mesures de quarantaine en Europe. Ainsi la lutte

actuelle repose sur l'arrachage obligatoire des arbres infectés ainsi que des arbres environnants, dès lors que la contamination atteint un seuil. Des études d'évaluation de la sensibilité semblent indiquer une résistance au niveau de variétés provenant d'Amérique du Nord mais dont les parents sont issus d'Asie Centrale. Des variétés issues du groupe d'Europe Centrale présentent également ce caractère de résistance (Martínez-Gómez *et al.*, 2000).

De nombreuses études de cartographie exploitant des croisements différents chez l'abricotier convergent sur l'identification d'un **QRL majeur *PPVres*** d'environ 196 kb, situé en haut du chromosome 1 (Hurtado *et al.*, 2002; Vilanova *et al.*, 2003; Lambert *et al.*, 2007; Marandel *et al.*, 2009; Dondini *et al.*, 2011). Ce QRL bien que majeur ne semble pas être suffisant afin de conférer la résistance (Decroocq *et al.*, 2014) et les recherches actuelles se consacrent à l'identification des régions génomiques complémentaires. Les études de cartographie de QRLs ont été par la suite complétées par une GWAS (Mariette *et al.*, 2016) et une étude transcriptomique (Rubio *et al.*, 2015) qui ont permis d'identifier une sélection de gènes candidats, certains sous-jacents au QRL *PPVres*. Très récemment, des travaux engageant des approches d'analyse d'expression (qRT-PCR), de transcriptomique et de génomique comparative entre cultivars résistants et sensibles ont mis en avant un mécanisme de *silencing* de deux gènes *ParPMC1* et *ParPMC2* comme conférant la résistance (Zuriaga *et al.*, 2018). Ces deux gènes, dont la protéine prédite correspond à une MATHd ('meprin and TRAF-C homology domain') font d'ailleurs partie des candidats des études de Mariette *et al.*, et Rubio *et al.* Les mécanismes cellulaires impliqués dans la résistance restent encore à élucider.

### **L'enroulement chlorotique de l'abricotier (ECA)**

L'ECA est une maladie de quarantaine à phytoplasme, très commune en Europe, pouvant causer l'obstruction des vaisseaux conducteurs et la mort des arbres contaminés (Morvan, 1977). Cette maladie est principalement transmise par un insecte vecteur, la cicadelle *Cacopsylla pruni* mais peut être également transmise par greffage. Cette maladie est très préoccupante puisqu'elle mène à près de 60 à 70% des phénomènes de mortalité en verger. Des travaux d'évaluation de la sensibilité au sein d'un panel variétal ont pu mettre en évidence des cultivars tolérants à la maladie (Audergon *et al.*, 1989).

### **La moniliose**

La moniliose causée par *Monilia (M.) laxa*, *M. fructicola* et *M. fructigena* correspond au principal facteur de perte de production de l'espèce au niveau mondial (Lichou and Jay, 2012). Cette maladie cryptogamique affecte à la fois les fleurs, les tiges et les fruits en cours de végétation mais également lors du stockage des fruits après récolte. Le développement de la croissance mycélienne et ainsi de la maladie est fortement corrélé à l'humidité relative lors de la période critique de floraison. Plusieurs sources

de résistance à *M. laxa* ont été identifiées et la résistance associée semble être quantitative (Conte *et al.*, 2004).

### **Autres maladies cryptogamiques**

L'oïdium causé par *Sphaerotheca pannosa*, la rouille, par *Tranzschelia discolor* et la tavelure par *Venturia carpophila* sont d'autres maladies cryptogamiques d'ordre secondaire pour lesquelles des études de caractérisation de la sensibilité ont été effectuées sans qu'aucune hypothèse n'ait été formulée sur l'architecture génétique sous-jacente.

### **La maladie des tâches bactériennes**

Les criblures bactériennes causées par *Xanthomonas arboricola* impliquent des défoliations et des dommages sur les fruits. Au-delà de l'état sanitaire, l'impact économique de cette maladie de quarantaine tient à l'état des fruits malades, devenant non commercialisables. Une première étude de cartographie a initié les activités de recherche de caractérisation de la résistance à la bactérie. Un QRL majeur contrôlant l'incidence de la maladie (PVE=53%) a été identifié sur le chromosome 5 de la carte génétique 'Harostar' x 'Rouge de Mauves' (Socquet-Juglard *et al.*, 2013b).

### **Le dépérissement bactérien**

Enfin, le dépérissement bactérien causé par *Psy* est une maladie majeure affectant l'état sanitaire des vergers, qui peut conduire à des phénomènes de mortalité des charpentières. Plusieurs sources de résistances ont été identifiées et l'identification des marqueurs sous-jacents fait l'objet de programmes de recherches en cours de réalisation.

## **3.2. Complexe d'espèces *Pseudomonas syringae***

### **3.2.1. Biologie, diversité et écologie**

Le dépérissement bactérien est une maladie causée par *Pseudomonas syringae* (*Psy*). Il s'agit d'un complexe d'espèces bactériennes, **hémibiotrophes**, appartenant à la classe des  $\gamma$ -Proteobactéries, aérobie stricte, à Gram négatif, en forme de bâtonnet et mobile grâce à l'existence d'un à trois flagelles polaires (Young, 2010). Il s'agit de l'espèce bactérienne qui est à l'origine des pertes économiques agricoles les plus importantes et sur laquelle porte actuellement le plus grand nombre de recherches scientifiques dans le monde de la phytopathologie (Mansfield *et al.*, 2012).

Actuellement, trois génomes ont été entièrement séquencés : en premier lieu *Pto* DC3000 (5 615 gènes) (Buell *et al.*, 2003), puis *P. syringae* pv. *syringae* (*Pss*) B728a (5 217 gènes) (Farrer *et al.*, 2009) et *P. syringae* pv. *phaseolicola* 1448A (5 144 gènes)

(Joardar *et al.*, 2005). Ils se posent comme les génomes de référence du complexe d'espèces et ont fait l'objet de nombreuses études d'interactions moléculaires plante x pathogène (O'Brien *et al.*, 2011). A ces trois génomes de référence s'ajoutent près de 300 ébauches de génomes assemblés accessibles sur des bases de données publiques telles que NCBI (<https://www.ncbi.nlm.nih.gov>), JGI (<https://genome.jgi.doe.gov/portal/>) et PPI (<http://www.pseudomonas-syringae.org>). La grande majorité des génomes de *Psy* ont une taille de 5 à 6 Mb (Baltrus *et al.*, 2017).

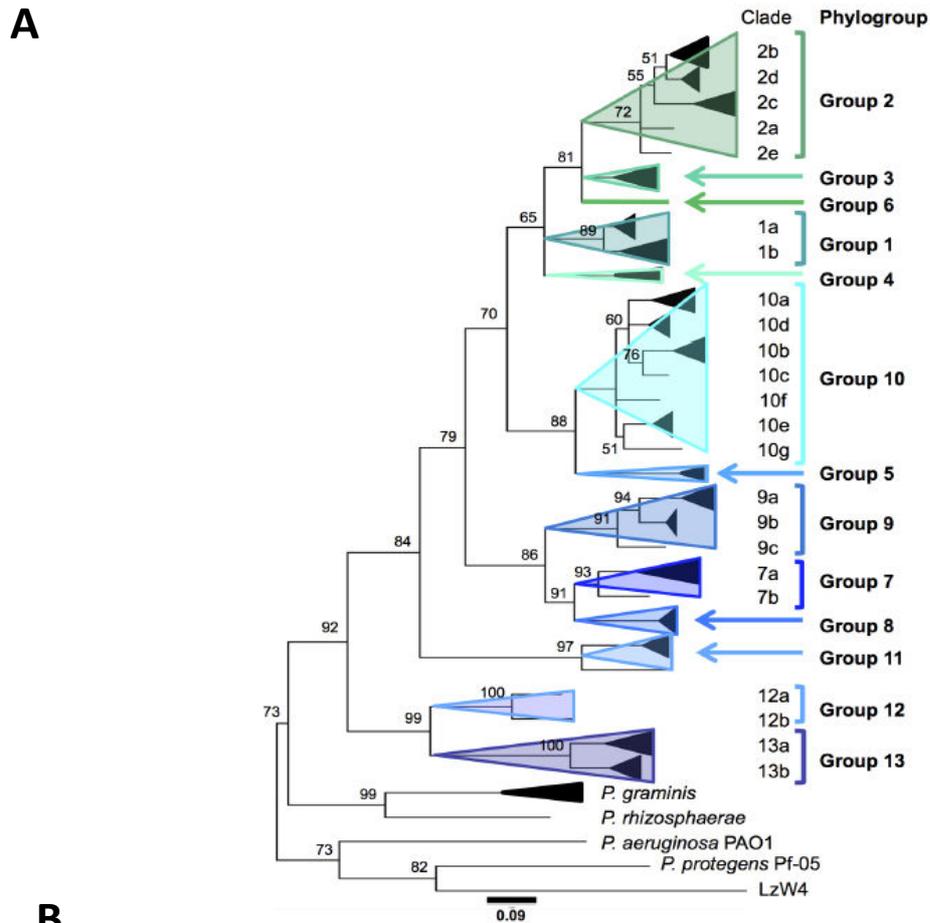
Le complexe *Psy* réunit près de soixante pathovars différents classifiés en fonction de leur pouvoir pathogène et de leur gamme d'hôte (Bull *et al.*, 2010; Young, 2010). Ces pathovars sont pathogènes sur plus de 400 espèces végétales parmi lesquelles figurent les *Prunus*, les *Malus* et d'autres espèces ligneuses telles que le saule (Gaignard and Luisetti, 1993; Kennelly *et al.*, 2007). La notion de **pathovar** (Young *et al.*, 1978) a été introduite afin de classier les bactéries phytopathogènes, dont la taxonomie est encore mal définie, en fonction de leur gamme d'hôte le plus souvent étroite et de comportements pathogènes semblables. Cependant, cette nomenclature est à nuancer car dans le cas de certaines espèces bactériennes, comme *Psy* plus particulièrement le pathovar *syringae* qui présente un large spectre d'hôtes, il n'existe pas de spécialisation entre souche bactérienne et espèce végétale hôte (Gaignard and Luisetti, 1993; Monteil *et al.*, 2013; Bartoli *et al.*, 2015*b,a*). Une importante variabilité génétique, biochimique et physiologique a été mise en évidence au sein de populations d'un même pathovar, prélevées au niveau de zones agricoles présentant des contextes pédo-climatiques et des productions végétales différentes (Roos and Hattingh, 1987). Aussi, et de façon plus surprenante, une importante variabilité phénotypique a également été mise en évidence entre des souches de la bactérie prélevées sur différents organes d'une même plante annuelle (Morris *et al.*, 2000).

La classification a récemment évolué pour se baser sur des analyses génétiques (hybridation ADN-ADN puis séquençage de gènes de ménage<sup>10</sup>), plus pertinentes afin de décrire les liens entre souches de *Psy* (Young, 2010). Une première classification basée sur la **phylogénie** d'un gène de ménage *rpoD* a reconnu une structuration en 7 **phylogroupes**<sup>11</sup> (PGs) (Parkinson *et al.*, 2010) auxquels ont été intégrés des espèces telles que *P. viridiflava*, *P. avellanae* et *P. savastanoi* dont la proximité génétique avec le complexe *Psy* avait déjà été démontrée (Gardan *et al.* 1999; Anzai *et al.* 2000; Mulet *et al.* 2010).

---

<sup>10</sup> Gènes conservés au sein d'une diversité (intra ou inter-spécifique) assurant des fonctions primordiales pour la survie de l'organisme.

<sup>11</sup> Groupe d'organismes apparentés du point de vue phylogénétique. Chez le complexe d'espèces *Psy*, les phylogroupes ont été délimités sur la base d'une distance génétique de moins de 5% au niveau des gènes de ménages.



**Figure 17 : Classification du complexe d'espèces *Psy*.**

A. Arbre phylogénétique de la classification actuelle de *Psy* en 13 phylo-groupes.

B. Tableau de profils caractéristiques des souches selon l'appartenance au phylo-groupe.

(Berge et al., 2014)

Abréviations: U, ubiquiste; P, plante; E, environnement; T3SS, système de sécrétion de type III; ina, activité glacogène; NA, non déterminé.

A ce jour, une classification plus récente du complexe d'espèces *Psy* se base sur une **division en 23 clades réparties au niveau de 13 PGs** (*Figure 17A*) dans lesquels s'inscrivent les 7 PGs précédemment caractérisés (Berge *et al.*, 2014). Cette classification résulte d'analyses phylogénétiques effectuées sur 4 gènes de ménage (*cts*, *gpaA*, *gyrB* et *rpoD*)<sup>12</sup> au sein d'une core-collection de 763 souches représentative de la diversité phénotypique et génétique du complexe d'espèces. Sur la base de cette classification, des outils moléculaires (amorces PCR) ont par la suite été développés afin d'identifier des souches de *Psy* et situer l'appartenance au PG, à moindre coût (Guilbaud *et al.*, 2016; Borschinger *et al.*, 2016).

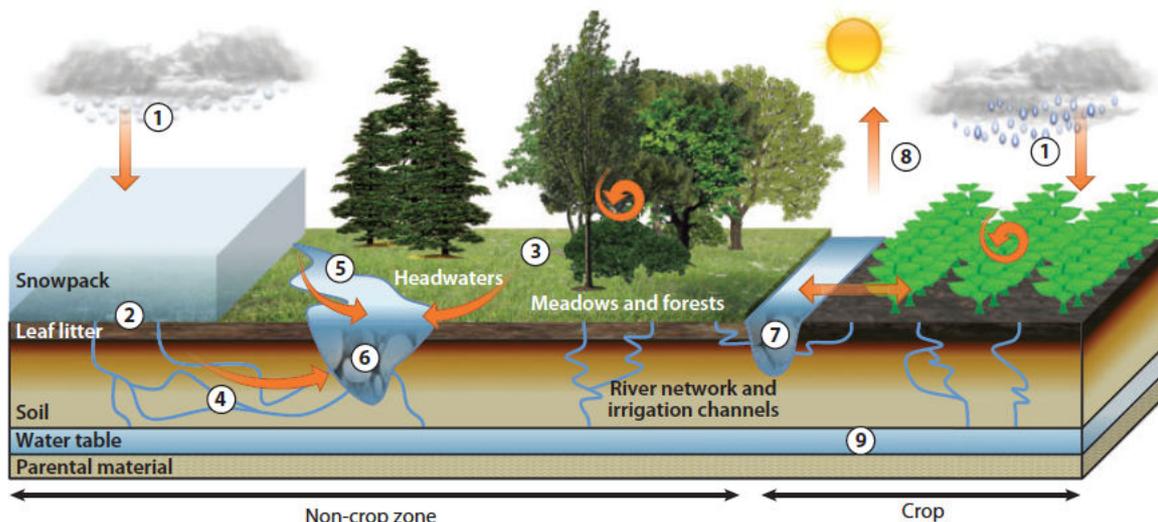
Le lien entre pouvoir pathogène et appartenance à un phylogroupe n'est pas encore élucidé (*Figure 17B*). Néanmoins, les PG 1,2,3, 7 et 8 regroupent une majorité de souches pathogènes et représentent ainsi un intérêt tout particulier dans les études d'épidémiologie de *Psy* (Berge *et al.*, 2014; Parisi *et al.*, 2016). Le PG 1 comprend de nombreuses souches isolées sur des plantes malades, parmi lesquelles *P.s. pv. tomato* (*Pto*) *P. avellanae* et *P.s. pv. actinidiae* (*Psa*), agents responsables de la mouchure bactérienne sur tomate, du chancre bactérien sur noisetier et kiwi, respectivement. Le PG 2 est le groupe le plus **ubiquiste**, comportant 5 clades, toutes rassemblant aussi bien des souches isolées au niveau de compartiments environnementaux (bassin d'irrigation, pluie, couvert végétal) qu'agricoles (plantes asymptomatiques et symptomatiques). On y trouve des *Pss* présentant une gamme d'hôte très large. Le PG 3 regroupe notamment des souches pathogènes des **espèces ligneuses**: *P.s. pv. morsprunorum* (*Psm*), *P. savastanoi* *pv. savastanoi*, *P.s. pvs. aesculi* et *mori*, agents responsables des chancres bactériens des *Prunus*, de l'olivier, du marronnier et du mûrier, respectivement. Enfin, les PGs 7 et 8 présentent une forte proximité génétique et rassemblent la plupart des souches de *P. viridiflava*, pathogènes sur une large gamme d'hôte.

*Psy* est une bactérie aérobie présentant une **grande diversité de niches écologiques** (*Figure 18*). Des travaux ont pu mettre en évidence la présence de souches bactériennes dans la neige et la pluie suggérant un lien étroit entre la bactérie et le cycle de l'eau (Morris *et al.*, 2008). Ceci pourrait alors expliquer la capacité de *Psy* à se propager sur de longues distances. Aussi, des souches bactériennes pathogènes ont été isolées au niveau de biofilms de rivières à l'extérieur des zones de production agricole suggérant alors une forte proximité entre les populations non-agricoles et agricoles de la bactérie (Morris *et al.*, 2007, 2013). Ainsi, les bassins versants alpins ont été mis en évidence comme moyens de dissémination de la bactérie vers les zones de culture en aval, dont la Vallée du Rhône fait partie (Monteil *et al.*, 2012). C'est

---

<sup>12</sup>*cts*: citrate synthase, *gapA*: glycéraldéhyde-3-phosphate déshydrogénase A, *rpoD*: RNA polymérase sigma70 factor, *gyrB* : gyrase B

pourquoi une bonne compréhension des modalités de colonisation, de développement et de survie de la bactérie sur les populations végétales semble être une condition nécessaire afin de maîtriser le risque de déclenchement de la maladie.



**Figure 18 : Représentation schématique de l'écologie de *Psy* à l'échelle du paysage.**

Les précipitations (pluie et neige) déposent *Psy* sur le sol (1). Dans les environnements alpins et subalpins, la bactérie survit au niveau du manteau neigeux (2) et dans les litières et le couvert végétal (3). Au cours de la fonte des neiges ou des pluies, une fraction des populations bactériennes migre vers les parties souterraines (4) et l'écoulement des eaux en surface qui alimente les cours d'eau alpin (5). *Psy* est ensuite transportée au niveau des colonnes d'eau et peut rejoindre les biofilms épilithiques (=attachés à un substrat non organique) (6). Les populations épiphytes de *Psy* sur des plantes sauvages peuvent également suivre cette voie après ruissellement lié à la pluie et à la rosée.

Dans les écosystèmes agricoles, des procédés similaires sont impliqués. En plus du déploiement provenant de plants, la bactérie peut migrer vers les plantes cultivées grâce aux précipitations (1) et l'eau d'irrigation des rivières (7). Les populations épiphytes sont ensuite transportées par masses d'air vers la troposphère (8). Des écoulements d'eau s'écoulent en profondeur et se nichent au niveau des nappes souterraines (9).

(Morris *et al.*, 2013)

Trois propriétés caractéristiques du complexe d'espèces *Psy* sont retrouvées partiellement selon l'appartenance au PG. Les différentes souches présentent en partie :

- (1) **un pouvoir pathogène** : elles sont susceptibles d'infecter le végétal et de déclencher une maladie grâce à leurs composantes de virulence et d'agressivité pouvant perturber la croissance et/ou la reproduction de la plante (Jones and Dangl, 2006).
- (2) **un pouvoir glaçogène** : elles sont capables de surélever le point de solidification de l'eau pure de -40°C à une température allant de -2 à -4°C grâce à leur propriété de noyau glaçogène codé par le **gène *ina*** (*ice nucleation activity*) (Maki *et al.*, 1974; Lindow, 1982). La protéine associée INP (*ice*

nucleation protein), ancrée au niveau de la membrane bactérienne permet l'interaction avec les molécules d'eau entourant la bactérie et leur orientation selon un maillage qui minimise la tension de surface, accélérant ainsi la prise en glace (Kajava and Lindow, 1993).

- (3) **un mode de vie épiphyte** : elles sont capables de survivre et de se multiplier à la surface des tissus végétaux (Hirano and Upper, 1990). Ces populations épiphytes sont souvent des prédicteurs de futures populations endophytes, à pouvoir pathogène latent, lors de transitions environnementales favorables (Rouse *et al.*, 1985; Hirano and Upper, 1990)

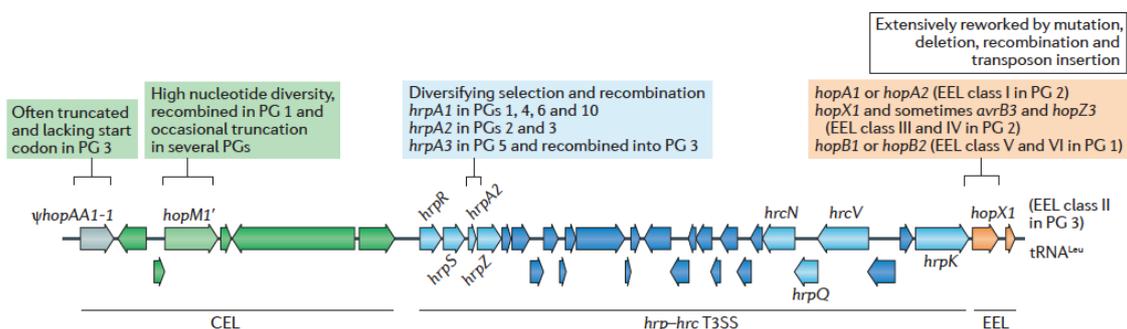
### 3.2.2. Composantes génétiques du pouvoir pathogène

Trois facteurs de virulence principaux sont à considérer dans le cas de *Psy*. Ces composantes du pouvoir pathogène ne semblent pas indépendantes de l'appartenance au PG (*Figure 17B*) (Berge *et al.*, 2014). Néanmoins, une variabilité est à noter entre souches d'un même PG, apportant une certaine complexité dans l'étude du lien entre évolution et acquisition de la virulence au sein du complexe *Psy*.

Le **pouvoir glaçogène** mis en association avec des températures de prise en glace des tissus a été mis en évidence comme étant le **principal facteur d'infection** chez plusieurs arbres fruitiers (pêcher, abricotier et poirier) en hiver (Weaver, 1978; Klement *et al.*, 1984; Montesinos and Vilardell, 1991). En présence de populations glaçogènes, le point de cristallisation de l'eau est augmenté permettant la prise en glace des tissus parenchymateux et corticaux à des températures supérieures d'environ 2-3°C aux valeurs habituelles figurant autour de -7°C (Prunier *et al.*, 2005; Lamichhane *et al.*, 2014). Les parois cellulaires végétales sont alors fragilisées et les épisodes de gel provoquent la lyse des cellules, offrant alors à la fois une porte d'entrée et une propagation de la bactérie au niveau des espaces intercellulaires et des vaisseaux conducteurs ainsi que des sucres favorables à sa multiplication (Klement *et al.*, 1984). Au-delà de l'intensité du froid, la fréquence et le nombre de chocs thermiques successifs serait vraisemblablement en relation avec le nombre et l'importance de l'infection des bourgeons chez l'abricotier (Prunier and Bordjiba, 1991). Le pouvoir glaçogène a été mis en évidence en tant que facteur de virulence chez *Pss* (PG 2) sur pêcher et abricotier (Weaver, 1978; Klement *et al.*, 1984). Il semble que cela soit également le cas chez *P. viridiflava* (PG 7) (Paulin and Luisetti, 1978). En revanche, *Psm* (PG 3) ne semble pas présenter de pouvoir glaçogène (Hirano and Upper, 1990; Scortichini *et al.*, 1995).

Le pouvoir pathogène de *Psy* passe également par la production de **protéines effectrices** intra ou extracellulaires issues du **système de sécrétion de type III** (T3SS) (Jin *et al.*, 2003). Ce système de sécrétion, similaire à une base annulaire

flanquée d'une aiguille, permet l'entrée des protéines effectrices de la bactérie dirigées contre le système de défense de la plante, directement dans les cellules de la plante hôte (Büttner and He, 2009). Ce T3SS est encodé par un complexe de gènes '*hypersensitive reaction and pathogenicity*' (*hrp*). Ces gènes, initialement identifiés chez *P.s. pv. phaseolicola* sont à l'origine du déclenchement de réactions d'hypersensibilité chez les plantes non-hôtes ou résistantes (mises en évidence sur des feuilles de tabac) et du déclenchement de la maladie dans le cas d'une interaction compatible (Lindgren *et al.*, 1986). Considérant *Psy*, le complexe *hrp* est constitué de 27 ORFs (Open-reading frames) organisés en 6 opérons sur le chromosome bactérien. On distingue 3 classes de gènes parmi les gènes *hrp* : (a) une classe de gènes qui codent pour les constituants de base du T3SS, comprenant **8 gènes *hrc*** ('*hrp genes conserved*') extrêmement bien conservés entre bactéries disposant d'un T3SS ; (b) des gènes (*hrpL*, *hrpR*, *hrpS*, *hrpV*) ayant des **fonctions régulatrices** de l'expression des gènes du T3SS ; et (c) des gènes codant pour des **protéines effectrices** (dont HrpA agissant au niveau extracellulaire) (Jin *et al.*, 2003). De plus, les gènes *hrp* sont entourés de deux régions complémentaires constituées d'effecteurs putatifs : '**conserved effector locus**' (CEL), conservés chez toutes les souches de *Psy* et '**exchangeable effector locus**' (EEL), variables selon les groupes génétiques. L'ensemble des 3 régions forme un îlot ('**pathogenicity island**') (Figure 19) dont la richesse du répertoire d'effecteurs chez une souche expliquerait en partie sa fitness et l'étendue de sa gamme d'hôte (Alfano *et al.*, 2000). 3 effecteurs *avrE*, *hopM1* et *hopAA-1* rentrent dans le cadre des motifs conservés formant le CEL (Lindeberg *et al.*, 2012). La région du EEL au niveau de l'opéron *hrp* a fait l'objet de multiples événements de mutations, délétion, recombinaison, et d'insertion de transposons sous pression des forces évolutives de co-adaptation aux récepteurs chez les différentes plantes hôtes (Xin *et al.*, 2018).



**Figure 19 : Représentation schématique de l'organisation et de la variabilité génétique de l'îlot de pathogénicité *hrp* chez *Psy*.**

(Xin *et al.*, 2018)

Abréviations: PG, phylogroupe; CEL, 'conserved effector locus'; T3SS, système de sécrétion de type III; EEL, 'exchangeable effector locus'

Enfin, la production de **phytotoxines** et d'**enzymes extracellulaires** est également un facteur de virulence contribuant au pouvoir pathogène de *Psy* (Kennelly *et al.*, 2007). Certaines phytotoxines, mises en évidence chez *Pss*, comme la **syringomycine** ou **syringopeptine** agissent en tant que toxine porogène de la membrane plasmique des cellules hôtes provoquant un déséquilibre des flux ioniques et plus particulièrement une entrée excessive des cations  $K^+$ ,  $H^+$ ,  $Ca^{2+}$ , fatale aux cellules (Hutchison *et al.*, 1995; Bender *et al.*, 1999). D'autres correspondent à des antimétabolites comme la **coronatine**, mise en évidence chez *Psm*. Cette molécule, interférant avec l'action du méthyl-jasmonate (MeJA) provoque la ré-ouverture des stomates au pathogène, dont la fermeture correspond à une des réactions initiales de défense physique de la plante face à l'agression (Bender *et al.*, 1999; Arrebola *et al.*, 2011). Après entrée de la bactérie dans l'apoplaste, il a été montré que la coronatine joue également un rôle dans l'inhibition des réactions de défense (dont celle propre au SA) et contribue aux symptômes chlorotiques externes (Mittal and Davis, 1995; Brooks *et al.*, 2004; Zheng *et al.*, 2012).

Récemment, des analyses de génomique comparative ont mis en évidence un opéron lié à la voie métabolique '**catechol oxygenase**', exclusif aux souches pathogènes des ligneux (Green *et al.*, 2010; Marcelletti *et al.*, 2011; Bartoli *et al.*, 2015b; Nowell *et al.*, 2016). Cet opéron permet à la bactérie de dégrader les composés phénoliques et composés aromatiques dérivés de la lignine, barrière structurale des espèces ligneuses, lui permettant ainsi de coloniser le système vasculaire. Cet opéron a d'ailleurs été identifié au niveau des génomes de *Psy* des PGs 1 et 3, rassemblant des souches connues pour être pathogènes d'espèces pérennes (*P.s. pvs. actinidiae*, *savastanoi*, *aesculi*, *morsprunorum* et *thaeae*) (Bartoli *et al.*, 2015b).

### 3.2.3. Evolution et patho-adaptation

Comme évoqué précédemment, le lien entre cumul des facteurs de virulence et gamme d'hôte ne semble pas évident car les composantes spécifiques du pouvoir pathogène très diverses ne semblent pas expliquer individuellement une contribution significative à l'étendue de la gamme d'hôte.

Il a été longtemps suggéré que l'étendue de la gamme d'hôte de *Psy* résulte de la variabilité qui a pu être mise en évidence au niveau des effecteurs (Alfano *et al.*, 2000; Greenberg and Vinatzer, 2003; Xin *et al.*, 2018). *Psy* se caractérise notamment par l'importante variabilité et richesse des répertoires d'effecteurs T3SS entre souches et la faible quantité d'effecteurs conservés (Baltrus *et al.*, 2011; Lindeberg *et al.*, 2012). Un inventaire complet des effecteurs du complexe *Psy* semble difficile à dresser compte-tenu de la grande diversité génétique existante. Néanmoins, les recherches

actuelles font état d'un **répertoire d'effecteurs du pan-génom**<sup>13</sup> de *Psy* comprenant **57 familles** avec une moyenne de 15 à 30 effecteurs au sein des génomes de souches individuelles (Lindeberg *et al.*, 2012).

A l'aide d'expérimentations de délétions d'effecteurs combinées de complémentations successives, un ensemble de seulement 8 effecteurs (*avrPtoB*, *hopM1*, *avrE*, *hopE1*, *hopG1*, *hopAM1-1*, *hopAA1-1*, *hopN1*) s'est avéré suffisant pour restituer la totalité de la virulence d'une souche mutante *effector-less* de *Pto* DC3000 sur des plants de tabac (Cunnac *et al.*, 2011). La délétion de nombreux effecteurs sans induire une perte de virulence de *Psy* suggère une **redondance fonctionnelle**, indiquant la potentialité que certains effecteurs ne soient utiles que face à certaines plantes hôtes (Lindeberg *et al.*, 2006). Une des hypothèses soulevées afin de justifier le maintien et la diversification de ces facteurs de virulence serait leur utilité face à l'**adaptation** à de nouveaux environnements fluctuants et à de nouvelles plantes hôtes. Ainsi de nombreuses souches environnementales, proches génétiquement de souches pathogènes peuvent présenter des répertoires d'effecteurs T3SS ou d'autres facteurs de virulence tels que l'opéron catéchol, très similaires, comme cela a pu être illustré avec *Pto* et *Psa* (Monteil *et al.*, 2013, 2016; Bartoli *et al.*, 2015*b*). Ceci illustre la potentialité d'**émergence de nouveaux pathotypes** suite à l'exposition de souches environnementales à de nouvelles plantes hôtes, en très peu d'évènements évolutifs (Monteil *et al.*, 2013; Morris *et al.*, 2013). Cette patho-adaptation peut être même accélérée via des transferts horizontaux de gènes ou plasmides (Diallo *et al.*, 2012; Nowell *et al.*, 2016).

Pour d'autres souches telles que *Pss* (PG 2), caractérisée par un répertoire d'effecteurs T3SS relativement pauvre pour une gamme d'hôte très large, la corrélation richesse du répertoire T3SS - gamme d'hôte semble être à nuancer (*Figure 17B*) (Berge *et al.*, 2014). *Pss* se révèle à l'inverse être remarquable par sa richesse en clusters liés à la production de toxines (syringomycine et syringopeptine) tout comme les souches membres du PG 10, se démarquant également avec le plus petit répertoire d'effecteurs du complexe (Hockett *et al.*, 2014; Berge *et al.*, 2014). La comparaison de souches à évolution indépendante, infectant la même plante hôte a mis en évidence l'existence de multiples configurations des facteurs de virulence pouvant provoquer la même maladie (Wang *et al.*, 2007*a*). Au-delà, de la virulence sur la même plante hôte, il est vraisemblable que cette multiplicité de configuration des facteurs de virulence pourrait également converger vers des gammes d'hôtes similaires.

---

<sup>13</sup> Gamme complète de gènes au sein d'une espèce donné.

Pan-genome = core-genome (gènes conservés présents chez toutes les souches) + flexible-genome (gènes propres à un nombre réduit de souches)

Récemment, deux études complémentaires ont permis de mettre en regard des données de caractérisation du pouvoir pathogène de souches sur cerisier (incluant des souches témoins non pathogènes) (Hulin *et al.*, 2018b) ainsi que les données génomiques respectives de ces souches afin d'identifier des gènes dont la présence/absence est significativement corrélée avec la virulence sur cerisier (Hulin *et al.*, 2018a). Ainsi il a été montré que le répertoire d'effecteurs de souches *Psy* pathogènes du cerisier était également réduit et que le pouvoir pathogène semblerait davantage résulter de la production de toxines (Hulin *et al.*, 2018a). La spécialisation des pathovars de *Psy* sur un certain nombre d'espèces, notamment ligneuses, semble résulter d'un processus de 'gain-loss' ou 'trade-off' au niveau de leurs répertoires d'effecteurs en particulier du EEL afin de maximiser la fitness en assurant leur croissance et transmission tout en évitant leur détection par le système de défense de la plante, notamment via les facteurs d'avirulence Avr (Arnold and Jackson, 2011; Nowell *et al.*, 2016).

Une sélection négative sur les effecteurs *avrPto*, *hopAB1* et *hopC1* a été relevée chez *Pss*, *P.s. pv. avii* et *Psm* races 1 et 2, pathogènes du cerisier alors que d'autres effecteurs *hopAR1*, *hopBB1*, *hopBF1* et *hopH1* ont été acquis via transfert horizontal de gènes, complétant ainsi l'action des toxines et augmentant alors l'agressivité de ces souches (Hulin *et al.*, 2018a). On notera que cette étude en plus des souches pathogènes du cerisier, a inclus près de 14 autres souches pathogènes d'autres *Prunus* (parmi *Psm*, *Pss*, *P.s. pvs. cerasicola*, *persica*, et *amygdali*). De manière très intéressante, excepté pour deux d'entre elles, *hopAB1* était également exclu au niveau de leur génome. Des expérimentations de suivi de développement bactérien sur des feuilles inoculées avec des transconjugants *Psm* exprimant individuellement *hopC1* et *hopAB1* ont indiqué une charge bactérienne significativement inférieure à celle présentée par le témoin ('empty vector'). Des réactions d'hypersensibilité induits par ces deux transconjugants ont également confirmé la fonction d'avirulence de *hopC1* et *hopAB1* au niveau de la plante.

La faculté généraliste que présentent les souches du PG 2 et PG 10 et leur nombre réduit d'effecteurs T3SS pourraient donc résulter d'une importante réduction au niveau du génome, des facteurs d'avirulence induisant l'ETI chez de nombreux hôtes. Une hypothèse complémentaire pourrait également consister en la préservation de la fonctionnalité de leur répertoire d'effecteurs, même minime, de par la faible occurrence des séquences d'insertions observée au niveau de leur génome. En effet, une étude portant sur dix souches des PGs 1, 2 et 3 (dont 2 du PG 2) semble indiquer une corrélation négative et significative liant le nombre de séquences d'insertion à l'étendue de la gamme d'hôte (Morris *et al.*, communication personnelle). La modification de la virulence sur feuilles de kiwi, résultant de désactivations de gènes (*hrpS*, *hrpR*) via des phénomènes d'insertion de séquences a déjà été observée au niveau de souches de *Psa* (Firrao *et al.*, 2018). Si cette hypothèse s'avère vérifiée, il serait intéressant d'apporter des éléments de compréhension sur la nature du (des) mécanisme(s) moléculaire(s)

permettant à de telles souches de se préserver de l'apparition de séquences d'insertion au niveau de leur génome.

### 3.2.4. Bilan : modèle du mécanisme d'infection de *Psy*

Sur la base des connaissances portant sur les interactions moléculaires plante x pathogène (explicitées en partie 1.2) et sur les facteurs de virulence précédemment abordés, un modèle général de transition de la phase épiphyte à la phase endophyte de *Psy* a été développé dans une revue très récente de Xin *et al.* 2018. La plupart des mécanismes abordés ont été mis en évidence lors d'études d'interaction entre *Pto* DC3000 et *A. thaliana* ou la tomate.

La *figure 20A* représente le mécanisme de l'entrée de *Psy* au niveau des chambres sous-stomatiques des feuilles via les barrières naturelles (stomates) ou les blessures.

Une reconnaissance de la flagelline (MAMP) par le récepteur FLS2 (PRR) va initier une réaction en cascade de type PTI menant à l'activation de la voie de signalisation liée au SA et la fermeture des stomates, en cas de résistance non-hôte de la plante (Melotto *et al.*, 2006). Afin de contourner la PTI, la coronatine qui est une toxine mimique du Me-JA va cibler le récepteur COI1 ou réprimer le récepteur JAZ afin d'activer la voie de signalisation antagoniste du JA (Katsir *et al.*, 2008). La répression de JAZ, également assurée par les protéines effectrices AvrB, HopBB1, HopX1 et HopZ1, va permettre de lever l'inhibition de l'expression d'un certain nombre de composantes intermédiaires : les facteurs de transcription MYC2, MYC3, MYC4, ANAC19, ANAC55 et ANAC72 (Zheng *et al.*, 2012; Jiang *et al.*, 2013; Gimenez-Ibanez *et al.*, 2014; Yang *et al.*, 2017).

L'expression des facteurs de transcription ANAC induit un contrôle négatif sur l'activité de ICS1 ('isochorismate synthase 1 chloroplastic protein') et positif sur BSMT1 ('salicylate/benzoate carboxyl methyltransferase 1'), tous deux impliqués dans la biosynthèse et le métabolisme du SA, respectivement. Cela a alors pour effet d'abaisser la concentration de SA au niveau de la chambre sous-stomatique, et d'inhiber la fermeture des stomates (Zheng *et al.*, 2012). L'effecteur AvrB agit également sur l'ouverture des stomates en activant la voie de signalisation du JA via la dégradation du récepteur JAZ, et en modulant la phosphorylation des protéines RIN4 (RPM1-interacting protein) et l'activité des ATPases membranaires (Lee *et al.*, 2015; Zhou *et al.*, 2015).

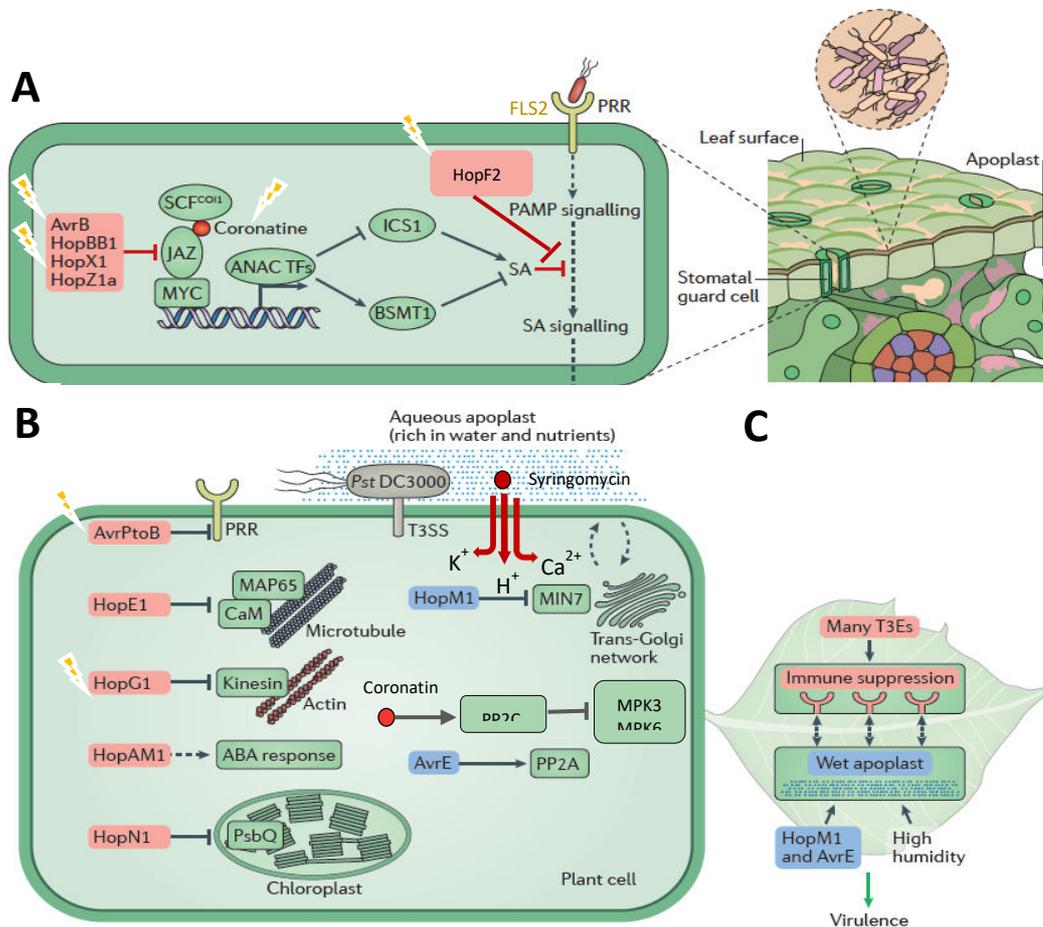
Enfin, les effecteurs HopF2 et HopM1 répriment la production de ROS et la fermeture des stomates (Hurley *et al.*, 2014; Lozano-Durán *et al.*, 2014).

Suite à l'entrée dans les feuilles, *Psy* se retrouve au niveau de l'apoplaste des cellules mésophylles<sup>14</sup> (*Figure 20B*). Des réactions associées à l'ETI et la PTI en réponse à la

---

<sup>14</sup> Cellules photosynthétiques localisées entre les épidermes supérieur et inférieur d'une feuille.

reconnaissance de la flagelline et des effecteurs T3SS permettent l'inhibition de la croissance bactérienne.



**Figure 20 :** Représentation schématique des mécanismes d'infection de *Psy* à l'entrée de la feuille via les stomates (A) et au niveau des cellules du mésophylle (B et C).

Les éclairs jaunes signifient l'existence d'interactions entre effecteurs et gènes *R* mis en évidence. Adapté de (Xin *et al.*, 2018)

Abréviations: PRR, 'pattern-recognition receptor'; SCF, 'Skp-Cullin-F-box protein complex'; COI, 'coronatine insensitive 1'; JAZ, 'jasmonate ZIM domain'; ANAC, 'NAM, ATAF and CUC (NAC)-domaine containing'; TF, 'transcription factor'; ICS1, 'isochorismate synthase 1 chloroplastic'; BSMT1, 'salicylate/benzoate carboxyl methyltransferase 1'; SA, acide salicylique; T3SS, système de sécrétion de type III; T3Es, effecteurs du T3SS; MAP65, '65kDa microtubule-associated protein'; CaM, calmoduline; ABA, acide abscissique; PsbQ, 'chloroplast protein photosystem II subunit Q'; MIN7, 'HopM1-interactor 7'; PP2A, 'protein phosphatase 2A'.

Bien que les mécanismes sous-jacents soient encore incertains, il semble que la production de phytoalexines, le contrôle négatif du T3SS et le renforcement des membranes cellulaires ainsi que la baisse de concentration des nutriments disponibles via l'activité des transporteurs de sucre SPT1 et SPT13 contribuent à la défense de la plante (Jones *et al.*, 2016; Yamada *et al.*, 2016). Afin de contrer ces réactions de

défense, la bactérie utilise à nouveau la diversité de ses effecteurs ainsi que la coronatine afin d'interférer avec la balance hormonale de la plante et inhiber la voie du SA. La coronatine exerce un contrôle positif sur la protéine phosphatase 2C (PP2C) qui déphosphorylise et désactive deux MAPKs, MPK3 et MPK6, régulatrices de la cascade de signalisation liée à la PTI (Mine *et al.*, 2017). La syringomycine, une autre toxine produite par certaines souches de la bactérie induit la formation de pores dans la membrane et agit comme biosurfactant afin de faciliter le mouvement bactérien (Hutchison *et al.*, 1995).

Sur les 8 effecteurs T3SS conservés de *Pto* DC3000, 5 sont impliqués dans la suppression des défenses de la plante. AvrPtoB inhibe à la fois la PTI et l'ETI en ciblant différents PRRs tels que le LRR FLS2, CERK1 ('chitin elicitor receptor kinase 1') et Bti9 ('LysM receptor-like kinase') et en provoquant la dégradation de MPK2 ainsi que l'inhibition des réactions d'hypersensibilité (Göhre *et al.*, 2008; Gimenez-Ibanez *et al.*, 2009; Zeng *et al.*, 2012). HopG1 et HopE1 ciblent deux constituants du cytosquelette : l'actine et les microtubules, en interagissant avec une kinésine et une calmoduline associée à la protéine MAP65 ('microtubule-associated protein 65'), respectivement (Guo *et al.*, 2016; Shimono *et al.*, 2016). Les deux interactions mènent à l'inhibition de différentes réactions de défenses dont le dépôt de callose sur les membranes cellulaires (Block *et al.*, 2010). HopAM1 induit une hypersensibilité à l'ABA au niveau de la plante et réprime également les dépôts de callose (Goel *et al.*, 2008). HopN1 cible la protéine chloroplastique PsbQ ('photosystem II subunit Q') et supprime la production de ROS et le dépôt de callose chez *A. thaliana* (Rodríguez-Herva *et al.*, 2012).

Les deux effecteurs HopM1 et AvrE ont également la capacité de supprimer les ROS et le dépôt de callose (Hauck *et al.*, 2003) mais il semble que leur principal rôle dans la virulence soit l'établissement d'un environnement aqueux au niveau de l'apoplaste par accumulation d'eau au niveau des espaces du parenchyme ('water soaking') (Figure 20C) (Xin *et al.*, 2016). HopM1 cible un facteur MIN7 impliqué dans le trafic de vésicules et localisé dans la face trans<sup>15</sup> de l'appareil de Golgi. AvrE interagit avec une protéine phosphatase PP2A. Le phénomène de 'water-soaking' favorise le développement de *Psy* de par la dilution des composés antimicrobiens et l'accès facilité aux nutriments (Xin *et al.*, 2018).

Près de 23 gènes *R* (Tableau 1) ont ainsi pu être mis en évidence et clonés dans le cadre d'études d'interactions entre diverses souches de *Psy*, essentiellement *Pto* DC3000 et diverses espèces modèles, essentiellement *A. thaliana* et tomate (Kourelis and van der Hoorn, 2018). Ces gènes *R* peuvent agir à différents étapes de l'infection selon les mécanismes qu'ils engendrent (Figures 20A et 20B).

---

<sup>15</sup> Face de sortie des vésicules sécrétées.

**Tableau 1 : Liste des 23 gènes de résistance *R* à *Psy* clonés.**

Certains gènes *R* peuvent interagir avec des effecteurs/co-récepteurs différents, et peuvent être impliqués dans la résistance face à plusieurs souches de *Psy* chez plusieurs plantes hôtes. - : donnée inconnue. D'après (Kourelis et van der Hoorn, 2018)

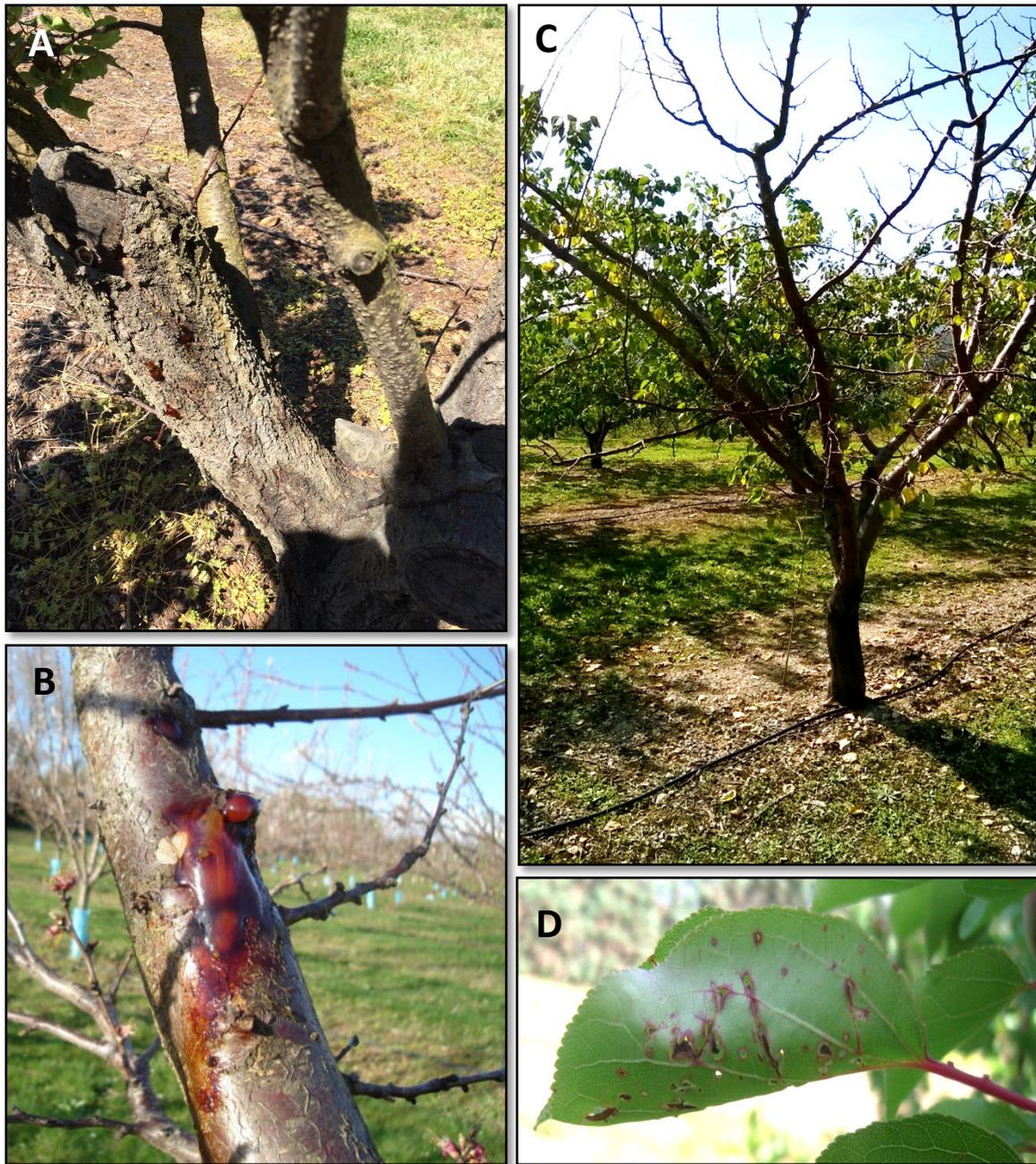
Récepteur	Effecteur	Souche	Co-récepteur	Hôte	Référence
AtRLP30	-	<i>pv. phaseolicola</i>	-	Arabidopsis	(Wang <i>et al.</i> , 2008)
Bti9	-	-	-	Tomate	(Zeng <i>et al.</i> , 2012)
COI1	Coronatine	-	-	Arabidopsis	(Zhang <i>et al.</i> , 2015)
Hth	-	-	-	Tabac	(Carmona <i>et al.</i> , 1993)
LecRK-IX.2	-	<i>pv. tomato</i> DC3000	-	Arabidopsis	(Luo <i>et al.</i> , 2017)
Prf	AvrPto	<i>pv. tomato</i> JL1065	Pto	Tabac	(Peart <i>et al.</i> , 2002)
Prf	AvrPtoB	<i>pv. syringae</i> B728a	Pto	Tomate	(Lin <i>et al.</i> , 2006)
Prf	AvrPtoB	<i>pv. tomato</i> JL1065	Pto	Tomate	(Lin <i>et al.</i> , 2006)
Prf	AvrPtoB	<i>pv. tomato</i> PT23	Pto	Tomate	(Lin <i>et al.</i> , 2006)
Prf	AvrPtoB	<i>pv. tomato</i> T1	Pto	Tomate	(Lin <i>et al.</i> , 2006)
Prf	HopPmaL/HopAB	<i>pv. maculicola</i>	Pto	Tomate	(Lin <i>et al.</i> , 2006)
Prf	HopAB3/AvrPtoB	<i>pv. syringae</i> B728a	Fen	Tomate	(Rosebrock <i>et al.</i> , 2007)
Prf	AvrPtoB	<i>pv. tomato</i> DC3000	Fen	Tomate	(Ntoukakis <i>et al.</i> , 2009)
Prf	AvrPto	<i>pv. tomato</i> JL1065	Pto	Tomate	(Salmeron <i>et al.</i> , 1994)
Prf	AvrPto	<i>pv. tomato</i> JL1065	LescPth5/LescPthF	Tomate	(Chang <i>et al.</i> , 2002)
Prf	HopAB2/AvrPtoB	<i>pv. tomato</i> DC3000	Pto	Tomate	(Kim <i>et al.</i> , 2002)
Prf	HopAB3/AvrPtoB	<i>pv. syringae</i> B728a	Fen	Tomate	(Chien <i>et al.</i> , 2013)
RBA1	HopBA1	<i>pv. aptata</i> str. DSM 50252	-	Arabidopsis	(Nishimura <i>et al.</i> , 2017)
Roq1	HopQ1-1	<i>pv. tomato</i> DC3000	-	Tabac	(Schultink <i>et al.</i> , 2017)
Roq1	XopQ	<i>pv. phaseolicola</i> 1448A	-	Tabac	(Schultink <i>et al.</i> , 2017)
Rpg1-b	AvrB	<i>pv. glycinea</i> race 0	GmRIN4	Soja	(Kessens <i>et al.</i> , 2014)
Rpg1r	AvrRpm1	<i>pv. maculicola</i> M6	GmRIN4	Soja	(Ashfield <i>et al.</i> , 2014)
Rpm1	AvrPpiA1	<i>pv. pisi</i>	RIN4	Arabidopsis	(Dangl, 1992)
Rpm1	AvrRpm1	<i>pv. maculicola</i> M6	RIN4	Arabidopsis	(Mackey <i>et al.</i> , 2002)
Rpm1	AvrB	<i>pv. glycinea</i> race 0	RIN4	Arabidopsis	(Mackey <i>et al.</i> , 2002)
RPS2	AvrRpm1	<i>pv. maculicola</i> M6	RIN4	Arabidopsis	(Kim <i>et al.</i> , 2009a)
RPS2	AvrRpt2	<i>pv. tomato</i> JL1065	RIN4	Arabidopsis	(Axtell and Staskawicz, 2003)
RPS5	AvrRpt2	<i>pv. tomato</i> JL1065	PBS1RCS2	Arabidopsis	(Kim <i>et al.</i> , 2016)
RPS5	AvrPphB/HopAR1/AvrPph3	<i>pv. phaseolicola</i> race3	PBS1	Arabidopsis	(Swiderski and Innes, 2001)
RPS6	HopA1/hrmA/hopP syA	<i>pv. syringae</i> strain 61	-	Arabidopsis	(Kim <i>et al.</i> , 2009b)
RRS1-S	AvrRps4/AvrPpiE	<i>pv. pisi</i> et <i>pv. tomato</i>	RPS4	Arabidopsis	(Narusaka <i>et al.</i> , 2009)
RTP1	-	-	-	Arabidopsis	(Pan <i>et al.</i> , 2016)
SILyk13	AvrPtoB	<i>pv. tomato</i> DC3000	-	Tomate	(Zeng <i>et al.</i> , 2012)
SNC2 (RLP51)	-	<i>pv. tomato</i> DC3000	-	Arabidopsis	(Zhang <i>et al.</i> , 2010b)
SUMM2	HopAI1	<i>pv. tomato</i> DC3000	CRCK3	Arabidopsis	(Zhang <i>et al.</i> , 2017)
Tao1	AvrB	<i>pv. glycinea</i> race 0	-	Arabidopsis	(Eitas <i>et al.</i> , 2008)
TN13	-	<i>pv. tomato</i> DC3000	MOS6	Arabidopsis	(Roth <i>et al.</i> , 2017)
TN2	-	<i>pv. tomato</i> DC3000	exo70b1	Arabidopsis	(Zhao <i>et al.</i> , 2015)
ZAR1	HopZ1a/HopPsyH	<i>pv. syringae</i> A2	ZED1	Arabidopsis	(Lewis <i>et al.</i> , 2013)
ZAR1	HopF2a	<i>pv. aceris</i> M302273PT	ZRK3	Arabidopsis	(Seto <i>et al.</i> , 2017)

### 3.3. Dépérissement bactérien de l'abricotier

#### 3.3.1. Symptômes

Le dépérissement bactérien, également appelé bactériose ou chancre bactérien, est une maladie touchant l'ensemble des organes aériens de l'arbre ainsi que l'appareil racinaire, dans une moindre mesure. La maladie se manifeste selon deux types de symptômes.

Certains se déclarent de **façon localisée** au niveau des tissus parenchymateux alors que d'autres atteignent la **vascularisation de l'arbre** (Lamichhane *et al.*, 2014). Dans le premier cas, les symptômes remarquables correspondent à des **lésions et criblures sur feuilles** (*Figure 21B*) et sur **fruits** apparaissant en fin de printemps et début d'été (Prunier and Cotta, 1987), à un **dessèchement des fleurs et bourgeons** (Hattingh *et al.*, 1989) ainsi qu'à des **nécroses apicales** (Golzar and Cother, 2008). Ces infections locales conduisent rarement à la mort de la plante mais portent davantage atteinte aux fleurs ainsi qu'à l'activité photosynthétique (Lamichhane *et al.*, 2014). Dans le second cas, lorsque les symptômes affectent le système vasculaire, ils sont d'ordre **systémique**. Cela traduit l'entrée et le déplacement de la bactérie à travers les vaisseaux conducteurs, la maladie peut alors devenir dévastatrice et conduire à la mort de l'arbre. Les symptômes caractéristiques correspondent à la **formation de chancre sur les rameaux et charpentières** (*Figure 21A*) associée à une **exsudation de gomme** (*Figure 21C*) (Kennelly *et al.*, 2007). Plus précisément, des nécroses et tâches d'aspect huileux apparaissent en hiver suivies de la formation au printemps d'un méplat correspondant à une zone d'aplatissement caractéristique du chancre. A l'intérieur du bois, les vaisseaux sont dégradés et brunissent en longueur à partir de la zone infectée. Lorsqu'un rameau est touché, sa survie dépend de sa capacité à ceinturer le chancre par la formation d'un **bourrelet cicatriciel**. Si la cicatrisation est complète, la vascularisation est maintenue (Erikson, 1945). Dans le cas contraire, un flétrissement puis un dessèchement complet du rameau (*Figure 21D*) fera suite à l'infection. Dans le cas des charpentières, l'infection mène le plus souvent à la mort de l'arbre lors de la réactivation des tissus, le printemps qui suit (Scortichini, 2002). Ainsi, au-delà de la perte de production fruitière que le dépérissement bactérien est susceptible d'engendrer, le problème majeur induit par cette maladie correspond à la perte de vergers associée au dépérissement des arbres. Cela concerne tout particulièrement les jeunes vergers, encore non productifs ou en début de production, âgés de un à cinq ans.



**Figure 21 : Photographies de symptômes de dépérissement bactérien sur abricotier.** Phénomène de mortalité de charpentière (A), exsudation de gomme sur rameaux (B), flétrissement de l'arbre (C) et criblures sur feuilles (D).

*Crédits photographiques: Jean-Marc Audergon & Mariem Omrani*

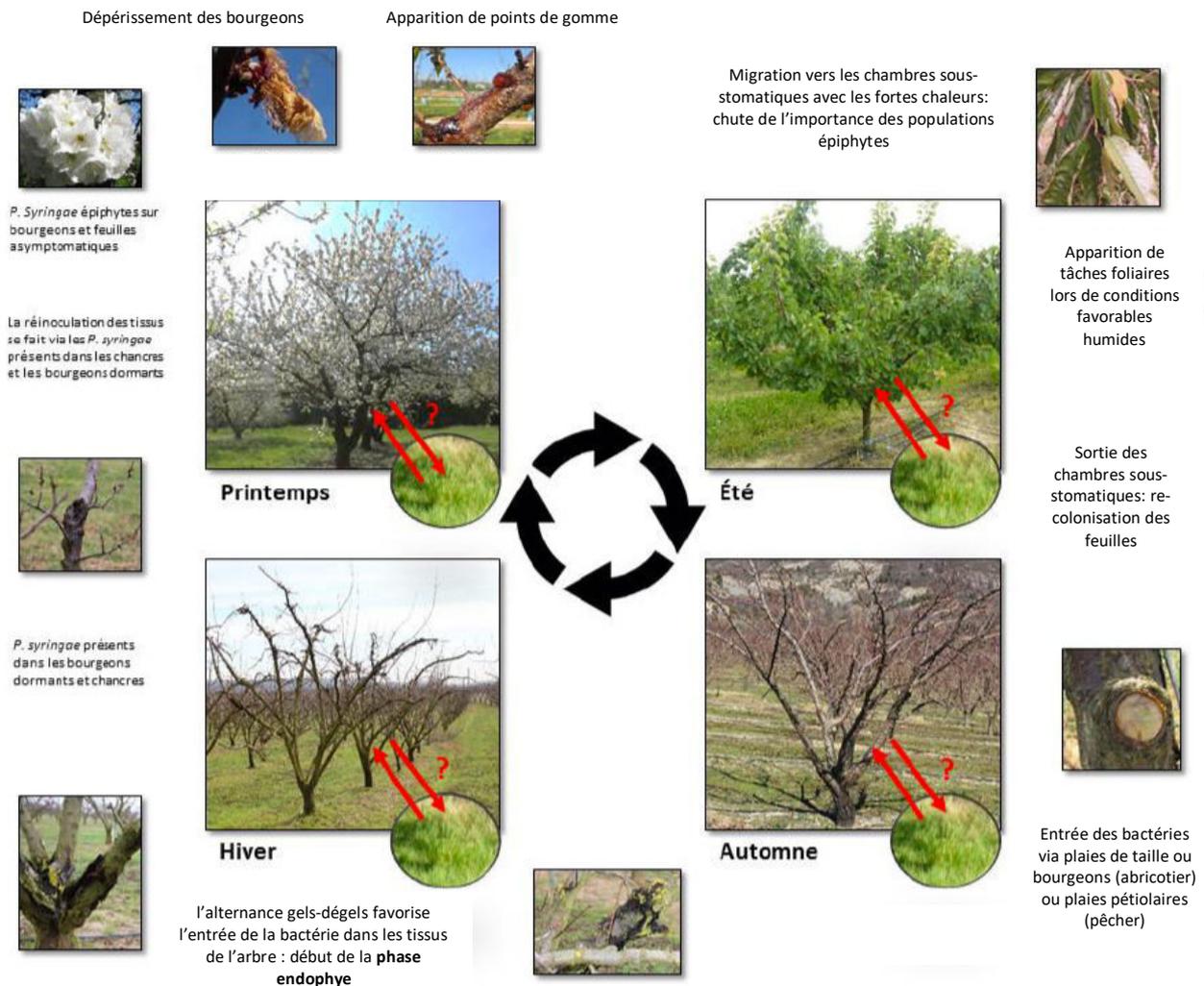
### 3.3.2. Epidémiologie

Les études sur le dépérissement bactérien ont été initiées par Brzezinski en Pologne au début du XX<sup>e</sup> siècle. Ce dernier a établi le lien entre les symptômes de gommose, le dépérissement de nombreux *Prunus* et une origine bactérienne de la maladie (Brzezinski, 1902). La maladie sévit sur de nombreux *Prunus* partout dans le monde: en Californie (Wilson, 1953) et en Nouvelle Zélande dès 1954 (Young, 1987). En France, elle est citée comme première cause de mortalité en Rhône-Alpes par Chabrolin dès 1922 (Lichou and Jay, 2012). Toutefois, les premiers dégâts majeurs et caractéristiques de la maladie sur abricotier ont été observés en France entre 1967 et 1973 au niveau de vergers de la vallée du Rhône (Prunier and Cotta, 1987).

Trois groupes génétiques différents dans le complexe d'espèces de *Psy* ont été identifiés historiquement comme étant responsables de la maladie sur les *Prunus*. Il s'agit de ***Pss*** (PG 2 clade b), ***Psm* races 1 et 2** (PG 1) (Vicente *et al.*, 2004; Kennelly *et al.*, 2007) et ***P. viridiflava*** (PG 7) (Prunier and Cotta, 1987). Néanmoins, du fait de sa large gamme d'hôte, il n'est pas exclu que d'autres groupes génétiques de *Psy* puissent également contribuer aux symptômes. Parmi les 13 PGs et 23 clades récemment caractérisés, les souches observées au niveau des vergers d'abricotiers appartiennent fréquemment à 4 PGs (PGs 1 clade a, 2 clade b, 3 et 7) avec une prévalence des souches pathogènes parmi le PG 2 (Parisi *et al.*, 2016). Le rôle des couvre-sols ainsi que des débris de culture en tant que réservoirs des PGs 2 et 7 est également à considérer comme source d'inoculum primaire.

*Psy* est une bactérie présente à la fois dans les tissus dans l'**apoplasme**, en tant qu'**endophyte** mais également sur les parties aériennes en tant qu'**épiphyte** (Hattingh *et al.*, 1989). Le cycle épidémiologique de *Psy* chez les arbres fruitiers est présenté au niveau de la *figure 22*. Des travaux conduits sur cerisier, abricotier, pêcher et amandier ont ainsi mis en évidence la présence des souches *Pss* et *Psm* à la surface de feuilles saines, de fruits ainsi que sur des brindilles et des adventices en verger (Crosse, 1959; English and Davis, 1960; Bordjiba and Prunier, 1991). Il semblerait alors que ces populations épiphytes observées dans la plupart des cas à la surface de plantes asymptomatiques puissent représenter la source d'**inoculum primaire** capable d'infecter les tiges et les branches (Crosse, 1959; Roos and Hattingh, 1986). Deux périodes de l'année sont propices au développement de la bactérie sur les feuilles: pendant la phase de croissance printanière lors de la sortie des feuilles ainsi qu'à l'automne (Cameron, 1970; Hattingh *et al.*, 1989; Bordjiba and Prunier, 1991). Les **chambres sous-stomatiques** semblent également représenter un habitat propice à la bactérie durant les périodes de forte chaleur. Elles s'y multiplieraient avant de sortir en masse à nouveau à la surface des feuilles (Hattingh *et al.*, 1989). Ces échanges dynamiques entre population « interne » et « externe » des feuilles sont à l'origine

d'une forte variation de la population épiphyte (Hattingh *et al.*, 1989). La bactérie subsiste ensuite au cours de l'hiver en s'installant au niveau des bourgeons dormants et des chancres (Sundin *et al.*, 1988).



**Figure 22 : Cycle épidémiologique du dépérissement bactérien des arbres fruitiers causé par *Psy*.**

Le rôle des couvre-sols en tant que réservoir d'inoculum primaire ainsi que les potentiels échanges entre populations sur la plante et celles du couvert végétal sont également représentés (flèches rouges).

*Illustration adaptée de la thèse de B. Borschinger (2016).*

Ainsi, les stades phénologiques et physiologiques de l'arbre sont particulièrement importants dans la dynamique de la population bactérienne et donc dans l'expression de la maladie. Les bactéries ont pour caractéristique d'emprunter le plus souvent un mode de pénétration opportuniste au niveau de la plante hôte. Ainsi elles utilisent le plus souvent des voies d'entrée naturelles telles que les lenticelles et les stomates ainsi que des lésions résultant de facteurs biotiques (opérations de taille, insectes) et

abiotiques (gel) (Crosse and Garrett, 1966). Les lésions sont d'ailleurs la voie la plus empruntée, elles permettent un accès direct à l'appareil vasculaire. Les voies naturelles, en revanche semblent ne pas être exploitées par toutes les sous-espèces de *Psy* (Lamichhane *et al.*, 2014). Bien que les plaies pétiolaires aient été mises en évidence comme étant une voie d'entrée commune au pêcher, cerisier et prunier (Cao *et al.*, 2013), il semble que cela ne s'applique pas à l'abricotier pour lequel la **porte d'entrée naturelle** privilégiée correspondrait aux **bourgeons** (Crosse, 1953; Bordjiba and Prunier, 1991).

### 3.3.3. Impact économique

Depuis les années 1970, la sévérité du chancre bactérien prend de plus en plus d'ampleur et limite le développement de la culture en France et plus généralement dans le monde. Ceci s'explique en partie par l'utilisation de nouvelles variétés, permettant une extension de cette culture rémunératrice au-delà des zones de production traditionnelles mais qui sont néanmoins plus sensibles, ainsi que par l'apport de nouvelles méthodes culturales (Prunier and Cotta, 1987; Scortichini, 2010).

En France, il s'agit de la **première cause de mortalité** des arbres en verger en zone septentrionale de la **Vallée du Rhône**, caractérisée par des hivers humides et froids, impliquant des pertes économiques importantes. Ces dernières années ont d'ailleurs été marquées par des attaques récurrentes: 2004, 2006, 2009, et plus particulièrement 2010, 2013 et 2015 au niveau des départements de la Drôme et de l'Ardèche. Une enquête 'bactériose' visant à recenser les vergers atteints par la maladie, a été menée en 2010 par la Chambre d'Agriculture de la Drôme et la FDSEA26 auprès de 126 arboriculteurs de la Drôme, de l'Ardèche et de l'Isère. Les résultats sont révélateurs de l'étendue et la gravité de la maladie dans la région: près de 49% de la surface totale (374 ha) de vergers d'abricotiers ont été touchés ; ceci ayant mené par la suite à l'arrachage volontaire de près de 6% de la surface totale des vergers (CA26, communication personnelle). Du point de vue économique, sur la base de données fournies par la Chambre d' Agriculture de la Drôme, les **pertes engendrées par la maladie en 2010** pouvaient atteindre **53 500€/ha** pour un verger d'abricotiers Bergeron en considérant les frais d'implantation et d'entretien du verger durant les trois premières années de non-production, la perte de production ainsi que le réinvestissement (CA26, communication personnelle).

### 3.3.4. Facteurs pédoclimatiques influençant le développement de la maladie

Les caractéristiques pédoclimatiques des régions sensibilisantes au dépérissement bactérien témoignent de conditions communes favorables au développement de la maladie. Les **températures négatives** interviennent dans le développement de la maladie soit en facilitant la pénétration bactérienne en lien avec **son activité glacogène**, soit en augmentant la sensibilité des tissus à la bactérie (Gaignard and Luisetti, 1993). Aussi les différents chocs thermiques associés à des **cycles de gel et dégel** favoriseraient la pénétration et l'infection bactériennes sur le pêcher et l'abricotier, en provoquant un **afflux d'eau** ou 'water soaking' au niveau des espaces intercellulaires du parenchyme (Klement *et al.*, 1974; Vigouroux, 1989, 1999).

D'autre part, des travaux antérieurs ont pu mettre en évidence le facteur aggravant causé par les **sols acides, peu profonds, de texture grossière** (Vigouroux and Huguet, 1980). En effet, ces sols présentant une forte oxygénation, retiennent une quantité importante d'eau en hiver induisant une hydratation excessive des arbres. Cela les prédispose alors à la maladie à l'arrivée du froid hivernal (Vigouroux and Bussi, 1994). A l'inverse, durant l'été, ces sols induisent un **stress hydrique** chez les arbres qui, associé à une **mauvaise alimentation calcique** représente également un facteur sensibilisant à la maladie. Une corrélation positive a d'ailleurs été mise en évidence entre le taux de calcium du sol et la résistance des pêchers (Vigouroux and Bussi, 1989). Il semble effectivement que les ions calcium  $\text{Ca}^{2+}$  aient un important pouvoir d'inhibition de la syringomycine (Backman and DeVay, 1971).

### 3.3.5. Lutte existante

Le contrôle de la maladie s'avère particulièrement limité de par l'absence de lutte chimique à la fois efficace, durable et respectueuse des considérations environnementales.

L'usage d'antibiotiques sur les cultures est interdit en Europe, mais peut être soumis à certaines mesures dérogatoires à l'occasion d'épisodes d'émergence, notamment de maladies de quarantaine. Ce fut le cas en Allemagne et Autriche dans les années 2000, avec l'autorisation de l'usage de la streptomycine pour lutter contre le feu bactérien (*E. amylovora*) sur pommier et poirier. L'occurrence de plusieurs épisodes de résistances en Asie et Océanie de souches de *Psa* suivant l'utilisation massive de streptomycine sur les cultures de kiwi dès les années 80, soulève les problèmes d'antibiorésistance (Nakajima *et al.*, 1995; Han *et al.*, 2003, 2004). Ces préoccupations sont d'autant plus importantes dans le cas de gènes de résistances plasmidiques du fait de la facilité de transmission via la voie de la conjugaison bactérienne, ce risque

de transfert de bactérie à bactérie étant avéré dans le cas de *Psy* (Gutiérrez-Barranquero *et al.*, 2017).

En Europe, la majorité des traitements chimiques se basent sur l'**utilisation du cuivre**, autorisé en agriculture biologique. Le cuivre étant un produit de contact, sans action systémique dans la plante, son utilisation s'est avérée davantage efficace afin de cibler les populations épiphytes de *Psy* sur abricotier et cerisier (Olson and Jones, 1983; Wimalajeewa *et al.*, 1991). Cela permet alors de limiter l'importance de l'inoculum sur la plante. Une bonne compréhension de l'épidémiologie de la maladie et de l'alternance des phases épiphyte/endophyte de la bactérie est alors primordiale afin de cibler la fenêtre optimale d'application. Toutefois, l'usage des traitements cupriques n'est pas exempté de problématiques de phytotoxicité (Kennelly *et al.*, 2007), de contournements de résistance (Cazorla *et al.*, 2002; Vanneste and Voyle, 2003) et de dépôts résiduels dans les fruits.

Ces limitations ont alors encouragé la recherche de méthodes de luttés alternatives tournées vers la **lutte biologique**, l'utilisation des **stimulateurs de défense naturelle** (SDN) et l'emploi de **mesures prophylactiques**.

La lutte biologique repose sur l'exploitation de **souches microbiennes antagonistes** des souches *Psy* pathogènes. L'exploitation de diverses souches bactériennes *E. carotovora*, *E. herbicola*, *P. fluorescens*, *P. putida* a permis de limiter la multiplication des populations de *Psy* pathogènes, réduisant de 10 jusqu'à 500 fois la charge bactérienne pathogène sur cerisier, amandier et poirier (Lindow, 1982; Cody *et al.*, 1987; Sobiczewski, 1987). Une stratégie complémentaire de lutte se base sur l'apport de bactéries ayant les mêmes besoins nutritionnels que les populations de *Psy* pathogènes afin d'induire une compétition pour les ressources. La pulvérisation de souches non glaçogènes de *Psy*, obtenues par mutation du gène *ina* a permis de réduire de façon significative les populations naturelles glaçogènes sur poirier et ainsi les dégâts induits par le gel (Lindow, 1987).

L'acibenzolar-*S*-methyl (ASM) est une molécule (commercialisée sous la marque Bion® en Europe), agissant comme un SDN en induisant les réactions de défenses liées à la SAR (Kennelly *et al.*, 2007). Son application préventive sur noisetier et manguier a permis une réduction modérée des dégâts liés à *Psy* (Scortichini and Liguori, 2003; Cazorla *et al.*, 2006).

Des études ont permis de mettre en évidence l'influence de la **nature du porte-greffe** et de la **hauteur de greffage** sur le développement et la sévérité de la maladie chez l'abricotier (Prunier *et al.*, 1999, 2005). Ces facteurs résultant de choix culturaux de l'arboriculteur, il est donc possible de limiter les dégâts liés à la maladie par la mise en œuvre de mesures prophylactiques. Les porte-greffes de type abricotier et prunier se sont avérés sensibilisants tandis que les abricotiers greffés sur pêcher Rubira®

notamment, présentaient un meilleur comportement. Le **greffage haut supérieur à 1.40 m** est à préconiser, il peut techniquement être mis en œuvre avec une taille favorisant un port retombant, de type parapluie. De plus, des mesures relativement simples, portant sur la désinfection systématique des outils et l'évitement des périodes de vulnérabilité de l'arbre pour effectuer une taille, se sont également révélées efficaces afin de limiter l'infection (Luisetti *et al.*, 1992).

Enfin, le **levier génétique** à travers la création de variétés résistantes présente également une alternative intéressante de lutte contre le dépérissement bactérien. **Aucune source de résistance totale** n'a été identifiée à ce jour parmi la diversité naturelle chez les *Prunus*. Seules des **résistances partielles** procurant un niveau de résistance modéré ont pu être identifiées chez le cerisier (Santi *et al.*, 2004; Mgbechi-Ezeri *et al.*, 2017), le prunier (Li *et al.*, 2015) et l'abricotier (Prunier and Bordjiba, 1991; Prunier *et al.*, 1993; Donmez *et al.*, 2009; Brun *et al.*, 2011), **sans que les mécanismes de résistance impliqués n'aient été précisés**. Chez le pêcher, une étude in-vitro de génération de variants soma-clonaux a permis de mettre en évidence un génotype, 122-1, présentant un haut niveau de résistance à *Pss* (Hammerschlag, 2000). Bien que la stabilité de la résistance de ce somaclone se soit maintenue après une période de 2 ans de croissance en serre et un cycle de 2 ans de micro-propagation, il n'est pas impossible que certains phénomènes de réversion aient lieu après plantation en verger comme ce fut le cas pour des soma-clones de pêcher sélectionnés précocement in vitro pour leur résistance à *X. campestris* pv. *pruni* (Hammerschlag *et al.*, 1994). La sensibilité variétale que présente l'abricotier à la maladie est très diverse et est assez remarquable (Prunier *et al.*, 1993). Certaines variétés, récemment introduites au niveau du Bassin méditerranéen et qui se sont avérées être extrêmement sensibles au dépérissement bactérien, pourraient être une cause de l'augmentation de la virulence de *Psy* (Scortichini, 2010). Des différences de sensibilité variétale ont été mises en évidence en verger sur rameaux (Brun *et al.* 2011) et après inoculation au sein des ressources génétiques cultivées (Prunier *et al.*, 1993). Ces travaux ont notamment pu mettre en évidence une très forte sensibilité présentée chez le cultivar Bergeron (note d'incidence 5 à 8 sur 9 selon les années), or il s'agit de la variété la plus cultivée dans la vallée du Rhône. A l'inverse une moindre sensibilité a été observée chez différents fonds génétiques issus de divers groupes génétiques: 'Bakour' (0 à 1 sur 9), 'Orangered' (1 sur 9) (Brun *et al.*, 2011), 'Hâtif Colomer', 'Luizet', 'Palsteyn' (Prunier *et al.*, 1993), 'Harlayne' (Layne, 1981).

La variabilité génétique observée au sein des ressources génétiques ouvre alors une perspective d'amélioration du caractère de sensibilité au dépérissement bactérien dans des programmes d'innovation variétale.

## Contexte et objectifs de la thèse

Les différents niveaux de résistance à *Psy* observés au sein des ressources génétiques d'abricotier semblent indiquer une résistance de type quantitative. Une perspective de développement de variétés partiellement résistantes à travers le pyramidage des QRLs au sein de fonds génétiques sensibles semblerait prometteuse sous réserve d'avoir une meilleure appréhension de l'interaction *Psy* x plante hôte et d'identifier les déterminants génétiques de résistance sous-jacents. Du fait du caractère partiel de la résistance, il semble néanmoins primordial de maintenir les pratiques culturales adaptées afin de s'assurer de la durabilité de la résistance en verger.

Ces pré-requis ont donc permis de proposer puis d'engager un projet destiné à évaluer l'impact de la diversité génétique de l'abricotier et de l'agent pathogène sur la sensibilité à la maladie : projet CASDAR (2013-2016) RésiBac 'Prédiction de l'apparition du chancre bactérien (*Pseudomonas syringae*) chez les arbres fruitiers – Preuve de concept : cas de l'abricotier'. Cette thèse s'inscrit dans le cadre de ce projet qui visait à (i) élaborer des stratégies de résistance génétique des arbres fruitiers au chancre bactérien en verger, (ii) à mettre au point des outils d'évaluation du potentiel pathogène des souches de *Psy* présentes dans l'environnement.

Il a été conduit en considérant l'expertise conjointe des unités : UR1052 'Génétique et Amélioration des Fruits et Légumes' (GAFL) sous la coordination de Jean-Marc Audergon pour la partie 'pré-breeding' et caractérisation de la sensibilité variétale de l'abricotier et UR407 'Pathologie Végétale' (PV) sous la coordination de Cindy Morris pour la partie caractérisation de la gamme d'hôte, des composantes du pouvoir pathogène au sein de la diversité de *Psy* et du potentiel d'émergence de nouveaux foyers.

Les objectifs scientifiques de ce travail de thèse ont consisté en :

- (1) la caractérisation phénotypique de la sensibilité de diverses populations d'abricotiers à *Psy* en verger,
- (2) la caractérisation des QRLs associés à la résistance quantitative de l'abricotier à *Psy* à travers la mobilisation d'approches d'analyse de liaison et d'association,
- (3) la validation des marqueurs associés à la résistance,
- (4) et l'étude des interactions génotype abricotier x souche *Psy* afin d'évaluer la durabilité des résistances déployées en verger.

L'UERI de Gotheron, le domaine expérimental de l'Amarine ainsi que le collectif CEP Innovation ont contribué à la caractérisation phénotypique de la sensibilité à *Psy* au sein de différentes populations d'abricotiers. Le laboratoire de biologie moléculaire du GAFL a défini les marqueurs moléculaires mobilisés pour la cartographie génétique de liaison et l'entreprise ADNid a été partenaire du projet en assurant la partie génotypage.



## Partie II - Matériel & Méthodes

Tableau 2 : Synthèse du matériel végétal, bactérien, des données phénotypiques et génotypiques exploitées au cours du travail de thèse.

Objectifs	Caractérisation des QRLs et gènes associés à la résistance à <i>Psy</i>			Caractérisation de l'interaction GxG	
Matériel végétal	Core-collection 73 accessions	Population biparentale pseudo-F1 'BerBa' 208 hybrides	3 Populations biparentales interconnectées 'BerBa 2 <sup>ème</sup> génération' – 307 hybrides (3 hybrides 'BerBa' x A4049)	3 accessions témoins	
					20 accessions (sensibles, intermédiaires, résistants)
Type d'analyse	GWAS	Analyse de liaison	Analyse de liaison : Etude des interactions QRL x fond génétique	Protocole in vitro : étude migration de <i>Psy</i> + recherches de corrélations avec les données verger	En verger : Analyse statistique du plan factoriel d'interaction
Données génotypiques	63,236 SNPs	Carte père-mère 149/139 SNPs	Carte individuelle par hybride + carte consensus (162 SNPs) sur les 3 hybrides parentaux	/	/
Location et années de l'essai	Domaine de l'Amarine 4 années (2013 à 2016) + validation 2018 sur génotypes	UERI de Gotheron 3 années (2012 à 2014)	UERI de Gotheron 3 années (2016 à 2018)	Laboratoire en 2016	UERI de Gotheron 2 années (2017 & 2018)
Souches <i>Psy</i> inoculées	<i>Pss</i> 41A (PG 2)	<i>Pss</i> 41A (PG 2)	<i>Pss</i> 41A (PG 2)	9 souches <i>Psy</i> (PG 1 et 2)	
Phénotypes	lgc, lge, bs, bi	lgc, lge, bs, bi		bs	lgc, bs
Partie de la thèse	III	IV	IV	V	V

Ce chapitre résume l'ensemble des matériels végétal, bactérien, et les méthodes utilisées afin de répondre aux objectifs scientifiques de cette thèse. Le détail des matériels et des protocoles sont explicités en détail lors des chapitres suivants.

## 1. Matériel végétal

Le matériel végétal exploité en cartographie correspond à 3 populations présentées au niveau du *tableau 2* et de la *figure 23*.



**Figure 23 : Photographies des dispositifs végétaux exploités suite à inoculation de *Psy*.** (A) Core-collection (73 accessions) à l'Amarine (30), (B) Descendance biparentale BerBa (207 hybrides) à Gotheron (26), (C) 3 descendances interconnectées BerBa (307 hybrides) à Gotheron (26),  
*Crédits photographiques: Morgane Roth, Audrey Ferro & Mariem Omrani*

### **Core-collection - partie III**

La core-collection est un sous-ensemble de la collection de ressources génétiques abricotier détenue par l'Unité GAFL au niveau du domaine expérimental de l'Amarine (30). Elle a été définie suite aux travaux conduits sur la diversité génétique à l'aide de marqueurs moléculaires SSRs (Bourguiba *et al.*, 2012), ayant permis par la suite de sélectionner un sous-ensemble de 120 accessions sur les 560 détenues en collection. Sur ces 120 accessions, un panel de 73 est entré dans le cadre des analyses de sensibilité au dépérissement bactérien. Ces 73 cultivars présentent une diversité génétique large, et sont issus de diverses origines géographiques (*Figure 24*). La core-collection a été greffée sur un porte-greffe abricotier 'Manicot'.



Figure 24 : Cartographie et répartition de l'origine géographique des 73 accessions de la core-collection.

#### Pseudo-F1 'BerBa' – partie IV

La population 'BerBa', implantée en 2004 à l'INRA de Gotheron (26) est une population biparentale (218 individus génotypés dont 206 phénotypés) issue du croisement entre deux cultivars fortement homozygotes: 'Bergeron', cultivar issu du groupe Nord-Européen (France) (Hagen *et al.*, 2002) et 'Bakour', variété du groupe Nord-Africain (Tunisie) (Bourguiba *et al.*, 2010). Ces cultivars ont été sélectionnés en partie, respectivement pour leur sensibilité et leur résistance partielle au dépérissement bactérien (Brun *et al.*, 2011), afin de créer une population de ségrégation, support d'études de cartographies de QTLs.

#### Interconnectées 'BerBa' de 2<sup>nde</sup> génération – partie IV

La première phase d'observation de la 'BerBa' conduite de 2012 à 2014 a mené à la pré-sélection de 3 hybrides, choisis parmi les plus résistants au dépérissement bactérien (mais aussi au monilia): C092, C109 et D099. Ces 3 hybrides ont par la suite été croisés individuellement avec un génotype 'testeur' commun A4049, auto-fertile et résistant à la sharka. Les trois sous-populations interconnectées, implantées depuis 2015 à l'INRA de Gotheron comportent 118, 110 et 77 individus, respectivement.

Différents génotypes, ont été sélectionnés afin de réaliser les **études d'interactions GxG**. (*partie V*).

Le **plan factoriel d'interaction GxG**, réalisé à la fois en verger et en laboratoire a consisté en partie en l'étude de la sensibilité de 20 cultivars d'abricotiers choisis parmi les plus résistants, sensibles et les niveaux intermédiaires de sensibilité (détails en *partie V*).

L'étude de **migration de la bactérie** dans les tissus, en lien avec la connaissance de la sensibilité a été effectuée chez trois génotypes : Bergeron (sensible), Bakour (résistant), et Colomer (résistant). Colomer est une variété française issue du groupe génétique de diversité adaptative, rassemblant des génotypes à potentiel d'adaptation large sur des zones géographiques diverses (Hagen *et al.*, 2002).

## 2. Matériel bactérien

La sensibilité des génotypes étudiés au sein des 3 populations a été testée suite à des expérimentations d'inoculations artificielles en verger (protocole détaillé en 4.1).

Une souche de *Psy* référencée comme 41A (PG 2) est utilisée comme référence pour les inoculations (Bartoli *et al.*, 2015a). Il s'agit d'une souche initialement isolée par le Centre technique interprofessionnel des fruits et légumes (CTIFL) au niveau d'un verger fortement contaminé en région Rhône-Alpes et dont le pouvoir pathogène s'est avéré être le plus important lors d'essais de criblage de différentes souches en laboratoire et en serre.

Le plan factoriel d'interaction GxG a exploité au total 9 souches choisies essentiellement au niveau des PGs 1 et 2, ces deux PGs étant les plus représentés au niveau de la diversité de *Psy* en verger (Parisi *et al.*, 2016) (détails en *partie V*). 7 de ces souches ont été isolées à Gotheron à partir d'arbres symptomatiques, 2 au niveau de débris de récolte et de la terre. Parmi les 9 souches, 7 se sont avérées être pathogènes sur Bergeron lors de tests d'évaluation du pouvoir pathogène effectués en laboratoire (Parisi, communication personnelle).

## 3. Données génotypiques et cartes génétiques

Les données de génotypage ont été en majorité obtenues préalablement au début de la thèse.

### Core-collection – *partie III*

Une collaboration avec l'UMR Biologie du Fruit et Pathologie de Bordeaux dans le cadre du projet ANR ABRIWG (2012-2014) a permis d'accéder au **re-séquençage de 93 accessions** (dont font partie le panel de GWAS) en utilisant la technologie Illumina HI-SEQ2000. La profondeur moyenne était de 15 à 25X selon l'accession. Le traitement des données bioinformatiques a été effectué dans le cadre de l'étude d'association sur la sharka (Mariette *et al.*, 2016). Pour résumer, l'alignement des SNPs s'est effectué sur le génome de référence *v1.0* du pêcher en excluant les régions inter-géniques sur les 8 chromosomes assemblés, représentant 99% du génome. Le re-séquençage a permis l'identification de plus de 200k SNPs au sein du panel re-

séquencé. Après filtre sur la fréquence de l'allèle mineur, et sur la position physique, la matrice de marquage exploitée pour la GWAS est constituée de près de 63 000 SNPs (détails en *chapitre 3*).

#### **Pseudo-F1 'BerBa' – partie IV**

Le génotypage de la descendance biparentale BerBa a débuté par le criblage du polymorphisme de 300 SSRs chez les deux parents Bakour et Bergeron. Ce génotypage a conduit à la sélection de 38 marqueurs chez Bergeron et 52 marqueurs chez Bakour. Le rendement du criblage n'a donc pas été très efficace. Des travaux complémentaires ont montré que ces variétés situées en extrémité de phylum de diversification sont caractérisées par des niveaux d'homozygotie particulièrement élevés, limitant les possibilités d'identification de SSRs polymorphes. Les données de re-séquençage ABRIWG ont permis de sélectionner 149 SNPs chez Bakour et 139 SNPs chez Bergeron afin de génotyper la descendance de 218 individus puis élaborer une carte de type père-mère. Certains marqueurs SSRs choisis parmi les plus informatifs ont été incorporés aux deux cartes parentales. Elles comportent 11 et 10 groupes de liaisons et mesurent 823.9 et 624.1 cM, respectivement ; cela correspond à une couverture de 81% et 87% de la version *v2.0* du génome pêcher (détails en *partie IV*).

#### **Interconnectées 'BerBa' de 2<sup>nd</sup>e génération – partie IV**

Sur la base des marqueurs développés pour la BerBa, le criblage des marqueurs sur l'ensemble des parents hybrides 'C092', 'C109', 'D099' et 'A4049' a permis de sélectionner 185 SNPs polymorphes chez au moins un des trois parents hybrides. Dans le cadre du travail de thèse, trois cartes parentales ont été construites pour chacun des hybrides et une carte consensus a été élaborée en fusionnant les cartes individuelles à l'aide des marqueurs 'ponts'. La carte comporte à l'heure actuelle 162 marqueurs répartis sur 10 groupes de liaison et couvre 970 cM représentant 95% de la version *v2.0* du génome pêcher (détails en *partie IV*).

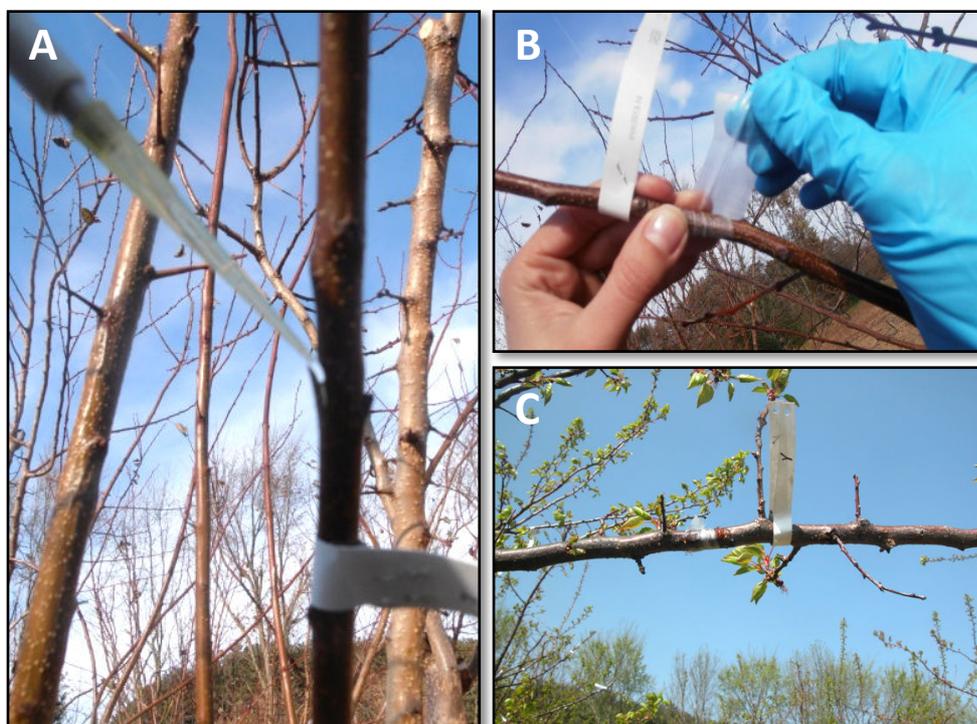
## **4. Design expérimental : protocole d'inoculations contrôlées**

### **4.1. En verger : conditions in-vivo**

Les différentes campagnes de phénotypage pour la sensibilité au dépérissement bactérien en verger ont été initiées préalablement au début de la thèse. Le phénotypage des 3 populations pour la sensibilité au dépérissement bactérien a été effectué en verger sur une période allant de 2012 à 2015 pour la BerBa, 2013 à 2016 pour la core-collection et 2016 à 2017 pour le dispositif des interconnectées BerBa ; une année correspondant à une répétition expérimentale. L'année 2018 a permis d'effectuer un

phénotypage à but de validation des résultats de GWAS, ainsi près de 650 accessions ont été inoculées à cette occasion (détails en *partie III*).

Le dispositif d'inoculations GxG en verger a été conduit lors de deux répétitions expérimentales, en 2017 et 2018.



**Figure 25 : Photographies du mode opératoire des inoculations artificielles en verger.**

*Crédits photographiques: Mouna Hadj Brahim & Morgane Roth*

Les données de sensibilité sont issues d'inoculations contrôlées effectuées chaque année, au cours de la fenêtre fin novembre-début décembre (avant expression du froid hivernal) sur tissus dormants, selon une procédure standardisée (Prunier and Bordjiba, 1991). Un inoculum de *Psy* est préparé à partir d'une souche cultivée sur milieu King's B pendant 48h à 24°C. La concentration de la suspension bactérienne est ajustée par ajout d'eau permutée stérilisée, à  $10^8$  CFU.ml<sup>-1</sup> à l'aide d'un spectrophotomètre. Un arbre par génotype, et de 3 à 5 réplicats techniques (selon le dispositif) sont ensuite inoculés. Une incision calibrée de 1 cm de large est effectuée superficiellement au niveau de l'écorce de rameaux d'1 an. Un volume de 25 µl d'inoculum est déposé au niveau de la blessure à l'aide d'une pipette (*Figure 25A*). Une bande parafilm est alors appliquée afin de protéger la blessure et maintenir une humidité favorable à l'efficacité de l'infection (*Figure 25B*). Pour chaque accession, un rameau supplémentaire servant de témoin négatif est inoculé avec le même volume d'eau stérilisée. Les symptômes se développent sur une durée de 5 à 6 mois avant le prélèvement des rameaux et la notation visuelle (*Figure 25C*).

## 4.1. En laboratoire : conditions in-vitro

### Expérimentation plan factoriel GxG

De façon similaire à la procédure en verger, des inoculations contrôlées en laboratoire des 9 souches de *Psy* référencées précédemment, sont effectuées sur des boutures de 18 cm, prélevées à l'état dormant. Ensuite, les boutures alternent des phases d'exposition à des températures différentes pendant une durée d'un mois, dont une exposition au froid à -12°C pendant 48h, permettant de mimer l'expression du gel en hiver. Ce protocole détaillé en *partie V* résulte d'essais de mise au point effectués avec l'équipe de Bactériologie. Les données in-vitro et in-vivo ont ensuite été analysées conjointement afin d'évaluer leur corrélation et de juger de la pertinence du protocole in-vitro afin de reproduire les différences variétales obtenues en verger.

### Etude de la migration de la bactérie dans les tissus et lien avec la sensibilité à la maladie

Trois géotypes précédemment caractérisés Bakour, Colomer ('résistants') et Bergeron ('sensible') ont été inoculés avec *Psy* 41A selon le protocole d'inoculations sur boutures. A l'issue de cette seconde expérimentation en conditions in-vitro, des isollements de tissus sur milieu KBC<sup>16</sup> ont permis d'évaluer le lien entre présence/migration de la bactérie dans les tissus et niveau de sensibilité.

## 5. Données phénotypiques

4 variables indicatrices de différents symptômes externes et internes de la maladie sont notées suite aux expérimentations d'inoculations contrôlées en verger (*Figure 26*) ou en laboratoire (*Figure 27*).

La **longueur de chancre** (lgc) correspond à une zone de symptômes visibles extérieurement englobant la zone cicatrisée et le méplat formé par la destruction des tissus du cambium.

La **longueur de la lésion sub-corticale** (lge) est la plaie ouverte au contact de l'air, formant une fente plus ou moins large, à partir du point d'inoculation.

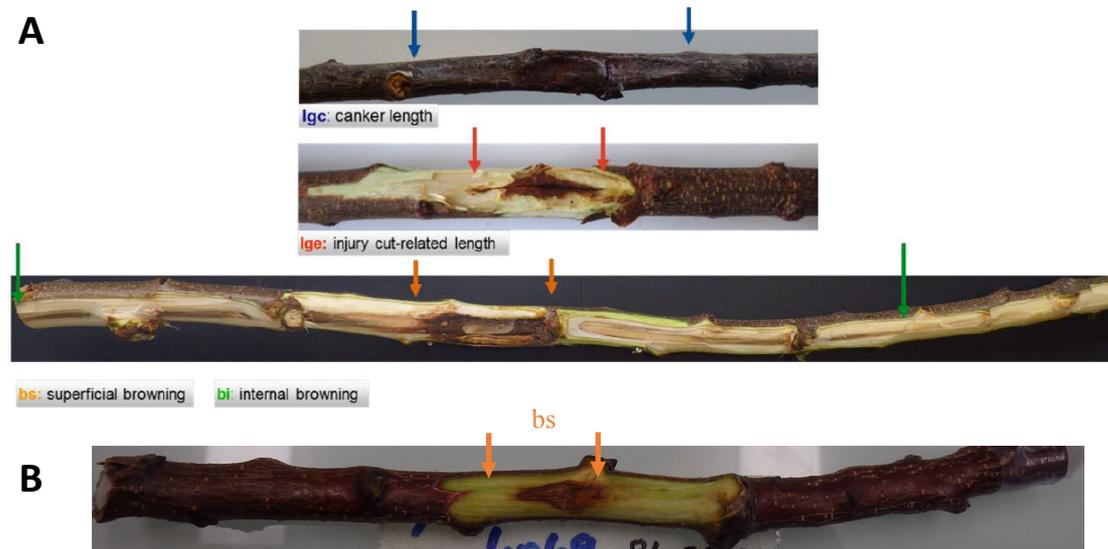
La **longueur du brunissement superficiel** (bs) est observée directement sous l'écorce dans les tissus verts de l'année, elle est mesurée après pelage de l'écorce.

Suite au pelage des tissus en profondeur, la **longueur du brunissement interne** (bi) correspond à un brunissement des vaisseaux conducteurs au cœur du bois.

---

<sup>16</sup> Milieu de culture facilitant l'isolation sélective de *Psy*. Il s'agit d'un milieu King's B complété avec de l'acide borique, et des composés antibiotiques (cycloheximide et céphalexine).

Ces données phénotypiques issues d'inoculations artificielles ont été complétées par des notations de mortalité en verger, liées à une forte pression naturelle de la maladie en 2017 et 2018.



**Figure 26 :** Photographies des phénotypes étudiés issus des expérimentations d'inoculations contrôlées sur rameaux en verger (A) sur et sur boutures en laboratoire (B).

lge, longueur de chancre; lge, longueur de la lésion sub-corticale; bs, brunissement superficiel; bi, brunissement interne.

Note: Seule la partie gauche du brunissement superficiel (bs) est visible sur la 3<sup>ème</sup> photographie A, la partie droite est encore recouverte d'écorce.

*Crédits photographiques: Mariem Omrani*



Partie III- Analyse de la diversité  
naturelle de la sensibilité  
variétale à *Psy* à travers une  
analyse d'association

# Introduction

Cette étude est sous la forme d'un article soumis en mai 2018, chez *BMC Plant Biology*.

L'article vise à caractériser la sensibilité variétale à *Psy*, observée au niveau phénotypique et génotypique, au sein d'une core-collection de 73 accessions.

Les phénotypes exploités sont issus de quatre campagnes d'inoculations artificielles réalisées en période hivernale en verger. Suite à des analyses statistiques préliminaires, sur les quatre phénotypes initialement évalués (*lgc*, *lge*, *bs* et *bi*), seuls la longueur de chancre (*lgc*) et le brunissement superficiel (*bs*) se sont avérés être discriminants de la sensibilité à la maladie entre les cultivars. En conséquence, seuls *lgc* et *bs* ont été exploités lors de la suite de l'étude.

Cette étude a permis d'identifier au sein du panel variétal, 7 SNPs significativement associés à la résistance partielle à la maladie dont 2 SNPs à effet majeur à la fois sur *lgc* et *bs* (41% et 26% de variance phénotypique expliquée, respectivement).

Afin d'apporter une validation à la fois sur le polymorphisme de l'ensemble des marqueurs candidats et sur la corrélation entre variants alléliques et niveau de sensibilité à la maladie, des travaux complémentaires ont été réalisés suite à la soumission de l'article :

Sur les 7 marqueurs identifiés, des amorces KASPar ont été dessinées et testées sur la core-collection dans le but de valider le polymorphisme sur les SNPs candidats.

Les couples d'amorces validés ont été par la suite mobilisés afin de génotyper un panel variétal de près de 335 accessions.

Ce panel variétal a été phénotypé lors de l'année 2018 afin d'évaluer la corrélation entre les génotypes obtenus au niveau des SNPs candidats et les différences de sensibilité à la maladie.

Ces différents éléments de validation sont abordés à la suite de l'article.

Les figures et tables additionnelles de l'article sont présentées au niveau des annexes I référencées à la fin du manuscrit.

# Genome-wide association multi-locus and multi-variate linear mixed models reveal two linked loci with major effects on partial resistance of apricot to bacterial canker

Mariem Omrani<sup>1,2,3</sup>

Email: [mariem.omrani@inra.fr](mailto:mariem.omrani@inra.fr)

Morgane Roth<sup>1</sup>

<sup>†</sup>Present address: Agroscope, Research Division Plant Breeding, Wädenswil, Switzerland

Email: [morgane.roth@agroscope.admin.ch](mailto:morgane.roth@agroscope.admin.ch)

Guillaume Roch<sup>1,4</sup>

Email : [guillaume.roch@inra.fr](mailto:guillaume.roch@inra.fr)

Alain Blanc<sup>1</sup>

Email : [alain.blanc@inra.fr](mailto:alain.blanc@inra.fr)

Cindy E.Morris<sup>2</sup>

Email: [cindy.morris@inra.fr](mailto:cindy.morris@inra.fr)

Jean-Marc Audergon<sup>1\*</sup>

\*Corresponding author

Email: [jean-marc.audergon@inra.fr](mailto:jean-marc.audergon@inra.fr)

<sup>1</sup>INRA, UR1052 Génétique et Amélioration des Fruits et Légumes, Centre de Recherche

PACA, Montfavet, France

<sup>2</sup>INRA, UR407 Pathologie Végétale, Centre de Recherche PACA, Montfavet, France

<sup>3</sup>ENGREF, AgroParisTech, Paris, France

<sup>4</sup>CEP Innovation, Lyon, France

## **Abstract**

### ***Background***

Diseases caused by *Pseudomonas syringae* (*Ps*) are recognized as the most damaging factors in fruit trees with a significant economic and sanitary impact on crops. Among them, bacterial canker of apricot is exceedingly difficult to control due to a lack of efficient prophylactic measures. Several sources of partial resistance have been identified among genetic resources but the underlying genetic pattern has not been elucidated thus far. In this study, we phenotyped bacterial canker susceptibility in an apricot core-collection of 73 accessions over 4 years by measuring canker and superficial browning lengths issued from artificial inoculations in the orchard.

In order to investigate the genetic architecture of partial resistance, we performed a genome-wide association study using best linear unbiased predictors on genetic (*G*) and genetic x year (*G* × *Y*) interaction effects extracted from linear mixed models.

Using a set of 63,236 single-nucleotide polymorphism markers genotyped in the germplasm over the whole genome, multi-locus and multi-variate mixed models aimed at mapping the resistance while controlling for relatedness between individuals.

### ***Results***

We detected 11 significant associations over 7 candidate loci linked to disease resistance under the two most severe years. Colocalizations between *G* and *G* × *Y* terms indicated a modulation on allelic effect depending on environmental conditions. Among the candidate loci, two loci on chromosomes 5 and 6 had a high impact on both canker and superficial browning, explaining 41% and 26% of the total phenotypic variance, respectively. We found an unexpected long-range linkage disequilibrium between these two markers revealing an inter-chromosomal LD block linking the two underlying genes. This result supports the hypothesis of a co-adaptation effect due to selection through population demography. Candidate genes annotations suggest a functional pathway involving abscisic acid, a hormone mainly known for mediating

abiotic stress responses but also reported as a potential factor in plant-pathogen interactions.

### ***Conclusions***

Our study contributed to the first detailed characterization of the genetic determinants of partial resistance to bacterial canker in a *Rosaceae* species. It provided tools for fruit tree breeding by identifying progenitors with favorable haplotypes and by providing major-effect markers for a marker-assisted selection strategy.

### ***Keywords***

GWAS, *Prunus*, partial resistance, bacterial canker, *Pseudomonas syringae*, candidate genes, apricot, multi-locus mixed model, multi-variate mixed model, linkage disequilibrium

## Introduction

For decades, breeding programs have notably focused on the introgression of major resistant genes into crop cultivars. However multiple episodes of resistance breakdown have led research efforts to target polygenic (quantitative or partial) resistances [1]. Although breeding strategies relying on pyramiding quantitative trait loci (QTLs) on top of monogenic genes might be costly and time-consuming, the long-term efficiency and sustainability of QTLs have been assumed if they are deployed parsimoniously in both space and time [2, 3]. Resistance sustainability has a particular interest for perennial woody crops considering their long generation time and life-span in orchards, and overall the long-term commitment to their breeding schemes.

Within the *Prunus* genus, apricot (*Prunus armeniaca* L.) is one of the most popular and typical crops of the Mediterranean Basin which provides 49% of the total production to the world market [4]. Among the biotic stresses affecting apricot crop durability and in a broader way stone fruit species, bacterial canker, caused by ubiquitous bacteria in the species complex *Pseudomonas syringae* (*Ps*), is one of the most damaging. This disease could potentially lead to the death of trees in the orchard, especially young trees within their five first years after planting. Three main groups - *Ps* pv. *syringae*, *Ps* *viridiflava* [5] and *Ps* pv. *morsprunorum* [6], associated with phylogroups 2, 7 and 3, respectively [7], have been historically associated with the disease on apricot. The symptoms affect aerial organs, resulting in lesions and shot holes in leaves, bud and blossom dieback, and apical necrosis, potentially leading to more severe symptoms when bacteria get to the vessels and spread through the vascular system [8, 9]. When the infection is systemic, cankers corresponding to necrosis and flattening of external tissues linked to a dissymmetric growth of the cambium in the spring, are visible on branches and scaffold limbs [10]. Several abiotic factors relying on soil and climatic conditions can favor bacterial canker severity. Among them, cold winter temperatures and especially high frequency of frost-defrost

episodes have been highlighted as major factors favoring bacterial canker [11, 12]. Integrated management practices such as the use of resistant and soil-adapted rootstock material and grafting at a tall height could lower bacterial canker incidence in orchards [13, 14]. These cultural recommendations are nevertheless, technically challenging for orchard management and provide only a partial protection in orchards. Therefore, the development of partially resistant cultivars seems to be a promising measure in addition to preventive practices to assure orchard durability. In this context, research efforts have focused on screening apricot genetic resources both under natural infections or controlled inoculations. Several genetic backgrounds with partial resistance to bacterial canker have been identified: ‘Bakour’ [14], ‘Hâtif Colomer’, ‘Luizet’, ‘Palsteyn’ [15] and ‘Orangered’ [16]. More globally, similar research initiatives have been conducted to phenotypically characterize differential susceptibilities to bacterial canker in *Prunus* rootstocks [17], sweet cherry [17-19] and plum [20], but to date the question of the genetic determinants of partial resistance to bacterial canker from the natural diversity in *Rosaceae* fruit trees has never been investigated.

Recent progress of high-throughput-sequencing technologies fostering the discovery of thousands and even millions of single nucleotide polymorphisms (SNPs) over whole genomes has opened up new opportunities to address the underlying architecture of quantitative traits [21]. Genome-Wide Association Studies (GWAS) have been extensively and successfully used in plant breeding thanks to the advancement of these novel genomics-based approaches [22]. GWAS enable mapping of QTLs and genes affecting trait variation in a wide collection of mostly unrelated individuals usually sampled from wild relative populations, breeding cultivars and landraces. This is in stark contrast to traditional linkage mapping methods that require the establishment of segregating populations beforehand [23, 24]. Association studies benefit from the numerous recombination events, which have occurred through species demographic

history, and rely on linkage disequilibrium (LD) caused in part by selection and population structure, to exploit natural allelic diversity and identify links between markers and causal loci underlying the trait of interest [25-27].

GWAS turns out to be a particularly suitable approach for plants and especially for perennial crops since costs associated with making and maintaining large progenies in the orchard can be spared. Therefore, the use of GWAS has gradually become more widespread for determining the genetic basis of variation in complex traits in *Rosaceae* fruit species [28-30]. More particularly, the apricot genome displays many advantages for GWAS applications due to: (i) its diploidy ( $2n=16$ ) and small size (294 Mb/n) [31], (ii) its high level of heterozygosity resulting from a general outcrossing mating system [32] and (iii) its high nucleotide diversity related both to the early seed propagation [33] and geographically broad distribution of the germplasm with a diversification from Central Asia [34, 35]. Moreover, the apricot genome has been characterized with a very fast LD decay within 100 bp in a large genetic diversity panel [30], allowing association mapping with a very precise resolution conditionally upon the use of a high density of markers [36].

In the present study, we investigate the underlying genetic architecture of partial resistance to bacterial canker in apricot using an association approach. The specific objectives of this research are to (i) identify the genetic and environmental components of bacterial canker susceptibility using a core-collection of 73 apricot accessions that has been phenotyped in the orchard over a 4-year campaign, (ii) develop an appropriate genome-wide association methodology taking into account the species genetic architecture and the stratification of our population set, and (iii) identify candidate genes and decipher the molecular basis of polygenic resistance to bacterial canker.

## Material & Methods

### *Plant Material*

A core-collection composed of 73 accessions of apricot trees was used for this study (60 accessions were previously investigated for genetic determinism of resistance to Plum Pox virus [30]). This population was grafted on Manicot apricot rootstock from 1995 to 2001 at the INRA experimental station 'Domaine de l'Amarine' (Gard, France). Name, geographical origin and further information of the plant material are presented in *Additional file 1*.

### *Experimental Design & Phenotyping*

Controlled inoculations over the core-collection (one tree for each accession, three replicates per tree) were performed annually at the end of November over four years (2013 to 2016) according to a previously described procedure [14]. Briefly, bacterial inoculum was prepared from a culture of strain (called '41A', highly aggressive and initially isolated from an infected apricot orchard) from phylogroup 2 of *Ps* [37] grown on King's B medium [38] for 48h at 24°C. The concentration of the suspension was adjusted with sterile demineralized water to  $10^8$  CFU.ml<sup>-1</sup> using a spectrophotometer. Three one-year old shoots per tree were injured by performing a standardized one-centimeter-long incision in the bark. A 25 µl aliquot of inoculum was injected by pipetting through the wound before wrapping with a parafilm band. For each accession, an additional shoot served as a negative control and was inoculated with the same volume of sterile water. Symptoms were allowed to develop for six months in the orchard before pruning of inoculated shoots for disease visual assessment. Evaluation of phenotypes was carried out by several operators according to a balanced design in order to be able to assess the effect of operators on phenotypic variability. Phenotypes consisted of the length of two different types of symptoms (*Figure 1*): canker length (*lgc*) and superficial browning length (*bs*, not evaluated in 2014). Canker length results from necrosis of the cambium that slowly extends on the outer tissues

through the vegetative period of growth [39]. In addition, discoloration of inner vascular tissues may also occur as vertical-brown streaks.



**Figure 1: Photograph of typical bacterial canker symptoms on branches after controlled inoculation.**

Canker length  $lgc$  (blue arrows) and superficial browning  $bs$  (red arrows) as observed 6 months after inoculation.

Furthermore, several covariates that might have been correlated with the phenotypes were assessed, including the diameter of the lower-part of the shoot as a measure of the vigor of the accession [40] and the height of the branch in the tree as a measure related to the local temperature close to the inoculation point.

### ***Statistical modelling of phenotypic data***

All statistical analysis were performed using R software (<https://www.r-project.org/>) and the ASReml-R package [41].

Analyses of Variance (ANOVA) were used to assess significance of fixed effects: year, operator, genotype, genotype-by-year interaction and the fixed covariates diameter and height of the branch on  $lgc$  and  $bs$  according to the following model:

$$P_{ijk} = \mu + Y_i + (Y/h)_i + (Y/\phi)_i + Y_i/O_j + G_k + G_k \times Y_i + e_{ijk}$$

where  $P_{ijk}$  is the phenotypic value of accession  $k$  noted by the  $j^{th}$  operator in year  $i$ ,  $\mu$  the overall mean,  $Y_i$  the effect of year  $i$ ,  $(Y/h)_i$  and  $(Y/\phi)_i$  the nested effects of the shoot diameter and height, respectively within the year  $i$ ,  $Y_i/O_j$  the nested effect of the  $j^{th}$  operator within the year  $i$ ,  $G_k$  the effect of the accession  $k$ ,  $G_k \times Y_i$  the interaction term between year  $i$  and genotype  $k$ , and  $e_{ijk}$  the random independent and identically distributed residual term. Prior to modelling, phenotype distribution was normalized

using a  $\ln(x+1)$  transformation in order to achieve a normal distribution of the residuals (Shapiro test  $p > 0.05$ ).

Multi-variate models were used to model correlations of phenotypes between years. Preliminary analysis revealed that genetic variances differed significantly between years. Thus, a homogenous correlation (corh) variance-covariance (vcov) matrix was fitted on the interaction term  $G_k \times Y_i$  in the model. The residual error was modeled either for all years or for each year independently. In order to select the optimum model for each phenotype, a goodness-of-fit comparison was made based on Restricted Maximum Likelihood (REML) statistics using the Akaike Information Criterion (AIC). Due to the possible annual effect on phenotypes and the missing data in our design (*Additional file 1*), a within-year model was performed for each year according to the following linear mixed model:

$$P_{jk} = \mu + h + \phi + O_j + G_k + e_{jk}$$

Moreover, by considering a within-year model for each phenotype, we ensured the biological relevance of missing data replacement in the multi-year model allowed by the imputation due to the corh vcov structure.

Best unbiased linear predictors (BLUPs) were subsequently assessed from both within-year and multi-year models on  $G$  and  $G \times Y$  random effects for the phenotypes *lgc* and *bs*. These predictors extracted from the multi-year model were then used as phenotypic input data for GWAS. A validation of the results was performed using the within-year BLUPs. In the case of  $G \times Y$  BLUPs, imputed data were removed for missing individuals in each year in order to avoid statistical redundancy with the  $G$  BLUP prediction. REML estimates of genetic ( $\sigma_G^2$ ) and residual ( $\sigma_e^2$ ) variances were also computed from the multi-year model. Broad-sense heritability  $H^2$  for each phenotype per year was calculated as:

$$H^2 = \frac{\sigma_G^2}{\sigma_G^2 + \frac{\sigma_e^2}{n}}$$

where  $n = 3$  is the number of replicates per accession.

### ***Genotyping and SNP marker data filtering***

Genotyping and SNPs alignment were previously described in [30]. Briefly, the set of plants for GWAS was genotyped using NGS-Illumina HiSeq 2000/2500 sequencing with estimated depths between 15 to 25 folds depending on the accession. SNPs alignment was performed using the *Prunus persica* (*P. persica*) v1.0 reference genome and considering gene-space regions over the eight chromosome-level scaffold assemblies covering 99% of the genome [42]. Genotype assignment was performed according to an inference method based on maximum likelihood for estimating allele frequencies by using base counts at each position [43].

For our study, a minor allele frequency (MAF) threshold of 5% (according to [44]) was applied on the initial set of SNPs in order to remove rare variants and thus avoid false-positive associations. Then a filter on physical map position, discarding one of a pair of consecutive SNPs whose pairwise distance was less than 10 bp, was performed to limit the number of multiple statistical tests. After conducting all bioinformatic filters, a selection of 63,236 SNPs was kept for association study. Markers were named according to their physical positions on the genome, 'LG1\_72881' for a SNP located on chromosome 1 at 72,881 bp, for example.

### ***Population stratification analysis***

An additional filter based on LD pairwise pruning, that involved discarding markers for which squared correlation  $r^2 > 0.2$  (window 50 - stepwise 5 SNPs), was executed in order to keep a set of 21,942 independent markers (under the assumption of linkage equilibrium) for inferring population structure.

The Admixture program [45] was used for computing maximum likelihood estimations of individual ancestry fractions while testing scenarii from  $k=2$  to  $k=10$  ancestral groups. The model choice criteria was based on minimization of the cross-validation error. In order to insure the reliability of the optimal choice of  $k$ , a complementary PCA analysis was performed on the genotype matrix. Inference of relatedness between

individuals resulted in the calculation of an identity-by-state (IBS) allele sharing matrix from the same subset of 21,942 SNPs using the *emmax-kin* function implemented in the EMMAX (Efficient Mixed-Model Association eXpedited) program [46]. Both the ancestral proportion fraction matrix (Q) calculated according to the optimum model from Admixture and the IBS pairwise matrix between individuals (K) were used in GWAS to correct inflation of p-values due to stratification artefacts in the population.

### ***Linkage disequilibrium estimation***

Pairwise LD from a sampling of 1,000 markers per chromosome was computed using  $r^2$ . A corrected procedure compacted in the *LDcorSV* R package [47] was used allowing removal of population stratification bias on LD. Considering a set of markers on each chromosome, both initial and corrected  $r^2$  estimates ( $r_{vs}^2$ ) were then plotted against physical distances in order to investigate intra-chromosomal LD decay according to the following model assuming drift-recombination equilibrium [48]  $r^2 = \frac{1}{1+4bd} + e$  where  $r^2$  is the square of loci correlation between a marker pair,  $d$  is the pairwise physical distance between the two markers,  $b$  is a decay coefficient calculated with least squares estimates in a non-linear regression (*nls* function in R software) and  $e$  refers to a residual estimate.

### ***Genome-Wide Association Analysis***

Cumulative effects of K and Q matrices over genotypic data were first tested with EMMAX [46] as a covariance genetic matrix and fixed ancestral covariates impacting expression of phenotypes. Quantile-quantile (Q-Q) plots were realized to select the best predictive model between K, Q and K+Q for each phenotype using the *qqman* R package [49]. Thus according to the best model choice, two supplementary mixed models were performed: a recently implemented multi-locus mixed-model [50] from the MLMM algorithm [51] (on both  $G$  and  $G \times Y$  BLUPs from the multi-year model) and a multi-variate linear mixed-model (mvLMM) algorithm (between *lgc* and *bs G*

BLUPs from the multi-year model) developed in GEMMA software [52]. Multi-locus models consist of forward stepwise mixed-model regressions with an increasing number of included markers (regressors) while re-estimating genetic and residual variance components at each step. The implementation from MLMM allows a more suitable handling of the so-called “high dimension issue” resulting from a low population size and a high number of possible regressors introduced into the model [50]. Instead of using a classical p-value threshold that considers only the number of multiple tests (number of SNPs), the model selection criterion is based on the calculation of a more permissive extended Bayesian Information Criterion eBIC [53]. P-values of marker-trait associations are given considering both the optimal model (with possibly several regressors) and each associated SNP as unique regressor avoiding p-value inflation due to forward regressor inclusion.

The multi-trait model takes advantage of the correlation structure between multiple phenotypes to increase power to detect not only pleiotropic genetic variants but also specific variants affecting only one of the correlated phenotypes [52]. Multi-testing correction on output p-values was performed considering False Discovery Rate (FDR) [54] control per chromosome (5% significance level).

For all significant associations, we computed the allelic effect  $\alpha$  as (Minor allele mean – Major allele mean)/2 and the individual percentage of phenotypic variance explained (PVE). Furthermore, for each trait, strict sense heritability was calculated as the ratio

$$h^2 = \frac{\sigma_{\Sigma SNP}^2}{\sigma_{\Sigma SNP}^2 + \sigma_E^2}$$

where  $\sigma_{\Sigma SNP}^2$  is the additive genetic variance explained by all associated

loci and  $\sigma_E^2$  is the remaining variance linked to both dominance, epistasis genetic effects and non-genetic residual error. In order to ensure their reliability, all detected associations were subjected to further validations (ANOVA tests) on BLUPs from the within-year model.

Estimates of  $r_{VS}^2$  around marker-trait-associated loci were computed to give a graphical representation of the pairwise LD using a modified version from the *snp.plotter* R

package [55]. Moreover, performing genome-wide sampling of 1,000 markers,  $r_{vs}^2$  distributions were prospected considering both inter-chromosomal and candidate loci SNP pairs, and compared to intra-chromosomal scale in order to detect any specific linkages.

### ***Haplotype analysis***

Groups of markers were constructed by considering all detected loci for each  $G$  and  $G \times Y$  BLUP phenotype. Marker haplotypes were identified among all the accessions according to the genotypes on the candidate loci for each group of markers. Mean phenotypic distributions of the different haplotypes comprising at least two accessions were compared using HSD post-hoc tests ( $\alpha=5\%$ ) for the associated  $G$  BLUP phenotypes considering both multi-year and within-year models. For haplotypes represented by only one individual, mean value was compared with the significantly differing groups from the HSD tests. Favorable (and unfavorable) haplotypes were defined for each multi-year and within-year model  $G$  BLUP phenotype according to the following criteria: (i) showing a significantly lower (or upper) group mean value than all other haplotypes groups and (ii) with stable performances among years.

### ***Candidate gene identification***

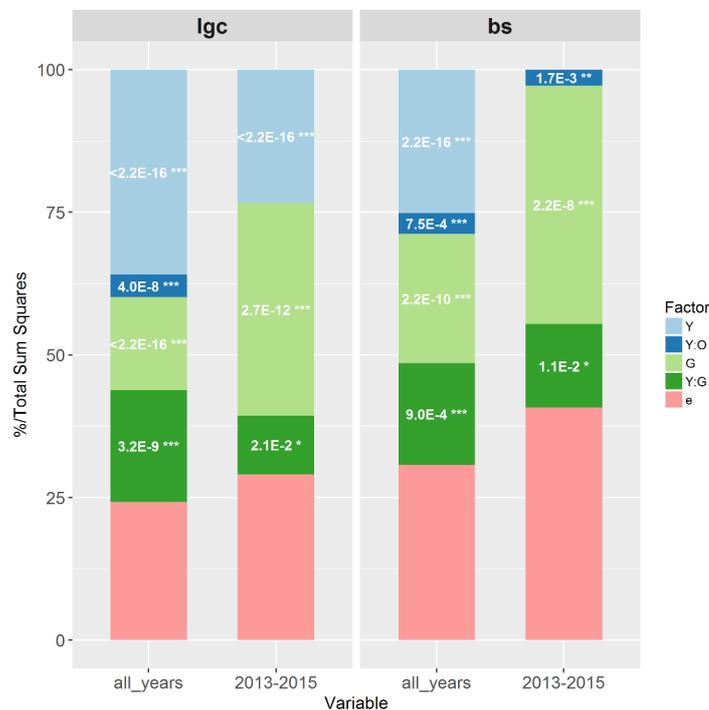
For each detected loci, candidate genes were identified and localized on the first and second version of the peach genome (*P. persica* v1.0 and v2.1, publicly available at <https://www.rosaceae.org/species/prunus/all>). Putative homolog protein annotations were obtained following Blastp searches upon the *Arabidopsis thaliana* (*A. thaliana*) genome using the TAIR10 database (<https://www.arabidopsis.org/>). SNP gene localization and effect on the protein sequence were determined using the GBrowse tool (<https://www.rosaceae.org/gb/gbrowse/>) on the *P. persica* v1.0 genome and ORFfinder (<https://www.ncbi.nlm.nih.gov/orffinder/>), respectively. Then, the predicted impact of non-synonymous SNP on the biological function of the protein

was evaluated with Provean v1.1.3 web-interface software (<http://provean.jcvi.org/index.php>) [56].

## Results

### *Phenotypic variance decomposition and heritability of partial resistance to bacterial canker*

Considering differences between all controls and inoculated shoots regardless of the genotype effect, a highly significant effect of bacterial inoculation was observed (Wilcoxon unilateral t.test  $p < 2.2E-16$ ) for *lgc* whatever the year and for *bs* in 2013 and 2015 (lesser extent in 2014  $p < 9.77E-12$  and 2016  $p < 9.06E-06$ ). In the variance analysis, both branch height and shoot diameter had non-significant effects on the variation for the two phenotypes and were thus discarded from the multi-year and within-year models.



**Figure 2: Phenotypic variance decomposition for *lgc* and *bs* considering all years and 2013 & 2015 data.**

For *lgc* and *bs* phenotypes, the chart displays the proportion of each fixed effect in the total sum of squares: year (Y), light blue; operator nested in year (Y:O), dark blue; genetic (G), light green; interaction year x genetic (Y:G), dark green; the random residuals, pink. P-values and their level of significance are indicated as following: \*\*\* $P < 0.001$ , \*\* $0.001 \leq P < 0.01$ , \* $0.01 \leq P < 0.05$ .

Considering all years, part of the total phenotypic variation was attributed in order of significance to year ( $p < 2.2E-16$ , 36% *lgc* and 25% *bs*), genotype (*lgc*  $p < 2.2E-16$ , 16% and *bs*  $p < 2.2E-10$ , 23%), genotype x year interaction (*lgc*  $p < 3.2E-09$ , 20% and *bs*  $p < 9.0E-04$ , 18%) and operator (nested in year) effects (*lgc*  $p < 4.0E-08$ , 4% and *bs*  $p < 1.7E-03$ , 4%) (*Figure 2*).

The  $G \times Y$  interaction term for *bs* was mainly due to scale-change (73% of the sum of squares) than cross-over interactions (method 1 from [57]). By contrast, for *lgc*, the sum of squares partitioning indicated a more equilibrated contribution from each type of interaction (58% and 42% for cross-over and scale-change, respectively). Cross-over interactions were most likely due to the re-ranking of individuals with non-extreme behaviours especially in 2014 and 2016. These two years were characterized as non-favorable for disease development compared to 2013 and 2015: ANOVA fixed effects for *lgc/bs* equaled 2.56/NA (2014) and 2.97/0.72 cm (2016), respectively compared to 3.90/5.62 (2013) and 7.94/4.87 cm (2015). Winters from 2014 and 2016 were characterized by relatively warmer temperatures especially in February. This could be shown in terms of degree days over the winters of 2013 to 2016 (see *Additional file 2*). Linear mixed models resulted in the selection of an optimal model (based on minimization of AIC, see Methods) for which, in addition to the heterogeneity of genetic variances, residual variance was either common (*lgc*) or specific (*bs*) to the year. Genetic variation displayed a quantitative distribution and broad-sense heritability  $H^2$  per trait depended therefore on yearly-winter expression (*Figure 3A & 3B*). For the year with maximum genetic variance (2015),  $H^2$  reached 59% for *lgc* and 78% for *bs* (in comparison, for 2013  $H^2_{lgc}=35\%$  and  $H^2_{bs}=53\%$ ).

Significant phenotypic correlations between 2013 and 2015 (*lgc*:  $r=0.97 \pm 0.32$ ,  $p=8.90E-03^{**}$  – *bs*:  $r=0.54 \pm 0.21$ ,  $p=2.91E-02^*$ ) were obtained from the corh vcov structure. *lgc* and *bs*  $G$  BLUPs were also highly correlated (Pearson correlation



observed aside from the slight proximity of covariate Q2 (the fraction of ancestry linked to Europe Continental group) with the susceptible phenotypes (*Additional file 4*).

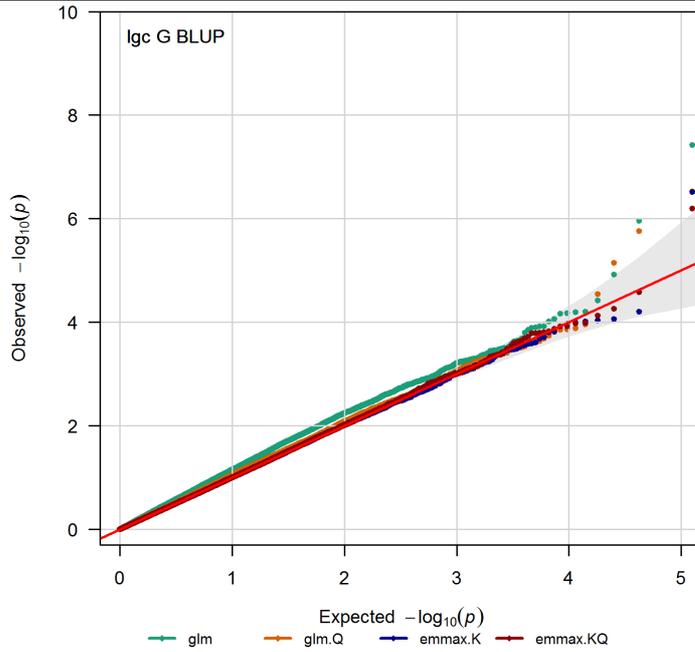
#### ***LD decay in the GWAS panel***

Pairwise mean LD estimates ( $r^2$ ) within each chromosome revealed a low level of dependence between markers. The range varied from  $2.88\text{E-}02 \pm 5.08\text{E-}02$  (chromosome 8) to  $4.13\text{E-}02 \pm 7.71\text{E-}02$  (chromosome 5) with a decrease of 20% (chromosome 5) to 31% (chromosome 4) when correcting for population structure and individual relatedness (for  $r_{vs}^2$  average values see Methods, *Additional file 5*). Moreover the intra-chromosomal LD in our population set decayed very quickly within 100 to 200 bp depending on the chromosome with no significant contribution of population stratification.

#### ***Association mapping and local LD patterns***

For all  $G$  and  $G \times Y$  BLUPs extracted from the multi-year model, QQ-plots revealed that the non-corrected ‘glm’ model did not markedly inflate p-values (see *Figure 4* for an illustration of *lgc*  $G$  BLUP – all other phenotypes are in *Additional file 6*). Controlling for Q yielded to a slight improvement in reduction of false-positive signals (type I error) while the K model (emmax K, see Methods) allowed the best prediction of p-values with a very limited deviation from the expected p-values resulting from the null model).

However, for some BLUP phenotypes (*lgc* and *bs* ( $G \times Y$ )<sub>2015</sub>), whatever the correction applied, both naïve ‘glm’ model, K and Q models led to equivalent estimations of p-values. Interestingly, controlling for both relatedness (K) and structure (Q) did not provide better results than the K model (*Figure 4*, *Additional file 6*).



**Figure 4: QQ-plot of the distribution of observed p-values for *lgc G* BLUP phenotype.** Comparison of non-corrected ‘glm’ model (green) and corrected glm.Q (orange), emmax.K (blue) and emmax.KQ (red) model prediction. The 95% confidence interval is indicated in grey.

The multi-locus mixed model approach (correcting for relatedness) detected 8 associations for all *G* and *G* × *Y* BLUP phenotypes with one colocalization for *lgc G*, (*G* × *Y*)<sub>2013</sub> and (*G* × *Y*)<sub>2015</sub> BLUPs and PVE ranging from 17.1 % to 45.5% (Table 1). Two main signals on the global *G* terms were localized on chromosomes 5 (LG5\_5394803) and 6 (LG6\_15273858), and associated respectively with *bs G* BLUP (p=8.10E-06) and *lgc G* BLUP (p=1.2E-07) with PVE reaching 25.9% and 41.4%, respectively. The colocalization between *G* and *G* × *Y* BLUPs on *lgc* indicated a modulation according to the year of both allelic effect ( $\alpha_{2013} = 0.14$ ,  $\alpha_{2015} = 0.26$ , ln(x+1) scale on the length in cm) and PVE on the phenotype (PVE<sub>2013</sub>= 40.1%, PVE<sub>2015</sub>= 45.5%). Additional associations were noticed on *bs (G* × *Y*)<sub>2015</sub> BLUP compared to *G* BLUP with loci detected on chromosomes 3 (LG3\_4322444, p=7.22E-06) and 5 (LG5\_4842835, p=7.00E-05). The number of associated loci per phenotype ranged from one (*lgc G*, (*G* × *Y*)<sub>2013</sub> and (*G* × *Y*)<sub>2015</sub> BLUPs) to three (*bs G* BLUP) with an important contribution of the additive genetic variance to the total phenotype variability ( $h_{lgc G}^2=0.41$ ,  $h_{bs G}^2=0.62$ ).

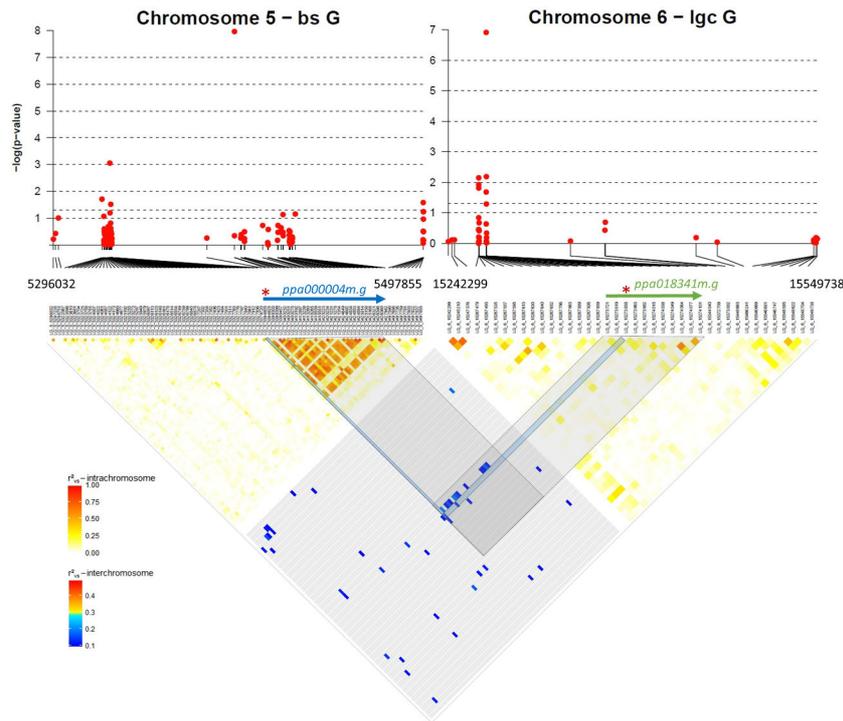
**Table 1: Detailed information about the 11 detected associations using  $G$  and  $G \times Y$  BLUPs for  $lgc$  and  $bs$ .**

SNP ID <sup>a</sup>	Maj/Min Allele <sup>b</sup>	Chr <sup>c</sup>	MAF <sup>d</sup>	Phenotype	GWAS model	P-val <sup>e</sup>	P-val <sub>reg</sub> <sup>f</sup>	P-val <sub>FDR</sub> <sup>g</sup>	$\alpha^h$	PVE <sup>i</sup>	SNP location	Gene ID <i>P. persica</i> v1.0 (v2.1)	TAIR10 BLASTp homology <sup>j</sup> Description	E-value
LG6_15273858	A/G	6	0.43	lgc $G$	multi-locus	1.22E-07	1.22E-07	-	0.20	41	Intron	<i>ppa018341m.g</i>	AT5G41980 Unknown Protein CONTAINS Harbinger transposase-derived nuclease domain	7E-51
				bs $G$	GEMMA	2.88E-07	-	2.21E-03	0.21	-				
			0.42	lgc ( $G \times Y$ ) <sub>2013</sub>	multi-locus	3.85E-07	3.85E-07	-	0.14	40				
			0.43	lgc ( $G \times Y$ ) <sub>2015</sub>	multi-locus	1.04E-06	1.04E-06	-	0.26	46				
LG5_5394803	G/A	5	0.19	lgc $G$	GEMMA	9.46E-06	-	4.88E-02	-0.15	-	Intron	<i>ppa000004m.g</i> (Prupe.5G050700)	AT4G17140 Predicted: Pleckstrin (PH) domain containing protein	0.0
				bs $G$	multi-locus	8.10E-06	1.05E-08	-	-0.32	26				
LG3_10897844	T/A	3	0.32	bs $G$	multi-locus	4.60E-03	2.73E-06	-	-0.31	19	CDS*	<i>ppa017602m.g</i> (Prupe.3G152500)	AT5G17460 Glutamyl-tRNA (Gln) amidotransferase subunit C	2E-66
LG2_22535054	T/C	2	0.30		multi-locus	8.08E-05	2.27E-06	-	-0.29	17	Intron	<i>ppa023961m.g</i> (Prupe.2G242100)	AT3G46850 Subtilisin-like protease SBT4.6	0.0
LG3_4322444	G/A	3	0.39	bs ( $G \times Y$ ) <sub>2015</sub>	multi-locus	7.22E-06	6.98E-06	-	0.46	43	CDS*	<i>ppa008956m.g</i> (Prupe.3G072200)	AT1G52360 Coatomer subunit beta'-2	6.8e-117
LG5_4842835	T/A	5	0.19		multi-locus	7.00E-05	1.68E-05	-	-0.25	20	Intron	<i>ppa020388m.g</i> (Prupe.5G044100)	AT4G18010 Type I inositol polyphosphate 5-phosphatase 2, IP5P2	0.0
LG7_18047191	C/T	7	0.45	lgc $G$	GEMMA	4.47E-07	-	2.72E-03	-0.05	3	Intron	<i>ppa001722m.g</i> (Prupe.7G179600)	AT5G66030 GRIP Protein	0.0
				bs $G$					0.19	12				

<sup>a</sup> SNP ID noted as LG[chromosome number]\_[physical position in bp]. <sup>b</sup> Major/Minor Allele. <sup>c</sup> Chromosome. <sup>d</sup> Minor Allele Frequency. <sup>e</sup> SNP p-value (considering as the only fixed regressor in the model). <sup>f</sup> SNP p-value from the multi-locus optimal-model (multiple regressors). <sup>g</sup> SNP p-value in mvLMM corrected for FDR (False Discovery Rate). <sup>h</sup> Allelic effect ( $\ln(x+1)$  scale). <sup>i</sup> Percentage of phenotypic variance explained by the SNP. <sup>j</sup> Functional annotation obtained using BLASTp homology on *A.thaliana* (TAIR10 database). CDS\*= Coding Sequence with SNP polymorphism impacting a non-synonymous substitution. *ppa018341m.g* is not referenced in *P.persica* v2.1

Despite the high phenotypic correlation ( $r=0.57\pm 0.10^{***}$ , Pearson correlation) between the two phenotypes, no marker colocalization was found.

Furthermore, 3 SNPs impacting both *lgc* *G* and *bs* *G* BLUPs were identified on chromosomes 5, 6 and 7 with the mvLMM (*Table 1*). Genetic correlation  $r_g(lgc-bs)$  reached 0.52 (MLE estimate) suggesting a common genetic pattern underlying *lgc* and *bs* variabilities. Among these markers, 2 SNPs were previously detected with the multi-locus mixed model for either *lgc* (LG6\_15273858,  $p_{FDR}=2.21E-03$ ) or *bs* (LG5\_5394803,  $p_{FDR}=4.88E-02$ ) with similar allelic effects. The association detected on chromosome 7 using mvLMM, accounted for a lower but significant part of the total phenotypic variability with PVE=3.0% for *lgc* and PVE=12.0% for *bs*.

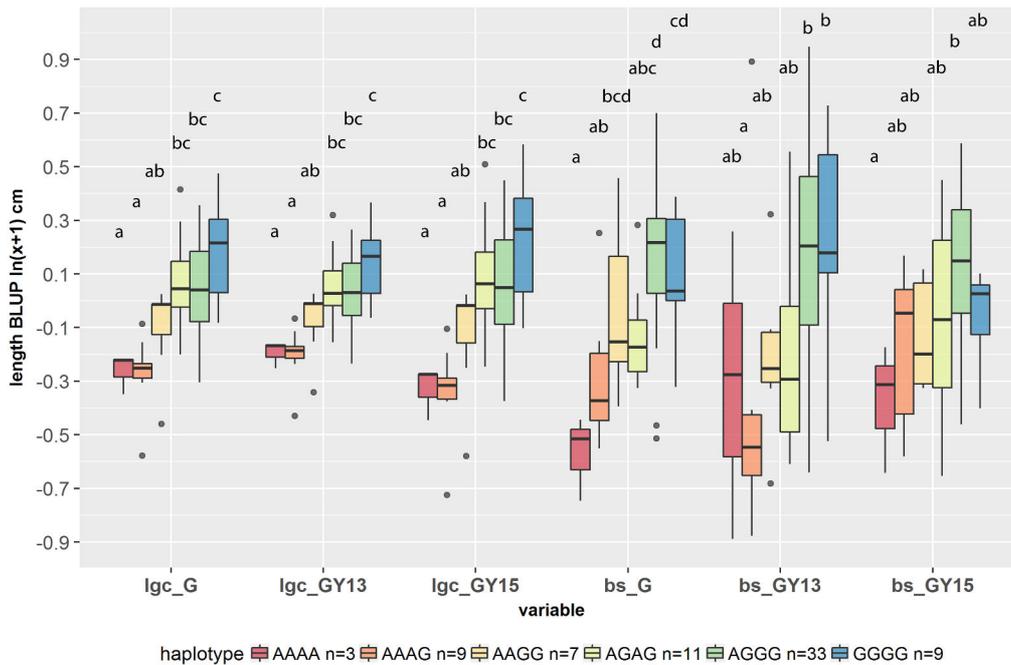


**Figure 5: Manhattan plots of the  $-\log_{10}(p\text{-values})$  over physical positions in windows surrounding LG6\_15273858 and LG5\_5394803 associations.**

Genomic windows are approximately 201.82 kb (chromosome 5) and 307.44 kb (chromosome 6). Significance level was determined with eBIC criterion. As an indication, Bonferroni threshold= $-\log_{10}(p\text{-value}_{thr})=6.10$ . Detected SNP are shown with red stars and overlapping candidate genes indicated with arrows. Heatmaps of pairwise LD estimates ( $r_{vs}^2$ ) were drawn both within the genomic window around each candidate and between the two genomic frames (inter-chromosome scale). Different colors and scales are used to represent the pairwise LD estimates between inter or intra SNP pairs. Only inter-chromosomal pairwise LD  $r_{vs}^2$  over the 99<sup>th</sup> percentile of the  $r_{vs}^2$  distribution are displayed.

Manhattan plots displaying p-values for the main loci LG6\_15273858 and LG5\_5394803 in multi-locus and multi-variate mixed models are shown in *Figure 5*. Intra-chromosomal LD of the two chromosomes have different patterns around each detected locus (*Figure 5*).

In the genomic region around LG5\_5394803, a relatively high LD block overlapped precisely with *pp00004m.g*, the underlying candidate gene (*cf Table 1*). This might be related to the above-mentioned higher mean  $r^2$  of chromosome 5 compared to all other chromosomes (*Additional file 5*). In contrast, SNP peak LG6\_15273858 and all other candidate loci (see *Additional file 7*) appeared to be located in a low LD region with no specific genomic structure.



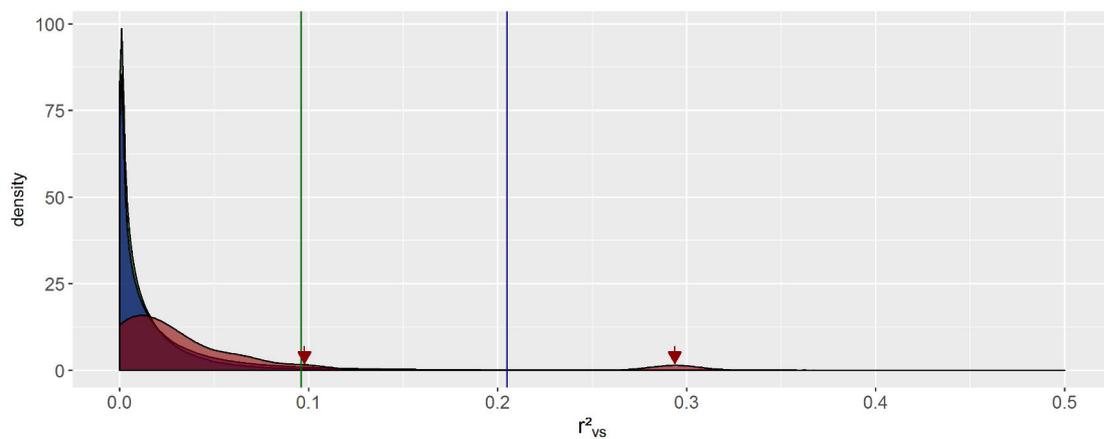
**Figure 6:  $G$  and  $G \times Y$  BLUP distributions according to the haplotype constituted by LG6\_15273858 and LG5\_5394803.**

Statistical differences between groups are indicated – HSD post-hoc tests on each  $G$  &  $G \times Y$  term. For *bs G* and  $(G \times Y)_{2015}$  BLUPs, tests were performed on residuals from an ANOVA model accounting the other associated loci from multi-locus model (LG3\_10897844 and LG52\_22535054 for *bs G* LG3\_4322444 and LG5\_4842835 for *bs (G × Y)*<sub>2015</sub>).

Y-axis values are indicated in  $\ln(x+1)$  scale. Haplotype class ‘AGAA’ (1 individual) was discarded in the boxplot.

As these two main SNPs were detected by the two mapping approaches, we focused on the effect of the haplotype, constituted by the genotypes on LG6\_15273858 and LG5\_5394803, on the total phenotypic variance of both *lgc* and *bs*. Both loci seemed to demonstrate an additive relation between allelic dose and the phenotype considering either  $G$  or  $G \times Y$  BLUPs (*Figure 6*) with no epistasis which was confirmed by performing an ANOVA (non-significant SNPxSNP interaction term  $p > 0.05$ ).

Individual effects of LG6\_15273858, LG5\_5394803 and all other detected loci on their respective  $G$  or  $G \times Y$  BLUP phenotypes are shown in *Additional File 7*.



**Figure 7: Distribution of  $r_{vs}^2$  unbiased estimates between intra-chromosomal, inter-chromosomal and detected loci random pairs of markers.**

Inter-chromosomal, intra-chromosomal and in between detected loci LD distributions are displayed in green, blue and red, respectively.

Calculations performed considering  $4.9E+06$  intra-chromosomal and  $4.3E+04$  inter-chromosomal random pairs of markers. 99<sup>th</sup> percentile lines of  $r_{vs}^2$  are drawn for intra-chromosomal and inter-chromosomal distributions. Correlations between candidate loci pairs LG6\_15273858-LG5\_5394803 (1<sup>st</sup> arrow) and LG5\_5394803-LG5\_4842835 (2<sup>nd</sup> arrow) are indicated.

Interestingly, a relatively high LD square correlation value ( $r_{vs}^2=9.76E-02$ ), higher than the 99<sup>th</sup> percentile of  $r_{vs}^2$  inter-chromosomal distribution, was revealed more specifically between the two main candidates LG6\_15273858 and LG5\_5394803 (*Figure 7*). This correlation between associated loci proved to be within a longer LD block (*Figure 5*) revealing an inter-chromosomal linkage between the underlying candidate genes *pp000004m.g* (chromosome 5) and *pp018341m.g* (chromosome 6). Likewise, the correlation between LG5\_5394803 and LG5\_4842835 ( $r_{vs}^2=2.94E-01$ )

detected both on chromosome 5 appeared to be significantly higher than the 99<sup>th</sup> percentile of  $r_{vs}^2$  intra-chromosomal distribution (*Figure 7*).

The genotype group distribution from all detected SNPs from the multi-year model BLUPs was tested on the within-year model BLUPs with heterogeneous significance levels according to the year and the marker PVE. Marker LG7\_18047191 was only significant for *bs2013* and *bs2015*. Significance of the two main loci was confirmed individually for *lgc2013*, *lgc2015*, *lgc2016* and *bs2015* with LG6\_15273858 and for *bs2013* and *bs2015* considering LG5\_5394803 (uni and multi-factorial ANOVA and post-hoc tests in *Additional file 8*). Strict-sense heritability estimates ranged from 10.85% (*lgc2014*) to 64.46% (*bs2015*) with non-negligible values even in the non-favorable years 2014 and 2016. Only *bs2016* led to non-significant results for all detected markers (*Additional file 8*).

### ***Haplotypes associated with partial resistance***

We found three groups of haplotypes among the core-collection having several SNPs associated to canker resistance; 18, 45 and 9 haplotypes shared 3, 5 and 2 markers associated to *lgc G*, *bs G* and *bs (G × Y)<sub>2015</sub>* BLUPs, respectively (*Additional file 9*). Mean comparison was performed on BLUP values from multi-year and within-year models after discarding low frequency classes in order to limit false-positive findings (type I error). Five haplotypes were identified with favorable alleles and stability of the resistance through years (1 with *lgc G*, 3 with *bs G*, and 1 with *bs (G × Y)<sub>2015</sub>* BLUP, see *Additional file 9*). Accessions A1814 and Harlayne showed two and all of the three defined favorable haplotypes, respectively. A0008 and A1314 were in the same way noticeable for their partial resistance (see *Figure 3C*) but their corresponding haplotypes were either not located in the most resistant groups or, if they were, they appeared to be unstable between years (see *Additional file 9*).

The mean comparison test was non-significant for the within-year model *lgc2015* BLUP for the first group of haplotypes (related to the 3 candidate loci associated to

*lgc G* BLUP) which might be due to a loss of power when haplotype number (18) is high compared to population size (ANOVA test  $p=0.1322$ ). None of the comparison tests between haplotypes were significant for years 2014 and 2016, which might be explained by the low genetic variance in these years where there was low disease severity.

### ***Candidate genes for resistance to bacterial canker***

Putative candidate gene annotations for all multi-locus and mvLMM associated SNPs were reported in Table 1. To date, none of the associated SNPs were reported for disease resistance in a *Rosaceae* species or colocalized with previously known QTL. It is noteworthy that, over all the 7 detected loci, only 2 candidates LG3\_10897844 and LG3\_4322444 targeted polymorphisms in coding sequences (CDS) of *ppa017602m.g* and *ppa008956m.g*, respectively. Mutations were both non-synonymous with one substitution (K49M in *ppa017602m.g*) leading to a deleterious predictive effect on the corresponding protein, a glutamyl-tRNA (Gln) amidotransferase subunit C protein. All other loci comprising the two most significant loci LG6\_15273858 (*ppa018341m.g*) and LG5\_5394803 (*ppa000004m.g*) were localized in introns. The associated functional annotations of the two main candidate genes were related to Harbinger transposase-derived nuclease and Pleckstrin homology-like (PH) domains, corresponding to a superfamily of transposons found in plants and animals [58] and a conserved domain particularly abundant in proteins involved in signal transduction pathways [59], respectively.

## **Discussion**

We aimed at investigating the genetic pattern of partial resistance of apricot to bacterial canker caused by *P. syringae* using a broad genetic basis in the species. We showed that, although being polygenic, bacterial canker partial resistance appeared to rest upon a limited number of additive components: 5 SNPs on chromosomes 2,3,5,6

and 7 partially explaining 45% and 62% of the total phenotypic variance of canker (*lgc*) and superficial browning (*bs*) lengths, respectively. A large part of the genotypic variance was due to a winter effect allowing the detection of two additional loci on chromosomes 3 and 5 in the year most favorable for the expression of canker symptoms. Among the 7 candidate loci, two main SNPs on chromosome 5 and 6 displayed additive contributions reaching 41% and 26% of the variation of *lgc* and *bs*. Furthermore, these two markers showed a long-range inter-chromosomal LD revealing a multi-locus-linked selection event through population history. Due to the short decay of LD in apricot, candidate loci were identified with a high resolution leading to the identification of promising candidate genes involving a potential signaling cross-talk between abscisic (ABA), jasmonic (JA), and salicylic (SA) acids, molecules known to mediate long-distance signaling in plant-pathogen interactions. Overall, this study contributed to the very first characterization of the genetic variation and genomic determinants of partial resistance to bacterial canker in a fruit tree species.

***Detection of genomic regions controlling variation of partial resistance to bacterial canker is dependent on winter frost intensity***

As highlighted by the heterogeneity of genetic variances across four years, genetic variation was highly dependent on winter-frost severity with low genetic variations in 2014 and 2016 while the highest were registered in 2013 and 2015. Considering this environmental dependency, the screening and characterization of cultivars for bacterial canker resistance was a difficult task. Other limits of bacterial canker phenotyping concern time-consuming efforts in the orchard, a very long period of symptom development, the need to calibrate the observations between operators and no guarantee of getting significant genetic variation between individuals. Marker-assisted selection (MAS) programs would be in that case an attractive, powerful and cost-efficient alternative to phenotyping. Among the two most severe years, broad-sense heritability estimates ranged from 35% (*lgc* in 2013) to 78% (*bs* in 2015) showing

a considerable level of variation of susceptibility in the core-collection. Previous studies on other patho-systems in fruit and forest trees have reported moderate to high heritability estimates ranging from 30 to 40% for pitch canker resistance in loblolly pine [60], from 70 to 89% in peach x *Prunus davidiana* [61] and 60 to 87% in apple [62] for powdery mildew resistance. Importantly, these studies reported the crucial need for controlling GxE interactions in order to estimate accurate genetic effects in the face of environmental noise in pluri-annual data.

The two phenotypes *lgc* and *bs* displayed a high correlation (Pearson correlation  $r=0.57\pm 0.10$ ,  $p=1.48E-07^{***}$ ). Regarding the important contribution of the genetic factor to the total variance of the phenotypes, this correlation suggests that a common genetic pattern may control both phenotypes.

We identified a total of 11 associations, over 7 candidate SNPs on chromosomes 2, 3, 5, 6 and 7, linked to the variance of *lgc* and *bs* *G* and *G* × *Y* BLUPs. Among the stable associations, two main SNPs were detected on chromosomes 5 (LG5\_5394803) and 6 (LG6\_15273858) and explained 41% and 26% of the variation of *lgc* and *bs*, respectively. In the case of *lgc*, a unique locus (LG6\_15273858) was detected both for *G* × *Y* interaction BLUPs and the overall *G* term with a modulation of the allelic effect and the PVE according to the year due to the heterogeneity of winter temperatures. Moreover, new associations for *bs* were detected in 2015 which was the most severe year in terms of disease expression, suggesting that data on specific years and climatic conditions have the potential to reveal additional regions linked to disease resistance. Similar results have been obtained in the case of QTL mapping on tree architecture [63] and phenological traits [64] in apple underlying the usefulness of focusing on *G* and *G* × *Y* terms to target both stable and environment-specific genetic determinants.

***Candidate genes can be found by exploiting the quick LD decay in apricot***

Mapping resolution in GWAS mostly depends on the decay of LD: the more the LD

decreases, the sharper the resolution appears around the detected locus [65]. Considering the original reproductive characteristics of the family, most *Rosaceae* species are cross-pollinating because of their self-incompatibly system. This results in a high number of effective recombination events and a rapid expected LD decay over the genome [66]. In our study, global mean LD over all chromosomes decayed over very short distances (100 to 200 bp) considering both corrections for population stratification and uncorrected estimates. Moreover, LD in the case of our population is not likely to rely on confounding structure effects related to differences in allele frequencies between groups resulting from non-random mating [67]. These results are congruent with those reported in the previous association study on apricot [30] based on a population comprising the major part of the material used in our work.

Regarding all these observations, GWAS is a very suitable approach to map the genetic regions linked to bacterial canker partial resistance in our diversity panel with a very precise resolution.

To date, none of the putative candidate genes or domains associated with the SNPs identified here have been reported for plant resistance function in the *Rosaceae* family. The two main SNPs LG6\_15273858 and LG5\_5394803 were located in genes with significant homologies to Harbinger transposase-derived nuclease (*ppa018341m.g*) and Pleckstrin Homology PH (*ppa000004m.g*) domains, respectively. These motifs are conserved domains found in a wide range of uncharacterized proteins with functions respectively in, chromatin remodeling [68] and signal transduction through interaction with membrane phosphoinositides [69, 70]. A role for phosphoinositides (binding to pleckstrin homology domain - *ppa000004m.g*) in regulating plant nuclear functions and possibly transcriptional activities through chromatin remodeling (Harbinger transposase-derived nuclease - *ppa018341m.g*) was revealed as a common response to a large range of both abiotic and biotic stresses [71]. It is noteworthy that the protein encoded by the candidate gene *ppa023961m.g* located on chromosome 2 (SNP:

LG2\_22535054) belongs to a large family of proteases: the subtilisin-like proteases (or subtilases) whose involvement in plant defense responses has been more and more described recently [72]. Two closely related members of the P69 subtilase family (*P69B* and *P69C*) were shown to be transcriptionally activated after *Ps* DC3000 infection with an elicitation from SA and JA [73, 74]. Although the link between *SBT4.6* encoded by *ppa023961m.g* and plant resistance remains to be investigated, transcripts from a close relative *SBT4.14* (also known as *AtXSP1*) were evidenced as contributing to xylem differentiation in *A. thaliana* [75]. Moreover, *SBT4.6* subtilase was predicted to be in the extracellular space (Uniprot database) suggesting a role in recognition of *Ps* before entry and colonization through the xylem, in the case of a compatible interaction. Interestingly, the G protein  $\beta$  WD-40 repeat, one of the domains of the protein encoded by another candidate gene (*ppa008956m.g* for *bs (G × Y)<sub>2015</sub>* association), had been previously shown to bind in-vitro with PH domains (*ppa000004m.g*) [69]. It has been shown in *A. thaliana* that myo-inositol polyphosphate 5-phosphatases, a large family englobing IP5P2 (*ppa020388m.g* including detected loci LG5\_4842835), hydrolyze a wide range of phosphoinositide phosphate substrates and is involved in stress responses through the abscisic acid (ABA) signaling pathway [76]. In addition, research that deployed transcriptional approaches demonstrated that the ABA pathway was one of the main targets of effectors secreted by *Ps* [77]. Regarding all their functions, most of the candidate genes or domains seem to support a role in a regulatory network involving signal transduction through membrane phosphoinositides in a cross talk involving ABA, with SA and JA, potentially. ABA mostly acts as negative regulator of disease resistance with an antagonistic effect on SA and JA [78, 79].

In addition, over all the seven candidate loci, only two single polymorphisms targeted CDS with non-synonymous substitutions. This observation emphasizes the importance of introns which can affect gene expression level or induce alternate splicing with an

impact on the phenotype. Moreover, the genotype data we used in the association mapping was restricted to variants located in gene-space regions. In turn, we may have missed inter-genic allelic diversity and regulatory variants potentially contributing to phenotype variation. The role of non-coding DNA such as promoters or/and enhancers in the variability of susceptibility to plant disease resistance has been demonstrated in several pathosystems including rice/*Xanthomonas oryzae* pv. *oryzae* [80] and maize/maize rough dwarf virus [81] and would be worth further investigation in the case of bacterial canker partial resistance.

These results open the door for subsequent validations to confirm the candidate loci polymorphisms over a pool of genetic resources and the correlations between the SNP genotype and the susceptibility level to the disease. To achieve this goal, pyrosequencing could be performed in order to confirm the haplotypes. Moreover, in order to take into account the significance of non-coding polymorphism candidates from our study, validation of the candidate genes could be performed with qRT-PCR across a pool of favorable and unfavorable haplotypes identified in our study.

***Combining genome-wide multi-locus and multi-variate association models is a powerful method to decipher the genomic pattern of partial resistance to bacterial canker***

There have been very few association studies using both genome-wide multi-locus and multi-variate mixed models [82, 83]. Our approach combining both methods proved to be powerful in limiting the number of candidates by controlling for LD, capturing a large part of broad-sense heritability and detecting common genetic variants impacting the two phenotypes.

The two main SNPs LG5\_5394803 and LG6\_15273858 were associated independently with multi-locus models on *lgc* (chromosome 6) and *bs* (chromosome 5) and co-detected on both phenotypes with the multi-variate model. The higher correlation between LG5\_5394803 and LG6\_15273858 ( $r_{vs}^2=9.76E-02$ ) compared to background

pairwise loci could explain the lack of co-localization with the multi-locus model. Indeed, the multi-locus model corrects for potential spurious associations due to intra or inter-chromosomal LD by re-estimating genetic variance at the stepwise inclusion of each associated locus as regressor in the model [51]. The drawback of such correction is that pleiotropic genetic variants cannot be detected, which in turn emphasizes the advantage of using multi-variate association models as a complementary approach. Noteworthy, multi-variate mixed models, by taking advantage of correlations between phenotypes can potentially also capture variants with smaller effects than those detected with traditional uni-variate analyses [84]. As an example, this allowed us to detect the small, but significant effect of LG7\_18047191 (chromosome 7) on both *lgc* and *bs G* BLUP (with PVE reaching 3 and 12%, respectively).

Despite the small size of our core-collection (73 accessions), we have successfully shown that our method has enough power to detect marker-trait associations for bacterial canker resistance. Nevertheless, our population might not have been large enough to capture SNPs with minor effects due to a lack of statistical power with type II (false-negative) errors. An interesting next analytic perspective would be the use of genomic selection (GS) models on a higher number of individuals [85]. This would assess the effect of all markers and in particular low-effect variants in the phenotypic variance regardless of their frequency in the population, allowing to validate the candidates while detecting new loci and thus giving an overall better estimation of the genetic effects [86]. In our case, due to genotypic sampling bias, these genetic effects could have been either overestimated for the detected loci [87] or underestimated for rare variants possibly linked to resistance. For instance, the cultivars A0008 and A1314 noticed for particular phenotype of partial resistance were not highlighted by the analysis of haplotypes in our study, but they might be relevant material to cross in progenies. GS models or linkage analysis using bi-parental, multi-parental or interconnected progenies [64, 88-90] derived from this material could thus map

additional rare variants which could have been removed in our study (filter  $MAF < 0.05$ ).

***Specific intra and inter-chromosomal LD patterns give insight into genomic variation through population history***

This study allowed us to reveal intra and inter-chromosomal patterns with either low or high LD structure around the candidate loci. More specifically, one of the main detected SNPs on chromosome 5 (LG5\_5394803) displayed a high LD block overlapping precisely with the underlying gene *ppa000004m.g*, while all other associations targeted low LD regions even considering the other detected loci (LG5\_4842835) located 552 kb downstream. We also detected a high level of inter-chromosomal LD ( $r_{vs}^2 = 9.76E-02$ ) between the two main detected SNPs included in the candidate genes of interest *ppa000004m.g* and *ppa018341m.g* on chromosomes 5 and 6.

A high level of LD may often be interpreted as a direct effect of selective breeding in a genomic area where genes of agronomic interest are gathered. It also could be possibly due to a population structure confounding effect but we have shown for both intra and inter-chromosomal LD patterns that high levels of LD remain after correcting for population structure and relatedness, indicating a biological relevance behind these allelic associations. Considering the broadly diverse germplasm we used through this study, resulting from numerous recombination episodes through evolution, such strong LD patterns, especially between unlinked loci (over different chromosomes) are unexpected. These observations could then suggest a possible linkage due to a strong multi-locus linked selection acting as an evolutionary counterforce preserving an efficient and functional synergy between *ppa000004m.g* and *ppa018341m.g*. The same hypothesis could be drawn from the long-range LD between LG5\_5394803 and LG5\_4842835 impacting either  $G$  or  $G \times Y$  variation of superficial browning.

High LD between physically unlinked loci is likely to happen when strong selection is applied in conjunction with epistasis, pleiotropy [91] and/or co-adaptation on the basis of biological mechanisms leading to a “functional linkage” as it has been described in *Arabidopsis thaliana* with the genes couple GS-OH and MAM1 [92]. These two genes located respectively on chromosomes 2 and 5, showed a significant long-range genome-wide correlation and epistasis effects with favorable allelic combinations impacting the biosynthetic pathway related to herbivore resistance [92]. While the biological relation between *ppa000004m.g* and *ppa018341m.g* needs further investigation, LG5\_5394803 and LG5\_4842835 were included in two genes (encoding PH domain protein and IP5P2) related to phosphoinositide substrates and membrane components. The proximity of ontologies between these two genes located 0.5 Mb apart suggests a functional linkage.

Both long-range and inter-chromosomal LD between some of the markers of interest and the annotation of their underlying genes emphasize the idea of a functional network involved in bacterial canker resistance. In addition, the particular LD block structure of *ppa000004m.g* could be explained by a prevalent role of this gene in this network, possibly as an immediate-early gene involved in the initial steps of *Ps* infection. The function of *ppa000004m.g* interacting with membrane phosphoinositides corroborates the key-gene hypothesis as the plasma membrane is one of the first battle grounds in the plant-pathogen interaction.

## Conclusion

We aimed at targeting the genetic pattern controlling susceptibility to bacterial canker in apricot by exploring the natural genetic diversity in a panel of 73 apricot accessions. To achieve this goal, we used a very dense set of 63,236 SNPs spread over the genome and performed association mapping using both multi-locus and multi-variate mixed

models on two correlated phenotypes resulting from controlled inoculations in the orchard during two years favorable for disease expression.

A total of 11 associations was detected with two main components located on chromosomes 5 and 6 contributing to a large part of additive variation of the phenotype (41% and 26% of PVE). Putative candidate genes linked to the abscisic acid pathway were reported with a very precise resolution given the extremely low LD shaping the apricot tree genome organization.

In terms of methodology, we also demonstrated the complementary of genome-wide association multi-locus and multi-variate mixed models allowing i) estimation of the additive individual contribution of all associated loci in the total phenotypic variation and ii) detection of common genetic variants impacting the two phenotypes. Subsequent analyses considering a bigger GWAS population and/or the use of bi or multi-parental QTL mapping approaches would allow a further validation of the associated loci.

In a breeding context, further research is required to assess the extent of resistance due to all identified loci relative to the diversity of *Ps* in the orchard. Combining resistance genes showing major and strain-specific effects with genes displaying minor and non-specific effects was suggested to be an efficient strategy to preserve sustainability of the resistance by generating contradictory selection pressures on pathogen evolution [93, 94]. In this perspective, the loci detected in this study could be used in MAS in order to complement the global multi-trait selection strategy [66]. Gene pyramiding would be a feasible approach not only for apricot trees but also considering in a larger perspective other *Prunus* species because of the synteny between their genomes [95].

## **Declarations**

### ***Ethics approval and consent to participate***

‘Not applicable’

### ***Consent for publication***

‘Not applicable’

### ***Availability of data and materials***

The datasets generated and/or analyzed during the current study are included in Additional files 10 and 11.

### ***Competing interests***

The authors declare that they have no competing interests.

## **Acknowledgements**

We acknowledge the experimental teams of UR-GAFL and Domaine de l’Amarine for their help in implementing the experimentations. We particularly thank Charlotte Chandeysson, Jean Leonetti, Eric Martin, Timmy Defert, Carlos Gil and Anne-Marie Ferreol for their highly valuable contribution in field experiment and phenotyping. Many thanks to Mouna Hadj Brahim for her work as Masters student. We are also grateful to Brigitte Mangin for sharing script and providing valuable scientific advice. We finally thank Christopher Sauvage, Bénédicte Quilot-Turion, Brigitte Mangin and Lucie Tamisier for reviewing the manuscript.

The Resibac CASDAR (2013-2016) and ABRIWG (2012-2014) projects supported this work.

MO was supported by a fellowship from the French Ministry of Agriculture and Agri Food and hosted as a PhD student.

All authors read, approved the final manuscript and declared that they have no conflict of interest in the publication of this document.

## References

1. St Clair DA: Quantitative disease resistance and quantitative resistance loci in breeding. *Annual Review of Phytopathology* 2010, 48(1):247-268.
2. Mundt CC, Cowger C, Garrett KA: Relevance of integrated disease management to resistance durability. *Euphytica* 2002, 124(2):245-252.
3. Mundt CC: Durable resistance: A key to sustainable management of pathogens and pests. *Infection, Genetics and Evolution* 2014, 27:446-455.
4. Crop production information: Apricots [<http://www.fao.org/faostat/en/#compare>]
5. Gardan L, Prunier JP, Luisetti J, Bezelgues JJ: Responsabilité de divers *Pseudomonas* dans le dépérissement bactérien de l'abricotier en France. *Revue de Zoologie Agricole et de Pathologie Végétale* 1973, 4:112-120.
6. Crosse JE: Bacterial diseases of stone-fruit trees in Britain: IX. Bacteriosis of apricot. *Transactions of the British Mycological Society* 1953, 36(1):38-45.
7. Berge O, Monteil CL, Bartoli C, Chandeysson C, Guilbaud C, Sands DC, Morris CE: A user's guide to a data base of the diversity of *Pseudomonas syringae* and its application to classifying strains in this phylogenetic complex. *Plos One* 2014, 9(9):e105547.
8. Scortichini M: Epidemiology and predisposing factors of some major bacterial diseases of stone and nut fruit trees species. *Journal of Plant Pathology* 2010, 92(1):S73-S78.
9. Lamichhane JR, Varvaro L, Parisi L, Audergon JM, Morris CE: Disease and frost damage of woody plants caused by *Pseudomonas syringae*: seeing the forest for the trees. *Advances in Agronomy* 2014, 126:235-295.
10. Kennelly MM, Cazorla FM, de Vicente A, Ramos C, Sundin GW: *Pseudomonas syringae* diseases of fruit trees: progress toward understanding and control. *Plant Disease* 2007, 91(1):4-17.
11. Vigouroux A: Ingress and spread of *Pseudomonas* in stems of peach and apricot promoted by frost-related water-soaking of tissues. *Plant Disease* 1989, 73(10):854-855.
12. Klement Z, Rozsnyay DS, Báló E, Pánczél M, Prileszky G: The effect of cold on development of bacterial canker in apricot trees infected with *Pseudomonas syringae* pv. *syringae*. *Physiological Plant Pathology* 1984, 24(2):237-246.
13. Young JM: Orchard management and bacterial diseases of stone fruit. *New Zealand Journal of Experimental Agriculture* 1987, 15(2):257-266.
14. Prunier JP, Bordjiba O: Effect of frost on bacterial necrosis of apricot buds. *Acta Horticulturae* 1991(293):495-502.
15. Prunier JP, Audergon JM, Cotta J: Etude de la sensibilité variétale de l'abricotier au chancre bactérien. In: *Programme de recherche Agrimed: Deuxièmes rencontres sur l'abricotier: 1993; Avignon: Office des publications officielles des communautés européennes; 1993: 111-112.*
16. Brun L, Warlop F, Mercier V, Broquaire J, Clauzel G, Gomez C, Parveaud CE, Audergon JM: Quelle sélection fruitière pour une production durable, à faible niveau d'intrants ? Méthodologie pour un réseau de sélection variétale décentralisée. *Innovations Agronomiques* 2011, 15:105-115.
17. Garrett CME: Screening *Prunus* rootstocks for resistance to bacterial canker, caused by *Pseudomonas morsprunorum*. *Journal of Horticultural Science* 1979, 54(3):189-193.

18. Mgbechi-Ezeri J, Porter L, Johnson KB, Oraguzie N: Assessment of sweet cherry (*Prunus avium* L.) genotypes for response to bacterial canker disease. *Euphytica* 2017, 213(7):145.
19. Bedford KE, Sholberg PL, Kappel F: Use of a detached leaf bioassay for screening sweet cherry cultivars for bacterial canker resistance. In: *XXVI International Horticultural Congress: Genetics and Breeding of Tree Fruits and Nuts: 2003*: International Society for Horticultural Science (ISHS), Leuven, Belgium; 2003: 365-368.
20. Li B, Hulin MT, Brain P, Mansfield JW, Jackson RW, Harrison RJ: Rapid, automated detection of stem canker symptoms in woody perennials using artificial neural network analysis. *Plant Methods* 2015, 11(57).
21. Davey JW, Hohenlohe PA, Etter PD, Boone JQ, Catchen JM, Blaxter ML: Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nature Reviews Genetics* 2011, 12(7):499-510.
22. Zhu C, Gore M, Buckler ES, Yu J: Status and prospects of association mapping in plants. *The Plant Genome Journal* 2008, 1(1):5.
23. Khan MA, Korban SS: Association mapping in forest trees and fruit crops. *Journal of Experimental Botany* 2012, 63(11):4045-4060.
24. Ogura T, Busch W: From phenotypes to causal sequences: using genome wide association studies to dissect the sequence basis for variation of plant development. *Current Opinion in Plant Biology* 2015, 23:98-108.
25. Nordborg M, Weigel D: Next-generation genetics in plants. *Nature* 2008, 456(7223):720-723.
26. Ingvarsson PK, Street NR: Association genetics of complex traits in plants. *New Phytologist* 2011, 189(4):909-922.
27. Huang X, Han B: Natural variations and genome-wide association studies in crop plants. *Annual Review of Plant Biology* 2014, 65(1):531-551.
28. Font i Forcada C, Oraguzie N, Reyes-Chin-Wo S, Espiau MT, Socias i Company R, Fernández i Martí A: Identification of genetic loci associated with quality traits in almond via association mapping. *Plos One* 2015, 10(6):e0127656.
29. Cao K, Zhou ZK, Wang Q, Guo J, Zhao P, Zhu GR, Fang WC, Chen CW, Wang XW, Wang XL *et al*: Genome-wide association study of 12 agronomic traits in peach. *Nature Communications* 2016, 7(13246).
30. Mariette S, Wong Jun Tai F, Roch G, Barre A, Chague A, Decroocq S, Groppi A, Laizet Yh, Lambert P, Tricon D *et al*: Genome-wide association links candidate genes to resistance to Plum Pox Virus in apricot (*Prunus armeniaca*). *New Phytologist* 2016, 209(2):773-784.
31. Arumuganathan K, Earle ED: Nuclear DNA content of some important plant species. *Plant Molecular Biology Reporter* 1991, 9(3):208-218.
32. Folta KM, Gardiner SE: Genomics-based opportunities in apricot. In: *Genetics and Genomics of Rosaceae*. Edited by Folta KM, Gardiner SE. New York, NY: Springer New York; 2009: 315-335.
33. Zaurov D, Molnar T, Eisenman S, M. Ford T, Mavlyanova R, M. Capik J, Reed Funk C, C. Goffreda J: Genetic resources of apricots (*Prunus armeniaca* L.) in central Asia. *HortScience* 2013, 48(6):681-691.

34. Yilmaz KU, Gurcan K: Genetic Diversity in Apricot. In: *Genetic Diversity in Plants Mahmut Caliskan*. Edited by IntechOpen; 2012: 249-270.
35. Bourguiba H, Audergon J-M, Krichen L, Trifi-Farah N, Mamouni A, Trabelsi S, D'Onofrio C, Asma BM, Santoni S, Khadari B: Loss of genetic diversity as a signature of apricot domestication and diffusion into the Mediterranean Basin. *BMC Plant Biology* 2012, 12(49).
36. Rikkerink EHA, Oraguzie NC, Gardiner SE: Prospects of association mapping in perennial horticultural crops. In: *Association Mapping in Plants*. Edited by Oraguzie NC, Rikkerink EHA, Gardiner SE, De Silva HN. New York: Springer; 2007: 249-269.
37. Bartoli C, Carrere S, Lamichhane JR, Varvaro L, Morris CE: Whole-genome sequencing of 10 *Pseudomonas syringae* strains representing different host range spectra. *Genome Announcements* 2015, 3(2):e00379-00315.
38. King EO, Ward MK, Raney DE: Two simple media for the demonstration of pyocyanin and fluorescein. *Journal of Laboratory and Clinical Medicine* 1954, 44(2):301-307.
39. Manners JG: Disease symptoms. In: *Principles of Plant Pathology*. Cambridge University Press; 1993: 51-65.
40. Cao T, Sayler RJ, DeJong TM, Kirkpatrick BC, Bostock RM, Shackel KA: Influence of stem diameter, water content, and freezing-thawing on bacterial canker development in excised stems of dormant stone fruit. *Phytopathology* 1999, 89(10):962-966.
41. Gilmour AR, Thompson R, Cullis BR: Average information REML: an efficient algorithm for variance parameter estimation in linear mixed models. *Biometrics* 1995, 51(4):1440-1450.
42. Verde I, Abbott AG, Scalabrin S, Jung S, Shu S, Marroni F, Zhebentyayeva T, Dettori MT, Grimwood J, Cattonaro F *et al*: The high-quality draft genome of peach (*Prunus persica*) identifies unique patterns of genetic diversity, domestication and genome evolution. *Nat Genet* 2013, 45(5):487-494.
43. Kim SY, Lohmueller KE, Albrechtsen A, Li Y, Korneliussen T, Tian G, Grarup N, Jiang T, Andersen G, Witte D *et al*: Estimation of allele frequency and association mapping using next-generation sequencing data. *BMC Bioinformatics* 2011, 12(1):231.
44. Aulchenko YS, Ripke S, Isaacs A, van Duijn CM: GenABEL: an R library for genome-wide association analysis. *Bioinformatics* 2007, 23(10):1294-1296.
45. Alexander DH, Novembre J, Lange K: Fast model-based estimation of ancestry in unrelated individuals. *Genome Research* 2009, 19(9):1655-1664.
46. Kang HM, Sul JH, Service SK, Zaitlen NA, Kong S-y, Freimer NB, Sabatti C, Eskin E: Variance component model to account for sample structure in genome-wide association studies. *Nat Genet* 2010, 42(4):348-357.
47. Mangin B, Siberchicot A, Nicolas S, Doligez A, This P, Cierco-Ayrolles C: Novel measures of linkage disequilibrium that correct the bias due to population structure and relatedness. *Heredity* 2011, 108:285-291.
48. Heifetz EM, Fulton JE, O'Sullivan N, Zhao H, Dekkers JCM, Soller M: Extent and consistency across generations of linkage disequilibrium in commercial layer chicken breeding populations. *Genetics* 2005, 171(3):1173-1181.
49. Turner SD: qqman: an R package for visualizing GWAS results using Q-Q and manhattan plots. *bioRxiv* 2014.

50. Bonnafous F, Fievet G, Blanchet N, Boniface M-C, Carrère S, Gouzy J, Legrand L, Marage G, Bret-Mestries E, Munos S *et al*: Comparison of GWAS models to identify non-additive genetic control of flowering time in sunflower hybrids. *Theoretical and Applied Genetics* 2018, 131(2):319–332.
51. Segura V, Vilhjálmsson BJ, Platt A, Korte A, Seren Ü, Long Q, Nordborg M: An efficient multi-locus mixed model approach for genome-wide association studies in structured populations. *Nat Genet* 2012, 44(7):825-830.
52. Zhou X, Stephens M: Efficient multivariate linear mixed model algorithms for genome-wide association studies. *Nature Methods* 2014, 11(4):407-409.
53. Chen J, Chen Z: Extended bayesian information criteria for model selection with large model spaces. *Biometrika* 2008, 95(3):759-771.
54. Benjamini Y, Hochberg Y: Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B (Methodological)* 1995, 57(1):289-300.
55. Luna A, Nicodemus KK: snp.plotter: an R-based SNP/haplotype association and linkage disequilibrium plotting package. *Bioinformatics* 2007, 23(6):774-776.
56. Choi Y, Chan AP: PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels. *Bioinformatics* 2015, 31(16):2745-2747.
57. Muir W, Nyquist WE, Xu S: Alternative partitioning of the genotype-by-environment interaction. *Theoretical and Applied Genetics* 1992, 84(1-2):193-200.
58. Kapitonov VV, Jurka J: Harbinger transposons and an ancient HARB1 gene derived from a transposase. *DNA and Cell Biology* 2004, 23(5):311-324.
59. Musacchio A, Gibson T, Rice P, Thompson J, Saraste M: The PH domain: a common piece in the structural patchwork of signalling proteins. *Trends in Biochemical Sciences* 1993, 18(9):343-348.
60. Quesada T, Gopal V, Cumbie WP, Eckert AJ, Wegrzyn JL, Neale DB, Goldfarb B, Huber DA, Casella G, Davis JM: Association mapping of quantitative disease resistance in a natural population of loblolly pine (*Pinus taeda* L.). *Genetics* 2010, 186(2):677-686.
61. Foulongne M, Pascal T, Pfeiffer F, Kervella J: QTLs for powdery mildew resistance in peach × *Prunus davidiana* crosses: consistency across generations and environments. *Molecular Breeding* 2003, 12(1):33-50.
62. Calenge F, Durel CE: Both stable and unstable QTLs for resistance to powdery mildew are detected in apple after four years of field assessments. *Molecular Breeding* 2006, 17(4):329-339.
63. Segura V, Durel C-E, Costes E: Dissecting apple tree architecture into genetic, ontogenetic and environmental effects: QTL mapping. *Tree Genetics & Genomes* 2009, 5(1):165-179.
64. Allard A, Bink MCAM, Martinez S, Kelner J-J, Legave J-M, di Guardo M, Di Pierro EA, Laurens F, van de Weg EW, Costes E: Detecting QTLs and putative candidate genes involved in budbreak and flowering time in an apple multiparental population. *Journal of Experimental Botany* 2016, 67(9):2875-2888.
65. Brachi B, Morris GP, Borevitz JO: Genome-wide association studies in plants: the missing heritability is in the field. *Genome Biology* 2011, 12(232).

66. Arús P, Gardiner S: Genomics for improvement of *Rosaceae* temperate tree fruit. In: *Genomics-Assisted Crop Improvement: Vol 2: Genomics Applications in Crops*. Edited by Varshney RK, Tuberosa R. Dordrecht: Springer Netherlands; 2007: 357-397.
67. Zhang W, Collins A, Gibson J, Tapper WJ, Hunt S, Deloukas P, Bentley DR, Morton NE: Impact of population structure, effective bottleneck time, and allele frequency on linkage disequilibrium maps. *Proceedings of the National Academy of Sciences of the United States of America* 2004, 101(52):18075-18080.
68. Duan C-G, Wang X, Xie S, Pan L, Miki D, Tang K, Hsu C-C, Lei M, Zhong Y, Hou Y-J *et al*: A pair of transposon-derived proteins function in a histone acetyltransferase complex for active DNA demethylation. *Cell Research* 2016, 27(2):226-240.
69. Rebecchi MJ, Scarlata S: Pleckstrin homology domains: a common fold with diverse functions. *Annual Review of Biophysics and Biomolecular Structure* 1998, 27(1):503-528.
70. van Leeuwen W, Ökrész L, Bögre L, Munnik T: Learning the lipid language of plant signalling. *Trends in Plant Science* 2004, 9(8):378-384.
71. Dieck CB, Boss WF, Perera IY: A role for phosphoinositides in regulating plant nuclear functions. *Frontiers in Plant Science* 2012, 3(50).
72. Figueiredo J, Sousa Silva M, Figueiredo A: Subtilisin-like proteases in plant defence: the past, the present and beyond. *Molecular Plant Pathology* 2018, 19(4):1017-1028.
73. Jordá L, Coego A, Conejero V, Vera P: A genomic cluster containing four differentially regulated subtilisin-like processing protease genes is in tomato plants. *Journal of Biological Chemistry* 1999, 274(4):2360-2365.
74. Zhao Y, Thilmony R, Bender CL, Schaller A, He SY, Howe GA: Virulence systems of *Pseudomonas syringae* pv. *tomato* promote bacterial speck disease in tomato by targeting the jasmonate signaling pathway. *The Plant Journal* 2003, 36(4):485-499.
75. Zhao C, Johnson BJ, Kositsup B, Beers EP: Exploiting secondary growth in *Arabidopsis*. Construction of xylem and bark cDNA libraries and cloning of three xylem endopeptidases. *Plant Physiology* 2000, 123(3):1185-1196.
76. Gunesekera B, Torabinejad J, Robinson J, Gillaspay GE: Inositol polyphosphate 5-phosphatases 1 and 2 are required for regulating seedling growth. *Plant Physiology* 2007, 143(3):1408-1417.
77. de Torres-Zabala M, Truman W, Bennett MH, Lafforgue G, Mansfield JW, Rodriguez Egea P, Bögre L, Grant M: *Pseudomonas syringae* pv. *tomato* hijacks the *Arabidopsis* abscisic acid signalling pathway to cause disease. *The EMBO Journal* 2007, 26(5):1434-1443.
78. Lievens L, Pollier J, Goossens A, Beyaert R, Staal J: Abscisic acid as pathogen effector and immune regulator. *Frontiers in Plant Science* 2017, 8(587).
79. Checker VG, Kushwaha HR, Kumari P, Yadav S: Role of phytohormones in plant defense: signaling and cross talk. In: *Molecular Aspects of Plant-Pathogen Interaction*. Edited by Singh A, Singh IK. Singapore: Springer Singapore; 2018: 159-184.
80. Cui H, Wang C, Qin T, Xu F, Tang Y, Gao Y, Zhao K: Promoter variants of Xa23 alleles affect bacterial blight resistance and evolutionary pattern. *Plos One* 2017, 12(10):e0185925.
81. Shi L, Weng J, Liu C, Song X, Miao H, Hao Z, Xie C, Li M, Zhang D, Bai L *et al*: Identification of promoter motifs regulating ZmeIF4E expression level involved in maize rough

dwarf disease resistance in maize (*Zea Mays* L.). *Molecular Genetics and Genomics* 2013, 288(3):89-99.

82. Albert E, Segura V, Gricourt J, Bonnefoi J, Derivot L, Causse M: Association mapping reveals the genetic architecture of tomato response to water deficit: focus on major fruit quality traits. *Journal of Experimental Botany* 2016, 67(22):6413-6430.

83. Jaiswal V, Gahlaut V, Meher PK, Mir RR, Jaiswal JP, Rao AR, Balyan HS, Gupta PK: Genome wide single locus single trait, multi-locus and multi-trait association mapping for some important agronomic traits in common wheat (*T. aestivum* L.). *Plos One* 2016, 11(7):e0159343.

84. Yang JJ, Williams LK, Buu A: Identifying pleiotropic genes in genome-wide association studies for multivariate phenotypes with mixed measurement scales. *Plos One* 2017, 12(1):e0169893.

85. Meuwissen THE, Hayes BJ, Goddard ME: Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 2001, 157(4):1819-1829.

86. Crossa J, Pérez-Rodríguez P, Cuevas J, Montesinos-López O, Jarquín D, de los Campos G, Burgueño J, González-Camacho JM, Pérez-Elizalde S, Beyene Y *et al*: Genomic selection in plant breeding: methods, models, and perspectives. *Trends in Plant Science* 2017, 22(11):961-975.

87. Utz HF, Melchinger AE, Schön CC: Bias and sampling error of the estimated proportion of genotypic variance explained by quantitative trait loci determined from experimental data in maize using cross validation and validation with independent samples. *Genetics* 2000, 154(4):1839-1849.

88. Di Pierro EA, Gianfranceschi L, Di Guardo M, Koehorst-van Putten HJJ, Kruisselbrink JW, Longhi S, Troggio M, Bianco L, Muranty H, Pagliarani G *et al*: A high-density, multi-parental SNP genetic map on apple validates a new mapping approach for outcrossing species. *Horticulture Research* 2016, 3(16057).

89. Billotte N, Jourjon MF, Marsillac N, Berger A, Flori A, Asmady H, Adon B, Singh R, Nouy B, Potier F *et al*: QTL detection by multi-parent linkage mapping in oil palm (*Elaeis guineensis* Jacq.). *Theoretical and Applied Genetics* 2010, 120(8):1673-1687.

90. Tisné S, Pomiès V, Riou V, Syahputra I, Cochard B, Denis M: Identification of ganoderma disease resistance loci using natural field infection of an oil palm multiparental population. *G3: Genes/Genomes/Genetics* 2017, 7(6):1683-1692.

91. Skelly DA, Magwene PM, Stone EA: Sporadic, global linkage disequilibrium between unlinked segregating sites. *Genetics* 2016, 202(2):427-437.

92. Brachi B, Meyer CG, Villoutreix R, Platt A, Morton TC, Roux F, Bergelson J: Coselected genes determine adaptive variation in herbivore resistance throughout the native range of *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences* 2015, 112(13):4032-4037.

93. Pilet-Nayel M-L, Moury B, Caffier V, Montarry J, Kerlan M-C, Fournet S, Durel C-E, Delourme R: Quantitative resistance to plant pathogens in pyramiding strategies for durable crop protection. *Frontiers in Plant Science* 2017, 8(1838).

94. Calenge F, Faure A, Goerre M, Gebhardt C, Van de Weg WE, Parisi L, Durel CE: Quantitative trait loci (QTL) analysis reveals both broad-spectrum and isolate-specific QTL

for scab resistance in an apple progeny challenged with eight isolates of *Venturia inaequalis*. *Phytopathology* 2004, 94(4):370-379.

95. Dirlewanger E, Graziano E, Joobeur T, Garriga-Caldere F, Cosson P, Howad W, Arus P: Comparative mapping and marker-assisted selection in *Rosaceae* fruit crops. *Proceedings of the National Academy of Sciences* 2004, 101(26):9891-9896.

## Éléments complémentaires de validation

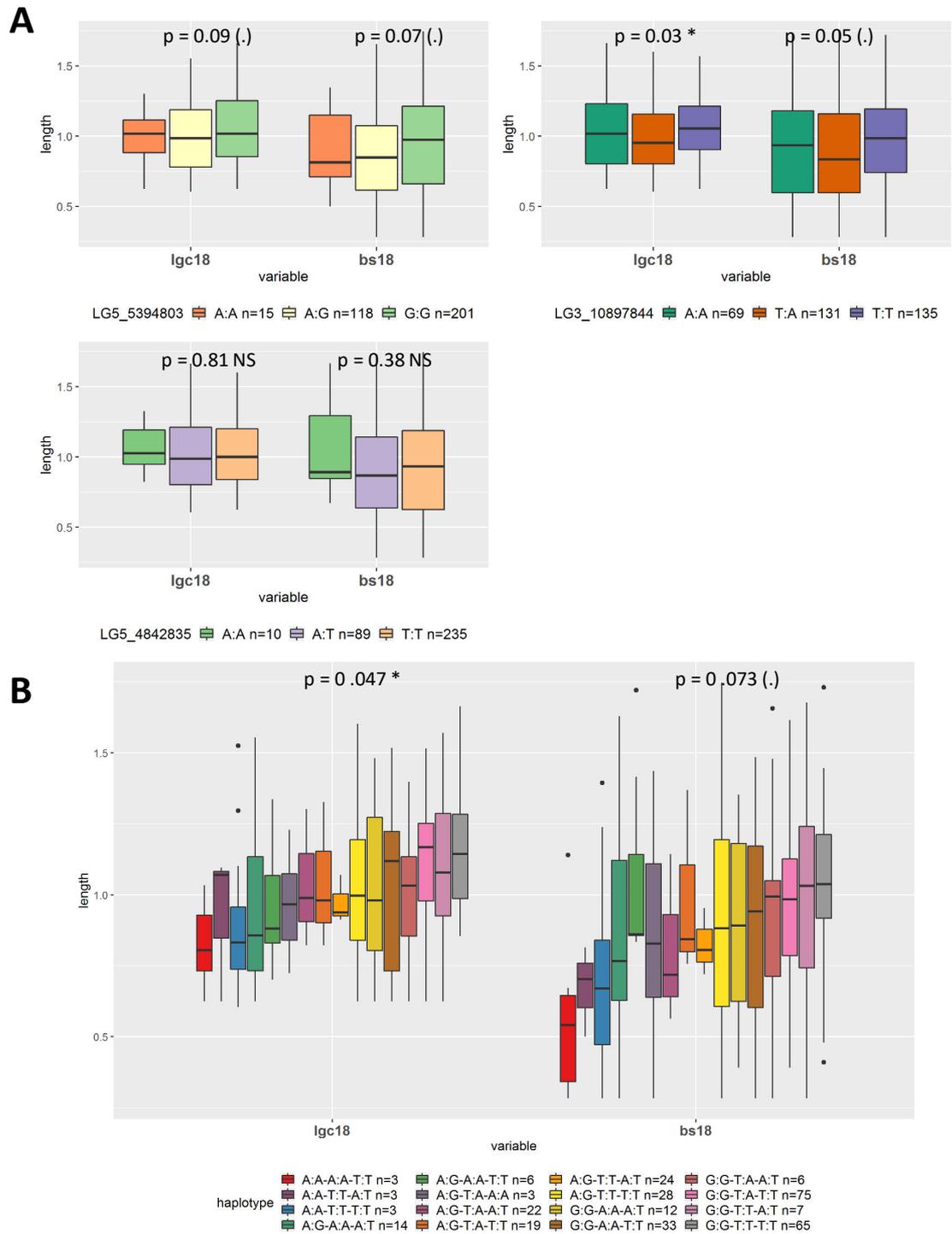
Sur les 7 marqueurs identifiés, 4 jeux d'amorces KASPar ont pu être développés ciblant spécifiquement les SNPs candidats LG5\_5394803 (SNP majeur), LG3\_10897844, LG5\_4842835 et LG7\_18047191 exprimant respectivement 26% et 19% de la variance de *bs G*, 20% de celle de *bs (G × Y)<sub>2015</sub>*, et enfin à la fois 3% et 12% des variances de *lgc G* et *bs G*. Il n'a pas été possible de développer des amorces spécifiques au SNP majeur LG6\_15273858 en raison de sa localisation à proximité immédiate d'une zone répétée du génome. Néanmoins du fait de sa liaison à l'échelle du génome avec LG5\_4842835, pour lequel des amorces ont été établies, il est attendu que la variance expliquée par ce marqueur soit en partie expliquée par LG5\_4842835.

La comparaison des données de marquage exploitées pour la GWAS avec celles issues de l'utilisation des amorces développées s'est effectuée sur la base d'un échantillonnage indépendant des accessions de la core-collection élargie (comprenant les 73 individus exploités en GWAS). Les résultats des deux approches sont globalement très cohérents avec un taux d'erreur de 1/94, 2/96, null et 1/95 pour LG5\_5394803 (SNP majeur), LG3\_10897844, LG5\_4842835 et LG7\_1804719, respectivement. Les données de polymorphisme exploitées en GWAS peuvent donc être considérées comme valides.

Le génotypage associé au phénotypage 2018 de près de 335 accessions sur la base des amorces ciblant 3 des 4 marqueurs nous a permis d'évaluer la corrélation entre les génotypes obtenus et les différences de sensibilité à la maladie. Les profils génotypiques ciblant le 4<sup>e</sup> marqueur (LG7\_18047191) sont à ce jour en cours de traitement par BioGEVES.

Les résultats des ANOVA à 1 facteur testant individuellement la significativité de chacun des 3 marqueurs sont indiqués au niveau de la *figure 27*.

Sur les 3 marqueurs, seul LG3\_10897844 présente un effet significatif ( $p=0.03^*$ ) sur la distribution de *lgc* avec un effet en limite de significativité sur *bs* ( $p=0.05$ ). Le marqueur LG5\_5394803 présente également des effets en limite de significativité sur les deux phénotypes (*lgc* :  $p=0.09$ , *bs*:  $p=0.07$ ). Enfin, le marqueur LG5\_4842835 ne présente aucune influence sur *lgc* et *bs*. Une étude analogue effectuée sur les haplotypes constitués par les génotypes aux 3 marqueurs indique un faible effet de l'haplotype sur *lgc* ( $p=0.047^*$ ) et un effet en limite de significativité sur *bs* ( $p=0.073$ ).



**Figure 27 :** Boxplots représentant la distribution de l'étendue des symptômes (BLUP *lgc* et *bs*) issus du phénotypage de validation en 2018 en fonction des génotypes aux marqueurs candidats LG5\_5394803, LG3\_10897844, LG5\_4842835.

- A. Comparaison des génotypes testés sur chacun des 3 marqueurs pris individuellement.  
 B. Comparaison des haplotypes constitués par les génotypes au niveau des 3 marqueurs.

Les p-values issues des ANOVA effectuées sur chaque marqueur et le niveau de significativité ( $\alpha=0.05$ ) du test sont indiqués sur chaque boxplot. NS : non significatif, (.) : en limite de significativité ( $0.05 < p < 0.1$ ), \* :  $0.01 < p < 0.05$ .

Cette faible significativité (voire nulle) des effets des marqueurs candidats peut s'expliquer par plusieurs facteurs :

- Un effet significatif de la parcelle (*lgc* :  $p=2.4 \cdot 10^{-3**}$ , *bs*:  $p=2.6 \cdot 10^{-3**}$ ), sur l'expression des symptômes, qui n'a pas pu être dissocié correctement de l'effet génétique dans la modélisation statistique, du fait du plan incomplet. L'ensemble des 335 géotypes sont en effet répartis sur 4 parcelles, pouvant présenter des conditions pédoclimatiques hétérogènes.
- La taille bien supérieure du panel testé (335), face à la population traitée en GWAS (73). En effet, il n'est pas irraisonnable de supposer que ces résultats résultent d'un bruit de fond dans les distributions de *lgc* et *bs* lié à l'existence de composantes génétiques à effet supérieur, non identifiées dans le panel de 73 accessions. Les héritabilités relativement importantes des phénotypes ( $H^2_{lgc} = 0.57$ ,  $H^2_{bs} = 0.62$ ) semblent supporter cette hypothèse.

Ce dernier point est à relier au problème de dépendance de l'estimation des effets avec la taille et l'échantillonnage de la population menant très souvent à des effets surestimés des QTLs dans le cas de populations à effectifs limités (Utz *et al.*, 2000). Cette limite est d'autant plus marquée, dans le cas d'études ciblant des traits fortement influencés par l'environnement à héritabilité faible ou moyenne.

Afin de mesurer la dépendance des effets des QTLs à la population étudiée, il a ainsi été suggéré d'effectuer des procédures de validation croisée, analogues aux modèles de sélection génomique, consistant en une recherche de QTLs sur une partie des géotypes, prise aléatoirement (population d'estimation) et en une procédure de test de la validité des QTLs sur les géotypes restants (population de test) (Utz *et al.*, 2000). Le calcul des corrélations entre les valeurs phénotypiques prédites des individus de la population de test (sur la base du modèle établi à l'aide de la population d'estimation) et leurs valeurs réelles observées permet de se rendre compte de l'effet de l'échantillonnage génétique sur l'estimation des effets des QTLs. A l'issue d'un nombre important de ré-échantillonnages aléatoires des populations d'estimation et de tests, la distribution des corrélations mène à l'estimation d'un biais permettant d'apprécier la dépendance des effets génétiques avec la constitution de la population.

Partie IV- Cartographie de la  
résistance à *Psy* par analyse de  
liaison au sein d'une descendance  
pseudo-F1 et de trois  
descendances interconnectées de  
2<sup>ème</sup> génération

# Introduction

Cette étude est sous la forme d'un manuscrit d'article en cours de préparation, pour lequel la revue ciblée est *Molecular Breeding*.

L'article vise à cartographier la sensibilité à *Psy* au sein d'une (i) population pseudo F1 'BerBa' (Bergeron x Bakour) de 218 individus et de (ii) 3 descendance interconnectées de 307 hybrides. Ces trois familles de 2<sup>ème</sup> génération sont issues de 3 hybrides de la BerBa présélectionnés pour leur faible sensibilité au chancre, croisés individuellement avec 'A4049', un géniteur cumulant auto-fertilité et résistance à la sharka.

Similairement à la GWAS, les phénotypes exploités au sein de chaque dispositif sont issus de 3 et 2 campagnes d'inoculations artificielles réalisées en période hivernale en verger. Ces notations continues ont été complétées dans le cas des populations interconnectées d'une notation globale de mortalité issue d'observations conduites en 2017 et 2018, 2 années à forte pression de la maladie en verger.

Cette étude a permis d'identifier 27 QRLs dont plusieurs colocalisations soit entre années, soit entre variables dans la même année.

Les perspectives immédiates issues du travail effectué consistent en une densification du marquage sur la carte consensus afin de :

combler le gap entre les deux groupes de liaison du chromosome 4 car le début du chromosome 4\_2 présente une région d'intérêt cumulant plusieurs QRLs relatifs aux symptômes issus des inoculations artificielles en 2016 et 2017,

améliorer la précision de cartographie autour du QRL majeur situé sur le chromosome 3 (pseudo- $R^2=25\%$ ) propre au phénomène de mortalité.

Les figures et tables additionnelles de l'article sont présentées au niveau des annexes II référencées à la fin du manuscrit.

# Detecting QTLs involved in partial resistance of apricot to *Pseudomonas syringae* in biparental and interconnected progenies

Mariem Omrani<sup>1,2,3</sup>

Carole Confolent<sup>1,4</sup>

<sup>1</sup>Present address: INRA, UMR 1095 Génétique, Diversité et Ecophysiologie des Céréales, F-63000 Clermont-Ferrand, France

Morgane Roth<sup>4</sup>

<sup>†</sup> Present address: Agroscope, Research Division Plant Breeding, Wädenswil, Switzerland

Guy Clauzel<sup>4</sup>

Guillaume Roch<sup>1,5</sup>

Laurent Brun<sup>4</sup>

Naïma Dlalah<sup>1</sup>

Patrick Lambert<sup>1</sup>

Cindy E. Morris<sup>2</sup>

Jean-Marc Audergon<sup>1\*</sup>

\*Corresponding author

Email: [jean-marc.audergon@inra.fr](mailto:jean-marc.audergon@inra.fr)

<sup>1</sup> INRA, UR1052 Génétique et Amélioration des Fruits et Légumes, F-84143 Montfavet, France

<sup>2</sup> INRA, UR407 Pathologie Végétale, F-84143 Montfavet, France

<sup>3</sup> ENGREF, AgroParisTech, F-75732 Paris, France

<sup>4</sup> INRA, UE0695, Unité Expérimentale Recherches Intégrées de Gotheron, F-26320 Saint-Marcel-Lès-Valence, France

<sup>5</sup> CEP Innovation, F-69364 Lyon, France

## Abstract

Bacterial canker of apricot caused by *Pseudomonas syringae* (*Psy*) is a particularly limiting factor of crop cultivation in cold and humid winter areas. In this study, a quantitative trait locus (QTL) analysis designed for four biparental populations was carried out and tested in order to map genetic determinants of apricot resistance to *Psy*. Among the four populations, three were considered as interconnected, meaning they were issued from a cross involving a common genitor. A complementary approach between a two way pseudo-test cross strategy on the first biparental progeny and a multi-parental mapping on the three second generation interconnected families was performed. The use of MCQTL outbred package allowed us to compare within-family analysis and across-family analysis in the three interconnected families, exploiting a consensus map produced with 162 SNPs. A set of 15 QTLs involved in both quantitative traits related to symptoms lengths resulting from artificial inoculations and binary mortality trait noted over two severe years, was identified. No overlapping has been noticed between these two traits which may suggest two distinct mechanisms involved in resistance. The underlying biologic hypothesis are discussed. A comparison between results from within and across family models showed that they are both complementary, across-family models being more powerful to detect common QTLs across families and within family models, population-specific QTLs.

## Introduction

Breeding of perennial *Rosaceae* species is challenged by multiple factors related to their long vegetative cycle, long period of juvenility, large plant size, high level of heterozygosity coupled with a strong phylogenetic differentiation for some species such as apricot (*Prunus armeniaca* L.) (Bourguiba *et al.*, 2012). In addition, due to the polygenic architecture of most agronomic traits, phenotypic evaluation is often hampered by the interaction with environmental factors in the orchard. Overall, this

laborious and costly phenotype characterization can be constrained for some traits to the need to maintain the trees until they reach maturity (Rikkerink *et al.*, 2007; Khan and Korban, 2012). Then the use of molecular markers to assist breeding through marker-assisted-breeding (MAB) overcomes these limitations by allowing the selection of genotypes with the favorable alleles at the seedling stage (Collard and Mackill, 2008). Application of MAB implies beforehand the detection of either major genes or quantitative trait loci (QTLs) involved in the variation of continuous traits. Classically, in perennial trees, these detections were mostly achieved through linkage analysis (LA) in single full-sib families derived from the cross between two contrasted genotypes. Due to the limited number of recombination events, linkage analysis takes benefits from long linkage disequilibrium (LD) and allows the use of a moderate density of markers to cover the genome. However, for this approach the main limitations are a low resolution of mapping (Bazakos *et al.*, 2017) and especially an instability of QTLs among different genetic backgrounds due to a poor diversity in the parental allelic pool (Kenis *et al.*, 2008). These problems can be addressed by an alternative approach, genome-wide association (GWA) which deals with highly diverse natural populations showing a higher number of recombination events and a wider allelic pool (Khan and Korban, 2012; Ogura and Busch, 2015). Nevertheless, GWA also shows some drawbacks with confounding effects due to population stratification (Sillanpää, 2011) or false negatives due to rare variants or minor-effect polymorphisms (Myles *et al.*, 2009). Recently, methods known as ‘LD-LA mapping’ aimed to reconcile GWAS and linkage analysis through development of multiparental progenies (Wu *et al.*, 2002). QTL mapping on pedigree-based populations or multiple bi-parental populations connected by common parents allows to extend polymorphism at the targeted QTLs while presenting more equilibrated allelic frequencies than GWA panels. This strategy has shown its efficiency in perennial trees with higher QTL detection power and better accuracies of QTL positions and effects than traditional

linkage analysis (van de Weg *et al.*, 2004; Billotte *et al.*, 2010; Allard *et al.*, 2016; Fresnedo-Ramírez *et al.*, 2016).

Bacterial canker caused by the ubiquitous bacterium *Pseudomonas syringae* (*Psy*) remains one of the most damaging biotic stress of *Prunus* species leading to potential death of trees in the orchard (Kennelly *et al.*, 2007). Apricot is more particularly impacted with a lack of effective protection measures in the orchard limited to the use of resistant and soil-adapted rootstocks and a tall-height grafting (Prunier *et al.*, 1999). Several sources of partial resistance have been identified (Prunier and Bordjiba, 1991; Prunier *et al.*, 1993; Brun *et al.*, 2011) and genetic improvement through selection of haplotypes with favorable alleles offers promising perspectives in disease management. Recently, we deciphered the genetic pattern of susceptibility to bacterial canker by performing a GWA study in a broadly diversified apricot core-collection (Omrani *et al.*, *submitted*). A total of 11 associations, mapped to the single nucleotide polymorphism (SNP) level, were identified on two traits, among which two main loci explained together up to 41% and 26% of the total phenotypic variance. Although favorable haplotypes were determined according to the detected SNPs, our study did not allow to identify some known resistant cultivars among them, possibly due to under-represented variants in the apricot panel.

The main purpose of this study was to map QTLs linked to bacterial canker resistance in an apricot multi-parental scheme composed of a biparental progeny BerBa and three interconnected populations derived from the cross of three pre-selected BerBa hybrids with a common genitor. Here we describe step by step: i) the phenotypic characterization of the genetic material achieved via artificial inoculation in the orchard and by exploiting natural disease mortality incidence, ii) construction of individual maps of all the progenies, iii) construction of a consensus genetic map of the interconnected families using both SNPs and single sequence repeats (SSRs), and iv) QTL mapping by several approaches. These approaches included: (1) a two-way

pseudo test-cross strategy (Grattapaglia and Sederoff, 1994) for BerBa progeny and for the three interconnected progenies a comparison between a (2) within-family model analyzing each family independently and (3) a connected model assuming the same effects of the detected QTLs across families.

## **Material & Methods**

### ***Plant Material***

Four families were used for mapping bacterial canker partial resistance in this study. An F1 family of 218 individuals, “Bergeron” x “Bakour” (BerBa) was derived from the cross between the susceptible cultivar “Bergeron” and the resistant cultivar “Bakour” (Brun *et al.*, 2011). Following a two-years period of multi-diseases evaluation, three hybrids from BerBa designated as C092, C109 and D099 were selected for their low level of susceptibility to monilia and bacterial canker, and were individually crossed to a self-fertile and plum pox virus resistant genitor ‘A4049’. The three interconnected progenies composed of 110, 118 and 77 individuals, respectively and the biparental population BerBa were observed and are maintained at the INRA experimental station ‘UERI de Gotheron’ (Drôme, France).

### ***DNA extraction, SNP genotyping and marker segregation type***

A total of 300 simple sequence repeat (SSR or microsatellite) marker loci were screened for polymorphism in “Bakour” and “Bergeron”. Genome sequences of the two parents allowed the search of SNPs polymorphisms to complement the use of SSRs markers. The criteria for choosing these SSRs and SNPs were (1) a good distribution along the genome, (2) the highest proportion of polymorphism between the two parents. For the interconnected progenies, SNPs markers were selected among the ones exploited for “Bakour” and “Bergeron” according to the genotype profiles of A4049 with additional SNPs, added in order to maximize the polymorphism of the hybrids. Extraction of total genomic DNA was carried out from fresh young leaves of a single individual

plant (150 mg) per genotype using the DNeasy Plant Mini Kit (Qiagen) and following the manufacturer's instructions with minor modifications: addition of 1% w/v of PVP-40 to the AP1 buffer solution. DNA concentrations were estimated using spectrofluorometry and aliquots were diluted to a final concentration of 20 ng.µL<sup>-1</sup>. After genotyping markers segregation types were affected to the three following categories <lmxll>, <nnxnp> or <hkxhk>.

### ***Linkage analysis and genetic maps construction***

Construction of linkage maps was performed using JoinMap® v4.0 software by setting the 'cross-pollinator' (CP) data type (Van Ooijen, 2006). Recombination frequencies were converted into map distances using the Haldane mapping function.

In the Berba population, two individual parental linkage maps were built for each parent according to a 'two-way pseudo test-cross' strategy (Grattapaglia and Sederoff, 1994). In the three interconnected progenies, an individual map for each of the parental hybrid was constructed. Individual maps of the parental hybrids were then integrated into a consensus linkage map using the 'combine maps' function in Joinmap® with genetic distances estimated from mean recombination frequencies of 'bridge markers' between the individual maps.

For each map, linkage groups were defined according to an independence LOD score higher than 3. Markers were ordered on each linkage group considering a regression mapping function using pairwise data of the segregating loci with a recombination frequency lower than 0.40 and a LOD score higher than 1. Linkage phases between markers were determined and marker segregation distortion was assessed comparing observed and expected segregation ratios (1:1 or 1:2:1) with a Chi2-test.

### ***Evaluation of bacterial canker susceptibility***

Evaluation of susceptibility to bacterial canker in the BerBa and the interconnected progenies was conducted annually on each individual in the orchard over a 3-year period (2012 to 2014) and a 2-year period (2016 & 2017), respectively, according to a

previously described procedure (*Omrani et al., submitted*). Briefly, artificial inoculations were performed in the orchard on three to five one-year old branches per genotype (one tree for each genotype), depending on the population using an inoculum of *Pss* strain 41A (Bartoli *et al.*, 2015). An additional shoot per genotype was inoculated with sterile water and consisted in a negative control. Bacterial canker susceptibility was scored as the length of four different symptoms (Additional file 1): canker length (*lgc*), sub-lesion length (*lge*), superficial (*bs*) and internal browning (*bi*) lengths (except *lge* and *bi* for the interconnected progenies). During symptom notation, attention was paid to the effect of the operator as it has been shown to impact the phenotypic variability (*Omrani et al., submitted*). This effect was controlled with a balanced design, each genotype being evaluated by all the operators. Furthermore, due to a high level of natural disease pressure on the years 2017 and 2018, death of trees occurred on the orchard. A mortality score (0/1) had been attributed then in the three interconnected progenies on the two years.

### ***Statistical modeling of phenotypic data***

All statistical analyses were conducted using R software (<https://www.r-project.org>). Due to the possible annual effect related to contrasted winter expression through years, each phenotype was considered annually as a subdivision into a fixed operator effect, a random genotype effect and a random residual effect with a linear mixed model. Prior to modeling, normalization of phenotype distribution was performed using a  $\ln(x+1)$  transformation. Best unbiased linear predictors (BLUPs) were subsequently predicted from each phenotype per year and were then used as input variable data for linkage analysis. Restricted maximum likelihood estimates of genetic ( $\sigma_G^2$ ) and residual ( $\sigma_e^2$ ) variances were also extracted from the model and broad-sense heritability  $H^2$  estimated per year as  $\sigma_G^2 / (\sigma_G^2 + \sigma_e^2 / n)$  where  $n = \{3, 5\}$  is the number of replicates per genotype. Pearson correlations coefficients were computed to assess within-year and between-years correlations between all studied phenotypes.

### ***QTL mapping***

The R package Rqtl was used for QTL mapping on the separate genetic maps of ‘Bergeron’ and ‘Bakour’ in the BerBa progeny, leading to the detection of additive QTLs in a pseudo-test cross strategy. Analysis of the interconnected progenies was realized using both the outbred module of MCQTL (Jourjon *et al.*, 2005) and an home-made script allowing both analysis of continuous and binary traits.

Precisely, missing data imputation and pseudo-marker inference with a step of 2 cM was performed using *fill.geno* and *sim.geno* functions of Rqtl for ‘Bergeron’ and ‘Bakour’ maps and *ProbaPop* function of MCQTL for the consensus map.

For the interconnected progenies analysis, output of *ProbaPop* function consisted of the estimation of genotype probabilities at each chromosome location for both genotyped and inferred markers in each family. Depending on the phenotype category, either continuous (*lgc*, *lge*, *bs* and *bi*) or binary (mortality), the output probabilities given marker information were then used as regressors for QTL testing in linear or generalized linear models (glm) in R, respectively (Haley and Knott, 1992). We implemented the two MCQTL models in R: a within-family model (‘disconnected model’) testing the presence of a QTL for each progeny, independently and an across-family model (‘connected model’) assuming same location and effects of the QTL whatever the family considered. Each model was tested under hypothesis of additive effects acting alone (additive hypothesis) or with dominance effects (additive + dominance hypothesis). Phenotypic value  $Y_{ck}$  of the  $k^{\text{th}}$  individual of family  $c$  was modeled according to the general formula:

$$Y_{ck} = \mu_c + \sum_{ij} p_{ck,ij}^l \theta_{c,ij}^l + \varepsilon_{ck}$$

with  $\mu_c$  is the global mean in the family  $c$ ;  $p_{ck,ij}^l$  is the probability of the  $k^{\text{th}}$  individual having genotype  $ij$  (with  $i, j$  are the alleles provided by parents 1 and 2) at the tested locus  $l$  given the marker information;  $\theta_{c,ij}^l$  is the mean of the  $ij$  genotype at locus  $l$  in family  $c$  and  $\varepsilon_{ck}$  the residual error. The mortality binary score was analyzed with a

binomial regression,  $Y_{ck}$  corresponding to a variable related to a *logit* function ( $\log(\frac{p}{1-p})$ ) with  $p$  as the probability of the individual death in the orchard.

Moreover, the locus  $l$  effect is decomposed in additive ( $\alpha$ ) and dominance ( $\delta$ ) effects:

$$\theta_{c,ij}^l = \alpha_{c,i}^l + \alpha_{c,j}^l + \delta_{c,ij}^l$$

where  $ij = \{ac, ad, bc, bd\}$  considering four different alleles segregating in a outbred cross between heterozygous parents. At any tested locus, the sum of the two allelic effects of each parent was null (*ie*  $\alpha_{c,i}^l + \alpha_{c,j}^l = 0$ ) by constraint of the model. Under the additive hypothesis, the dominance parameters are all assumed to be null. In the connected model, genetic parameters are assumed to be identical over the families (*ie*  $\alpha_{c,i}^l = \alpha_i^l \forall c \forall i$ ;  $\delta_{c,ij}^l = \delta_{ij}^l \forall c \forall ij$ ).

Fisher and Chi-square tests of comparison of stacked models (with or without QTL) were computed to test single QTL locations in analysis of continuous phenotypes (*lge*, *lge*, *bs*, *bi*) and mortality, respectively. Statistical significance of QTLs was assessed from the  $-\log_{10}(\text{p-value})$  according to MCQTL procedure. Model choice criteria between either connected or disconnected model, and ‘additive’ or ‘additive+dominance’ hypothesis was based on comparison of stacked models using analysis of variance (ANOVA). Differences in effects among pairs of alleles was tested using a t test ( $\alpha=5\%$ ).

For all mapping populations, LOD threshold values ( $p < 0.05$ ) were estimated trait by trait from a 1,000 permutations test considering intra-family permutations in the case of the interconnected progenies (Churchill and Doerge, 1994). LOD threshold values were very similar between traits for each population analysis and averaged 2.5 for the two way pseudo-testcross strategy and ranged from 3 to 3.47 for the interconnected progenies. Therefore, QTL search was based on LOD threshold of 2.5 and 3.47 for BerBa and interconnected progenies analysis, respectively. For each significant QTL, a support interval defined as a LOD score drop off equal to 1 on both sides of the LOD peak was defined. After a single QTL scan, the closest marker of the significant

LOD peak was put into the cofactor. Subsequent additional QTL scans and cofactor setting were performed until no additional peaks were observed. In addition to the chromosome location and LOD score, additive (and possible dominance) effects and phenotypic variation explained by individual (PVE) QTL and the model ( $R^2$ ) were computed for each trait. In the case of mortality notation, we used Nagelkerke's pseudo- $R^2$  estimated from the *glm* (Nagelkerke, 1991). Linkage maps and QTL positions were drawn using MapChart software.

## Results and Discussion

### *SSR and SNP linkage mapping*

Over the initial set of 300 SSRs, only 38 and 52 markers were polymorphic on 'Bergeron' and 'Bakour', respectively, possibly due to the position of these two cultivars at the end of their phylum of diversification, North European and Mediterranean groups, respectively (Bourguiba *et al.*, 2012). Individual parental maps of 'Bergeron' and 'Bakour' had 10 and 11 linkage groups (LG) over the 8 chromosomes, including 122 and 128 SNPs with 17 and 21 SSRs, respectively (Additional files 2 and 3). These SSRs markers were chosen according to the minimum amount of missing data. Parental maps measured 624.1 cM in 'Bergeron' and 823.9 cM in 'Bakour' with an average marker density of  $4.48 \pm 3.42$  and  $5.52 \pm 4.51$  cM and a genome coverage representing 87% and 81% of the *Prunus persica* v2.0 reference, respectively. Leaving aside the respective gaps splitting chromosome 1 and 2 of 'Bergeron' and 2, 4 and 5 of 'Bakour', maximum gap length was 17.7 (chromosome 8) and 34.9 (beginning of chromosome 5\_2) cM in 'Bergeron' and 'Bakour' maps, respectively. Lengths of individual parental maps of C092, C109 and D099 were 742.1, 810.3 and 750 cM for 16, 13 and 15 linkage groups, respectively (Additional file 4)

The consensus map of the three interconnected progenies consisted of 162 SNPs distributed over 10 linkage groups, chromosomes 2 and 4 being split into two parts

(Additional files 2 and 5). The total map measured 970.11 cM covering 95% of the *P. persica* v2.0 genome with an average marker density of  $5.99 \pm 5.50$  cM. Among the 162 SNPs, only 33 were polymorphic in the three interconnected progenies (76 in two of them and the remaining 53 in a single progeny). The 10 linkage groups spanned from 21.10 (LG4\_1) to 233.32 cM (LG1) with an average marker density ranging from  $2.64 \pm 3.33$  cM (LG4\_1) to  $8.33 \pm 9.65$  cM (LG2\_2).

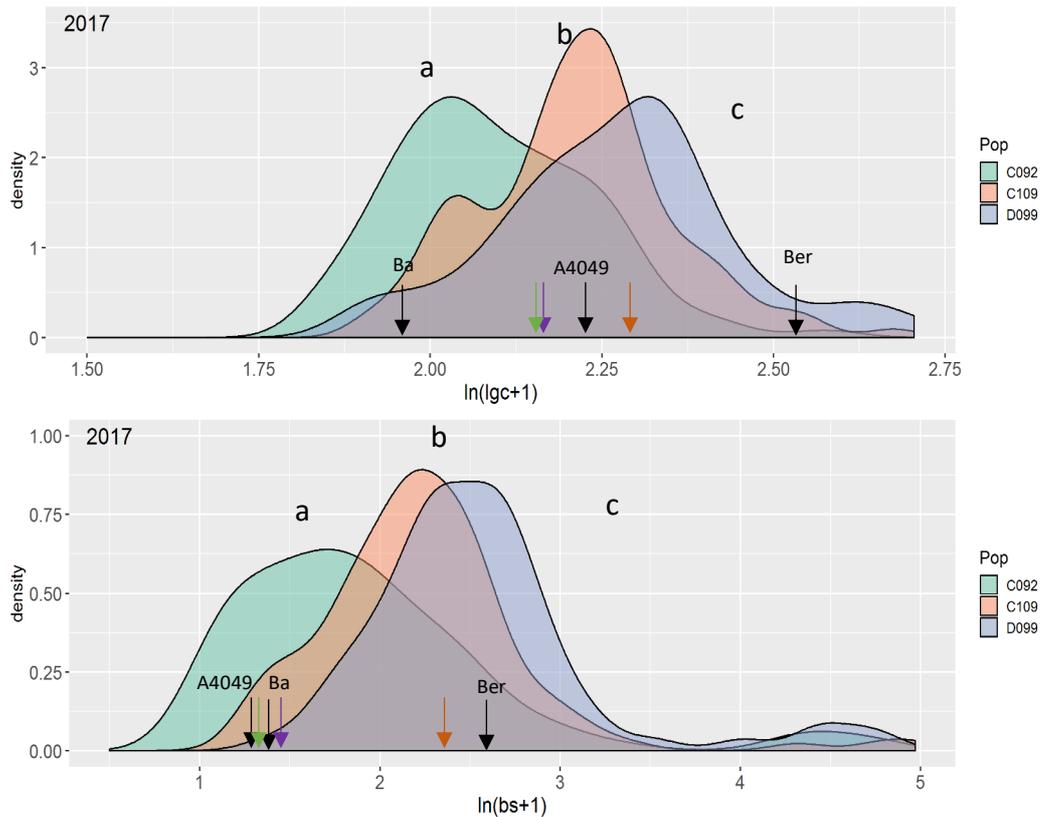
### ***Statistical characterization of the phenotypes***

Comparison between all controls and inoculated shoots regardless of the genotype effect revealed a highly significant effect of bacterial inoculation (Wilcoxon unilateral t.test  $p < 2.2E-16$ ) for all studied variables through years in both Berba and the interconnected progenies.

For each of the four interconnected families, genetic distribution of the symptoms related to the artificial inoculations displayed a quantitative distribution with variances depending on the year and the cross providing insight into differences on winter expressions and on genetic backgrounds related to each population, respectively (*Figure 1, Table 1*). Globally, 2017 was the year with a maximum genetic variance (except for *lgc* in C092 x A4049). A highly significant effect of the population was observed on this year for *lgc* ( $p = 1.975E-13^{***}$ ) and *bs* ( $p = 4.088E-9^{***}$ ) with the three populations classified into three different groups (SNK post-hoc mean comparison test): C092 x A4049 being the least susceptible (*lgc*:  $2.09 \pm 0.14^a$ , *bs*:  $1.93 \pm 0.76^a$ ), C109 x A4049 (*lgc*:  $2.21 \pm 0.14^b$ , *bs*:  $2.24 \pm 0.59^b$ ) an intermediate group and D099 x A4049 the most susceptible (*lgc*:  $2.27 \pm 0.17^c$ , *bs*:  $2.56 \pm 0.64^c$ ).

In 2017, considering the two symptoms, D099 x A4049 appeared to be the population with the largest part of genetic variance, with broad-sense heritability estimates reaching 36% for *lgc* and 75% for *bs*. Significant phenotypic correlations were obtained between *lgc*<sub>2016</sub> and *bs*<sub>2017</sub> (Pearson correlation  $r = 0.13^*$ ) with no correlation between the same variable through the two years. *lgc* and *bs* G BLUPs were also

highly correlated for each year (2016:  $r=0.25^{***}$ , 2017:  $r=0.68^{***}$  Pearson correlations corrected with a FDR  $\alpha=5\%$  for multiple-testing).



**Figure 1: Genetic distributions of *lgc* and *bs* BLUPs in the three interconnected progenies for the year 2017.**

Density plots of symptoms genetic BLUPs represented in C092 x A4049 (green), C109 x A4049 (orange) and D099 x A4049 (purple) progenies. All BLUP values are represented on a  $\ln(x+1)$  scale. Arrows on the x-axis are positioned on the average value of the parents C092, D099 and C109, respectively in green, purple and orange, and on the genitor A4049 and the grand-parents 'Bakour' (Ba) and 'Bergeron' (Ber) in black. Group of means are indicated according to a Student-Newman-Keuls test ( $\alpha=0.05$ ).

For Berba, a differential susceptibility between the two parents 'Bergeron' and 'Bakour' was dependent on the year with 2012 and 2014 not being as discriminant as 2013 (Additional file 5) Broad-sense heritability estimates ranged from 27% (*bs* 2014) to 65% (*bi*2014 and *lgc*2013) (Table 1).. Correlations between phenotypes in the same year ranged from 0.19\* (*bs* and *bi* in 2014) to 0.82\*\*\* (*lgc-lge* and *lgc-bs* in 2013) and through years from 0.16\* to 0.42\*\* (Additional file 6).

**Table 1: Broad-sense heritability estimates of BerBa and individual interconnected progenies for the traits related to symptoms post-inoculations in the orchard.**

Population	Year	lgc	lge	bs	bi
BerBa	2012	0.55	-	-	0.43
	2013	0.65	0.55	0.63	0.63
	2014	0.51	0.57	0.27	0.65
C092 x A4049	2016	0.35	-	0.33	-
	2017	0.21	-	0.78	-
C109 x A4049	2016	0.19	-	0.23	-
	2017	0.23	-	0.66	-
D099 x A4049	2016	0.15	-	0.36	-
	2017	0.36	-	0.75	-

***QTL analysis revealing complementary use of within and across-family models.***

A total of 12, 5 and 10 QTLs were mapped for all the studied phenotypes, except *bi2012* and *lgc2016*, using the pseudo-test cross strategy in ‘Bergeron’ and ‘Bakour’ parental maps, the across and within-family models in the interconnected progenies, respectively (Table 2 and Additional file 7).

All detected QTLs using BerBa progeny were mapped exclusively on ‘Bergeron’ parental map with  $R^2$  explained by all associated QTLs per trait ranging from 5% (*bs2013*) to 20% (*lgc2014*). Considering all analyses together, from one to three QTLs were detected for each trait. In the interconnected progenies, individual QTL had PVE estimated from 7% (*bs2016* – ‘connected model’) to 27% (*lgc2017* – ‘connected model’). Several QTLs found on different phenotypes and/or years either colocalized or had overlapping support intervals (*Figure 2*).

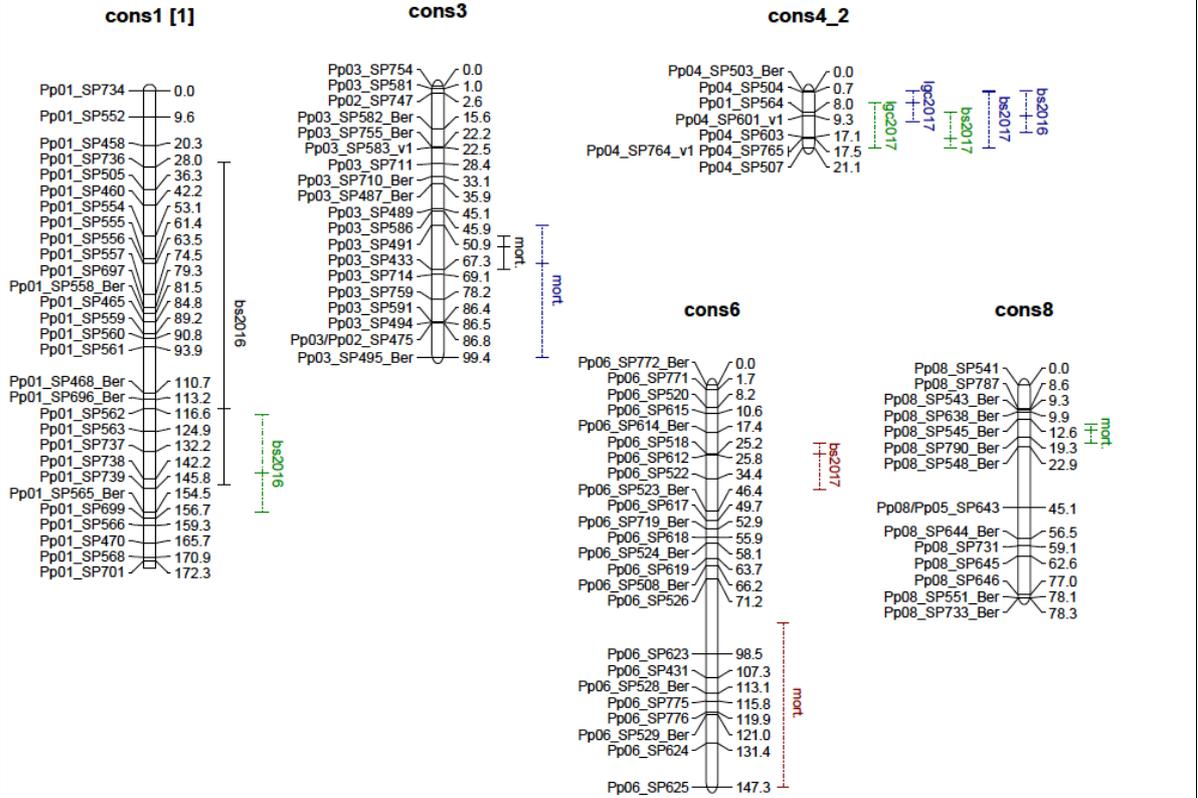
**Table 2: QTLs detected on ‘Bergeron’ genetic map with a pseudo test-cross strategy.**

For traits with multiple detected QTLs,  $-\log(\text{pvalue})$  and  $R^2$  of the model are indicated in addition to individual contribution of QTLs (PVE). Results shaded in grey and indicated in bold are related to colocalization or overlapping QTLs of either the same phenotype in different years or different phenotypes in the same year, respectively.

		model		marker					Support	
Trait	Year	$-\log\text{Pvalue}$	$R^2$	$-\log\text{Pvalue}$	PVE	LG	Locus	Position (cM)	interval (cM)	Additive Effects
lgc	2012	3.34	0.06	3.34	0.06	1_2	SP702_1	64.3	40-67.7	$-0.082 \pm 0.023$
lgc	2013	3.67	0.067	<b>3.67</b>	<b>0.067</b>	<b>6</b>	<b>SP526_6</b>	<b>45</b>	<b>35-56.3</b>	<b><math>0.082 \pm 0.021</math></b>
lge	2013	4.71	0.088	<b>4.71</b>	<b>0.088</b>	<b>6</b>	<b>SP526_6</b>	<b>45</b>	<b>38.9-50</b>	<b><math>0.078 \pm 0.018</math></b>
bs	2013	2.96	0.053	2.96	0.053	4	SP499_4	24	10-35	$0.105 \pm 0.032$
bi	2013	3.34	0.060	3.34	0.060	8	SP543_8	2.2	0-6	$0.164 \pm 0.046$
				4.68	0.081	4	SP504_4	64	55-71.3	$-0.055 \pm 0.013$
lgc	2014	8.50	0.202	<b>4.57</b>	<b>0.079</b>	<b>1_2</b>	<b>SP472_1</b>	<b>50</b>	<b>40-56.4</b>	<b><math>-0.054 \pm 0.013</math></b>
				4.02	0.068	7	SP536_7	30	24.4-40	$0.049 \pm 0.012$
lge	2014	5.82	0.132	<b>4.18</b>	<b>0.076</b>	<b>1_2</b>	<b>SP471_1</b>	<b>44</b>	<b>30-50</b>	<b><math>-0.067 \pm 0.016</math></b>
				3.96	0.072	7	SP538_7	50	40-57.9	$0.065 \pm 0.016$
bi	2014	7.03	0.158	<b>5.94</b>	<b>0.113</b>	<b>1_2</b>	<b>SP472_1</b>	<b>52</b>	<b>45-56.4</b>	<b><math>-0.207 \pm 0.04</math></b>
				3.05	0.051	8	MP340_8	11.9	5-25	$0.133 \pm 0.039$

A region located from the beginning of chromosome 4\_2 to position 21.09 cM was detected both on *lgc2017*, *bs2017* and *bs2016* using both within and across-family models in the interconnected families with estimation of QTL position varying slightly between connected and disconnected models. QTLs had an average support interval region of  $18.40 \pm 6.33$  cM on ‘Bergeron’ map and  $10.37 \pm 2.73$  cM on the consensus map where two particular regions on chromosome 1 and 4\_2 (exceeding 100 cM and located on an exact position, respectively) were excluded from the calculation. Considering that overlapping QTLs between within-family analyses and across-family

analyses are the same, all detected QTLs in the within-family analyses were also found in the across-family analyses with differences on estimated PVE.



**Figure 2: Map of the QTLs detected with the within-family and across-family analyses in the interconnected progenies.**

Dashed and continuous support intervals refer to QTLs detected with the within-family and across-family analyses, respectively. Green, blue and brown colors indicate that the detected QTL were specific to C092 x A4049, C109 x A4049 and D099 x A4049 families, respectively.

‘Mort’ is related to the mortality incidence, resulting from 2017 and 2018 notations in the orchard.

When a QTL was identified by both analyses, its PVE estimated by the within-family model was from 18 to 42% lower than that estimated by the across-family analysis. For instance, a QTL underlying mortality trait was detected on chromosome 3 (LOD peak at 58.91 cM) with PVE of 13, 16 and 15% in C092 x A4049, C109 x A4049 and D099 x A4049 families respectively. This QTL explained a major part of the genetic variance considering the across-family analyses with PVE reaching 25%. Moreover, across-family analyses allowed to detect 4 additional QTLs on chromosomes 8, 6, 4\_2 and 1 that were not detected with the within-family analyses. The complementary use

of within and across-family models proved to be efficient in order to map genetic determinants of resistance to bacterial canker in apricot. Comparing to the two way-pseudo test cross strategy applied in ‘Bergeron’ and ‘Bakour’ parental maps, a larger number of QTLs have been detected through the study of the interconnected progenies, taking advantage of a higher number of recombinants. The exploitation of second generation populations allowed detection of QTLs which could have been at an homozygous state in ‘Bakour’ making them impossible to be revealed using BerBa progeny. While for QTLs detected by both within and across-family models, support intervals and effects were estimated with a higher precision for the across-family model, some QTLs were only detected on some specific progenies, as it could have been observed in previous studies (Billotte *et al.*, 2010; Mérot-L’Anthoëne *et al.*, 2014). C092, D099 and C109 had significant effects on the means of their respective progenies, revealing differences in their polygenic backgrounds issued from different associations patterns of ‘Bakour’ resistance determinants. Therefore, it was reasonable to expect numerous progeny-specific QTLs. Another explanation could be due to under-representation of some specific individual progeny variants which could have been diluted in the whole population (Mérot-L’Anthoëne *et al.*, 2014). This justifies the interest of using both approaches in order to maximize the chance of capturing both general and specific population QTLs.

***Blocking Psy infection and migration in the tissues are complementary defense mechanisms governed by different genetic regions.***

In the interconnected progenies, no co-localization or overlap have been noticed between QTL related to the mortality notation (chromosomes 3,6 and 8 depending on within or across-family models) and the ones linked to *lgc* and *bs* symptoms (chromosomes 1, 4 and 6 depending on within or across-family models). These results seem to demonstrate the existence of two underlying mechanisms involved in different levels, according to the phenotypes they are related to. It has been suggested in apricot

that buds are the main doors of entry into the plant (Prunier and Bordjiba, 1991). The first mechanism would act in a systemic way, by blocking the infection at the entry of the buds. The second mechanism relying on the symptoms obtained after controlled inoculations would be related to the limitation of the bacterium migration in the tissues. Support intervals of these QTLs were too long on the consensus map to allow an identification of putative candidate genes in order to consolidate this hypothesis. A future perspective of densification over the detected regions would be useful to shorten these intervals and to get a more precise list of candidate genes. Several pathosystems have been presented with underlying resistance mechanisms resulting from different genetic factors, affecting either different organs (Werner *et al.*, 1986) or acting on several ways of pathogen control. For instance, in the case of apricot resistance to plum pox virus, locus *PPVres* had been shown to control the level of symptoms while another one is assumed to modulate the viral accumulation (Decroocq *et al.*, 2014).

## References

- Allard A, Bink MCAM, Martinez S, Kelner J-J, Legave J-M, di Guardo M, Di Pierro EA, Laurens F, van de Weg EW, Costes E. 2016. Detecting QTLs and putative candidate genes involved in budbreak and flowering time in an apple multiparental population. *Journal of Experimental Botany* **67**(9), 2875–2888.
- Bartoli C, Carrere S, Lamichhane JR, Varvaro L, Morris CE. 2015. Whole-genome sequencing of 10 *Pseudomonas syringae* strains representing different host range spectra. *Genome Announcements* **3**, 2–3.
- Bazakos C, Hanemian M, Trontin C, Jiménez-Gómez JM, Loudet O. 2017. New strategies and tools in quantitative genetics: how to go from the phenotype to the genotype. *Annual Review of Plant Biology* **68**(1), 435–455.
- Billotte N, Jourjon M-F, Marseillac N, Berger A, Flori A, Asmady H, Adon B, Singh R, Nouy B, Potier F, Cheah SC, Rohde W, Ritter E, Courtois B, Charrier A, Mangin B. 2010. QTL detection by multi-parent linkage mapping in oil palm (*Elaeis guineensis* Jacq.). *Theoretical and Applied Genetics* **120**(8), 1673–1687.
- Bourguiba H, Audergon J-M, Krichen L, Trifi-Farah N, Mamouni A, Trabelsi S, D’Onofrio C, Asma BM, Santoni S, Khadari B. 2012. Loss of genetic diversity as a signature of apricot domestication and diffusion into the Mediterranean Basin. *BMC Plant Biology* **12**(49).

- Brun L, Warlop F, Mercier V, Broquaire J-M, Clauzel G, Gomez C, Parveaud C-E, Audergon J-M.** 2011. Quelle sélection fruitière pour une production durable, à faible niveau d'intrants? Méthodologie pour un réseau de sélection variétale décentralisée. *Innovations Agronomiques* **15**, 105–115.
- Churchill GA, Doerge RW.** 1994. Empirical threshold values for quantitative trait mapping. *Genetics* **138**(3), 963–971.
- Collard BCY, Mackill DJ.** 2008. Marker-assisted selection : an approach for precision plant breeding in the twenty-first century. *Phil. Trans. R. Soc. B* **363**(1491), 557–572.
- Decroocq S, Chague A, Lambert P, Roch G, Audergon JM, Geuna F, Chiozzotto R, Bassi D, Dondini L, Tartarini S, Salava J, Krška B, Palmisano F, Karayiannis I, Decroocq V.** 2014. Selecting with markers linked to the PPVres major QTL is not sufficient to predict resistance to Plum Pox Virus (PPV) in apricot. *Tree Genetics & Genomes* **10**(5), 1161–1170.
- Fresnedo-Ramírez J, Frett TJ, Sandefur PJ, Salgado-Rojas A, Clark JR, Gasic K, Peace CP, Anderson N, Hartmann TP, Byrne DH, Bink MCAM, van de Weg EW, Crisosto CH, Gradziel TM.** 2016. QTL mapping and breeding value estimation through pedigree-based analysis of fruit size and weight in four diverse peach breeding programs. *Tree Genetics & Genomes* **12**(25).
- Grattapaglia D, Sederoff R.** 1994. Genetic linkage maps of *Eucalyptus grandis* and *Eucalyptus urophylla* using a pseudo-testcross: mapping strategy and RAPD markers. *Genetics* **137**(4), 1121–1137.
- Haley CS, Knott SA.** 1992. A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. *Heredity* **69**(4), 315–324.
- Jourjon M-F, Jasson S, Marcel J, Ngom B, Mangin B.** 2005. MCQTL: multi-allelic QTL mapping in multi-cross design. *Bioinformatics* **21**(1), 128–130.
- Kenis K, Keulemans J, Davey MW.** 2008. Identification and stability of QTLs for fruit quality traits in apple. *Tree Genetics & Genomes* **4**(4), 647–661.
- Kennelly MM, Cazorla FM, de Vicente A, Ramos C, Sundin GW.** 2007. *Pseudomonas syringae* diseases of fruit trees: progress toward understanding and control. *Plant Disease* **91**(1), 4–17.
- Khan MA, Korban SS.** 2012. Association mapping in forest trees and fruit crops. *Journal of Experimental Botany* **63**(11), 4045–4060.
- Mérot-L'Anthoëne V, Mangin B, Lefebvre-Pautigny F, Jasson S, Rigoreau M, Husson J, Lambot C, Crouzillat D.** 2014. Comparison of three QTL detection models on biochemical, sensory, and yield characters in *Coffea canephora*. *Tree Genetics and Genomes* **10**(6), 1541–1553.
- Myles S, Peiffer J, Brown PJ, Ersoz ES, Zhang Z, Costich DE, Buckler ES.** 2009. Association mapping: critical considerations shift from genotyping to experimental design. *The Plant Cell* **21**(8), 2194–2202.
- Nagelkerke NJD.** 1991. A note on a general definition of the coefficient of determination. *Biometrika* **78**(3), 691–692.
- Ogura T, Busch W.** 2015. From phenotypes to causal sequences: using genome wide association studies to dissect the sequence basis for variation of plant development. *Current Opinion in Plant Biology* **23**, 98–108.

- Van Ooijen JW.** 2006. JoinMap ® 4, software for the calculation of genetic linkage maps in experimental populations. Kyazma BV, Wageningen.
- Prunier J-P, Audergon J-M, Cotta J.** 1993. Etude de la sensibilité variétale de l'abricotier au chancre bactérien. Programme de recherche Agrimed: Deuxiemes rencontres sur l'abricotier. Recueil des communications. Luxembourg, 111–112.
- Prunier J-P, Bordjiba O.** 1991. Effect of frost on bacterial necrosis of apricot buds. *Acta Horticulturae* **293**, 495–502.
- Prunier J-P, Jullian J-P, Audergon J-M.** 1999. Influence of rootstock and the height of grafting on the susceptibility of apricot cultivars to bacterial canker. *Acta Horticulturae* **488**, 643–648.
- Rikkerink EHA, Oraguzie NC, Gardiner SE.** 2007. Prospects of association mapping in perennial horticultural crops. In: Oraguzie NC, In: Rikkerink EHA, In: Gardiner SE, In: De Silva HN, eds. *Association Mapping in Plants*. New York, NY: Springer New York, 249–269.
- Sillanpää MJ.** 2011. Overview of techniques to account for confounding due to population stratification and cryptic relatedness in genomic data association analyses. *Heredity* **106**(4), 511–519.
- van de Weg EW, Voorrips RE, Finkers R, Kodde LP, Jansen J, Bink MCAM.** 2004. Pedigree genotyping: a new pedigree-based approach of QTL identification and allele mining by exploiting breeding material. *Acta Horticulturae* **663**, 45–50.
- Werner DJ, Ritchie DF, Cain DW, Zehr EI.** 1986. Susceptibility of peaches and nectarines plant introductions and other *Prunus* species to bacterial spot. *HortScience* **21**(1), 127–130.
- Wu R, Ma C-X, Casella G.** 2002. Joint linkage and linkage disequilibrium mapping of quantitative trait loci in natural populations. *Genetics* **160**(2), 779–792.



Partie V - Caractérisation de  
l'interaction abricotier x *Psy*

## Introduction

Les approches de cartographie développées au niveau des parties III et IV ont exploité des données de sensibilités collectées sur des populations suite à inoculations contrôlées à partir d'une souche agressive donnée de *Psy* 41A appartenant au phylogroupe (PG) 2 clade a. En considérant une perspective de sélection du matériel végétal, il est primordial d'étudier l'étendue de l'interaction GxG (plante x pathogène) afin d'assurer un spectre d'action large et durable des résistances déployées.

Des études citées préalablement, se sont intéressées à évaluer la validité et l'effet de QRLs face à la diversité de l'agent pathogène, à travers des études de cartographie dans le cadre des pathosystèmes : *V. inaequalis* x pommier (Calenge *et al.*, 2004; Soufflet-Freslon *et al.*, 2008; Van *et al.*, 2013), *Xanthomonas arboricola* x abricotier (Socquet-Juglard *et al.*, 2013b) et *Erwinia amylovora* x poirier (Montanari *et al.*, 2016).

Préalablement au screening à large échelle de populations avec un certain nombre de souches pathogènes dans un but de cartographie de QRLs, des dispositifs factoriels de taille plus modeste permettent déjà d'apporter des éléments de réponse sur l'importance de ces interactions GxG.

Au cours de cette partie, il est proposé d'étudier :

- au niveau statistique, le plan factoriel d'interaction à la fois en conditions in-vivo (inoculations en verger) et in-vitro (protocole d'inoculation sur organes détachés) entre 20 accessions d'abricotiers présentant des niveaux de sensibilité variable à la bactériose et 9 souches de *Psy* affiliées majoritairement au PG 2 et issues de prélèvement sur des arbres symptomatiques ou sur des débris de culture en verger (Parisi *et al.*, 2016).
- au niveau microbiologique, la présence et la migration de la bactérie dans les tissus de 3 génotypes témoins et le lien avec la sensibilité à la maladie.

# Matériel & Méthodes

## 1. Plan factoriel d'interaction GxG

La liste des 20 accessions d'abricotier exploitées au niveau du plan factoriel d'interaction GxG figure en annexe III-1.

Les 9 souches de *Psy* (Tableau 3) ont été échantillonnées essentiellement au niveau du PG 2, ce PG étant le plus représenté au niveau de la diversité de *Psy* en verger (Parisi *et al.*, 2016). 7 de ces souches ont été isolées à Gotheron à partir d'arbres symptomatiques, 2 au niveau de débris de récolte et de la terre. Deux souches du PG 1-a (groupe de *P. avellenae* et *Psa*, responsables des chancres bactériens du noisetier et du kiwi) et 7 (groupe de *P. viridiflava*) ont été ajoutées au dispositif afin de comparer son effet relativement aux souches du PG 2. La souche *Pss* 41A, utilisée lors des inoculations à large échelle des populations à but de cartographie génétique, fait également partie du dispositif.

**Tableau 3 : Caractéristiques des 9 souches exploitées lors du dispositif d'inoculations GxG**

PG-clade	Nom	Comportement	Isolée de	Code souche
1-a	19B	Agressive sur Bergeron	Arbre symptomatique	A
2-b	P3D7	Agressive sur Bergeron	Débris de récolte	B
2-b	P4.01.01.C3	Peu pathogène sur Bergeron	Arbre symptomatique	C
2-b	41A	Agressive sur Bergeron	Arbre symptomatique	D
2-b	P9T3	Agressive sur Bergeron	Terre	E
2-d	B728A	Symptômes internes sur Manicot	Arbre symptomatique	F
2-d	1B	Agressive sur Bergeron	Arbre symptomatique	G
2-d	7C	Agressive sur Bergeron	Arbre symptomatique	H
7	DG8.15	Agressive sur Bergeron	Arbre symptomatique	I

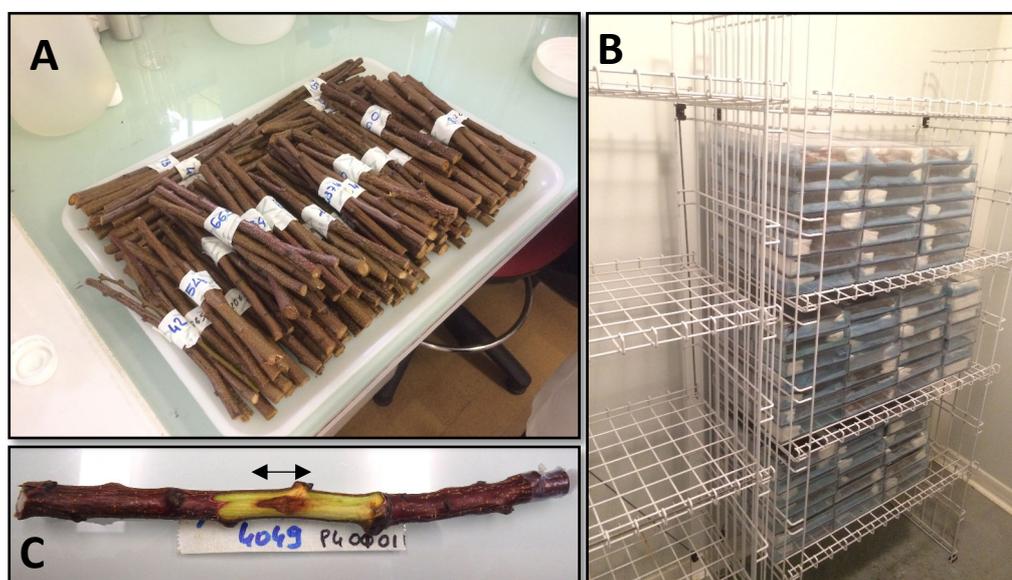
### 1.1. Conditions in-vivo en verger

Les inoculations en verger ont été effectuées selon le même protocole opératoire que les parties III et IV. Les variables étudiées correspondent à la longueur de chancre (*lgc*) et au brunissement superficiel (*bs*).

Les géotypes ont été inoculés à raison de 4 répétitions par souche avec un témoin négatif 'eau' par géotype lors de la première année d'expérimentation en 2017. Suite à des phénomènes d'infections naturelles observées au niveau des rameaux témoins négatifs en 2017, le traitement eau a été appliqué sur 4 rameaux en 2018, de façon analogue au facteur souche. En raison de l'important nombre de rameaux inoculés par géotype, un plan équilibré a été instauré afin de répartir équitablement le nombre de répétitions par souche sur plusieurs arbres de chaque géotype (2 à 4 arbres).

### 1.2. Conditions in-vitro en laboratoire

Parallèlement à la procédure en verger, des inoculations contrôlées en laboratoire des 9 souches de *Psy* référencées précédemment, sont effectuées sur des boutures de 18 cm (5 répétitions par combinaison souche x géotype), prélevées à l'état dormant selon le protocole établi en annexe III-2 (*Figure 28A*).



**Figure 28: Photographies du protocole GxG réalisé en conditions contrôlées in-vitro.**

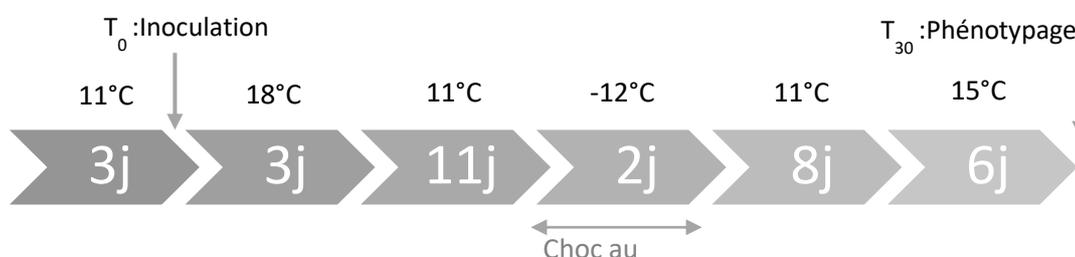
- Les boutures taillées à 18 cm, puis désinfectées sont regroupées par lot de géotype.
- Les boîtes (chacune correspondant à une combinaison souche x géotype) introduites en chambre climatique et soumises à des séquences de températures différentes.
- Le phénotypage consiste en la lecture de la variable 'nécrose superficielle', analogue au 'brunissement superficiel' issu des expérimentations en verger.

*Crédits photographiques: Mariem Omrani*

Un traitement témoin eau est également effectué de manière analogue aux autres souches. Un stockage des boutures dans des boîtes fermées et humidifiées à la base, dans une chambre climatique à 11°C se fait pendant 3 jours préalablement à l'inoculation.

Suite à l'inoculation, les boutures suivent ensuite une période d'incubation d'une durée totale de 30 jours, caractérisée par la mise en œuvre de cycles de températures différentes, préalablement établies dans le cadre de tests de pouvoir pathogène sur la variété Bergeron (travaux de *L. Parisi*) (*Figure 29*). Un choc au froid d'une durée de 48h a lieu à  $T_{14}$ , mimant l'action du froid hivernal en verger.

Ce protocole est inspiré d'études visant à reproduire les symptômes de dépérissement bactérien en conditions contrôlées (Prunier and Bordjiba, 1991; Sobiczewski and Jones, 1992; Santi *et al.*, 2004)



**Figure 29** : Schéma de la séquence temporelle d'incubation des boutures en conditions in-vitro dans le cadre du dispositif d'étude GxG.

Les boîtes sont randomisées dans la chambre climatique au cours de l'incubation (*Figure 28B*). Un suivi de la température de la chambre et de l'humidité des boîtes est réalisé tous les deux jours. Si le papier filtre à la base des boîtes est trop sec, un ajout de 3 mL d'eau stérile est réalisé au niveau des boîtes sous une hotte stérile, afin d'éviter la déshydratation des tissus.

A  $T_{30}$ , la séquence d'incubation est terminée et le phénotypage a lieu dans les mêmes conditions de lecture que le dispositif en verger (répartition des boutures selon un plan complet prenant compte des différents opérateurs). Le phénotype considéré correspond à la nécrose superficielle, analogue à *bs* en verger (*Figure 28C*).

## 2. Analyses statistiques

Les données phénotypiques sont ensuite analysées selon la procédure développée en parties III et IV. Une ANOVA considérant les effets génétiques individuels et les interactions GxG (plante x pathogène), en plus de l'effet notateur est appliquée à la fois sur les données verger 2017 & 2018 et les données in-vitro. L'analyse sur les

données verger prend également en compte l'effet année et les potentiels effets d'interaction GxY.

Les différentes variables ont été transformées afin d'approcher la normalité des résidus : transformation  $1/(x+1)$  pour la nécrose superficielle, et racine carrée pour *lgc* et *bs*. Après extraction des coefficients estimés GxG et rétro-transformation (pour revenir vers la variable brute) pour la nécrose superficielle, une analyse multivariée a été effectuée sur l'ensemble des variables considérées. Une étude de corrélations (coefficients de corrélation de Pearson) entre variables de chaque dispositif (verger et laboratoire) et entre les dispositifs est effectuée afin de juger de la répétabilité des données de sensibilité entre souches et accessions.

### **3. Étude de la migration bactérienne dans les tissus**

Trois géotypes précédemment caractérisés Bakour, Colomer ('résistants') et Bergeron ('sensible') ont été inoculés avec *Psy* 41A selon le protocole d'inoculations sur boutures (5 réplicats biologiques). Un traitement contrôle consistant en des boutures inoculées avec de l'eau stérilisée a été réalisé parallèlement sur le même nombre de réplicats.

A l'issue de cette seconde expérimentation en conditions in-vitro, des isolements de tissus sur milieu KBC ont été effectués au niveau du point d'inoculation (PI), à 5 cm de part et d'autre du PI (D5 et G5) et au niveau du cœur du bois. Les boîtes de Pétri sont ensuite mises à l'étuve pendant 48h à 25°C puis observées sous rayonnement d'ultra-violets afin de détecter de la fluorescence, démontrant la présence/absence de *Psy* au niveau des 4 points de prélèvement.

Les réactions de nécrose superficielle et de brunissement interne ont été également relevées et mises en regard avec la migration de la bactérie dans les tissus.

## Résultats et Discussion

Pour le dispositif en verger (conditions in-vivo), la structuration de la variation globale de *lgc* et *bs* est expliquée par ordre d'importance décroissante: par l'effet de l'année (Y,  $p < 2.2e-16^{***}$ ), de la souche (S,  $p < 2.2e-16^{***}$ ), de l'opérateur (pour *lgc*,  $p < 2.2e-16^{***}$  et *bs*,  $p = 3.89e-13^{***}$ ), du génotype abricotier (G,  $p < 2.2e-16^{***}$ ) (Tableau 4).

**Tableau 4: Résultats des ANOVA testant les effets génotype plante (G), souche (S), opérateur, année (Y), et les divers effets d'interaction G :Y, G :S, Y :S, G :Y :S au niveau des dispositifs factoriels GxG déployés en verger et en laboratoire.**

Les différentes variables ont été transformées afin d'approcher la normalité des résidus : transformation  $1/(x+1)$  pour la nécrose superficielle, et racine carrée pour *lgc* et *bs*.

Abréviations : SS, somme des carrés des écarts ; F value : valeur du test de Fisher.

p-values : NS, non significatif ; \* $0.01 < p < 0.05$  ; \*\*\* $< 0.001$ .

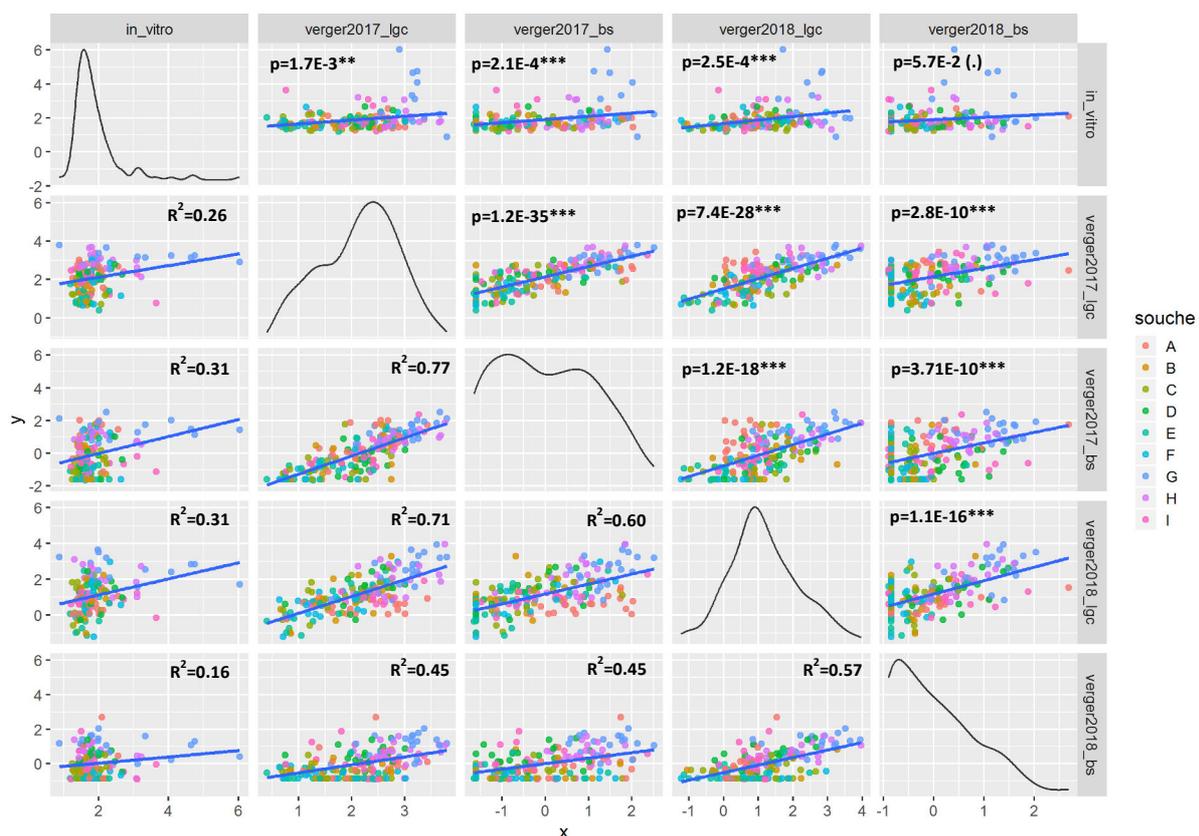
	In-vitro			In-vivo (2017&2018)					
	necrose superficielle			lgc			bs		
Factor	SS	F value	p-value	SS	F value	p-value	SS	F value	p-value
genotype (G)	0.68	5.9	4.89e-13***	68.5	6.62	<2.2e-16***	158.07	8.52	<2.2e-16***
strain (S)	0.46	7.98	4.20e-11***	360.02	82.61	<2.2e-16***	525.37	67.28	<2.2e-16***
operator	0.05	4.27	1.45e-2*	161.16	49.3	<2.2e-16***	70.36	12.01	3.89e-13***
Y	-	-	-	208.27	382.34	<2.2e-16***	195.89	200.7	<2.2e-16***
G:Y	-	-	-	-	-	-	165.69	8.94	<2.2e-16***
G:S	NS	NS	NS	-	-	-	53.26	1.6	2.06e-05***
Y:S	-	-	-	-	-	-	236.73	6.82	8.70e-09***
G:Y:S	-	-	-	268.15	1.4872	2.11e-06***	-	-	-
Adjusted R <sup>2</sup>	0.22			0.53			0.47		

Des effets d'interactions intervenant en particulier avec l'année bien que mineurs face aux facteurs individuels G, Y et S nécessitent une analyse ultérieure du dispositif année par année. Un effet d'interaction GxG mineur est à noter pour *bs* et interagit avec lui-même avec l'année pour *lgc*. L'effet de l'interaction GxG n'est pas significatif pour la nécrose superficielle (conditions in-vitro). La part de variance phénotypique globale expliquée par le modèle est comprise entre 22% et 53% avec le modèle le plus prédictif concernant la variable *lgc* dans le dispositif en verger en 2017 et 2018.

Suite à extraction des coefficients prédits GxY, pris annuellement dans le cas du dispositif en verger, des études de corrélations entre variables des deux dispositifs, et entre années ont été effectuées (Figure 30). Les variables *lgc* et *bs* sont aussi bien corrélées au sein d'une année (2017 :  $0.77^{***}$ , 2018 :  $0.57^{***}$ ) qu'entre années avec des corrélations allant de  $0.45^{***}$  (*lgc2017-bs2018*) à  $0.71^{***}$  (*lgc2017-lgc2018*). Ces résultats semblent indiquer que les effets d'interaction GxY ou YxS résultent

vraisemblablement davantage d'hétérogénéité de la variance que de changements de rangs.

Aussi, la nécrose superficielle issue du dispositif in-vitro bien que présentant des corrélations plus faibles avec les données in-vivo (corrélations allant de 0.16, en limite de significativité, avec *bs2018* à 0.31\*\*\* avec *bs2017*) semble être représentative des différences de sensibilité obtenues en verger.

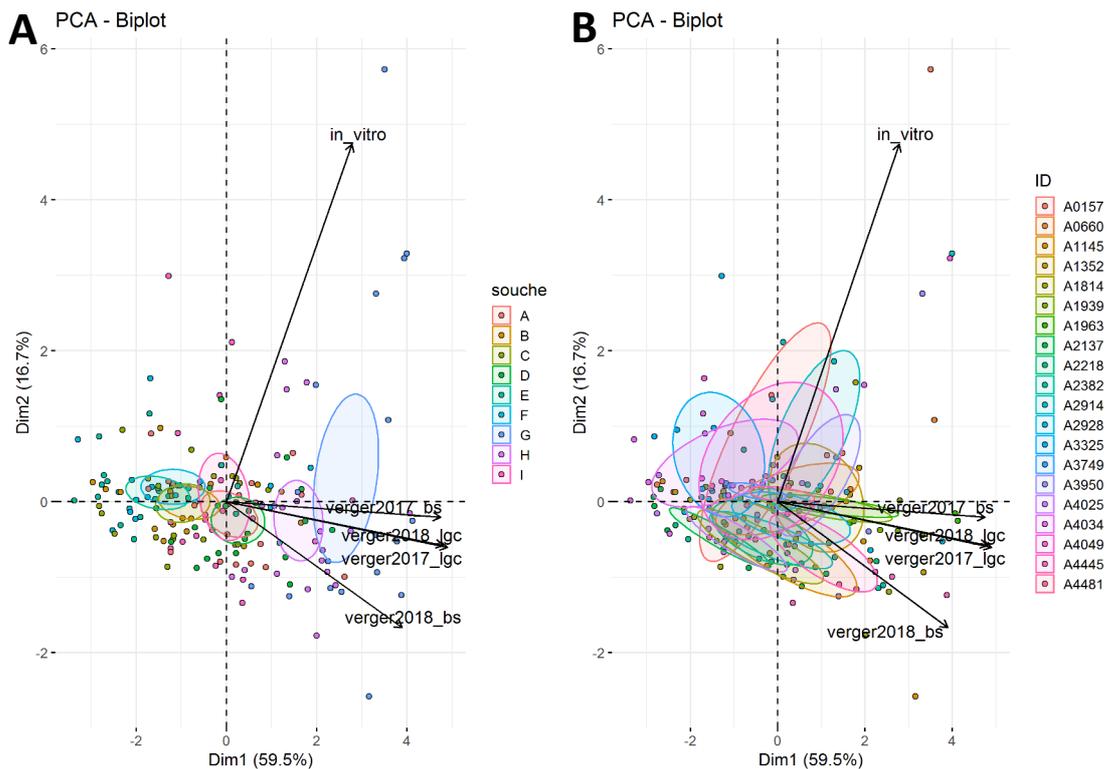


**Figure 30: Distribution et corrélations entre les variables étudiées dans le dispositif factoriel GxG - la nécrose superficielle mesurée en conditions in-vitro et les variables *lgc* et *bs* obtenues en verger en 2017 et 2018.**

L'échelle de couleurs correspond à une distinction des points selon les souches inoculées.

Les coefficients de corrélation de Pearson ( $R^2$ ) et les p-values de signification des tests de corrélation (corrégés des tests multiples avec un FDR à 5%) sont indiquées en-dessous et au-dessus de la diagonale, respectivement.

La variabilité totale de la nécrose superficielle, et des variables *lgc* et *bs* en 2017 et 2018, structurée majoritairement selon l'effet souche est visible au niveau de l'analyse bidimensionnelle en *figure 31A*. 76% de la variabilité totale est expliquée par une projection sur 2 axes principaux dont le premier rend compte de près de 60% de la variabilité. Les variables relatives au dispositif en verger ont une contribution principale au niveau de l'axe 1 alors que la nécrose superficielle, variable du dispositif in-vitro est majoritairement structurée au niveau de l'axe 2 (comme confirmé par les faibles corrélations, bien que significatives, entre les deux dispositifs).



**Figure 31:** Graphiques bidimensionnels représentant la répartition de l'effet G souche (A) et de l'effet G abricotier (B) au niveau de la nécrose superficielle, mesurée en conditions in-vitro, et *lgc* et *bs*, mesures obtenues en verger en 2017 et 2018.

Des ellipses de confiance permettent de mesurer la variance intra-groupe pour chaque souche (A) et chaque accession (B).

Au niveau de la projection des individus (combinaison génotype x souche) en fonction d'une structuration établie par les souches (*Figure 31A*), deux souches ressortent significativement plus agressives que les autres : les souches 1B (G) et 7C (H), toutes deux appartenant au PG2-d. Le pouvoir pathogène significativement supérieur de ces deux souches a été confirmé par des tests de comparaison post-hoc (SNK  $\alpha=0.05$ ) conduits sur *lgc* et *bs* chaque année en verger et sur la nécrose superficielle en laboratoire. Quel que soit le dispositif ou la variable, les souches 41A (D, PG 2b), DG8.15 (I, PG 7) et 19B (A, PG 1a) sont également pathogènes alors que les réactions provoquées par les souches B728A (F), P4.01.01.C3 (C), P3D7 (B) et P9T3 (E) ne se distinguent pas des réactions induites par l'inoculation 'eau' (sauf en 2017 du fait de phénomènes de contaminations naturelles des témoins négatifs en verger). Au niveau de la structuration de la variabilité des différents traits selon le génotype abricotier (*Figure 31B*), le classement est moins net, ce qui est illustré par le chevauchement des ellipses de confiance.

**Tableau 5 : Relation entre la présence et la migration de *Psy* souche 41A et l'apparition de la nécrose superficielle au niveau de tissus prélevés chez 3 génotypes – Bakour et Colomer (résistants en vert) et Bergeron (sensible en rouge).**

Les prélèvements de tissus ont été effectués sur le point d'inoculation (PI), 5 cm en-dessous du PI (G5) et 5 cm au-dessus du PI (D5), la bouture étant dirigée bourgeons vers le haut, et enfin au niveau du cœur du bois. Un traitement contrôle négatif consistant en l'inoculation d'eau stérile a également été effectué. Les cases orange indiquent une présence de symptômes de nécrose superficielle au niveau des points PI, D5 et G5 et de brunissement interne au niveau de BI. Les cases 0/+ indiquent l'absence/présence de la bactérie. Les cases 'ch' indiquent des cas de contamination fongique.

		Colomer				Bakour				Bergeron			
		Position des points				Position des points				Position des points			
Traitement	Répétition	G5	PI	D5	cœur	G5	PI	D5	cœur	G5	PI	D5	cœur
Eau	1	+	0	0	0	0	0	0	0	0	0	0	0
	2	+	0	0	0	0	0	0	0	0	ch	0	0
	3	+	+	0	+	0	0	0	0	ch	ch	0	0
	4	0	0	0	0	0	0	0	0	0	0	0	0
	5	0	0	0	0	0	0	0	0	0	0	0	0
41A	1	+	+	+	0	+	+	+	0	+	+	+	0
	2	+	+	+	0	+	+	+	0	+	+	+	+
	3	+	+	+	0	+	+	+	0	+	+	+	+
	4	+	+	+	0	+	+	+	+	+	+	+	0
	5	+	+	+	+	+	+	+	0	+	+	+	+

Dans le cas du traitement '41A', la présence de la bactérie au niveau du PI et de part et d'autre est systématique quel que soit la variété considérée, et ce, indépendamment de la présence de nécrose superficielle. Une fois inoculée, la bactérie semble se déplacer en longueur dans les tissus, mais pas en profondeur comme l'atteste l'absence globale de bactéries au niveau du cœur du bois quel que soit la variété. L'expérimentation ne nous permet pas de distinguer une quelconque différence variétale au niveau de la charge bactérienne, ni de l'expression des symptômes mis à part pour le brunissement interne avec une sensibilité supérieure chez Bergeron (5/5 réplicats) que chez Bakour (0/5) et Colomer (1/5). Quelques cas de présence de *Psy* ont été observés dans le cas du traitement 'eau'. Ceci peut attester de la présence de *Psy* à l'état endophyte dans les tissus, hypothèse plausible au vu de certains résultats issus d'un échantillonnage au niveau d'arbres pris aléatoirement à Gothon (Parisi, communication orale).

Bien que notre expérimentation de plan factoriel GxG ait ciblé une faible diversité du côté abricotier et *Psy*, la faible part relative des effets d'interaction GxG face aux facteurs individuels G et S semble indiquer des composantes d'interaction moléculaires GxG à effet principalement additif.

Deux souches 1B (G) et 7C (H), toutes deux appartenant au PG2-d, présentent un pouvoir pathogène significativement supérieur aux 7 autres souches exploitées. B728A bien qu'étant également affiliée au PG2-d ne présente en revanche pas de pouvoir pathogène. Cette observation semble alors confirmer la diversité d'agressivité au sein d'un même PG, pouvant résulter de combinaisons différentes de facteurs de virulence entre souches.

Notre étude de plan factoriel 20 accessions x 9 souches offre également un support intéressant d'analyse de matrices de compatibilité GxG dont l'agencement (modularité ou emboîtement/'nestedness') peut être révélateur des mécanismes de co-évolution entre plante et pathogène (Weitz *et al.*, 2013).

Cette approche, inspirée de l'étude des réseaux en écologie (Fortuna *et al.*, 2010), a été transposée dans le cadre des interactions bactérie-phage (Weitz *et al.*, 2013) et utilisée afin de caractériser de façon structurelle les motifs de compatibilité au niveau des pathosystèmes Solanacée x Potato virus Y (Moury *et al.*, 2014) et *Linum marginale* x *Melampsora lini* (Barrett *et al.*, 2015), notamment. Ces études portent majoritairement sur l'étude de données d'infection (0/1=interaction incompatible/compatible) d'un set de variants ou souche intra-spécifique d'une espèce pathogène donnée, confrontée à une diversité d'hôtes en intra-spécifique mais aussi en inter-spécifique (Moury *et al.*, 2017). Une structure 'nestedness' de la matrice de compatibilité reflète un emboîtement de la gamme d'hôte de variants/souches spécialistes dans la gamme d'hôte de variants/souches plus généralistes. A l'inverse, des structures de modularité indiquent des pattern d'interaction préférentiels (modules) entre groupes de variants/souches et génotypes de plante apparentés entre eux.

Une étude similaire a été conduite suite à des tests phénotypiques de caractérisation de la gamme d'hôte de 37 souches *Psy* testées sur 21 espèces végétales dont 10 étaient ligneuses (Morris *et al.*, travaux soumis). La structure emboîtée s'est révélée être significative tandis que l'hypothèse de modularité a été rejetée. Classiquement, le modèle emboîté indiquerait une patho-adaptation résultant d'occurrence successives de mutations bénéfiques chez le pathogène entrant dans le cadre de relations gène-pour-gène et permettant d'élargir sa gamme d'hôte (Weitz *et al.*, 2013; Moury *et al.*, 2014). Selon, cette hypothèse, l'évolution des agents pathogènes irait dans le sens de génotypes spécialisés vers des génotypes généralistes. Or, pour *Psy*, l'hypothèse la plus vraisemblable serait à l'inverse une évolution des gammes d'hôtes larges à plus étroites. Au niveau de l'étude du plan factoriel *Psy* x 21 espèces végétales, une corrélation significative négative entre nombre de séquences d'insertion et étendue de la gamme d'hôte a été mise en évidence. Plutôt que de considérer des mutations à avantage sélectif positif pour *Psy*, l'occurrence des séquences d'insertion a été ainsi avancée afin d'expliquer cette patho-adaptation 'à contre-courant' pouvant résulter de pertes de fonction de gènes (Morris *et al.*, communication personnelle).

Notre étude consacrée au cas des interactions GxG en explorant la diversité intra-spécifique chez l'abricotier permet d'adresser un regard complémentaire, en se focalisant sur des données quantitatives, attestant de différences d'agressivité entre souches de *Psy*. Les variables quantitatives sont largement utilisés pour décrire la résistance partielle du côté de l'hôte, elles sont à l'inverse beaucoup moins étudiées du côté de l'agent pathogène, pour caractériser leur agressivité. Pourtant, à l'heure actuelle, il est de plus en plus admis que le pouvoir pathogène résulte d'une composante qualitative (virulence : capacité d'infection) et d'une composante quantitative d'agressivité (Lannou, 2012).

<b>lgc 2018</b>	<b>D</b>	<b>A</b>	<b>G</b>	<b>H</b>	<b>B</b>	<b>C</b>	<b>E</b>	<b>F</b>	<b>I</b>	<b>lgc 2018</b>	<b>A</b>	<b>G</b>	<b>H</b>	<b>I</b>	<b>B</b>	<b>C</b>	<b>D</b>	<b>E</b>	<b>F</b>
A0660	8	4	9	8	0	0	0	1	6	A0157	3	7	4	4	0	0	3	0	0
A1145	6	2	9	3	0	1	0	0	5	A0660	4	9	8	6	0	0	8	0	1
A3325	6	3	6	2	0	0	0	3	3	A1145	2	9	3	5	0	1	6	0	0
A1814	6	5	9	8	0	4	0	0	4	A1814	5	9	8	4	0	4	6	0	0
A0157	3	3	7	4	0	0	0	0	4	A1939	4	9	8	5	0	0	8	2	0
A1963	5	1	7	6	3	0	0	0	1	A1963	1	7	6	1	3	0	5	0	0
A2137	6	4	6	6	1	1	1	1	5	A2137	4	6	6	5	1	1	6	1	1
A2218	4	4	6	4	0	2	0	0	4	A2218	4	6	4	4	0	2	4	0	0
A2382	7	0	9	6	0	2	1	0	5	A2382	0	9	6	5	0	2	7	1	0
A4034	4	1	7	6	0	0	2	0	5	A3325	3	6	2	3	0	0	6	0	3
A4049	4	0	7	6	0	0	0	0	2	A3749	0	5	6	2	0	0	3	0	0
A4445	7	7	8	7	4	3	2	0	5	A4025	0	6	5	5	0	1	3	0	0
A2914	2	2	7	3	0	0	0	0	4	A4034	1	7	6	5	0	0	4	2	0
A2928	3	1	6	5	0	0	0	0	3	A4049	0	7	6	2	0	0	4	0	0
A1352	1	3	7	6	0	0	0	0	2	A4445	7	8	7	5	4	3	7	2	0
A3950	4	0	7	4	0	0	1	0	3	A4481	2	6	4	4	0	0	4	1	0
A4481	4	2	6	4	0	0	1	0	4	A1352	3	7	6	2	0	0	1	0	0
A3749	3	0	5	6	0	0	0	0	2	A2914	2	7	3	4	0	0	2	0	0
A1939	8	4	9	8	0	0	2	0	5	A2928	1	6	5	3	0	0	3	0	0
A4025	3	0	6	5	0	1	0	0	5	A3950	0	7	4	3	0	0	4	1	0

**Figure 32 : Structuration de l'effet de modularité en 2 ou 4 modules détectés sur la variation de *lgc 2018* au niveau du plan factoriel d'interaction en verger 20 accessions x 9 souches.**

Analyse effectuée par Benoit Moury. Les données ont été arrondies et normalisées sur une échelle de 0 à 9 afin d'assurer une conformité avec les méthodes statistiques développées.

Un trait quantitatif (période de latence, taille de lésion, la production de toxines) mesuré pour caractériser des différences entre génotypes de l'agent pathogène est également sous contrôle du génotype de la plante et de l'interaction GxG.

Dans le cadre d'une démarche de validation de la transposition des méthodes d'analyse des matrices de compatibilité aux données quantitatives, notre plan factoriel d'interaction en verger 20 accessions x 9 souches sur les années 2017 et 2018 a été analysé par Benoit Moury (Pathologie Végétale). Quel que soit la variable (*lgc*, *bs*) et l'année considérée, la structure emboîtée ressort de façon significative ( $p < 0.01^{**}$ )

En complément, un effet de modularité a été également détecté sur *lgc* en 2018 avec une structuration en 2 ou 4 modules (*Figure 32*).

Dans le cas de la répartition en 2 modules, l'opposition du module de souches 19B (A, PG1-a), 1B (G, PG2-d), 7C (H, PG2-d) et DG8.15 (I, PG7) à celui comprenant P3D7 (B, PG2-b), P4.01.01.C3 (C, PG2-b), 41A (D, PG2-b), P9T3 (E, PG2-b) et B728A (F, PG2-d) est globalement cohérente avec les différences de pouvoir pathogène étudié précédemment, le premier groupe étant pathogène (en particulier G et H) et le second, non différent du traitement 'eau' (à l'exception de 41A). Dans le cas de la répartition en 4 groupes, les souches I et D forment 2 groupes supplémentaires, pathogènes mais dans une moindre mesure que les souches A, G et H. La proximité phylo-génétique des souches ne semble pas influencer la structuration de manière très sensible. Nous pouvons néanmoins noter dans le cas de la division en 2 modules, l'opposition du module avec les souches du PG2-b, sauf F (B728A du PG2-d), aux autres souches des PGs 1-a, 2-d et 7.

La structuration du côté génotype chez l'abricotier est plus délicate à justifier, ceci est vraisemblablement à relier à la significativité bien plus importante de l'effet souche face au facteur génotype plante, comme explicité précédemment. Les données de caractérisation de la diversité de la collection abricotier de 560 accessions (travail de H. Bourguiba) dont font partie les 20 accessions étudiées, ne nous ont pas permis de distinguer un quelconque lien entre appartenance à un module, et appartenance à un groupe de structure.

Globalement, l'analyse de notre matrice de données quantitatives d'interaction GxG, semble confirmer la structure 'nestedness' déjà illustrée sur les gammes d'hôtes de *Psy* (Morris *et al.*, travaux soumis). Deux souches 1B (G) et 7C (H), particulièrement agressives sur l'ensemble des génotypes sont responsables plus particulièrement de cet effet d'emboîtement. L'analyse de la structure de leur génome pourrait alors évaluer l'impact du nombre de séquences d'insertion sur le niveau d'agressivité en intra-spécifique, de façon analogue à l'étude sur les données d'infection sur les hôtes en inter-spécifique.

L'effet significatif de la modularité nécessite en revanche de plus amples recherches afin d'expliquer les hypothèses biologiques sous-jacentes à la distinction de ces 2 et 4 modules. De plus, la détection de ces modules dans le cas isolé de *lgc* en 2018, nous invite à mesurer la validité de ce résultat avec précaution.

Une étude plus approfondie, analogue à l'étude de Hulin *et al.* 2018, pourrait permettre d'identifier des éléments complémentaires de structure du génome (plasmides) ou des gènes dont le polymorphisme pourrait être significativement corrélé avec les différences d'agressivité. Ces analyses apparentées à des études d'association sur une diversité de souches pathogènes, sont connues pour être limitantes pour analyser les composantes de pouvoir pathogène de part 1) l'importance de la structure dans les populations réservoirs (comme explicité chez *Psy* au niveau de l'étude bibliographique), 2) le DL long observé chez beaucoup de populations pathogènes, limitant alors la résolution de

cartographie (Bartoli and Roux, 2017). La première contrainte nécessite d'analyser de façon indépendante chaque phylo-groupe afin d'éviter les faux-positifs, stratégie employée avec succès lors de la première GWAS effectuée sur *Psy* (Monteil *et al.*, 2016). A ce jour, aucune étude n'a encore été conduite afin de caractériser l'importance et l'étendue du DL sur le génome de *Psy*. Un marquage moléculaire dense pourrait permettre de limiter une baisse de résolution liée à un DL particulièrement étendu. Notre matériel d'étude offre une base de travail afin d'envisager ces approches sur génome complet, sous réserve de compléter les souches étudiées des PG1-a, PG2-b et PG2-d avec un nombre plus important de souches aussi bien agressives que non-agressives. Cela nécessite alors, comme réalisé dans notre cadre, une caractérisation phénotypique de l'agressivité de ces souches pouvant se limiter à une évaluation sur un génotype témoin abricotier comme Bergeron, pour lequel les tests d'agressivité se sont avérés être particulièrement discriminants entre souches.

## Partie VI - Discussion et perspectives

Ce travail de thèse avait pour principal objectif de caractériser la variabilité génétique de l'abricotier pour la résistance au dépérissement bactérien causé par *Psy*, à travers la mobilisation d'approches conjointes d'analyse de liaison et d'association. Il a été complété par, une étude préliminaire visant à préciser en verger et en laboratoire la nature des interactions entre les diversités propres à l'hôte et à la bactérie (interaction GxG), la quantification statistique du niveau d'interaction permettant d'adresser la problématique de durabilité des résistances. Les aboutissants de ces recherches apportent des éléments de compréhension sur l'architecture génétique de la résistance de l'abricotier à *Psy*, à travers l'identification des QRLs associés et des mécanismes sous-jacents (*partie 1*). Aussi, au vu des résultats portés sur les régions génomiques identifiées au niveau des parties III et IV, notre étude a également mené à la mise en évidence d'une certaine complémentarité entre la cartographie de liaison au sein du dispositif de populations biparentales et l'étude d'association sur la core-collection (*partie 2*). Après une revue critique des facteurs limitants de notre étude et des propositions de pistes d'amélioration (*partie 3*), l'impact du travail réalisé sur les programmes d'innovation variétale de l'abricotier et plus généralement des *Prunus* sera discuté (*partie 4*). Enfin, sur la base des travaux préliminaires sur l'interaction GxG en partie V, des perspectives de recherches futures sont discutées afin d'approfondir la caractérisation de l'interaction plante ligneuse-*Psy* et des facteurs de risque au développement de la maladie, et au contournement des composantes de résistance identifiées (*partie 5*).

## 1. Architecture génétique de la résistance partielle de l'abricotier au dépérissement bactérien

Un nombre important de QRLs a pu être identifié aussi bien au niveau des populations biparentales BerBa, des trois populations interconnectées de 2<sup>ème</sup> génération ainsi que de la core-collection de 73 accessions pour chacun des symptômes issus des inoculations artificielles : *lgc*, *bs* principalement, *lge* et *bi* dans une moindre mesure, et l'aptitude au blocage de l'entrée de *Psy* (évalué avec la notation de mortalité) suite à infection naturelle en verger.

Les deux grands types d'approches de cartographies de liaison et d'association ont permis de détecter près de 20 régions génomiques couvrant la totalité des chromosomes dont en particulier : 2 QRLs majeurs sur les chromosomes 5 et 6 expliquant conjointement 41 et 26% de la variation de *lgc* et *bs* sur deux années favorables à l'expression de la maladie, 1 QRL majeur sur le bas du chromosome 3 expliquant 25% de l'aptitude au blocage de l'infection et 1 QRL sur le chromosome 4, responsable de 20 et 27% de la variation de *lgc* et *bs* en 2017.

Le déterminisme génétique de la résistance au dépérissement bactérien s'est avéré être polygénique, en accord avec les hypothèses préalables aux travaux. Néanmoins, la part élevée de variance expliquée par certains marqueurs majeurs indique une architecture polygénique reposant sur un nombre limité de QRLs, résultat relativement inattendu de nos travaux. Ainsi en génétique d'association, les héritabilités au sens-strict de *lgc* et *bs* en 2013 et 2015 atteignent près de 0.45 et 0.62 et reposent sur 2 et 5 des 7 SNPs candidats détectés, respectivement.

Considérant l'étude effectuée sur les populations interconnectées, aucune colocalisation ou chevauchement d'intervalles de confiance n'est à noter entre les QRLs (population spécifique ou non spécifique) associés à la notation de mortalité (chromosomes 3, 6 et 8) et les QRLs relatifs aux symptômes issus d'inoculations artificielles (chromosomes 1, 4 et 6). Ces résultats semblent indiquer deux mécanismes de résistance sous-jacents intervenant à deux niveaux, à travers l'exploitation de la notation effectuée en conditions d'infection naturelle et des données phénotypiques issus des inoculations. Le premier mécanisme agirait de façon systémique, au niveau de la porte d'entrée de la bactérie, vraisemblablement au niveau des bourgeons afin de bloquer l'infection et le second, de manière plus localisée en limitant la migration de la bactérie dans les tissus.

Les intervalles supports de ces QRLs se sont avérés être relativement longs au niveau de la carte consensus, et n'ont pas permis d'engager une recherche des annotations des gènes contenus dans ces intervalles. Une perspective de densification des zones candidates permettra à l'avenir de réduire ces intervalles et de dresser une liste potentiellement plus réduite de gènes candidats, pouvant apporter des éléments de compréhension supplémentaires sur la complémentarité de ces deux mécanismes.

Certaines études ont pu mettre en évidence les faibles corrélations obtenues au niveau de données de sensibilité évaluées sur plusieurs organes d'une même espèce: sur fruit et feuilles dans le cas de *Xanthomonas campestris* sur pêcher (Werner *et al.*, 1986), et sur fruits et tiges de cerisier (Hulin *et al.*, 2018b), et feuilles et rameaux d'abricotier (travaux non publiés) dans le cas de *Psy*. Il est très probable au-delà des considérations expérimentales d'évaluation, que ces différences de sensibilité entre tests résultent de l'expression de facteurs génétiques différents contrôlant des mécanismes complémentaires.

Aussi, la complémentarité de mécanismes a été également avancée dans le cas de la résistance à la sharka chez l'abricotier : le locus *PPVres* étant impliqué dans le développement des symptômes (les haplotypes ayant les allèles procurant la résistance étant tolérants), et un autre locus encore non identifié et *a priori* non lié à *PPVres*, pourrait contrôler la charge virale et moduler l'effet de *PPVres* (Decroocq *et al.*, 2014).

## 2. Comparaison des approches de cartographie : convergences et particularités

De plus en plus d'études avancent la pertinence de la validation croisée entre les études d'association et les analyses de liaison (Pascual *et al.*, 2015; Albert *et al.*, 2016; Bazakos *et al.*, 2017). De par les différences au niveau de la diversité ciblée et de l'étendue du DL au sein des descendances de cartographie et des panels variétaux, les QTLs identifiés peuvent dépendre de la population exploitée et la résolution, de l'importance de la recombinaison. Il est intéressant de mettre en regard ces deux approches, afin d'évaluer la dépendance des QTLs détectés aux fonds génétiques utilisés.

Sur la vingtaine de QRLs détectés, deux régions de colocalisations entre les deux approches sont à noter : 1) une région sur le chromosome 6 entre 35 et 56.3 cM sur la carte de Bergeron, détectée avec les variables *lgc* et *lge* en 2013, et sur laquelle s'ancre physiquement (à proximité du SP525 à 39.8 cM) le SNP candidat à effet majeur LG6\_1527385 détecté en analyse d'association avec les variables *lgc* et *bs* en 2013 et 2015 et 2) une région sur le chromosome 7 entre 40 et 57.9 cM sur la carte de Bergeron, détectée avec la variable *lge* en 2014, et sur laquelle s'ancre physiquement le SNP candidat LG7\_18047191 (à proximité du SP725 à 40 cM) détecté en analyse d'association avec les variables *lgc* et *bs* en 2013 et 2015.

Aussi, l'étude d'association a permis de mettre en évidence des génotypes présentant une résistance intéressante à exploiter (Colomer, Harlayne) mais pour lesquels les haplotypes au niveau des SNPs candidats ne correspondaient pas aux haplotypes à combinaison allélique procurant la résistance. De même, Bakour ne figurait pas non plus parmi les haplotypes considérés comme les plus résistants. Ceci semble indiquer l'existence de variants procurant une résistance qui n'ont pas pu être capturés en GWAS, vraisemblablement en raison de leur faible fréquence au sein du panel de 73 accessions.

Ainsi d'autres QRLs sont propres à chacune des approches avec des colocalisations notées entre années (exemple de *bs* en 2016 et 2017 sur le groupe de liaison 4\_2 de la carte consensus) et entre phénotypes pour la même année (exemple de *lgc* et *lge* en 2013 sur le chromosome 6 de Bergeron).

L'utilisation des modèles multi-locus et multi-variés en GWAS s'est avéré particulièrement intéressante à mettre en œuvre afin de 1) limiter le nombre de SNPs détectés à ceux qui sont potentiellement causaux, en contrôlant les effets de DL à l'échelle du génome, 2) augmenter la puissance de détection de marqueurs à effet pléiotrope contrôlant la variation à la fois de *lgc* et *bs*. De manière surprenante et non-prévisible, l'usage conjoint de ces deux modèles nous a permis de mettre en évidence

un effet de liaison à l'échelle du génome entre les deux SNPs candidats à effet majeur LG6\_1527385 et LG5\_5394803, justifiant le fait que le modèle multi-locus n'ait pas permis de détecter conjointement les deux marqueurs en exploitant chacune des deux variables.

Concernant les analyses conduites au niveau des populations interconnectées, l'exploitation du modèle déconnecté, estimant les effets des QRLs par famille s'est avéré intéressant à mettre en œuvre et a permis de détecter près de 10 QRLs supplémentaires par rapport aux 5 QRLs détectés sous l'hypothèse connectée. Cette observation pourrait s'expliquer par un effet de dilution allélique de certains variants sous-représentés au niveau du modèle connecté, nécessitant alors une analyse par famille comme cela a pu déjà être mis en évidence chez le caféier (Mérot-L'Anthoëne *et al.*, 2014) ou le palmier à huile (Billotte *et al.*, 2010). Un effet de la population a été mis en évidence sur l'ensemble des variables exploitées soulignant les différences au niveau des fonds polygéniques des géniteurs C092, D099 et C109 ayant capturé des QRLs différents. Aussi pour la même variable, les QRLs détectés peuvent être situés sur des chromosomes différents, c'est le cas de l'aptitude à la survie en 2018, dont 3 QRLs ont été détectés sur les chromosomes 8, 3 et 6 chez les populations issues de C092, C109 et D099, respectivement. Pour les QRLs détectés par les deux modèles, les résultats ont montré des parts de variance plus élevées et des intervalles support moins étendus dans le cas du modèle connecté, illustrant la meilleure puissance de ce modèle liée à l'exploitation d'un nombre de recombinants plus élevé.

### **3. Regard critique et perspectives de recherches futures**

Afin d'étudier la sensibilité à la maladie des géotypes exploités en cartographie, le choix a été d'effectuer des inoculations contrôlées réalisées en verger. Les conditions en verger permettent d'assurer la réactivation printanière des tissus, déterminante dans le développement des symptômes. Toutefois, le revers de cette approche correspond à une forte dépendance de l'expression des symptômes aux conditions hivernales pouvant également mener à l'estimation de variances génétiques très faibles, les années non favorables au dépérissement bactérien. De plus, le domaine de l'Amarine, situé dans le Gard, ne se présente pas comme un site aux conditions particulièrement sensibilisantes à la maladie en raison de la douceur relative des périodes hivernales. L'hétérogénéité environnementale nécessite alors d'être prise en compte dans les dispositifs d'analyse pluri-annuels, ce qui a pu être assuré au cours de l'étude et conforté par une analyse année par année.

L'essai complémentaire réalisé en conditions contrôlées en laboratoire dans le cadre du dispositif GxG s'est avéré pertinent afin de discriminer l'agressivité des différentes souches étudiées. Il doit en revanche encore être consolidé afin de mettre en évidence de façon plus marquée les différences de sensibilité variétale avant d'offrir une perspective de transposition des données obtenues en verger.

Les résultats reposent sur l'exploitation de populations en verger dont la taille a pu apparaître comme étant un facteur limitant, en particulier dans le cas de la core-collection (73 accessions). Les modes opératoires d'inoculations contrôlées et de phénotypage dans le cas du dépérissement bactérien sont des protocoles extrêmement difficiles à réaliser, nécessitant des moyens humains conséquents et se posent comme la principale limite technique de notre étude. Le nombre d'individus traités en est alors réduit. Néanmoins, de nombreux QRLs ont pu être identifiés avec une précision de l'ordre du SNP candidat en GWAS grâce à l'attention portée à la prise en compte des effets environnementaux dans la modélisation des phénotypes.

En raison de la forte structuration de la core-collection ( $K=3$  groupes pour 73 individus), une perspective d'ajout d'accessions supplémentaires, en particulier d'individus issus du croisement de génotypes échantillonnés au sein de groupes différents, pourrait permettre d'accroître la puissance statistique tout en limitant l'effet d'artefact lié à la structure de population.

Au niveau des données génotypiques, une perte d'information au niveau de la cartographie réside dans l'utilisation de cartes génétiques séparées entre Bergeron et Bakour. Cette méthodologie de 'two-way pseudo-testcross' fréquemment utilisée au niveau des espèces pérennes ne permet de détecter les QTLs qu'à l'état hétérozygote chez chacun des parents. Une carte consensus construite entre les parents pourrait dépasser cette limite et détecter des QTLs génériques aux deux parents ainsi qu'évaluer les interactions possibles entre allèles (dominance). Etant donné qu'aucun QRL n'a été détecté sur la carte de Bakour, les QRLs potentiels s'ils existent seraient alors à l'état homozygote chez Bakour nécessitant alors l'élaboration d'une carte consensus sous réserve d'une couverture suffisante et exhaustive de marqueurs ponts. Or le faible polymorphisme observé entre Bergeron et Bakour complique cette perspective.

Le recours aux trois populations interconnectées de 2<sup>ème</sup> génération permet de lever en partie ces limitations de par l'augmentation du niveau de polymorphisme au sein des hybrides de 2<sup>ème</sup> génération, notamment au niveau des QRLs qui auraient pu être à l'état homozygote chez Bakour.

Toutefois, deux zones de 'gap' nécessitent un marquage supplémentaire afin de combler ces zones et relier les chromosomes 2 et 4, séparés en deux groupes de liaison. L'élaboration de la carte consensus et la cartographie de QRLs associée aux trois

populations interconnectées a également été limitée par le faible nombre de marqueurs polymorphes sur l'ensemble des populations, engendrant alors beaucoup de données manquantes au niveau de la matrice de génotypes. Bien qu'une inférence des génotypes sur ces données manquantes est possible en se basant sur la structure du DL, il est difficile d'apprécier la qualité de l'inférence et son impact dans les résultats de cartographie. Une densification des régions candidates se basant sur la recherche des marqueurs les plus polymorphes possible entre les populations pourrait permettre une meilleure précision de détection et d'estimation des effets.

Les individus C092, D099, et C109 géniteurs des populations interconnectées ont été sélectionnés préalablement à nos travaux, sur la base de données de susceptibilité au chancre bactérien et au monilia, collectées en verger sur une période de 2 ans, en 2012 et 2013. Ce choix consensus par rapport aux deux caractères, a été raisonné par rapport aux échelles de temps longues au niveau des croisements pour les espèces pérennes. Toutefois ce choix d'individus peut être contesté et n'est vraisemblablement pas optimal pour cartographier de façon exhaustive l'ensemble des déterminants issus de Bakour. Il est possible qu'un certain nombre de QRLs à plus faible effet aient pu échapper à notre analyse. Le choix d'individus tels que D045, D054 et D061, dont la faible susceptibilité s'est avérée stable sur les années 2012 et 2013 aurait pu être judicieux afin de maximiser le cumul des facteurs de résistances recherchés.

Enfin, l'adossement des données de marquage au génome de *P. persica*, seul génome de référence actuellement disponible amène un certain nombre d'approximations liées notamment à la perte d'information de séquences (le génome abricotier étant en théorie plus long de 4Mb/n), à des phénomènes de réarrangements chromosomiques pouvant avoir lieu entre espèces du genre *Prunus* même si une synténie globale est observée (Dirlewanger *et al.*, 2004). Ces limitations pourront être levées avec l'assemblage abricotier en cours de réalisation et qui pourra être mobilisable prochainement.

## 4. Applications pour la sélection variétale

Au cours de ces travaux, plusieurs sources de résistance provenant de plusieurs groupes phylo-géographiques ont pu être identifiées d'une part. Bien que la plupart des sources de résistances, semblent provenir du groupe d'Asie Centrale/de l'Est (Colomer – A0008, Harlayne) et du groupe Irano-Caucasien/Bassin Méditerranéen (Bakour – A2137, Arrogante – A1314, Manicot - A1236), le groupe d'Europe Continentale, présentant la plus grande variance de sensibilité aux inoculations contrôlées comporte également quelques géniteurs moins sensibles (sur la base

des travaux de H. Bourguiba sur la caractérisation de la diversité) qui ont pu être repérés suite au phénotypage à large échelle de 2018 : Royer (A0664), Malice (A2241). Il serait utile d'indiquer que Malice résulte d'un croisement (Canino x Hamidi) x Bergeron, Hamidi étant une accession proche de Bakour, toutes deux, d'origine tunisienne. Bien qu'ancrée dans le groupe d'Europe Continentale, Malice intègre ainsi la résistance du groupe Nord-Africaine.

Ces accessions du groupe d'Europe Continentale, partiellement résistantes à *Psy*, pourraient présenter un intérêt tout particulier en tant que géniteurs dans les programmes d'introgession de la résistance au dépérissement bactérien de par leur proximité génétique avec les fonds génétiques élites autofertiles, résistants à la sharka et adaptés aux zones continentales. La proximité génétique, en particulier entre fonds génétiques cultivés, favorise l'association de combinaisons alléliques favorables non seulement pour le trait étudié mais également pour d'autres traits en raison de la liaison entre gènes sur des segments chromosomiques plus ou moins longs.

Un résultat également intéressant porte sur la pertinence de deux variables en particulier (longueur de chancre et brunissement superficiel) afin de caractériser les différences de sensibilité en verger en réponse à des inoculations contrôlées. L'expérimentation de mise au point d'un protocole d'inoculations sur boutures en laboratoire s'est avéré être relativement intéressante à mettre en œuvre afin de reproduire les différences de sensibilité variétale obtenues en verger face à une diversité de souches. Ce protocole est d'autant plus pertinent que la phase de développement de symptômes post-inoculations est réduite à 1 mois (face à 5 en verger), permettant ainsi la réalisation de plusieurs répétitions expérimentales dans une année, sous réserve d'entreposer en prévision au froid, une quantité suffisante de boutures à l'état dormant.

L'ensemble de ces indicateurs et protocoles d'évaluation de la sensibilité variétale et du pouvoir pathogène de souches pourrait également présenter un intérêt allant au-delà de l'exemple de l'abricotier et être transposable aux espèces ligneuses.

Un ensemble de QRLs liés à la résistance partielle a été mis en évidence dont certains pourraient être qualifiés de 'majeurs' : 2 QRLs identifiés très précisément à l'échelle du SNP, sur les chromosomes 6 et 5 avec 41% et 26% de PVE sur *lgc* et *bs* au sein de la core-collection, 2 régions couvrant une distance d'environ 12 et 22 cM sur les chromosomes 3 et 4 avec des PVE s'élevant à 25% et de 10 à 27% sur la survie de l'arbre et les phénotypes résultants des inoculations contrôlées *lgc* et *bs* en 2016 et 2017, respectivement au sein des populations interconnectées 'Berba'. Ces marqueurs pourraient alors compléter la SAM existante visant la résistance à la sharka (locus *PPVres*) et l'autofertilité (locus *S*) dans le cadre de programmes d'innovations ayant pour objectif la conception d'idéotypes résistants aux maladies.

Au-delà de l'application d'une SAM sur l'abricotier, la transposition des marqueurs sur l'ensemble des *Prunus* semble être une perspective réalisable, du fait de la synténie importante entre les génomes du genre (Dirlewanger *et al.*, 2004). Une collaboration avec l'UMR BFP de l'INRA de Bordeaux est actuellement en cours, afin de caractériser le polymorphisme des 7 marqueurs identifiés en GWAS au niveau d'un panel variétal de cerisiers. De façon analogue à la démarche de validation, la confrontation des données génotypiques aux marqueurs avec des données de caractérisation de la sensibilité à *Psy* (déjà acquises en verger) permettra d'évaluer la validité de ces marqueurs candidats chez d'autres espèces fruitières.

Toutefois, il est nécessaire avant toute perspective d'application en sélection, d'apprécier l'étendue de la validité des marqueurs ciblés sur la diversité génétique existante. Les colocalisations limitées entre QRLs (exceptées pour 2 d'entre eux) révélés par la cartographie de liaison sur les populations biparentales et ceux détectés sur le panel d'association semble indiquer des QRLs spécifiques d'une part à Bakour et/ou Bergeron et d'autre part à la core-collection. Aussi, la démarche de validation des SNPs candidats détectés en GWAS vis-à-vis d'un panel variétal plus large que la core-collection semble indiquer des limites de la transposition de ces marqueurs sur une diversité plus exhaustive (en dehors des limites liées à un effet spatial chevauchant partiellement avec l'effet génétique). La dépendance des effets des QTLs à la taille et à la composition de la population est une limite très souvent soulevée dans les études de cartographie (Utz *et al.*, 2000; Korte and Farlow, 2013). Ceci est à mettre en regard avec la taille de population de GWAS limitée (73 accessions) face au panel étendu de 335 accessions présentant vraisemblablement des individus à variants non ou sous-représentés au sein du panel utilisé en cartographie d'association. Une perspective de séquençage par capture ou de génome complet sur le panel variétal étendu pourrait ainsi permettre de lever les limitations rencontrées en favorisant la détection de QRLs complémentaires et l'estimation des effets de manière plus précise.

Une perspective offerte par l'acquisition de données de marquage SNP à haute densité sur une vaste collection consiste à évaluer le poids des marqueurs détectés en GWAS à travers des modèles de sélection génomique. La comparaison d'un modèle de régression linéaire multiple établissant les SNPs candidats en effet fixe, avec un modèle traitant tous les SNPs en effet aléatoire permettrait d'évaluer si la sélection peut se baser sur ces SNPs candidats à travers une application de SAM ou bien sur tout un fond génétique, comme cela a déjà pu être évalué chez le citronnier (Minamikawa *et al.*, 2017). Ces modèles de sélection génomique pourraient également permettre de détecter des QRLs à effet plus faible de par leur propriété à s'affranchir des effets de seuil de détection.

## 5. Vers une résistance durable des vergers à *Psy* ?

L'architecture génétique de la résistance au chancre bactérien de l'abricotier repose d'après les résultats des travaux, sur un nombre limité de QRLs à effets majeurs complémenté par des QRLs à plus faible effet. Une perspective de pyramidage de l'ensemble de ces QRLs au niveau de fonds génétiques cultivés pourrait contrôler de manière durable la maladie en verger comme cela a pu être mis en évidence dans le cas de la tavelure du pommier (Laloi *et al.*, 2017). L'approche de cartographie des QRLs de résistance au sein des différents types de populations s'est reposée sur la collecte de données de sensibilité issues d'inoculations à partir d'une souche donnée (41A – PG2b). Cette souche caractérisée par une agressivité importante n'est représentative que d'une partie de la diversité du complexe *Psy* en verger correspondant à des souches appartenant aux PGs 1 clade a, 2 clade b, 3 et 7 principalement (Parisi *et al.*, 2016). Avant d'initier des inoculations à large échelle explorant une plus grande diversité de souches, il peut être utile d'explorer l'étendue de l'interaction GxG au travers un plan factoriel avec des effectifs plus modestes.

Les travaux ont alors évalué au plan statistique l'interaction GxG au niveau d'un plan factoriel d'interaction 20 accessions abricotier x 9 souches *Psy*, échantillonnées au niveau des PGs 1 et 2, principalement. Cette expérimentation a été conduite en verger sur 2 années et en laboratoire, durant une campagne. Les effets d'interaction GxG dans le cas où ils étaient significatifs se sont avérés très faibles face aux facteurs génétiques individuels en particulier l'effet de la souche, ce dernier structurant de manière très significative le jeu de données. Ces résultats semblent indiquer des déterminismes génétiques de pouvoir pathogène d'ordre additif, non dépendant du fond génétique de l'hôte. Ceci est particulièrement intéressant, compte-tenu que la plupart des traits quantitatifs mesurés dans le cadre d'un pathosystème, sont sous contrôle des fonds génétiques de l'hôte, de l'agent pathogène mais également des composantes moléculaires de leur interaction (Lambrechts *et al.*, 2006; Lannou, 2012). Cependant, ces interactions GxG ont tendance à résulter de la dynamique de co-évolution entre hôte et pathogène (Lambrechts *et al.*, 2006; Gallet *et al.*, 2013). De par l'évolution de généraliste à spécialiste, la co-évolution dans le cas de *Psy* et ses espèces hôtes en particulier, les espèces ligneuses est potentiellement contestable. De plus, les variétés et les souches exploitées ont été échantillonnées indépendamment, étayant alors cette hypothèse. L'agressivité des souches 1B (G) et 7C (H) semblerait davantage provenir de leur faculté préalablement acquise à infecter sévèrement leurs hôtes.

Ces résultats, encore préliminaires, méritent d'être consolidés. Ils offrent néanmoins une perspective de construction de résistances à large spectre, à effet généraliste face à plusieurs souches.

Il est également important de considérer un maintien des pratiques prophylactiques en appui des résistances partielles mises en évidence. Ces pratiques relativement simples à mettre en œuvre permettent ainsi de réduire l'importance des réservoirs d'inoculum primaires en réduisant les populations épiphytes sur les arbres ou le couvre-sol. Ceci a pour effet finalement de baisser les pressions de sélection exercées sur les populations de *Psy* et ainsi de prolonger la durabilité des QRLs déployés en verger.

## Références bibliographiques

- Ade J, DeYoung BJ, Golstein C, Innes RW.** 2007. Indirect activation of a plant nucleotide binding site-leucine-rich repeat protein by a bacterial protease. *Proceedings of the National Academy of Sciences* **104**(7), 2531–2536.
- Albar L, Lorieux M, Ahmadi N, Rimbault I, Pinel A, Sy AA, Fargette D, Ghesquière A.** 1998. Genetic basis and mapping of the resistance to rice yellow mottle virus. I. QTLs identification and relationship between resistance and plant morphology. *Theoretical and Applied Genetics* **97**(7), 1145–1154.
- Albert E, Segura V, Gricourt J, Bonnefoi J, Derivot L, Causse M.** 2016. Association mapping reveals the genetic architecture of tomato response to water deficit: focus on major fruit quality traits. *Journal of Experimental Botany* **67**(22), 6413–6430.
- Alburquerque N, Egea J, Pérez-Tornero O, Burgos L.** 2002. Genotyping apricot cultivars for self-(in)compatibility by means of RNases associated with *S* alleles. *Plant Breeding*.
- Alfano JR, Charkowski AO, Deng W-L, Badel JL, Petnicki-Ocwieja T, van Dijk K, Collmer A.** 2000. The *Pseudomonas syringae* Hrp pathogenicity island has a tripartite mosaic structure composed of a cluster of type III secretion genes bounded by exchangeable effector and conserved effector loci that contribute to parasitic fitness and pathogenicity. *Proceedings of the National Academy of Sciences* **97**(9), 4856–4861.
- Allard A, Bink MCAM, Martinez S, Kelner J-J, Legave J-M, di Guardo M, Di Pierro EA, Laurens F, van de Weg EW, Costes E.** 2016. Detecting QTLs and putative candidate genes involved in budbreak and flowering time in an apple multiparental population. *Journal of Experimental Botany* **67**(9), 2875–2888.
- Anzai Y, Kim H, Park JY, Wakabayashi H, Oyaizu H.** 2000. Phylogenetic affiliation of the *Pseudomonas* based on 16S rRNA sequence. *International Journal of Systematic and Evolutionary Microbiology* **50**(4), 1563–1589.
- Arnold DL, Jackson RW.** 2011. Bacterial genomes: evolution of pathogenicity. *Current Opinion in Plant Biology* **14**(4), 385–391.
- Arrebola E, Cazorla FM, Perez-García A, de Vicente A.** 2011. Chemical and metabolic aspects of antimetabolite toxins produced by *Pseudomonas syringae* pathovars. *Toxins* **3**(9), 1089–1110.
- Arumuganathan K, Earle ED.** 1991. Nuclear DNA content of some important plant species. *Plant Molecular Biology Reporter* **9**(3), 208–218.
- Ashfield T, Redditt T, Russell A, Kessens R, Rodibaugh N, Galloway L, Kang Q, Podicheti R, Innes RW.** 2014. Evolutionary relationship of disease resistance genes in soybean and *Arabidopsis* specific for the *Pseudomonas syringae* effectors AvrB and AvrRpm1. *Plant Physiology* **166**(1), 235–251.
- Asselbergh B, De Vleeschauwer D, Höfte M.** 2008. Global switches and fine-tuning—ABA modulates plant pathogen defense. *Molecular Plant-Microbe Interactions* **21**(6), 709–719.
- Astle W, Balding DJ.** 2009. Population structure and cryptic relatedness in genetic association studies. *Statistical Science* **24**(4), 451–471.
- Atwell S, Huang YS, Vilhjálmsson BJ, Willems G, Horton M, Li Y, Meng D,**

- Platt A, Tarone AM, Hu TT, Jiang R, Mulyati NW, Zhang X, Amer MA, Baxter I, Brachi B, Chory J, Dean C, Debieu M, de Meaux J, Ecker JR, Faure N, Kniskern JM, Jones JDG, Michael T, Nemri A, Roux F, Salt DE, Tang C, Todesco M, Traw MB, Weigel D, Marjoram P, Borevitz JO, Bergelson J, Nordborg M. 2010. Genome-wide association study of 107 phenotypes in *Arabidopsis thaliana* inbred lines. *Nature* **465**(7298), 627–631.
- Audergon J-M. 1995. Variety and breeding. *Acta Horticulturae* **384**, 35–46.
- Audergon J-M, Castelain C, Morvan G, Chastelliere M-G. 1989. Apricot varietal sensibility and genetic variability to apricot chlorotic leaf roll disease. *Acta Horticulturae* **235**, 205–213.
- Axtell MJ, Staskawicz BJ. 2003. Initiation of RPS2-specified disease resistance in *Arabidopsis* is coupled to the AvrRpt2-directed elimination of RIN4. *Cell* **112**(3), 369–377.
- Bacanu S-A, Devlin B, Roeder K. 2002. Association studies for quantitative traits in structured populations. *Genetic Epidemiology* **22**(1), 78–93.
- Backman PA, DeVay JE. 1971. Studies on the mode of action and biogenesis of the phytotoxin syringomycin. *Physiological Plant Pathology* **1**(2), 215–233.
- Bailey LH, Hough FL. 1975. Apricot. *Advances in Fruit Breeding*. Purdue University Press, 367–386.
- Balasubramanian S, Schwartz C, Singh A, Warthmann N, Kim MC, Maloof JN, Loudet O, Trainer GT, Dabi T, Borevitz JO, Chory J, Weigel D. 2009. QTL mapping in new *Arabidopsis thaliana* advanced intercross-recombinant inbred lines. *PLoS ONE* **4**(2), e4318.
- Baltrus DA, McCann HC, Guttman DS. 2017. Evolution, genomics and epidemiology of *Pseudomonas syringae*: Challenges in bacterial molecular plant pathology. *Molecular Plant Pathology* **18**(1), 152–168.
- Baltrus DA, Nishimura MT, Romanchuk A, Chang JH, Mukhtar MS, Cherkis K, Roach J, Grant SR, Jones CD, Dangl JL. 2011. Dynamic evolution of pathogenicity revealed by sequencing and comparative genomics of 19 *Pseudomonas syringae* isolates. *PLoS Pathogens* **7**, e1002132.
- Barbary A, Djian-Caporalino C, Marteu N, Fazari A, Caromel B, Castagnone-Sereno P, Palloix A. 2016. Plant genetic background increasing the efficiency and durability of major resistance genes to root-knot nematodes can be resolved into a few resistance QTLs. *Frontiers in Plant Science* **7**(632), 56–64.
- Barrett LG, Encinas-Viso F, Burdon JJ, Thrall PH. 2015. Specialization for resistance in wild host-pathogen interaction networks. *Frontiers in Plant Science* **6**(761).
- Bartoli C, Carrere S, Lamichhane JR, Varvaro L, Morris CE. 2015a. Whole-genome sequencing of 10 *Pseudomonas syringae* strains representing different host range spectra. *Genome Announcements* **3**, 2–3.
- Bartoli C, Lamichhane JR, Berge O, Guilbaud C, Varvaro L, Balestra GM, Vinatzer BA, Morris CE. 2015b. A framework to gauge the epidemic potential of plant pathogens in environmental reservoirs: the example of kiwifruit canker. *Molecular Plant Pathology* **16**(2), 137–149.
- Bartoli C, Lamichhane JR, Berge O, Varvaro L, Morris CE. 2015c. Mutability in

- Pseudomonas viridiflava* as a programmed balance between antibiotic resistance and pathogenicity. *Molecular Plant Pathology* **16**(8), 860–869.
- Bartoli C, Roux F.** 2017. Genome-wide association studies in plant pathosystems: toward an ecological genomics approach. *Frontiers in Plant Science* **8**(763).
- Bazakos C, Hanemian M, Trontin C, Jiménez-Gómez JM, Loudet O.** 2017. New strategies and tools in quantitative genetics: how to go from the phenotype to the genotype. *Annual Review of Plant Biology* **68**(1), 435–455.
- Bender CL, Alarcón-Chaidez F, Gross DC.** 1999. *Pseudomonas syringae* phytotoxins: mode of action, regulation, and biosynthesis by peptide and polyketide synthetases. *Microbiology and Molecular Biology Reviews* **63**(2), 266–292.
- Berge O, Monteil CL, Bartoli C, Chandeysson C, Guilbaud C, Sands DC, Morris CE.** 2014. A user's guide to a data base of the diversity of *Pseudomonas syringae* and its application to classifying strains in this phylogenetic complex. *PLoS ONE* **9**(9), e105547.
- Van Der Biezen EA, Jones JDG.** 1998. Plant disease-resistance proteins and the gene-for-gene concept. *Trends in Biochemical Sciences* **23**(12), 454–456.
- Billotte N, Jourjon M-F, Marseillac N, Berger A, Flori A, Asmady H, Adon B, Singh R, Nouy B, Potier F, Cheah SC, Rohde W, Ritter E, Courtois B, Charrier A, Mangin B.** 2010. QTL detection by multi-parent linkage mapping in oil palm (*Elaeis guineensis* Jacq.). *Theoretical and Applied Genetics* **120**(8), 1673–1687.
- Bink MCAM, Jansen J, Madduri M, Voorrips RE, Durel C-E, Kouassi AB, Laurens F, Mathis F, Gessler C, Gobbin D, Rezzonico F, Patocchi A, Kellerhals M, Boudichevskaia A, Dunemann F, Peil A, Nowicka A, Lata B, Stankiewicz-Kosyl M, Jeziorek K, Pitera E, Soska A, Tomala K, Evans KM, Fernández-Fernández F, Guerra W, Korbin M, Keller S, Lewandowski M, Plochanski W, Rutkowski K, Zurawicz E, Costa F, Sansavini S, Tartarini S, Komjanc M, Mott D, Antofie A, Lateur M, Rondia A, Gianfranceschi L, van de Weg EW.** 2014. Bayesian QTL analyses using pedigreed families of an outcrossing species, with application to fruit firmness in apple. *Theoretical and Applied Genetics* **127**(5), 1073–1090.
- Biscarini F, Nazzicari N, Bink M, Arús P, Aranzana MJ, Verde I, Micali S, Pascal T, Quilot-Turion B, Lambert P, da Silva Linge C, Pacheco I, Bassi D, Stella A, Rossini L.** 2017. Genome-enabled predictions for fruit weight and quality from repeated records in European peach progenies. *BMC Genomics* **18**(432).
- Block A, Guo M, Li G, Elowsky C, Clemente TE, Alfano JR.** 2010. The *Pseudomonas syringae* type III effector HopG1 targets mitochondria, alters plant development and suppresses plant innate immunity. *Cellular Microbiology* **12**(3), 318–330.
- Boller T, Felix G.** 2009. A renaissance of elicitors: perception of Microbe-Associated Molecular Patterns and danger signals by Pattern-Recognition Receptors. *Annual Review of Plant Biology* **60**, 379–406.
- Bonnafous F, Fievet G, Blanchet N, Boniface M-C, Carrère S, Gouzy J, Legrand L, Marage G, Bret-Mestries E, Munos S, Pouilly N, Vincourt P, Langlade N, Mangin B.** 2018. Comparison of GWAS models to identify non-additive genetic control of flowering time in sunflower hybrids. *Theoretical and Applied Genetics* **131**(2), 319–

- Bordjiba O, Prunier J-P.** 1991. Establishment of an epiphytic phase by three species of *Pseudomonas* on apricot trees. *Acta Horticulturae* **293**, 487–494.
- Borschinger B, Bartoli C, Chandeysson C, Guilbaud C, Parisi L, Bourgeay J-F, Buisson E, Morris CE.** 2016. A set of PCRs for rapid identification and characterization of *Pseudomonas syringae* phylogroups. *Journal of Applied Microbiology* **120**(3), 714–723.
- Bourguiba H, Audergon J-M, Krichen L, Trifi-Farah N, Mamouni A, Trabelsi S, D’Onofrio C, Asma BM, Santoni S, Khadari B.** 2012. Loss of genetic diversity as a signature of apricot domestication and diffusion into the Mediterranean Basin. *BMC Plant Biology* **12**(49).
- Bourguiba H, Khadari B, Krichen L, Trifi-Farah N, Santoni S, Audergon J-M.** 2010. Grafting versus seed propagated apricot populations: Two main gene pools in Tunisia evidenced by SSR markers and model-based Bayesian clustering. *Genetica* **138**(9–10), 1023–1032.
- Brachi B, Faure N, Horton M, Flahauw E, Vazquez A, Nordborg M, Bergelson J, Cuguen J, Roux F.** 2010. Linkage and association mapping of *Arabidopsis thaliana* flowering time in nature. *PLoS Genetics* **6**(5), e1000940.
- Brachi B, Morris GP, Borevitz JO.** 2011. Genome-wide association studies in plants: the missing heritability is in the field. *Genome Biology* **12**(10), 232.
- Bradley DJ, Gilbert GS, Parker IM.** 2003. Susceptibility of clover species to fungal infection: The interaction of leaf surface traits and environment. *American Journal of Botany* **90**(6), 857–864.
- Brodny U, Nelson RR, Gregory L V.** 1986. The residual and interactive expressions of ‘defeated’ wheat stem rust resistance genes. *Phytopathology* **76**(5), 546–549.
- Brooks DM, Hernández-Guzmán G, Kloek AP, Alarcón-Chaidez F, Sreedharan A, Rangaswamy V, Peñaloza-Vázquez A, Bender CL, Kunkel BN.** 2004. Identification and characterization of a well-defined series of coronatine biosynthetic mutants of *Pseudomonas syringae* pv. *tomato* DC3000. *Molecular Plant-Microbe Interactions* **17**(2), 162–174.
- Brun L, Warlop F, Mercier V, Broquaire J-M, Clauzel G, Gomez C, Parveaud C-E, Audergon J-M.** 2011. Quelle sélection fruitière pour une production durable, à faible niveau d’intrants? *Méthodologie pour un réseau de sélection variétale décentralisée. Innovations Agronomiques* **15**, 105–115.
- Brzezinski J.** 1902. Etiologie du chancre et de la gomme des arbres fruitiers. *Comptes rendus de l’Académie des Sciences* **134**, 1170–1173.
- Buckler ES, Thornsberry JM.** 2002. Plant molecular diversity and applications to genomics. *Current Opinion in Plant Biology* **5**(2), 107–111.
- Buell CR, Joardar V, Lindeberg M, Selengut J, Paulsen IT, Gwinn ML, Dodson RJ, Deboy RT, Durkin AS, Kolonay JF, Madupu R, Daugherty S, Brinkac L, Beanan MJ, Haft DH, Nelson WC, Davidsen T, Zafar N, Zhou L, Liu J, Yuan Q, Khouri H, Fedorova N, Tran B, Russell D, Berry K, Utterback T, Van Aken SE, Feldblyum T V, D’Ascenzo M, Deng WL, Ramos AR, Alfano JR, Cartinhour S, Chatterjee AK, Delaney TP, Lazarowitz SG, Martin GB,**

- Schneider DJ, Tang X, Bender CL, White O, Fraser CM, Collmer A.** 2003. The complete genome sequence of the *Arabidopsis* and tomato pathogen *Pseudomonas syringae* pv. *tomato* DC3000. *Proceedings of the National Academy of Sciences* **100**(18), 10181–10186.
- Bull CT, De Boer SH, Denny TP, Firrao G, Fisher-Le-Saux M, Saddler GS, Scortichini M, Stead DE, Takikawa Y.** 2010. Comprehensive list of names of plant pathogenic bacteria, 1980-2007. *Journal of Plant Pathology* **92**(3), 551–592.
- Burgos L, Pérez-Tornero O, Ballester J, Olmos E.** 1998. Detection and inheritance of stylar ribonucleases associated with incompatibility alleles in apricot. *Sexual Plant Reproduction* **11**(3), 153–158.
- Büttner D, He SY.** 2009. Type III protein secretion in plant pathogenic bacteria. *Plant Physiology* **150**(4), 1656–1664.
- Byrne DH, Littleton TG.** 1989. Characterization of isozyme variability in apricots. *Journal of the American Society for Horticultural Science* **114**, 674–678.
- Calenge F, Faure A, Goerre M, Gebhardt C, Van de Weg WE, Parisi L, Durel C-E.** 2004. Quantitative trait loci (QTL) analysis reveals both broad-spectrum and isolate-specific QTL for scab resistance in an apple progeny challenged with eight isolates of *Venturia inaequalis*. *Phytopathology* **94**(4), 370–379.
- Cameron HR.** 1970. *Pseudomonas* content of cherry trees. *Phytopathology* **60**, 1343–1346.
- Campoy JA, Ruiz D, Alderman L, Cook N, Egea J.** 2012. The fulfilment of chilling requirements and the adaptation of apricot (*Prunus armeniaca* L.) in warm winter climates: an approach in Murcia (Spain) and the Western Cape (South Africa). *European Journal of Agronomy* **37**(1), 43–55.
- Campoy JA, Ruiz D, Egea J, Rees DJG, Celton J-M, Martínez-Gómez P.** 2011. Inheritance of flowering time in apricot (*Prunus armeniaca* L.) and analysis of linked quantitative trait loci (QTLs) using simple sequence repeat (SSR) markers. *Plant Molecular Biology Reporter* **29**(2), 404–410.
- Canonne J, Froidure-Nicolas S, Rivas S.** 2011. Phospholipases in action during plant defense signaling. *Plant Signaling and Behavior* **6**(1), 13–18.
- Cao T, Kirkpatrick BC, Shackel KA, DeJong TM.** 2013. Effect of leaf scar age, chilling and freezing-thawing on infection of *Pseudomonas syringae* pv. *syringae* through leaf scars and lenticels in stone fruits. *Fruits* **68**(2), 159–169.
- Cao K, Wang L, Zhu G, Fang W, Chen C, Luo J.** 2012. Genetic diversity, linkage disequilibrium, and association mapping analyses of peach (*Prunus persica*) landraces in China. *Tree Genetics & Genomes* **8**(5), 975–990.
- Cao K, Zhou Z, Wang Q, Guo J, Zhao P, Zhu G, Fang W, Chen C, Wang X, Wang X, Tian Z, Wang L.** 2016. Genome-wide association study of 12 agronomic traits in peach. *Nature Communications* **7**(13246).
- Caranta C, Lefebvre V, Palloix A.** 1997. Polygenic resistance of pepper to potyviruses consists of a combination of isolate-specific and broad-spectrum quantitative trait loci. *Molecular Plant-Microbe Interactions* **10**(7), 872–878.
- Carmona MJ, Molina A, Fernández JA, López-Fando JJ, García-Olmedo F.** 1993. Expression of the  $\alpha$ -thionin gene from barley in tobacco confers enhanced resistance to bacterial pathogens. *Plant Journal* **3**(3), 457–462.

- Catchen JM, Amores A, Hohenlohe PA, Cresko W, Postlethwait JH. 2011. Stacks: building and genotyping loci de novo from short-read sequences. *G3: Genes|Genomes|Genetics* **1**(3), 171–182.
- Cazorla FM, Arrebola E, Olea F, Velasco L, Hermoso JM, Pérez-García A, Torés JA, Farré JM, de Vicente A. 2006. Field evaluation of treatments for the control of the bacterial apical necrosis of mango (*Mangifera indica*) caused by *Pseudomonas syringae* pv. *syringae*. *European Journal of Plant Pathology* **116**(4), 279–288.
- Cazorla FM, Arrebola E, Sesma A, Pérez-García A, Codina JC, Murillo J, de Vicente A. 2002. Copper resistance in *Pseudomonas syringae* strains isolated from mango is encoded mainly by plasmids. *Phytopathology* **92**(8), 909–916.
- Cervellati C, Paetz C, Dondini L, Tartarini S, Bassi D, Schneider B, Masia A. 2012. A qNMR approach for bitterness phenotyping and QTL identification in an F1 apricot progeny. *Journal of Biotechnology* **159**(4), 312–319.
- Cevik V, Ryder CD, Popovich A, Manning K, King GJ, Seymour GB. 2010. A FRUITFULL-like gene is associated with genetic variation for fruit flesh firmness in apple (*Malus domestica* Borkh.). *Tree Genetics & Genomes* **6**(2), 271–279.
- Chagné D, Crowhurst RN, Troggio M, Davey MW, Gilmore B, Lawley C, Vanderzande S, Hellens RP, Kumar S, Cestaro A, Velasco R, Main D, Rees JD, Iezzoni A, Mockler TC, Wilhelm L, Van de Weg EW, Gardiner SE, Bassil N, Peace CP. 2012. Genome-wide SNP detection, validation, and development of an 8K SNP array for apple. *PLoS ONE* **7**(2), e31745.
- Chang JH, Tai Y-S, Bernal AJ, Lavelle DT, Staskawicz BJ, Michelmore RW. 2002. Functional analyses of the *Pto* resistance gene family in tomato and the identification of a minor resistance determinant in a susceptible haplotype. *Molecular Plant-Microbe Interactions* **15**(3), 281–291.
- Checker VG, Kushwaha HR, Kumari P, Yadav S. 2018. Role of phytohormones in plant defense: signaling and cross talk. In: Singh A,, In: Singh IK, eds. *Molecular Aspects of Plant-Pathogen Interaction*. Singapore: Springer Singapore, 159–184.
- Chien C-F, Mathieu J, Hsu C-H, Boyle P, Martin GB, Lin N-C. 2013. Nonhost resistance of tomato to the bean pathogen *Pseudomonas syringae* pv. *syringae* B728a is due to a defective E3 ubiquitin ligase domain in AvrPtoB B728a. *Molecular Plant-Microbe Interactions* **26**(4), 387–393.
- Choi HW, Klessig DF. 2016. DAMPs, MAMPs, and NAMPs in plant innate immunity. *BMC Plant Biology* **16**(232).
- Cody YS, Gross DC, Proebsting Jr. EL, Spotts RA. 1987. Suppression of ice nucleation-active *Pseudomonas syringae* by antagonistic bacteria in fruit tree orchards and evaluations of frost control. *Phytopathology* **77**, 1036–1044.
- Collard BCY, Jahufer MZZ, Brouwer JB, Pang ECK. 2005. An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: the basic concepts. *Euphytica* **142**(1), 169–196.
- Collard BCY, Mackill DJ. 2008. Marker-assisted selection : an approach for precision plant breeding in the twenty-first century. *Phil. Trans. R. Soc. B* **363**(1491), 557–572.
- Conte L, Nicotra A, Corazza L. 2004. New apricot selections resistant to *Monilinia laxa* (Aderh. et Ruhl.). *Acta Horticulturae* **663**, 245–250.

- Cook DE, Mesarich CH, Thomma BPHJ.** 2015. Understanding plant immunity as a surveillance system to detect invasion. *Annual Review of Phytopathology* **53**, 541–563.
- Corder EH, Saunders AM, Risch NJ, Strittmatter WJ, Schmechel DE, Gaskell PC, Rimmler JB, Locke PA, Conneally PM, Schmechel KE, Small GW, Roses AD, Haines JL, Pericak-Vance MA.** 1994. Protective effect of apolipoprotein E type 2 allele for late onset Alzheimer disease. *Nature Genetics* **7**(2), 180–184.
- Corwin JA, Kliebenstein DJ.** 2017. Quantitative resistance: More than just perception of a pathogen. *The Plant Cell* **29**(4), 655–665.
- Costa F.** 2015. MetaQTL analysis provides a compendium of genomic loci controlling fruit quality traits in apple. *Tree Genetics & Genomes* **11**(819).
- Crossa-Raynaud P.** 1960. *Problèmes d'arboriculture fruitière en Tunisie* .
- Crossa J, Pérez-Rodríguez P, Cuevas J, Montesinos-López O, Jarquín D, de los Campos G, Burgueño J, González-Camacho JM, Pérez-Elizalde S, Beyene Y, Dreisigacker S, Singh R, Zhang X, Gowda M, Roorkiwal M, Rutkoski J, Varshney RK.** 2017. Genomic selection in plant breeding: methods, models, and perspectives. *Trends in Plant Science* **22**(11), 961–975.
- Crosse JE.** 1953. Bacterial diseases of stone-fruit trees in Britain: IX. Bacteriosis of apricot. *Transactions of the British Mycological Society* **36**, 38–45.
- Crosse JE.** 1959. Bacterial canker of stone fruits : IV. Investigation of a method for measuring the inoculum potential of cherry trees. *Annals of Applied Biology* **47**(2), 306–317.
- Crosse JE, Garrett CME.** 1966. Bacterial canker of stone-fruits: VII. Infection experiments with *Pseudomonas mors-prunorum* and *P.syringae*. *Annals of Applied Biology* **58**(1), 31–41.
- Cunnac S, Chakravarthy S, Kvitko BH, Russell AB, Martin GB, Collmer A.** 2011. Genetic disassembly and combinatorial reassembly identify a minimal functional repertoire of type III effectors in *Pseudomonas syringae*. *Proceedings of the National Academy of Sciences* **108**(7), 2975–2980.
- Dangl JL.** 1992. Functional homologs of the *Arabidopsis* RPM1 disease resistance gene in bean and pea. *Plant Cell* **4**(11), 1359–1369.
- Dangl JL, Jones JD.** 2001. Plant pathogens and integrated defence responses to infection. *Nature* **411**(6839), 826–833.
- Davey JW, Hohenlohe PA, Etter PD, Boone JQ, Catchen JM, Blaxter ML.** 2011. Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nature Reviews Genetics* **12**(7), 499–510.
- Debieu M, Huard-Chauveau C, Genissel A, Roux F, Roby D.** 2016. Quantitative disease resistance to the bacterial pathogen *Xanthomonas campestris* involves an *Arabidopsis* immune receptor pair and a gene of unknown function. *Molecular Plant Pathology* **17**(4), 510–520.
- Decroocq S, Chague A, Lambert P, Roch G, Audergon JM, Geuna F, Chiozzotto R, Bassi D, Dondini L, Tartarini S, Salava J, Krška B, Palmisano F, Karayiannis I, Decroocq V.** 2014. Selecting with markers linked to the PPVres major QTL is not sufficient to predict resistance to Plum Pox Virus (PPV) in apricot. *Tree Genetics & Genomes* **10**(5), 1161–1170.
- Dekkers JCM, Hospital F.** 2002. The use of molecular genetics in the improvement of

- agricultural populations. *Nature Reviews Genetics* **3**(1), 22–32.
- Denby KJ, Kumar P, Kliebenstein DJ.** 2004. Identification of *Botrytis cinerea* susceptibility loci in *Arabidopsis thaliana*. *Plant Journal*.
- Deslandes L, Olivier J, Peeters N, Feng DX, Khounlotham M, Boucher C, Somssich I, Genin S, Marco Y.** 2003. Physical interaction between RRS1-R, a protein conferring resistance to bacterial wilt, and PopP2, a type III effector targeted to the plant nucleus. *Proceedings of the National Academy of Sciences* **100**(13), 8024–8029.
- Develey-Rivière MP, Galiana E.** 2007. Resistance to pathogens and host developmental stage: A multifaceted relationship within the plant kingdom. *New Phytologist* **175**, 405–416.
- Diallo MD, Monteil CL, Vinatzer BA, Clarke CR, Glaux C, Guilbaud C, Desbiez C, Morris CE.** 2012. *Pseudomonas syringae* naturally lacking the canonical type III secretion system are ubiquitous in nonagricultural habitats, are phylogenetically diverse and can be pathogenic. *The ISME Journal* **6**(7), 1325–1335.
- Dimitrijevic A, Horn R.** 2018. Sunflower hybrid breeding: from markers to genomic selection. *Frontiers in Plant Science* **8**(2238).
- Dirlewanger E, Graziano E, Joobeur T, Garriga-Caldere F, Cosson P, Howad W, Arús P.** 2004. Comparative mapping and marker-assisted selection in *Rosaceae* fruit crops. *Proceedings of the National Academy of Sciences* **101**(26), 9891–9896.
- Dirlewanger E, Quero-García J, Le Dantec L, Lambert P, Ruiz D, Dondini L, Illa E, Quilot-Turion B, Audergon J-M, Tartarini S, Letourmy P, Arús P.** 2012. Comparison of the genetic determinism of two key phenological traits, flowering and maturity dates, in three *Prunus* species: peach, apricot and sweet cherry. *Heredity* **109**(5), 280–292.
- Dodds PN, Lawrence GJ, Catanzariti A-M, Teh T, Wang C-IA, Ayliffe MA, Kobe B, Ellis JG.** 2006. Direct protein interaction underlies gene-for-gene specificity and coevolution of the flax resistance genes and flax rust avirulence genes. *Proceedings of the National Academy of Sciences* **103**(23), 8888–8893.
- Doerge RW.** 2002. Mapping and analysis of quantitative trait loci in experimental populations. *Nature Reviews Genetics* **3**(1), 43–52.
- Dondini L, Lain O, Vendramin V, Rizzo M, Vivoli D, Adami M, Guidarelli M, Gaiotti F, Palmisano F, Bazzoni A, Boscia D, Geuna F, Tartarini S, Negri P, Castellano M, Savino V, Bassi D, Testolin R.** 2011. Identification of QTL for resistance to plum pox virus strains M and D in Lito and Harcot apricot cultivars. *Molecular Breeding* **27**(3), 289–299.
- Donmez MF, Karlidag H, Esitken A.** 2009. Identification of resistance to bacterial canker (*Pseudomonas syringae* pv. *syringae*) disease on apricot genotypes grown in Turkey. *European Journal of Plant Pathology* **126**(2), 241–247.
- Dunning FM, Sun W, Jansen KL, Helft L, Bent AF.** 2007. Identification and mutational analysis of *Arabidopsis* FLS2 leucine-rich repeat domain residues that contribute to flagellin perception. *The Plant Cell* **19**(10), 3297–3313.
- Eitas TK, Nimchuk ZL, Dangl JL.** 2008. *Arabidopsis* TAO1 is a TIR-NB-LRR protein that contributes to disease resistance induced by the *Pseudomonas syringae* effector AvrB. *Proceedings of the National Academy of Sciences* **105**(17), 6475–6480.

- El-Soda M, Kruijer W, Malosetti M, Koornneef M, Aarts Mark GM.** 2014. Quantitative trait loci and candidate genes underlying genotype by environment interaction in the response of *Arabidopsis thaliana* to drought. *Plant, Cell & Environment* **38**(3), 585–599.
- Ellis JG, Lagudah ES, Spielmeier W, Dodds PN.** 2014. The past, present and future of breeding rust resistant wheat. *Frontiers in Plant Science* **5**(641).
- English H, Davis JR.** 1960. The source of inoculum for bacterial canker and blast of stone fruit trees. *Phytopathology* **50**(9), 634 (Abstr.).
- Erikson D.** 1945. Certain aspects of resistance of plum trees to bacterial canker: Part II. On the nature of the bacterial invasion of *Prunus* sp. by *Pseudomonas mors-prunorum* Wormald. *Annals of Applied Biology* **32**(2), 112–117.
- Fabre F, Montarry J, Coville J, Senoussi R, Simon V, Moury B.** 2012a. Modelling the evolutionary dynamics of viruses within their hosts: A case study using high-throughput sequencing. *PLoS Pathogens* **8**(4), e1002654.
- Fabre F, Rousseau E, Mailleret L, Moury B.** 2012b. Durable strategies to deploy plant resistance in agricultural landscapes. *New Phytologist* **193**(4), 1064–1075.
- Fan S, Bielenberg DG, Zhebentyayeva TN, Reighard GL, Okie WR, Holland D, Abbott AG.** 2010. Mapping quantitative trait loci associated with chilling requirement, heat requirement and bloom date in peach (*Prunus persica*). *New Phytologist* **185**(4), 917–930.
- Fan X, Mattheis JP, Roberts RG.** 2000. Biosynthesis of phytoalexin in carrot root requires ethylene action. *Physiologia Plantarum* **110**(4), 450–454.
- FAO.** 2018. FAOSTAT - Crops.
- Farrer RA, Kemen E, Jones JDG, Studholme DJ.** 2009. De novo assembly of the *Pseudomonas syringae* pv. *syringae* B728a genome using Illumina/Solexa short sequence reads. *FEMS Microbiology Letters* **291**(1), 103–111.
- Faust M, Surányi D, Nyujtó F.** 1998. Origin and dissemination of apricot. In: Janick J, ed. *Horticultural Reviews*. Leuven: John Wiley & Sons, 225–266.
- Firrao G, Torelli E, Polano C, Ferrante P, Ferrini F, Martini M, Marcelletti S, Scortichini M, Ermacora P.** 2018. Genomic structural variations affecting virulence during clonal expansion of *Pseudomonas syringae* pv. *actinidiae* biovar 3 in Europe. *Frontiers in Microbiology* **9**(656).
- Fisher RA.** 1919. XV.—The correlation between relatives on the supposition of mendelian inheritance. *Transactions of the Royal Society of Edinburgh* **52**(2), 399–433.
- Flint-Garcia SA, Thornsberry JM, Buckler ES.** 2003. Structure of linkage disequilibrium in plants. *Annual Review of Plant Biology* **54**(1), 357–374.
- Flor HH.** 1955. Host-parasite interactions in flax rust: its genetics and other implications. *Phytopathology* **45**, 680–685.
- Flor HH.** 1971. Current status of the gene-for-gene concept. *Annual Review of Phytopathology* **9**, 275–296.
- Folta KM, Gardiner SE.** 2009. Genomics-based opportunities in apricot. In: Folta KM,, In: Gardiner SE, eds. *Genetics and Genomics of Rosaceae. Plant Genetics and Genomics: Crops and Models*. New York, NY: Springer-Verlag New York, 315–335.
- Fortuna MA, Stouffer DB, Olesen JM, Jordano P, Mouillot D, Krasnov BR,**

- Poulin R, Bascompte J.** 2010. Nestedness versus modularity in ecological networks: Two sides of the same coin? *Journal of Animal Ecology* **21**(2), 82–91.
- French E, Kim BS, Iyer-Pascuzzi AS.** 2016. Mechanisms of quantitative disease resistance in plants. *Seminars in Cell and Developmental Biology* **56**, 201–208.
- Fresnedo-Ramírez J, Frett TJ, Sandefur PJ, Salgado-Rojas A, Clark JR, Gasic K, Peace CP, Anderson N, Hartmann TP, Byrne DH, Bink MCAM, van de Weg EW, Crisosto CH, Gradziel TM.** 2016. QTL mapping and breeding value estimation through pedigree-based analysis of fruit size and weight in four diverse peach breeding programs. *Tree Genetics & Genomes* **12**(25).
- Fukuoka S, Yamamoto SI, Mizobuchi R, Yamanouchi U, Ono K, Kitazawa N, Yasuda N, Fujita Y, Thi Thanh Nguyen T, Koizumi S, Sugimoto K, Matsumoto T, Yano M.** 2014. Multiple functional polymorphisms in a single disease resistance gene in rice enhance durable resistance to blast. *Scientific Reports* **4**(4550).
- Gaignard J-L, Luisetti J.** 1993. *Pseudomonas syringae*, bactérie épiphyte, glaçogène et pathogène. *Agronomie* **13**(5), 333–370.
- Gallet R, Bonnot F, Milazzo J, Tertois C, Adreit H, Ravigné V, Tharreau D, Fournier E.** 2013. The variety mixture strategy assessed in a  $G \times G$  experiment with rice and the blast fungus *Magnaporthe oryzae*. *Frontiers in Genetics* **4**(312).
- García-Arenal F, McDonald BA.** 2003. An analysis of the durability of resistance to plant viruses. *Phytopathology* **93**(8), 941–952.
- Gardan L, Shafik H, Belouin S, Broch R, Grimont F, Grimont PAD.** 1999. DNA relatedness among the pathovars of *Pseudomonas syringae* and description of *Pseudomonas tremae* sp. nov. and *Pseudomonas cannabina* sp. nov. (ex Sutic and Dowson 1959). *International Journal of Systematic Bacteriology* **49**, 469–478.
- Gasic K, Gonzalez DO, Thimmapuram J, Liu L, Malnoy M, Gong G, Han Y, Vodkin LO, Aldwinckle HS, Carroll NJ, Orvis KS, Goldsbrough P, Clifton S, Pape D, Fulton L, Martin J, Theising B, Wisniewski ME, Fazio G, Feltus FA, Korban SS.** 2009. Comparative analysis and functional annotation of a large expressed sequence tag collection of apple. *The Plant Genome* **2**(1), 23–38.
- Gebhardt C, Valkonen JPT.** 2001. Organization of genes controlling disease resistance in the potato genome. *Annual Review of Phytopathology* **39**, 79–102.
- Gimenez-Ibanez S, Boter M, Fernández-Barbero G, Chini A, Rathjen JP, Solano R.** 2014. The bacterial effector HopX1 targets JAZ transcriptional repressors to activate jasmonate signaling and promote infection in *Arabidopsis*. *PLoS Biology* **12**, e1001792.
- Gimenez-Ibanez S, Hann DR, Ntoukakis V, Petutschnig E, Lipka V, Rathjen JP.** 2009. AvrPtoB targets the LysM receptor kinase CERK1 to promote bacterial virulence on plants. *Current Biology* **19**(5), 423–429.
- Glazebrook J.** 2005. Contrasting mechanisms of defense against biotrophic and necrotrophic pathogens. *Annual Review of Phytopathology* **43**, 205–227.
- Goel AK, Lundberg D, Torres MA, Matthews R, Akimoto-Tomiyama C, Farmer L, Dangl JL, Grant SR.** 2008. The *Pseudomonas syringae* type III effector HopAM1 enhances virulence on water-stressed plants. *Molecular Plant-Microbe Interactions* **21**(3), 361–370.
- Goffinet B, Gerber S.** 2000. Quantitative trait loci: a meta-analysis. *Genetics* **155**(1), 463–

- Göhre V, Spallek T, Häweker H, Mersmann S, Mentzel T, Boller T, de Torres M, Mansfield JW, Robatzek S.** 2008. Plant pattern-recognition receptor FLS2 is directed for degradation by the bacterial ubiquitin ligase AvrPtoB. *Current Biology* **18**(23), 1824–1832.
- Golzar H, Cother EJ.** 2008. First report of bacterial necrosis of mango caused by *Pseudomonas syringae* pv. *syringae* in Australia. *Australasian Plant Disease Notes* **3**(1), 107–109.
- Gozzo F, Faoro F.** 2013. Systemic acquired resistance (50 years after discovery): Moving from the lab to the field. *Journal of Agricultural and Food Chemistry* **61**(51), 12473–12491.
- Grattapaglia D, Sederoff R.** 1994. Genetic linkage maps of *Eucalyptus grandis* and *Eucalyptus urophylla* using a pseudo-testcross: mapping strategy and RAPD markers. *Genetics* **137**(4), 1121–1137.
- Green S, Studholme DJ, Laue BE, Dorati F, Lovell H, Arnold DL, Cottrell JE, Bridgett S, Blaxter M, Huitema E, Thwaites R, Sharp PM, Jackson RW, Kamoun S.** 2010. Comparative genome analysis provides insights into the evolution and adaptation of *Pseudomonas syringae* pv. *aesculion* *Aesculus hippocastanum*. *PLoS ONE* **5**(4), e10224.
- Greenberg JT, Vinatzer BA.** 2003. Identifying type III effectors of plant pathogens and analyzing their interaction with plant cells. *Current Opinion in Microbiology* **6**(1), 20–28.
- Guilbaud C, Morris CE, Barakat M, Ortet P, Berge O.** 2016. Isolation and identification of *Pseudomonas syringae* facilitated by a PCR targeting the whole *P. syringae* group. *FEMS Microbiology Ecology* **92**(1), fiv146.
- Guo M, Kim P, Li G, Elowsky CG, Alfano JR.** 2016. A bacterial effector co-opts calmodulin to target the plant microtubule network. *Cell Host and Microbe* **19**(1), 67–78.
- Gutiérrez-Barranquero JA, Cazorla FM, de Vicente A, Sundin GW.** 2017. Complete sequence and comparative genomic analysis of eight native *Pseudomonas syringae* plasmids belonging to the pPT23A family. *BMC Genomics* **18**(365).
- Haas BJ, Wortman JR, Ronning CM, Hannick LI, Smith RK, Maiti R, Chan AP, Yu C, Farzad M, Wu D, White O, Town CD.** 2005. Complete reannotation of the *Arabidopsis* genome: methods, tools, protocols and the final release. *BMC Biology* **3**(7).
- Hagen LS, Khadari B, Lambert P, Audergon J-M.** 2002. Genetic diversity in apricot revealed by AFLP markers: species and cultivar comparisons. *Theoretical and Applied Genetics* **105**(2–3), 298–305.
- Hammerschlag FA.** 2000. Resistant responses of peach somaclone 122-1 to *Xanthomonas campestris* pv. *pruni* and to *Pseudomonas syringae* pv. *syringae*. *HortScience* **35**(1), 141–143.
- Hammerschlag FA, Werner DJ, Ritchie DF.** 1994. Stability of bacterial leaf spot resistance in peach regenerants under in vitro, greenhouse and field conditions. *Euphytica* **76**(1), 101–106.
- Han HS, Koh YJ, Hur JS, Jung JS.** 2004. Occurrence of the *strA-strB* streptomycin

- resistance genes in *Pseudomonas* species isolated from kiwifruit plants. *The Journal of Microbiology* **42**(4), 365–368.
- Han HS, Nam HY, Koh YJ, Hur JS, Jung JS.** 2003. Molecular bases of high-level streptomycin resistance in *Pseudomonas marginalis* and *Pseudomonas syringae* pv. *actinidiae*. *The Journal of Microbiology* **41**(1), 16–21.
- Hattingh MJ, Roos IMM, Mansvelt EL.** 1989. Infection and systemic invasion of deciduous fruit-trees by *Pseudomonas syringae* in South-Africa. *Plant Disease* **73**(10), 784–789.
- Hauck P, Thilmony R, He SY.** 2003. A *Pseudomonas syringae* type III effector suppresses cell wall-based extracellular defense in susceptible *Arabidopsis* plants. *Proceedings of the National Academy of Sciences* **100**(14), 8577–8582.
- Henderson CR.** 1984. *Applications of linear models in animal breeding*. Guelph: University of Guelph.
- Hirano SS, Upper CD.** 1990. Population biology and epidemiology of *Pseudomonas syringae*. *Annual Review of Phytopathology* **28**(1), 155–177.
- Hockett KL, Nishimura MT, Karlsrud E, Dougherty K, Baltrus DA.** 2014. *Pseudomonas syringae* CC1557: a highly virulent strain with an unusually small type III effector repertoire that includes a novel effector. *Molecular Plant-Microbe Interactions* **27**(9), 923–932.
- van der Hoorn RAL, Kamoun S.** 2008. From guard to decoy: A new model for perception of plant pathogen effectors. *The Plant Cell* **20**(8), 2009–2017.
- Hormaza JI, Yamane H, Rodrigo J.** 2007. *Apricot. Fruits and Nuts*. Springer Berlin Heidelberg, 171–187.
- Howad W, Yamamoto T, Dirlewanger E, Testolin R, Cosson P, Cipriani G, Monforte AJ, Georgi L, Abbott AG, Arús P.** 2005. Mapping with a few plants: using selective mapping for microsatellite saturation of the *Prunus* reference map. *Genetics* **171**(3), 1305–1309.
- Huang X, Han B.** 2014. Natural variations and genome-wide association studies in crop plants. *Annual Review of Plant Biology* **65**(1), 531–551.
- Huang J, Yang M, Lu L, Zhang X.** 2016. Diverse functions of small RNAs in different plant-pathogen communications. *Frontiers in Microbiology* **7**(1552).
- Hulin MT, Armitage AD, Vicente JG, Holub EB, Baxter L, Bates HJ, Mansfield JW, Jackson RW, Harrison RJ.** 2018a. Comparative genomics of *Pseudomonas syringae* reveals convergent gene gain and loss associated with specialization onto cherry (*Prunus avium*). *New Phytologist* **219**(2), 672–696.
- Hulin MT, Mansfield JW, Brain P, Xu X, Jackson RW, Harrison RJ.** 2018b. Characterization of the pathogenicity of strains of *Pseudomonas syringae* towards cherry and plum. *Plant Pathology* **67**(5), 1177–1193.
- Hurley B, Lee D, Mott A, Wilton M, Liu J, Liu YC, Angers S, Coaker G, Guttman DS, Desveaux D.** 2014. The *Pseudomonas syringae* type III effector HopF2 suppresses *Arabidopsis* stomatal immunity. *PLoS ONE* **9**(12), e114921.
- Hurtado M, Romero C, Vilanova S, Abbott AG, Llácer G, Badenes ML.** 2002. Genetic linkage maps of two apricot cultivars (*Prunus armeniaca* L.), and mapping of PPV (sharka) resistance. *Theoretical and Applied Genetics* **105**(2), 182–191.

- Hutchison ML, Tester MA, Gross DC.** 1995. Role of biosurfactant and ion channel-forming activities of syringomycin in transmembrane ion flux: a model for the mechanism of action in the plant-pathogen interaction. *Molecular Plant-Microbe Interactions* **8**(4), 610–620.
- Illa E, Lambert P, Quilot-Turion B, Audergon J-M, Dirlewanger E, Howad W, Dondini L, Tartarini S, Lain O, Testolin R, Bassi D, Arús P.** 2009. Linkage map saturation, construction, and comparison in four populations of *Prunus*. *The Journal of Horticultural Science and Biotechnology* **84**(6), 168–175.
- Ingvarsson PK, Street NR.** 2011. Association genetics of complex traits in plants. *New Phytologist* **189**(4), 909–922.
- Interfel.** 2013. L'abricot en chiffres.
- Iwata H, Hayashi T, Terakami S, Takada N, Sawamura Y, Yamamoto T.** 2013. Potential assessment of genome-wide association study and genomic selection in Japanese pear *Pyrus pyrifolia*. *Breeding Science* **63**(1), 125–140.
- Jansen RC, Stam P.** 1994. High resolution of quantitative traits into multiple loci via interval mapping. *Genetics* **136**(4), 1447–1455.
- Jia Y, McAdams SA, Bryan GT, Hershey HP, Valent B.** 2000. Direct interaction of resistance gene and avirulence gene products confers rice blast resistance. *The EMBO journal* **19**(15), 4004–4014.
- Jiang S, Yao J, Ma KW, Zhou H, Song J, He SY, Ma W.** 2013. Bacterial effector activates jasmonate signaling by directly targeting JAZ transcriptional repressors. *PLoS Pathogens* **9**, e1003715.
- Jin Q, Thilmony R, Zwiesler-Vollick J, He SY.** 2003. Type III protein secretion in *Pseudomonas syringae*. *Microbes and Infection* **5**(4), 301–310.
- Joardar V, Lindeberg M, Jackson RW, Selengut J, Dodson R, Brinkac LM, Daugherty SC, DeBoy R, Durkin AS, Giglio MG, Madupu R, Nelson WC, Rosovitz MJ, Sullivan S, Crabtree J, Creasy T, Davidsen T, Haft DH, Zafar N, Zhou L, Halpin R, Holley T, Khouri H, Feldblyum T, White O, Fraser CM, Chatterjee AK, Cartinhour S, Schneider DJ, Mansfield J, Collmer A, Buell CR.** 2005. Whole-genome sequence analysis of *Pseudomonas syringae* pv. *phaseolicola* 1448A reveals divergence among pathovars in genes involved in virulence and transposition. *Journal of Bacteriology* **187**(18), 6488–6498.
- Johnson R.** 1979. Letter to the editor: The concept of durable resistance. *Phytopathology* **69**, 198–199.
- Johnson R, Law CN.** 1975. Genetic control of durable resistance to yellow rust (*Puccinia striiformis*) in the wheat cultivar Hybride de Bersée. *Annals of Applied Biology* **81**, 385–391.
- Jones JDG, Dangl JL.** 2006. The plant immune system. *Nature* **444**(7117), 323–329.
- Jones JDG, Vance RE, Dangl JL.** 2016. Intracellular innate immune surveillance devices in plants and animals. *Science* **354**(6316), aaf6395.
- Jourjon M-F, Jasson S, Marcel J, Ngom B, Mangin B.** 2005. MCQTL: multi-allelic QTL mapping in multi-cross design. *Bioinformatics* **21**(1), 128–130.
- Kajava A V, Lindow SE.** 1993. A model of the three-dimensional structure of ice nucleation proteins. *Journal of Molecular Biology* **232**(3), 709–717.

- Kamoun S, Huitema E, Vleeshouwers VGAA.** 1999. Resistance to oomycetes: A general role for the hypersensitive response? *Trends in Plant Science* **4**(5), 196–200.
- Kang HM, Sul JH, Service SK, Zaitlen NA, Kong S, Freimer NB, Sabatti C, Eskin E.** 2010. Variance component model to account for sample structure in genome-wide association studies. *Nature Genetics* **42**(4), 348–356.
- Kang HM, Zaitlen NA, Wade CM, Kirby A, Heckerman D, Daly MJ, Eskin E.** 2008. Efficient control of population structure in model organism association mapping. *Genetics* **178**(3), 1709–1723.
- Kao CH, Zeng ZB, Teasdale RD.** 1999. Multiple interval mapping for quantitative trait loci. *Genetics* **152**(3), 1203–1216.
- Katsir L, Schillmiller AL, Staswick PE, He SY, Howe GA.** 2008. COI1 is a critical component of a receptor for jasmonate and the bacterial virulence factor coronatine. *Proceedings of the National Academy of Sciences* **105**(19), 7100–7105.
- Keen NT.** 1990. Gene-for-gene complementarity in plant-pathogen interactions. *Annual Review of Genetics* **24**, 447–463.
- Kenis K, Keulemans J, Davey MW.** 2008. Identification and stability of QTLs for fruit quality traits in apple. *Tree Genetics & Genomes* **4**(4), 647–661.
- Kennelly MM, Cazorla FM, de Vicente A, Ramos C, Sundin GW.** 2007. *Pseudomonas syringae* diseases of fruit trees: progress toward understanding and control. *Plant Disease* **91**(1), 4–17.
- Kessens R, Ashfield T, Kim SH, Innes RW.** 2014. Determining the GmRIN4 requirements of the soybean disease resistance proteins Rpg1b and Rpg1r using a *Nicotiana glutinosa*-based agroinfiltration system. *PLoS ONE* **9**(9), e108159.
- Khan MA, Korban SS.** 2012. Association mapping in forest trees and fruit crops. *Journal of Experimental Botany* **63**(11), 4045–4060.
- Kim MG, Geng X, Lee SY, Mackey D.** 2009a. The *Pseudomonas syringae* type III effector AvrRpm1 induces significant defenses by activating the *Arabidopsis* nucleotide-binding leucine-rich repeat protein RPS2. *Plant Journal* **57**(4), 645–653.
- Kim SH, Kwon SI, Saha D, Anyanwu NC, Gassmann W.** 2009b. Resistance to the *Pseudomonas syringae* effector HopA1 is governed by the TIR-NBS-LRR protein RPS6 and is enhanced by mutations in SRFR1. *Plant Physiology* **150**(4), 1723–1732.
- Kim YJ, Lin NC, Martin GB.** 2002. Two distinct *Pseudomonas* effector proteins interact with the Pto kinase and activate plant immunity. *Cell* **109**(5), 589–598.
- Kim SH, Qi D, Ashfield T, Helm M, Innes RW.** 2016. Using decoys to expand the recognition specificity of a plant disease resistance protein. *Science* **351**(6274), 684–687.
- Klement Z, Rozsnyay DS, Arsenijevic M.** 1974. Apoplexy of apricots. II. Relationship of winter-frost and the bacterial canker and dieback of apricots. *Acta Phytopathologica Academiae Scientiarum Hungaricae* **9**, 35–45.
- Klement Z, Rozsnyay DS, Báló E, Pánczél M, Prileszky G.** 1984. The effect of cold on development of bacterial canker in apricot trees infected with *Pseudomonas syringae* pv. *syringae*. *Physiological Plant Pathology* **24**(2), 237–246.
- Kliebenstein DJ, Rowe HC, Denby KJ.** 2005. Secondary metabolites influence *Arabidopsis*/*Botrytis* interactions: variation in host production and pathogen sensitivity. *Plant Journal* **44**(1), 25–36.

- Korte A, Farlow A.** 2013. The advantages and limitations of trait analysis with GWAS: a review. *Plant Methods* **9**(29).
- Korte A, Vilhjálmsson BJ, Segura V, Platt A, Long Q, Nordborg M.** 2012. A mixed-model approach for genome-wide association studies of correlated traits in structured populations. *Nature Genetics* **44**(9), 1066–1071.
- Kostina KF.** 1969. The use of varietal resources of apricots for breeding. *Trud. Nikit. Bot. Sad* **40**, 45–63.
- Kourelis J, van der Hoorn RAL.** 2018. Defended to the nines: 25 years of resistance gene cloning identifies nine mechanisms for R protein function. *The Plant Cell* **30**(2), 285–299.
- Kover PX, Valdar W, Trakalo J, Scarcelli N, Ehrenreich IM, Purugganan MD, Durrant C, Mott R.** 2009. A multiparent advanced generation inter-cross to fine-map quantitative traits in *Arabidopsis thaliana*. *PLoS Genetics* **5**(7), e1000551.
- Kumar S, Garrick DJ, Bink MCAM, Whitworth C, Chagné D, Volz RK.** 2013. Novel genomic approaches unravel genetic architecture of complex traits in apple. *BMC Genomics* **14**(393).
- Laloi G, Vergne E, Durel CE, Le Cam B, Caffier V.** 2017. Efficiency of pyramiding of three quantitative resistance loci to apple scab. *Plant Pathology* **66**(3), 412–422.
- Lambert P, Dicenta F, Rubio M, Audergon J-M.** 2007. QTL analysis of resistance to sharka disease in the apricot (*Prunus armeniaca* L.) ‘Polonais’ x ‘Stark Early Orange’ F1 progeny. *Tree Genetics & Genomes* **3**(4), 299–309.
- Lambrechts L, Fellous S, Koella JC.** 2006. Coevolutionary interactions between host and parasite genotypes. *Trends in Parasitology* **22**(1), 12–16.
- Lamichhane JR, Varvaro L, Parisi L, Audergon J-M, Morris CE.** 2014. Disease and frost damage of woody plants caused by *Pseudomonas syringae*: seeing the forest for the trees. *Advances in Agronomy* **126**, 235–295.
- Lander ES, Botstein D.** 1989. Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* **121**(1), 185–199.
- Lannou C.** 2012. Variation and selection of quantitative traits in plant pathogens. *Annual Review of Phytopathology* **50**, 319–338.
- Laurens F, Aranzana MJ, Arus P, Bassi D, Bink M, Bonany J, Caprera A, Corelli-Grappadelli L, Costes E, Durel C-E, Mauroux J-B, Muranty H, Nazzicari N, Pascal T, Patocchi A, Peil A, Quilot-Turion B, Rossini L, Stella A, Troggio M, Velasco R, van de Weg EW.** 2018. An integrated approach for increasing breeding efficiency in apple and peach in Europe. *Horticulture Research* **5**(11).
- Layne REC.** 1981. Harlayne apricot. *HortScience* **16**, 97–98.
- Leach JE, Vera Cruz CM, Bai J, Leung H.** 2001. Pathogen fitness penalty as a predictor of durability of disease resistance genes. *Annual Review of Phytopathology* **39**, 187–224.
- Lee D, Bourdais G, Yu G, Robotzek S, Coaker G.** 2015. Phosphorylation of the plant immune regulator RPM1-INTERACTING PROTEIN4 enhances plant plasma membrane H<sup>+</sup>-ATPase activity and inhibits flagellin-triggered immune responses in *Arabidopsis*. *The Plant Cell* **27**(7), 2042–2056.
- Lespinnasse Y, Pinet C, Laurens F, Durel C-E, Parisi L.** 2002. European research for durable resistance to scab on apple: The D.A.R.E. project. *Acta Horticulturae* **595**, 17–

- Lewis JD, Lee AH-Y, Hassan JA, Wan J, Hurley B, Jhingree JR, Wang PW, Lo T, Youn J-Y, Guttman DS, Desveaux D.** 2013. The *Arabidopsis* ZED1 pseudokinase is required for ZAR1-mediated immunity induced by the *Pseudomonas syringae* type III effector HopZ1a. *Proceedings of the National Academy of Sciences* **110**(46), 18722–18727.
- Li ZK, Arif M, Zhong DB, Fu BY, Xu JL, Domingo-Rey J, Ali J, Vijayakumar CHM, Yu SB, Khush GS.** 2006. Complex genetic networks underlying the defensive system of rice (*Oryza sativa* L.) to *Xanthomonas oryzae* pv. *oryzae*. *Proceedings of the National Academy of Sciences* **103**(21), 7994–7999.
- Li B, Hulin MT, Brain P, Mansfield JW, Jackson RW, Harrison RJ.** 2015. Rapid, automated detection of stem canker symptoms in woody perennials using artificial neural network analysis. *Plant Methods* **11**(57).
- Li ZK, Luo LJ, Mei HW, Paterson AH, Zhao XH, Zhong DB, Wang YP, Yu XQ, Zhu L, Tabien R, Stansel JW, Ying CS.** 1999. A ‘defeated’ rice resistance gene acts as a QTL against a virulent strain of *Xanthomonas oryzae* pv. *oryzae*. *Molecular and General Genetics* **261**(1), 58–63.
- Li B, Xu X.** 2002. Infection and development of apple scab (*Venturia inaequalis*) on old leaves. *Journal of Phytopathology* **150**(11–12), 687–691.
- Lichou J, Jay M.** 2012. *Abricot*. CTIFL.
- Lin NC, Abramovitch RB, Kim YJ, Martin GB.** 2006. Diverse AvrPtoB homologs from several *Pseudomonas syringae* pathovars elicit Pto-dependent resistance and have similar virulence activities. *Applied and Environmental Microbiology* **72**(1), 702–712.
- Lindeberg M, Cartinhour S, Myers CR, Schechter LM, Schneider DJ, Collmer A.** 2006. Closing the circle on the discovery of genes encoding hrp regulon members and type III secretion system effectors in the genomes of three model *Pseudomonas syringae* strains. *Molecular Plant-Microbe Interactions* **19**(11), 1151–1158.
- Lindeberg M, Cunnac S, Collmer A.** 2012. *Pseudomonas syringae* type III effector repertoires: last words in endless arguments. *Trends in Microbiology* **20**(4), 199–208.
- Lindgren PB, Peet RC, Panopoulos NJ.** 1986. Gene cluster of *Pseudomonas syringae* pv. ‘*phaseolicola*’ controls pathogenicity of bean plants and hypersensitivity on nonhost plants. *Journal of Bacteriology* **168**(2), 512–522.
- Lindow SE.** 1982. Population dynamics of epiphytic ice nucleation active bacteria on frost sensitive plants and frost control by means of antagonistic bacteria. In: Li PH., In: Sakai A, eds. *Plant Cold Hardiness and Freezing Stress. Mechanisms and Crop Implications*. Academic Press, 395–416.
- Lindow SE.** 1987. Competitive exclusion of epiphytic bacteria by ice(–)*Pseudomonas syringae* mutants. *Applied and Environmental Microbiology* **53**(10), 2520–2527.
- Loake G, Grant M.** 2007. Salicylic acid in plant defence—the players and protagonists. *Current Opinion in Plant Biology* **10**(5), 466–472.
- Lozano-Durán R, Bourdais G, He SY, Robatzek S.** 2014. The bacterial effector HopM1 suppresses PAMP-triggered oxidative burst and stomatal immunity. *New Phytologist* **202**(1), 259–269.
- Lu D, Wu S, Gao X, Zhang Y, Shan L, He P.** 2010. A receptor-like cytoplasmic kinase,

- BIK1, associates with a flagellin receptor complex to initiate plant innate immunity. *Proceedings of the National Academy of Sciences* **107**(1), 496–501.
- Luisetti J, Gaignard J-L, Vigouroux A, Saunier R, Lafuste J-P, Charras J.** 1992. Pêcher : le déperissement bacterien. *Arboriculture Fruitière* **447**, 20–28.
- Luo X, Xu N, Huang J, Gao F, Zou H, Boudsocq M, Coaker GL, Liu J.** 2017. A lectin receptor-like kinase mediates pattern-triggered salicylic acid signaling. *Plant Physiology* **174**(4), 2501–2514.
- Lynch M, Walsh B.** 1998. *Genetics and analysis of quantitative traits*. Sinauer.
- Mackay TFC.** 2001. The genetic architecture of quantitative traits. *Annual Review of Genetics* **35**(1), 303–339.
- Mackey D, Holt BF, Wiig A, Dangl JL.** 2002. RIN4 interacts with *Pseudomonas syringae* type III effector molecules and is required for RPM1-mediated resistance in *Arabidopsis*. *Cell* **108**(6), 743–754.
- Maghuly F, Borroto Fernandez E, Zelger R, Marschall K, Katinger H, Laimer M.** 2006. Genetic differentiation of apricot (*Prunus armeniaca* L.) cultivars with SSR markers. *European Journal of Horticultural Science* **71**(3), 129–134.
- Maki LR, Galyan EL, Chang-Chien MM, Caldwell DR.** 1974. Ice nucleation induced by *Pseudomonas syringae*. *Applied Microbiology* **28**(3), 456–459.
- Mansfield JW, Genin S, Magori S, Citovsky V, Sriariyanum M, Ronald P, Dow M, Verdier V, Beer S V, Machado MA, Toth I, Salmond G, Foster GD.** 2012. Top 10 plant pathogenic bacteria in molecular plant pathology. *Molecular Plant Pathology* **13**(6), 614–629.
- Marandel G, Salava J, Abbott AG, Candresse T, Decroocq V.** 2009. Quantitative trait loci meta-analysis of Plum pox virus resistance in apricot (*Prunus armeniaca* L.): new insights on the organization and the identification of genomic resistance factors. *Molecular Plant Pathology* **10**(3), 347–360.
- Marcel TC, Gorguet B, Ta MT, Kohutova Z, Vels A, Niks RE.** 2008. Isolate specificity of quantitative trait loci for partial resistance of barley to *Puccinia hordei* confirmed in mapping populations and near-isogenic lines. *New Phytologist* **177**(3), 743–755.
- Marcelletti S, Ferrante P, Petriccione M, Firrao G, Scortichini M.** 2011. *Pseudomonas syringae* pv. *actinidiae* draft genomes comparison reveal strain-specific features involved in adaptation and virulence to Actinidia species. *PLoS ONE* **6**(11), e27297.
- Mariette S, Wong Jun Tai F, Roch G, Barre A, Chague A, Decroocq S, Groppi A, Laizet Y, Lambert P, Tricon D, Nikolski M, Audergon J-M, Abbott AG, Decroocq V.** 2016. Genome-wide association links candidate genes to resistance to Plum Pox Virus in apricot (*Prunus armeniaca*). *New Phytologist* **209**(2), 773–784.
- Martínez-Gómez P, Dicenta F, Audergon J-M.** 2000. Behaviour of apricot (*Prunus armeniaca* L.) cultivars in the presence of sharka (plum pox potyvirus): a review. *Agronomie* **20**, 407–422.
- McClure KA, Gardner KM, Douglas GM, Song J, Forney CF, DeLong J, Fan L, Du L, Toivonen PMA, Somers DJ, Rajcan I, Myles S.** 2018. A genome-wide association study of apple quality and scab resistance. *The Plant Genome* **11**(1).

- McDonald BA, Linde C.** 2002. Pathogen population genetics, evolutionary potential, and durable resistance. *Annual Review of Phytopathology* **40**, 348–379.
- Mehlenbacher SA, Cociu V, Hough FL.** 1991. Apricots (*Prunus*). *Acta Horticulturae* **290**, 65–110.
- Melotto M, Underwood W, Koczan J, Nomura K, He SY.** 2006. Plant stomata function in innate immunity against bacterial invasion. *Cell* **126**(5), 969–980.
- Mérot-L’Anthoëne V, Mangin B, Lefebvre-Pautigny F, Jasson S, Rigoreau M, Husson J, Lambot C, Cruzillat D.** 2014. Comparison of three QTL detection models on biochemical, sensory, and yield characters in *Coffea canephora*. *Tree Genetics and Genomes* **10**(6), 1541–1553.
- Meuwissen THE, Hayes BJ, Goddard ME.** 2001. Prediction of total genetic value using genome-wide dense marker maps. *Genetics* **157**(4), 1819–1829.
- Mgbechi-Ezeri J, Porter L, Johnson KB, Oraguzie NC.** 2017. Assessment of sweet cherry (*Prunus avium* L.) genotypes for response to bacterial canker disease. *Euphytica* **213**(145).
- Micheletti D, Dettori MT, Micali S, Aramini V, Pacheco I, Da Silva Linge C, Foschi S, Banchi E, Barreneche T, Quilot-Turion B, Lambert P, Pascal T, Iglesias I, Carbó J, Wang L, Ma R, Li X, Gao Z, Nazzicari N, Troglio M, Bassi D, Rossini L, Verde I, Laurens F, Arús P, Aranzana MJ.** 2015. Whole-genome analysis of diversity and SNP-major gene association in peach germplasm. *PLoS ONE* **10**(9), e0136803.
- Miklas PN, Kelly JD, Beebe SE, Blair W.** 2006. Common bean breeding for resistance against biotic and abiotic stresses: From classical to MAS breeding. *Euphytica* **147**(1–2), 105–131.
- Minamikawa MF, Nonaka K, Kaminuma E, Kajiya-Kanegae H, Onogi A, Goto S, Yoshioka T, Imai A, Hamada H, Hayashi T, Matsumoto S, Katayose Y, Toyoda A, Fujiyama A, Nakamura Y, Shimizu T, Iwata H.** 2017. Genome-wide association study and genomic prediction in citrus: potential of genomics-assisted breeding for fruit quality traits. *Scientific Reports* **7**(4721).
- Mine A, Berens ML, Nobori T, Anver S, Fukumoto K, Winkelmüller TM, Takeda A, Becker D, Tsuda K.** 2017. Pathogen exploitation of an abscisic acid- and jasmonate-inducible MAPK phosphatase and its interception by *Arabidopsis* immunity. *Proceedings of the National Academy of Sciences* **114**(28), 7456–7461.
- Ministère de l’Agriculture et de l’Alimentation.** 2018. Agreste, Infos rapides Abricot.
- Mittal S, Davis KR.** 1995. Role of the phytotoxin coronatine in the infection of *Arabidopsis thaliana* by *Pseudomonas syringae* pv. *tomato*. *Molecular Plant-Microbe Interactions* **8**(1), 165–171.
- Monaghan J, Zipfel C.** 2012. Plant pattern recognition receptor complexes at the plasma membrane. *Current Opinion in Plant Biology* **15**(4), 349–357.
- Montanari S, Perchepped L, Renault D, Frijters L, Velasco R, Horner M, Gardiner SE, Chagné D, Bus VGM, Durel C-E, Malnoy M.** 2016. A QTL detected in an interspecific pear population confers stable fire blight resistance across different environments and genetic backgrounds. *Molecular Breeding* **36**(47).
- Monteil CL, Cai R, Liu H, Mechan Llontop ME, Leman S, Studholme DJ, Morris**

- CE, Vinatzer BA.** 2013. Nonagricultural reservoirs contribute to emergence and evolution of *Pseudomonas syringae* crop pathogens. *New Phytologist* **199**(3), 800–811.
- Monteil CL, Guilbaud C, Glaux C, Lafolie F, Soubeyrand S, Morris CE.** 2012. Emigration of the plant pathogen *Pseudomonas syringae* from leaf litter contributes to its population dynamics in alpine snowpack. *Environmental Microbiology* **14**(8), 2099–2112.
- Monteil CL, Yahara K, Studholme DJ, Mageiros L, Méric G, Swingle B, Morris CE, Vinatzer BA, Sheppard SK.** 2016. Population-genomic insights into emergence, crop adaptation and dissemination of *Pseudomonas syringae* pathogens. *Microbial Genomics* **2**(10).
- Montesinos E, Vilardell P.** 1991. Relationships among population-levels of *Pseudomonas syringae*, amount of ice nuclei, and incidence of blast of dormant flower buds in commercial pear orchards in Catalunya, Spain. *Phytopathology* **81**, 113–119.
- Morris CE, Glaux C, Latour X, Gardan L, Samson R, Pitrat M.** 2000. The relationship of host range, physiology, and genotype to virulence on Cantaloupe in *Pseudomonas syringae* from Cantaloupe blight epidemics in France. *Phytopathology* **90**(6), 636–646.
- Morris CE, Kinkel LL, Xiao K, Prior P, Sands DC.** 2007. Surprising niche for the plant pathogen *Pseudomonas syringae*. *Infection, Genetics and Evolution* **7**(1), 84–92.
- Morris CE, Monteil CL, Berge O.** 2013. The life history of *Pseudomonas syringae*: linking agriculture to Earth system processes. *Annual Review of Phytopathology* **51**(1), 85–104.
- Morris CE, Sands DC, Vinatzer BA, Glaux C, Guilbaud C, Buffière A, Yan S, Dominguez H, Thompson BM.** 2008. The life history of the plant pathogen *Pseudomonas syringae* is linked to the water cycle. *The ISME Journal* **2**(3), 321–334.
- Morvan G.** 1977. Apricot chlorotic leaf roll. *EPPO Bulletin* **7**(1), 37–55.
- Moury B, Fabre F, Hébrard E, Froissart R.** 2017. Determinants of host species range in plant viruses. *Journal of General Virology* **98**(4), 862–873.
- Moury B, Janzac B, Ruellan Y, Simon V, Ben Khalifa M, Fakhfakh H, Fabre F, Palloix A.** 2014. Interaction patterns between potato virus Y and eIF4E-mediated recessive resistance in the *Solanaceae*. *Journal of virology* **88**(17), 9799–9807.
- Moury B, Simon V.** 2011. dN/dS-based methods detect positive selection linked to trade-offs between different fitness traits in the coat protein of potato virus Y. *Molecular Biology and Evolution* **28**(9), 2707–2717.
- Mulet M, Lalucat J, García-Valdés E.** 2010. DNA sequence-based analysis of the *Pseudomonas* species. *Environmental Microbiology* **12**(6), 1513–1530.
- Mundt CC.** 2014. Durable resistance: A key to sustainable management of pathogens and pests. *Infection, Genetics and Evolution* **27**, 446–455.
- Muranty H, Troglio M, Ben Sadok I, Al Rifai M, Auwerkerken A, Banchi E, Velasco R, Stevanato P, van de Weg EW, di Guardo M, Kumar S, Laurens F, Bink MCAM.** 2015. Accuracy and responses of genomic selection on key traits in apple breeding. *Horticulture Research* **2**(15060).
- Myles S, Chia J-M, Hurwitz B, Simon C, Zhong GY, Buckler ES, Ware D.** 2010. Rapid genomic characterization of the genus *Vitis*. *PLoS ONE* **5**(1), e8219.
- Myles S, Peiffer J, Brown PJ, Ersoz ES, Zhang Z, Costich DE, Buckler ES.** 2009.

- Association mapping: critical considerations shift from genotyping to experimental design. *The Plant Cell* **21**(8), 2194–2202.
- Nakajima M, Yamashita S, Takikawa Y, Shinji T, Hibi T, Goto M.** 1995. Similarity of streptomycin resistance gene(s) in *Pseudomonas syringae* pv. *actinidiae* with *strA* and *strB* of plasmid RSF1010. *Japanese Journal of Phytopathology* **61**(5), 489–492.
- Nakashita H, Yasuda M, Nitta T, Asami T, Fujioka S, Arai Y, Sekimata K, Takatsuto S, Yamaguchi I, Yoshida S.** 2003. Brassinosteroid functions in a broad range of disease resistance in tobacco and rice. *Plant Journal* **33**(5), 887–898.
- Narusaka M, Shirasu K, Noutoshi Y, Kubo Y, Shiraishi T, Iwabuchi M, Narusaka Y.** 2009. RRS1 and RPS4 provide a dual resistance-gene system against fungal and bacterial pathogens. *Plant Journal* **60**(2), 218–226.
- Nass HA, Pedersen WL, Mackenzie DR, Nelson RR.** 1981. The residual effects of some defeated powdery mildew *Erysiphe-graminis-f-sp-tritici* resistance genes in isolines of winter wheat. *Phytopathology* **71**(12), 1315–1318.
- Navarro L, Bari R, Achard P, Lisón P, Nemri A, Harberd NP, Jones JDG.** 2008. DELLAs control plant immune responses by modulating the balance of jasmonic acid and salicylic acid signaling. *Current Biology* **18**(9), 650–655.
- Newman M-A, Dow JM, Molinaro A, Parrilli M.** 2007. Priming, induction and modulation of plant defence responses by bacterial lipopolysaccharides. *Journal of Endotoxin Research* **13**(2), 69–84.
- Newman M-A, Sundelin T, Nielsen JT, Erbs G.** 2013. MAMP (microbe-associated molecular pattern) triggered immunity in plants. *Frontiers in Plant Science* **4**(139).
- Niks RE, Qi X, Marcel TC.** 2015. Quantitative resistance to biotrophic filamentous plant pathogens: concepts, misconceptions, and mechanisms. *Annual Review of Phytopathology* **53**, 445–470.
- Nishimura MT, Anderson RG, Cherkis KA, Law TF, Liu QL, Machius M, Nimchuk ZL, Yang L, Chung E-H, El Kasmi F, Hyunh M, Osborne Nishimura E, Sondak JE, Dangl JL.** 2017. TIR-only protein RBA1 recognizes a pathogen effector to regulate cell death in *Arabidopsis*. *Proceedings of the National Academy of Sciences* **114**(10), E2053–E2062.
- Nordborg M, Donnelly P.** 1997. The coalescent process with selfing. *Genetics* **146**(3), 1185–1195.
- Nordborg M, Weigel D.** 2008. Next-generation genetics in plants. *Nature* **456**, 720–723.
- Nowell R, Laue BE, Sharp PM, Green S.** 2016. Comparative genomics reveals genes significantly associated with woody hosts in the plant pathogen *Pseudomonas syringae*. *Molecular Plant Pathology* **17**(9), 1409–1424.
- Ntoukakis V, Mucyn TS, Gimenez-Ibanez S, Chapman HC, Gutierrez JR, Balmuth AL, Jones AME, Rathjen JP.** 2009. Host inhibition of a bacterial virulence effector triggers immunity to infection. *Science* **324**(5928), 784–787.
- Nürnberg T, Brunner F, Kemmerling B, Piater L.** 2004. Innate immunity in plants and animals: striking similarities and obvious differences. *Immunological Reviews* **198**, 249–266.
- O'Brien HE, Thakur S, Guttman DS.** 2011. Evolution of plant pathogenesis in *Pseudomonas syringae*: a genomics perspective. *Annual Review of Phytopathology* **49**,

269–289.

- Ogura T, Busch W.** 2015. From phenotypes to causal sequences: using genome wide association studies to dissect the sequence basis for variation of plant development. *Current Opinion in Plant Biology* **23**, 98–108.
- Ogut F, Bian Y, Bradbury PJ, Holland JB.** 2015. Joint-multiple family linkage analysis predicts within-family variation better than single-family analysis of the maize nested association mapping population. *Heredity* **114**, 552–563.
- Olson BD, Jones AL.** 1983. Reduction of *Pseudomonas syringae* pv. *morsprunorum* on Montmorency sour cherry with copper and dynamics of the copper residues. *Phytopathology* **73**(11), 1520–1525.
- Olukolu BA, Trainin T, Fan S, Kole C, Bielenberg DG, Reighard GL, Abbott AG, Holland D.** 2009. Genetic linkage mapping for molecular dissection of chilling requirement and budbreak in apricot (*Prunus armeniaca* L.). *Genome* **52**(10), 819–828.
- Palloix A, Ayme V, Moury B.** 2009. Durability of plant major resistance genes to pathogens depends on the genetic background, experimental evidence and consequences for breeding strategies. *New Phytologist* **183**(1), 190–199.
- Pan Q, Cui B, Deng F, Quan J, Loake GJ, Shan W.** 2016. RTP1 encodes a novel endoplasmic reticulum (ER)-localized protein in *Arabidopsis* and negatively regulates resistance against biotrophic pathogens. *New Phytologist* **209**(4), 1641–1654.
- Papaïx J, Goyeau H, Du Cheyron P, Monod H, Lannou C.** 2011. Influence of cultivated landscape composition on variety resistance: An assessment based on wheat leaf rust epidemics. *New Phytologist* **191**(4), 1095–1107.
- Parisi L, Lespinasse Y, Guillaumes J, Krüger J.** 1993. A new race of *Venturia inaequalis* virulent to apples with resistance due to the *Vf* gene. *Phytopathology* **83**(5), 533–537.
- Parisi L, Morgaint B, Borschinger B, Guilbaud C, Chandeysson C, Combe F, Bourgeay J-F, Berge O, Brun L, Brachet M-L, Morris CE.** 2016. Diversité et structuration des populations de *Pseudomonas syringae* en vergers d'abricotier. 12èmes Rencontres Plantes-Bactéries. Aussois, France, .
- Parkinson N, Bryant R, Bew J, Elphinstone J.** 2010. Rapid phylogenetic identification of members of the *Pseudomonas syringae* species complex using the *rpoD* locus. *Plant Pathology* **60**(2), 338–344.
- Pascual L, Desplat N, Huang Bevan E, Desgroux A, Bruguier L, Bouchet J-P, Le Quang H, Chauchard B, Verschave P, Causse M.** 2015. Potential of a tomato MAGIC population to decipher the genetic control of quantitative traits and detect causal variants in the resequencing era. *Plant Biotechnology Journal* **13**(4), 565–577.
- Paulin J-P, Luisetti J.** 1978. Ice nucleation activity among phytopathogenic bacteria. *Proceedings of the Fourth International Conference on Plant Pathogenic Bacteria*, Angers, France. 725–731.
- Peart JR, Cook G, Feys BJ, Parker JE, Baulcombe DC.** 2002. An *EDS1* orthologue is required for N-mediated resistance against tobacco mosaic virus. *Plant Journal* **29**(5), 569–579.
- Petersen M, Brodersen P, Naested H, Andreasson E, Lindhart U, Johansen B, Nielsen HB, Lacy M, Austin MJ, Parker JE, Sharma SB, Klessig DF, Martienssen R, Mattsson O, Jensen AB, Mundy J.** 2000. *Arabidopsis* MAP kinase

- 4 negatively regulates systemic acquired resistance. *Cell* **103**(7), 1111–1120.
- Picañol R, Eduardo I, Aranzana MJ, Howad W, Batlle I, Iglesias I, Alonso JM, Arús P.** 2013. Combining linkage and association mapping to search for markers linked to the flat fruit character in peach. *Euphytica* **190**(2), 279–288.
- Pieterse CM, Leon-Reyes A, van der Ent S, van Wees SCM.** 2009. Networking by small-molecule hormones in plant immunity. *Nature Chemical Biology* **5**(5), 308–316.
- Pilet-Nayel M-L, Moury B, Caffier V, Montarry J, Kerlan M-C, Fournet S, Durel C-E, Delourme R.** 2017. Quantitative resistance to plant pathogens in pyramiding strategies for durable crop protection. *Frontiers in Plant Science* **8**(1838).
- Poland JA, Balint-Kurti PJ, Wisser RJ, Pratt RC, Nelson RJ.** 2009. Shades of gray: the world of quantitative disease resistance. *Trends in Plant Science* **14**(1), 21–29.
- Porta H.** 2002. Plant lipoxygenases. Physiological and molecular features. *Plant Physiology* **130**(1), 15–21.
- Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D.** 2006. Principal components analysis corrects for stratification in genome-wide association studies. *Nature Genetics* **38**(8), 904–909.
- Pritchard L, Birch PRJ.** 2014. The zigzag model of plant-microbe interactions: Is it time to move on? *Molecular Plant Pathology* **15**(9), 865–870.
- Pritchard JK, Stephens M, Donnelly P.** 2000. Inference of population structure using multilocus genotype data. *Genetics* **155**(2), 945–959.
- Prunier J-P, Audergon J-M, Cotta J.** 1993. Etude de la sensibilité variétale de l'abricotier au chancre bactérien. Programme de recherche Agrimed: Deuxiemes rencontres sur l'abricotier. Recueil des communications. Luxembourg, 111–112.
- Prunier J-P, Bordjiba O.** 1991. Effect of frost on bacterial necrosis of apricot buds. *Acta Horticulturae* **293**, 495–502.
- Prunier J-P, Cotta J.** 1987. Le dépérissement bacterien (*Pseudomonas* spp.) de l'abricotier : situation en France. *Bulletin OEPP* **17**(2), 251–256.
- Prunier J-P, Jullian J-P, Audergon J-M.** 1999. Influence of rootstock and the height of grafting on the susceptibility of apricot cultivars to bacterial canker. *Acta Horticulturae* **488**, 643–648.
- Prunier J-P, Jullian J-P, Minodier R, Clauzel G, Martins J.** 2005. Chancre bactérien sur abricotier. Eviter les dégâts par la mise en oeuvre de pratiques simples et raisonnées. *Arboriculture Fruitière* **595**, 23–30.
- Quenouille J, Paulhiac E, Moury B, Palloix A.** 2014. Quantitative trait loci from the host genetic background modulate the durability of a resistance gene: A rational basis for sustainable resistance breeding in plants. *Heredity* **112**, 579–587.
- Rehder A.** 1949. *Bibliography of cultivated trees and shrubs hardy in the cooler temperate regions of the northern hemisphere.*
- Rikkerink EHA, Oraguzie NC, Gardiner SE.** 2007. Prospects of association mapping in perennial horticultural crops. In: Oraguzie NC, In: Rikkerink EHA, In: Gardiner SE, In: De Silva HN, eds. *Association Mapping in Plants*. New York, NY: Springer New York, 249–269.
- Robert-Seilaniantz A, Grant M, Jones JDG.** 2011. Hormone crosstalk in plant disease and defense: More than just JASMONATE-SALICYLATE antagonism. *Annual Review*

of Phytopathology **49**, 317–343.

- Rocherieux J, Glory P, Giboulot A, Boury S, Barbeyron G, Thomas G, Manzanares-Dauleux MJ.** 2004. Isolate-specific and broad-spectrum QTLs are involved in the control of clubroot in *Brassica oleracea*. *Theoretical and Applied Genetics* **108**(8), 1555–1563.
- Rodríguez-Herva JJ, González-Melendi P, Cuartas-Lanza R, Antúnez-Lamas M, Río-Alvarez I, Li Z, López-Torrejón G, Díaz I, Del Pozo JC, Chakravarthy S, Collmer A, Rodríguez-Palenzuela P, López-Solanilla E.** 2012. A bacterial cysteine protease effector protein interferes with photosynthesis to suppress plant innate immune responses. *Cellular Microbiology* **14**(5), 669–681.
- Ronin Y, Korol A, Shtemberg M, Nevo E, Soller M.** 2003. High-resolution mapping of quantitative trait loci by selective recombinant genotyping. *Genetics* **164**(4), 1657–1666.
- Roos IMM, Hattingh MJ.** 1986. Resident populations of *Pseudomonas syringae* on stone fruit tree leaves in South Africa. *Phytophylactica* **18**(2), 55–58.
- Roos IMM, Hattingh MJ.** 1987. Pathogenicity and numerical-analysis of phenotypic features of *Pseudomonas syringae* strains isolated from deciduous fruit-trees. *Phytopathology* **77**(6), 900–908.
- Rosebrock TR, Zeng L, Brady JJ, Abramovitch RB, Xiao F, Martin GB.** 2007. A bacterial E3 ubiquitin ligase targets a host protein kinase to disrupt plant immunity. *Nature* **448**(7151), 370–374.
- Rosyara UR, Bink MCAM, van de Weg EW, Zhang G, Wang D, Sebolt A, Dirlewanger E, Quero-Garcia J, Schuster M, Iezzoni AF.** 2013. Fruit size QTL identification and the prediction of parental QTL genotypes and breeding values in multiple pedigreed populations of sweet cherry. *Molecular Breeding* **32**(4), 875–887.
- Roth C, Lüdke D, Klenke M, Quathammer A, Valerius O, Braus GH, Wiermer M.** 2017. The truncated NLR protein TIR-NBS13 is a MOS6/IMPORTIN- $\alpha$ 3 interaction partner required for plant immunity. *Plant Journal* **92**(5), 808–821.
- Rouse DI, Nordheim E V, Hirano SS, Upper CD.** 1985. A model relating the probability of foliar disease incidence to the population frequencies of bacterial plant pathogens. *Phytopathology* **75**(5), 505–509.
- Rousseau E, Moury B, Mailleret L, Senoussi R, Palloix A, Simon V, Valière S, Grognaud F, Fabre F.** 2017. Estimating virus effective population size and selection without neutral markers. *PLoS Pathogens* **13**(11), e1006702.
- Roussos PA, Denaxa N-K, Tsafouros A, Efstathios N, Intidhar B.** 2016. Apricot (*Prunus armeniaca* L.). In: Simmonds MSJ, In: Preedy VR, eds. *Nutritional Composition of Fruit Cultivars*. San Diego: Academic Press, 19–48.
- Rubio M, Ballester AR, Olivares PM, De Moura MC, Dicenta F, Martínez-Gómez P.** 2015. Gene expression analysis of plum pox virus (sharka) susceptibility/resistance in apricot (*Prunus armeniaca* L.). *PLoS ONE* **10**(12), e0144670.
- Ruiz D, Campoy JA, Egea J.** 2007. Chilling and heat requirements of apricot cultivars for flowering. *Environmental and Experimental Botany* **61**(3), 254–263.
- Ruiz D, Lambert P, Audergon J-M, Dondini L, Tartarini S, Adami M, Gennari F, Cervellati C, De Franceschi P, Sansavini S, Testolin R, Bureau S, Gouble**

- B, Reich M, Renard CMGC, Bassi D.** 2010. Identification of QTLs for fruit quality traits in apricot. *Acta Horticulturae* **862**, 587–592.
- Ryals JA, Neuenschwander UH, Willits MG, Molina A, Steiner HY, Hunt MD.** 1996. Systemic acquired resistance. *The Plant cell* **8**(10), 1809–1819.
- Salazar JA, Ruiz D, Campoy JA, Tartarini S, Dondini L, Martínez-Gómez P.** 2016. Inheritance of reproductive phenology traits and related QTL identification in apricot. *Tree Genetics & Genomes* **12**(71).
- Salmeron JM, Barker SJ, Carland FM, Mehta AY, Staskawicz BJ.** 1994. Tomato mutants altered in bacterial disease resistance provide evidence for a new locus controlling pathogen recognition. *Plant Cell* **6**(4), 511–520.
- Salzman RA, Tikhonova I, Bordelon BP, Hasegawa PM, Bressan RA.** 1998. Coordinate accumulation of antifungal proteins and hexoses constitutes a developmentally controlled defense response during fruit ripening in grape. *Plant physiology* **117**(2), 465–472.
- Santi F, Russell K, Ménard M, Dufour J.** 2004. Screening wild cherry (*Prunus avium*) for resistance to bacterial canker by laboratory and field tests. *Forest Pathology* **34**(6), 349–362.
- Sapoukhina N, Durel CE, Le Cam B.** 2009. Spatial deployment of gene-for-gene resistance governs evolution and spread of pathogen populations. *Theoretical Ecology* **2**(4), 229–238.
- Sapoukhina N, Paillard S, Dedryver F, de Vallavieille-Pope C.** 2013. Quantitative plant resistance in cultivar mixtures: Wheat yellow rust as a modeling case study. *New Phytologist* **200**(3), 888–897.
- Schultink A, Qi T, Lee A, Steinbrenner AD, Staskawicz B.** 2017. Roq1 mediates recognition of the *Xanthomonas* and *Pseudomonas* effector proteins XopQ and HopQ1. *Plant Journal* **92**(5), 787–795.
- Scortichini M.** 2002. Bacterial canker and decline of European hazelnut. *Plant Disease* **86**(7), 704–709.
- Scortichini M.** 2010. Epidemiology and predisposing factors of some major bacterial diseases of stone and nut fruit trees species. *Journal of Plant Pathology* **92**(1), S73–S78.
- Scortichini M, Biocca M, Rossi MP.** 1995. *Pseudomonas syringae* pv. *morsprunorum* on wild cherry for timber production: outbreak and field susceptibility. *European Journal of Forest Pathology* **25**(6-7), 343–350.
- Scortichini M, Liguori R.** 2003. Integrated management of bacterial decline of hazelnut, by using Bion as an activator of systemic acquired resistance (SAR). In: Iacobellis NS,, In: Collmer A,, In: Hutcheson SW,, In: Mansfield JW,, In: Morris CE,, In: Murillo J,, In: Schaad NW,, In: Stead DE,, In: Surico G,, In: Ullrich MS, eds. *Pseudomonas syringae* and related pathogens. Dordrecht: Springer Netherlands, 483–487.
- Scortichini M, Marcelletti S, Ferrante P, Firrao G.** 2013. A genomic redefinition of *Pseudomonas avellanae* species. *PLoS ONE* **8**(9), e75794.
- Segura V, Vilhjálmsson BJ, Platt A, Korte A, Seren Ü, Long Q, Nordborg M.** 2012. An efficient multi-locus mixed-model approach for genome-wide association studies in structured populations. *Nature Genetics* **44**(7), 825–830.
- Senthil-Kumar M, Mysore KS.** 2013. Nonhost resistance against bacterial pathogens:

- retrospectives and prospects. *Annual Review of Phytopathology* **51**, 407–427.
- Seto D, Koulena N, Lo T, Menna A, Guttman DS, Desveaux D.** 2017. Expanded type III effector recognition by the ZAR1 NLR protein using ZED1-related kinases. *Nature Plants* **3**(17027).
- Shimono M, Lu Y-J, Porter K, Kvitko BH, Henty-Ridilla J, Creason A, He SY, Chang JH, Staiger CJ, Day B.** 2016. The *Pseudomonas syringae* type III effector HopG1 induces actin remodeling to promote symptom development and susceptibility during infection. *Plant Physiology* **171**(3), 2239–2255.
- Sillanpää MJ.** 2011. Overview of techniques to account for confounding due to population stratification and cryptic relatedness in genomic data association analyses. *Heredity* **106**(4), 511–519.
- Sobiczewski P.** 1987. Antagonistic bacteria in relation to *Pseudomonas syringae* pv. *syringae* occurring in necroses and cankers of sour cherry trees. *Fruit Science Reports* **14**(4), 179–185.
- Sobiczewski P, Jones AL.** 1992. Effect of exposure to freezing temperatures on necrosis in sweet cherry shoots inoculated with *Pseudomonas syringae* pv. *syringae* or *P. s. morsprunorum*. *Plant Disease* **76**(5), 447–451.
- Sochor J, Zitka O, Skutkova H, Pavlik D, Babula P, Krska B, Horna A, Adam V, Provaznik I, Kizek R.** 2010. Content of phenolic compounds and antioxidant capacity in fruits of apricot genotypes. *Molecules* **15**(9), 6285–6305.
- Socquet-Juglard D, Christen D, Devènes G, Gessler C, Duffy B, Patocchi A.** 2013a. Mapping architectural, phenological, and fruit quality QTLs in apricot. *Plant Molecular Biology Reporter* **31**(2), 387–397.
- Socquet-Juglard D, Duffy B, Pothier JF, Christen D, Gessler C, Patocchi A.** 2013b. Identification of a major QTL for *Xanthomonas arboricola* pv. *pruni* resistance in apricot. *Tree Genetics & Genomes* **9**(2), 409–421.
- Soufflet-Freslon V, Gianfranceschi L, Patocchi A, Durel C-E.** 2008. Inheritance studies of apple scab resistance and identification of *Rvi14*, a new major gene that acts together with other broad-spectrum QTL. *Genome* **51**(8), 657–667.
- St Clair DA.** 2010. Quantitative disease resistance and quantitative resistance loci in breeding. *Annual Review of Phytopathology* **48**, 247–268.
- Stall RE, Jones JB, Minsavage G V.** 2009. Durability of resistance in tomato and pepper to *Xanthomonads* causing bacterial spot. *Annual Review of Phytopathology* **47**, 265–284.
- Stephens M.** 2013. A unified framework for association analysis with multiple related phenotypes. *PLoS ONE* **8**(7), e65245.
- Sun J, Huang G, Fan F, Wang S, Zhang Y, Han Y, Zou Y, Lu D.** 2017. Comparative study of *Arabidopsis* PBS1 and a wheat PBS1 homolog helps understand the mechanism of PBS1 functioning in innate immunity. *Scientific Reports* **7**(5487).
- Sundin GW, Jones AL, Olson BD.** 1988. Overwintering and population dynamics of *Pseudomonas syringae* pv. *syringae* and *Ps. s. morsprunorum* on sweet and sour cherry trees. *Canadian Journal of Plant Pathology* **10**(4), 281–288.
- Swiderski MR, Innes RW.** 2001. The *Arabidopsis* PBS1 resistance gene encodes a member of a novel protein kinase subfamily. *Plant Journal* **26**(1), 101–112.

- Tamisier L, Rousseau E, Barraillé S, Nemouchi G, Szadkowski M, Mailleret L, Grognard F, Fabre F, Moury B, Palloix A.** 2017. Quantitative trait loci in pepper control the effective population size of two RNA viruses at inoculation. *Journal of General Virology* **98**(7), 1923–1931.
- Thabuis A, Lefebvre V, Bernard G, Daubèze AM, Phaly T, Pochard E, Palloix A.** 2004. Phenotypic and molecular evaluation of a recurrent selection program for a polygenic resistance to *Phytophthora capsici* in pepper. *Theoretical and Applied Genetics* **109**(2), 342–351.
- Thoen Manus PM, Davila Olivas Nelson H, Kloth Karen J, Coolen S, Huang P-P, Aarts Mark GM, Bac-Molenaar JA, Bakker J, Bouwmeester Harro J, Broekgaarden C, Bucher J, Busscher-Lange J, Cheng X, Fradin Emilie F, Jongma Maarten A, Julkowska Magdalena M, Keurentjes Joost JB, Ligterink W, Pieterse CM, Ruyter-Spira C, Smant G, Testerink C, Usadel B, Loon Joop JA, Pelt Johan A, Schaik Casper C, Wees Saskia CM, Visser Richard GF, Voorrips RE, Vosman B, Vreugdenhil D, Warmerdam S, Wiegers Gerrie L, Heerwaarden J, Kruijer W, Eeuwijk Fred A, Dicke M.** 2016. Genetic architecture of plant stress resistance: multi-trait genome-wide association mapping. *New Phytologist* **213**(3), 1346–1362.
- Thomma BPHJ, Nürnberger T, Joosten MHAJ.** 2011. Of PAMPs and effectors: The blurred PTI-ETI dichotomy. *The Plant Cell* **23**(1), 4–15.
- Tisné S, Pomiès V, Riou V, Syahputra I, Cochard B, Denis M.** 2017. Identification of *Ganoderma* disease resistance loci using natural field infection of an oil palm multiparental population. *G3: Genes|Genomes|Genetics* **7**(6), 1683–1692.
- Tornero P, Gadea J, Conejero V, Vera P.** 1997. Two *PR-1* genes from tomato are differentially regulated and reveal a novel mode of expression for a pathogenesis-related gene during the hypersensitive response and development. *Molecular Plant-Microbe Interactions* **10**(5), 624–634.
- Urrestarazu J, Muranty H, Denancé C, Leforestier D, Ravon E, Guyader A, Guisnel R, Feugey L, Aubourg S, Celton J-M, Daccord N, Dondini L, Gregori R, Lateur M, Houben P, Ordidge M, Paprstein F, Sedlak J, Nybom H, Garkava-Gustavsson L, Troglio M, Bianco L, Velasco R, Poncet C, Théron A, Moriya S, Bink MCAM, Laurens F, Tartarini S, Durel C-E.** 2017. Genome-wide association mapping of flowering and ripening periods in apple. *Frontiers in Plant Science* **8**(1923).
- Utz HF, Melchinger AE, Schön CC.** 2000. Bias and sampling error of the estimated proportion of genotypic variance explained by quantitative trait loci determined from experimental data in maize using cross validation and validation with independent samples. *Genetics* **154**(4), 1839–1849.
- van Valen L.** 1973. A new evolutionary law. *Evolutionary Theory* **1**, 1–30.
- Van AL, Caffier V, Lasserre-Zuber P, Chauveau A, Brunel D, Le Cam B, Durel CE.** 2013. Differential selection pressures exerted by host resistance quantitative trait loci on a pathogen population: A case study in an apple × *Venturia inaequalis* pathosystem. *New Phytologist* **197**(3), 899–908.
- Vance V.** 2001. RNA Silencing in Plants--Defense and Counterdefense. *Science* **292**(5525),

2277–2280.

- Vanneste JL, Voyle MD.** 2003. Genetic basis of copper resistance in New Zealand strains of *Pseudomonas syringae*. *New Zealand Plant Protection* **56**, 109–112.
- Vavilov NI.** 1992. *Origin and geography of cultivated plants*. Cambridge University Press.
- Vera Cruz CM, Bai J, Oña I, Leung H, Nelson RJ, Mew T-W, Leach JE.** 2000. Predicting durability of a disease resistance gene based on an assessment of the fitness loss and epidemiological consequences of avirulence gene mutation. *Proceedings of the National Academy of Sciences* **97**(25), 13500–13505.
- Verde I, Abbott AG, Scalabrin S, Jung S, Shu S, Marroni F, Zhebentyayeva T, Dettori MT, Grimwood J, Cattonaro F, Zuccolo A, Rossini L, Jenkins J, Vendramin E, Meisel LA, Decroocq V, Sosinski B, Prochnik S, Mitros T, Policriti A, Cipriani G, Dondini L, Ficklin S, Goodstein DM, Xuan P, Fabbro CD, Aramini V, Copetti D, Gonzalez S, Horner DS, Falchi R, Lucas S, Mica E, Maldonado J, Lazzari B, Bielenberg D, Pirona R, Miculan M, Barakat A, Testolin R, Stella A, Tartarini S, Tonutti P, Arús P, Orellana A, Wells C, Main D, Vizzotto G, Silva H, Salamini F, Schmutz J, Morgante M, Rokhsar DS.** 2013. The high-quality draft genome of peach (*Prunus persica*) identifies unique patterns of genetic diversity, domestication and genome evolution. *Nature Genetics* **45**(5), 487–494.
- Verde I, Bassil N, Scalabrin S, Gilmore B, Lawley CT, Gasic K, Micheletti D, Rosyara UR, Cattonaro F, Vendramin E, Main D, Aramini V, Blas AL, Mockler TC, Bryant DW, Wilhelm L, Troglio M, Sosinski B, Aranzana MJ, Arús P, Iezzoni A, Morgante M, Peace CP.** 2012. Development and evaluation of a 9K SNP array for peach by internationally coordinated SNP detection and validation in breeding germplasm. *PLoS ONE* **7**(4), e35668.
- Verde I, Jenkins J, Dondini L, Micali S, Pagliarani G, Vendramin E, Paris R, Aramini V, Gazza L, Rossini L, Bassi D, Troglio M, Shu S, Grimwood J, Tartarini S, Dettori MT, Schmutz J.** 2017. The peach v2.0 release: high-resolution linkage mapping and deep resequencing improve chromosome-scale assembly and contiguity. *BMC Genomics* **18**(225).
- Veyrieras J-B, Goffinet B, Charcosset A.** 2007. MetaQTL: a package of new computational methods for the meta-analysis of QTL mapping experiments. *BMC Bioinformatics* **8**(49).
- Vicente JG, Alves JP, Russell K, Roberts SJ.** 2004. Identification and discrimination of *Pseudomonas syringae* isolates from wild cherry in England. *European Journal of Plant Pathology* **110**(4), 337–351.
- Vigouroux A.** 1989. Ingress and spread of *Pseudomonas* in stems of peach and apricot promoted by frost-related water-soaking of tissues. *Plant Disease* **73**(10), 854–855.
- Vigouroux A.** 1999. Bacterial canker of peach: effect of tree winter water content on the spread of infection through frost-related water soaking in stems. *Journal of Phytopathology* **147**(9), 553–559.
- Vigouroux A, Bussi C.** 1989. Importance of water consumption on calcium content and protection of peaches predisposed to bacterial dieback by growing in acid soils. *Acta Horticulturae* **254**, 291–296.

- Vigouroux A, Bussi C.** 1994. Une action possible des sols sur la prédisposition des pêchers au dépérissement bactérien par modification de la teneur en eau hivernale des tiges. *Agronomie* **14**(5), 319–326.
- Vigouroux A, Huguet C.** 1980. Prédisposition du pêcher au dépérissement bactérien sur sol d'arène et de diluvium. *Ann Phytopathol* **12**(4), 312.
- Vilanova S, Badenes ML, Burgos L, Martinez-Calvo J, Llácer G, Romero C.** 2006. Self-compatibility of two apricot selections is associated with two pollen-part mutations of different nature. *Plant Physiology* **142**(2), 629–641.
- Vilanova S, Romero C, Abbott AG, Llácer G, Badenes ML.** 2003. An apricot (*Prunus armeniaca* L.) F2 progeny linkage map based on SSR and AFLP markers, mapping plum pox virus resistance and self-incompatibility traits. *Theoretical and Applied Genetics* **107**(2), 239–247.
- Walters DR, McRoberts N.** 2006. Plants and biotrophs: a pivotal role for cytokinins? *Trends in Plant Science* **11**(12), 581–586.
- Wang G, Ellendorff U, Kemp B, Mansfield JW, Forsyth A, Mitchell K, Bastas K, Liu C-M, Woods-Tor A, Zipfel C, de Wit PJGM, Jones JDG, Tor M, Thomma BPHJ.** 2008. A genome-wide functional investigation into the roles of receptor-like proteins in *Arabidopsis*. *Plant Physiology* **147**(2), 503–517.
- Wang PW, Morgan RL, Scortichini M, Guttman DS.** 2007a. Convergent evolution of phytopathogenic *Pseudomonads* onto hazelnut. *Microbiology* **153**(7), 2067–2073.
- Wang D, Pajerowska-Mukhtar K, Culler AH, Dong X.** 2007b. Salicylic acid inhibits pathogen growth in plants through repression of the auxin signaling pathway. *Current Biology* **17**(20), 1784–1790.
- Warren RF, Holub E, Innes RW.** 1999. Identification of three putative signal transduction genes involved in *R* gene-specified disease resistance in *Arabidopsis*. *Genetics* **152**(1), 401–412.
- Weaver DJ.** 1978. Interaction of *Pseudomonas syringae* and freezing in bacterial canker on excised peach twigs. *Phytopathology* **68**, 1460–1463.
- van de Weg EW, Di Guardo M, Jänsch M, Socquet-Juglard D, Costa F, Baumgartner I, Broggini GAL, Kellerhals M, Troglio M, Laurens F, Durel C-E, Patocchi A.** 2018. Epistatic fire blight resistance QTL alleles in the apple cultivar 'Enterprise' and selection X-6398 discovered and characterized through pedigree-informed analysis. *Molecular Breeding* **38**(5).
- van de Weg EW, Voorrips RE, Finkers R, Kodde LP, Jansen J, Bink MCAM.** 2004. Pedigree genotyping: a new pedigree-based approach of QTL identification and allele mining by exploiting breeding material. *Acta Horticulturae* **663**, 45–50.
- Weiberg A, Wang M, Lin FM, Zhao H, Zhang Z, Kaloshian I, Huang HD, Jin H.** 2013. Fungal small RNAs suppress plant immunity by hijacking host RNA interference pathways. *Science* **342**(6154), 118–123.
- Weinberger JH.** 1950. Chilling requirements of peach varieties. *Proceedings. American Society for Horticultural Science* **56**, 122–128.
- Weitz JS, Poisot T, Meyer JR, Flores CO, Valverde S, Sullivan MB, Hochberg ME.** 2013. Phage-bacteria infection networks. *Trends in Microbiology* **21**(2), 82–91.
- Werner DJ, Ritchie DF, Cain DW, Zehr EI.** 1986. Susceptibility of peaches and

- nectarines plant introductions and other *Prunus* species to bacterial spot. *HortScience* **21**(1), 127–130.
- Wilde F, Schön CC, Korzun V, Ebmeyer E, Schmolke M, Hartl L, Miedaner T.** 2008. Marker-based introduction of three quantitative-trait loci conferring resistance to *Fusarium* head blight into an independent elite winter wheat breeding population. *Theoretical and Applied Genetics* **117**(1), 29–35.
- Wilson EE.** 1953. Bacterial Canker of Stone Fruits. *Yearbook of Agriculture*. 722–729.
- Wimalajeewa DLS, Cahill R, Hepworth G, Schneider HG, Washbourne JW.** 1991. Chemical control of bacterial canker (*Pseudomonas syringae* pv. *syringae*) of apricot and cherry in Victoria. *Australian Journal of Experimental Agriculture* **31**(5), 705–708.
- Wisser RJ, Balint-Kurti PJ, Nelson RJ.** 2006. The genetic architecture of disease resistance in maize: A synthesis of published studies. *Phytopathology* **96**(2), 120–129.
- Wisser RJ, Sun Q, Hulbert SH, Kresovich S, Nelson RJ.** 2005. Identification and characterization of regions of the rice genome associated with broad-spectrum, quantitative disease resistance. *Genetics* **169**(4), 2277–2293.
- Wu CH, Derevnina L, Kamoun S.** 2018. Receptor networks underpin plant immunity. *Science* **360**(6395), 1300–1301.
- Wu R, Ma C-X, Casella G.** 2002. Joint linkage and linkage disequilibrium mapping of quantitative trait loci in natural populations. *Genetics* **160**(2), 779–792.
- Xiao W, Zhao J, Fan S, Li L, Dai J, Xu M.** 2007. Mapping of genome-wide resistance gene analogs (RGAs) in maize (*Zea mays* L.). *Theoretical and Applied Genetics* **115**(4), 501–508.
- Xin XF, Kvitko BH, He SY.** 2018. *Pseudomonas syringae*: what it takes to be a pathogen. *Nature Reviews Microbiology* **16**(5), 316–328.
- Xin XF, Nomura K, Aung K, Velásquez AC, Yao J, Boutrot F, Chang JH, Zipfel C, He SY.** 2016. Bacteria establish an aqueous living space in plants crucial for virulence. *Nature* **539**(7630), 524–529.
- Yamada K, Saijo Y, Nakagami H, Takano Y.** 2016. Regulation of sugar transporter activity for antibacterial defense in *Arabidopsis*. *Science* **354**(6318), 1427–1430.
- Yamamoto T, Terakami S, Takada N, Nishio S, Onoue N, Nishitani C, Kunihisa M, Inoue E, Iwata H, Hayashi T, Itai A, Saito T.** 2014. Identification of QTLs controlling harvest time and fruit skin color in Japanese pear (*Pyrus pyrifolia* Nakai). *Breeding Science* **64**(4), 351–361.
- Yang J.** 2004. Genome-wide association studies to dissect the genetic architecture of yield-related traits in maize and the genetic basis of heterosis (IS University, Ed.). Graduate Theses and Dissertations.
- Yang L, Teixeira PJPL, Biswas S, Finkel OM, He Y, Salas-Gonzalez I, English ME, Epple P, Mieczkowski P, Dangl JL.** 2017. *Pseudomonas syringae* type III effector HopBB1 promotes host transcriptional repressor degradation to regulate phytohormone responses and virulence. *Cell Host and Microbe* **21**(2), 156–168.
- Yilmaz KU, Gurcan K.** 2012. Genetic Diversity in Apricot. *Genetic Diversity in Plants*. In Tech, 249–270.
- Young JM.** 1987. Orchard management and bacterial diseases of stone fruit. *New Zealand Journal of Experimental Agriculture* **15**(2), 257–266.

- Young JM.** 2010. Taxonomy of *Pseudomonas syringae*. *Journal of Plant Pathology* **92**(1), S5–S14.
- Young JM, Dye DW, Bradbury JF, Panagopoulos CG, Robbs CF.** 1978. A proposed nomenclature and classification for plant pathogenic bacteria. *New Zealand Journal of Agricultural Research* **21**(1), 153–177.
- Yu J, Holland JB, McMullen MD, Buckler ES.** 2008. Genetic design and statistical power of nested association mapping in maize. *Genetics* **178**(1), 539–551.
- Yu J, Pressoir G, Briggs WH, Vroh Bi I, Yamasaki M, Doebley JF, McMullen MD, Gaut BS, Nielsen DM, Holland JB, Kresovich S, Buckler ES.** 2005. A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nature Genetics* **38**(2), 203–208.
- Yue JX, Meyers BC, Chen JQ, Tian D, Yang S.** 2012. Tracing the origin and evolutionary history of plant nucleotide-binding site-leucine-rich repeat (NBS-LRR) genes. *New Phytologist* **193**(4), 1049–1063.
- Zaurov D, Molnar T, Eisenman S, Ford TM, Mavlyanova R, Capik JM, Reed Funk C, Goffreda JC.** 2013. Genetic resources of apricots (*Prunus armeniaca* L.) in central Asia. *HortScience* **48**(6), 681–691.
- Zenbayashi-Sawata K, Fukuoka S, Katagiri S, Fujisawa M, Matsumoto T, Ashizawa T, Koizumi S.** 2007. Genetic and physical mapping of the partial resistance gene, *pi34*, to blast in rice. *Phytopathology* **97**(5), 598–602.
- Zeng ZB.** 1994. Precision mapping of quantitative trait loci. *Genetics* **136**(4), 1457–1468.
- Zeng ZB, Kao CH, Basten CJ.** 1999. Estimating the genetic architecture of quantitative traits. *Genetics Research* **74**(3), 279–289.
- Zeng L, Velásquez AC, Munkvold KR, Zhang J, Martin GB.** 2012. A tomato LysM receptor-like kinase promotes immunity and its kinase activity is inhibited by AvrPtoB. *Plant Journal* **69**(1), 92–103.
- Zhang J, Li W, Xiang T, Liu Z, Laluk K, Ding X, Zou Y, Gao M, Zhang X, Chen S, Mengiste T, Zhang Y, Zhou JM.** 2010a. Receptor-like cytoplasmic kinases integrate signaling from multiple plant immune receptors and are targeted by a *Pseudomonas syringae* effector. *Cell Host and Microbe* **7**(4), 290–301.
- Zhang Z, Liu Y, Huang H, Gao M, Wu D, Kong Q, Zhang Y.** 2017. The NLR protein SUMM2 senses the disruption of an immune signaling MAP kinase cascade via CRCK3. *EMBO reports* **18**(2), 292–302.
- Zhang J, Shao F, Li Y, Cui H, Chen L, Li H, Zou Y, Long C, Lan L, Chai J, Chen S, Tang X, Zhou J.** 2007. A *Pseudomonas syringae* effector inactivates MAPKs to suppress PAMP-induced immunity in plants. *Cell Host and Microbe* **1**(3), 175–185.
- Zhang Y, Yang Y, Fang B, Gannon P, Ding P, Li X, Zhang Y.** 2010b. *Arabidopsis* snc2-1D activates receptor-like protein-mediated immunity transduced through WRKY70. *The Plant Cell* **22**(9), 3153–3163.
- Zhang L, Yao J, Withers J, Xin X-F, Banerjee R, Fariduddin Q, Nakamura Y, Nomura K, Howe GA, Boland W, Yan H, He SY.** 2015. Host target modification as a strategy to counter pathogen hijacking of the jasmonate hormone receptor. *Proceedings of the National Academy of Sciences* **112**(46), 14354–14359.
- Zhao T, Rui L, Li J, Nishimura MT, Vogel JP, Liu N, Liu S, Zhao Y, Dangl JL,**

- Tang D.** 2015. A truncated NLR protein, TIR-NBS2, is required for activated defense responses in the *exo70B1* mutant. *PLoS Genetics* **11**(1), e1004945.
- Zhebentyayeva T, Ledbetter C, Burgos L, Ll acer G.** 2012. Apricot. In: Media SS& B, ed. *Fruit Breeding*. Springer New York, 415–449.
- Zhebentyayeva T, Reighard G, Gorina V, Abbott AG.** 2003. Simple sequence repeat (SSR) analysis for assessment of genetic variability in apricot germplasm. *Theoretical and Applied Genetics* **106**(3), 435–444.
- Zheng Z, Qamar SA, Chen Z, Mengiste T.** 2006. *Arabidopsis* WRKY33 transcription factor is required for resistance to necrotrophic fungal pathogens. *Plant Journal* **48**(4), 592–605.
- Zheng XY, Spivey NW, Zeng W, Liu PP, Fu ZQ, Klessig DF, He SY, Dong X.** 2012. Coronatine promotes *Pseudomonas syringae* virulence in plants by activating a signaling cascade that inhibits salicylic acid accumulation. *Cell Host and Microbe* **11**(6), 587–596.
- Zhou G, Chen Y, Yao W, Zhang C, Xie W, Hua J, Xing Y, Xiao J, Zhang Q.** 2012. Genetic composition of yield heterosis in an elite rice hybrid. *Proceedings of the National Academy of Sciences* **109**(39), 15847–15852.
- Zhou X, Stephens M.** 2014. Efficient multivariate linear mixed model algorithms for genome-wide association studies. *Nature Methods* **11**(4), 407–409.
- Zhou Z, Wu Y, Yang Y, Du M, Zhang X, Guo Y, Li C, Zhou J-M.** 2015. An *Arabidopsis* plasma membrane proton ATPase modulates JA signaling and is exploited by the *Pseudomonas syringae* effector protein AvrB for stomatal invasion. *The Plant Cell* **27**(7), 2032–2041.
- Zhu H, Gilchrist L, Hayes P, Kleinhofs A, Kudrna D, Liu Z, Prom L, Steffenson B, Toojinda T, Vivar H.** 1999. Does function follow form? Principal QTLs for *Fusarium* head blight (FHB) resistance are coincident with QTLs for inflorescence traits and plant height in a doubled-haploid population of barley. *Theoretical and Applied Genetics* **99**(7–8), 1221–1232.
- Zipfel C.** 2008. Pattern-recognition receptors in plant innate immunity. *Current Opinion in Immunology* **20**(1), 10–16.
- Zipfel C, Robatzek S, Navarro L, Oakeley EJ, Jones JDG, Felix G, Boller T.** 2004. Bacterial disease resistance in *Arabidopsis* through flagellin perception. *Nature* **428**, 764–767.
- Zuriaga E, Romero C, Blanca JM, Badenes ML.** 2018. Resistance to Plum Pox Virus (PPV) in apricot (*Prunus armeniaca* L.) is associated with down-regulation of two *MATHd* genes. *BMC Plant Biology* **18**(25).

# Table des annexes

ANNEXE I : FIGURES ET TABLES ADDITIONNELLES DE L'ARTICLE GWAS DE LA PARTIE III. I	
ANNEXE II : FIGURES ET TABLES ADDITIONNELLES DE L'ARTICLE DE L'ANALYSE DE LIAISON SUR LA BERBA ET LES TROIS POPULATIONS INTERCONNECTEES DE LA PARTIE IV.....	xvi
ANNEXE III : DONNEES ADDITIONNELLES DE L'ETUDE DE L'INTERACTION GxG DE LA PARTIE V.....	xxvi

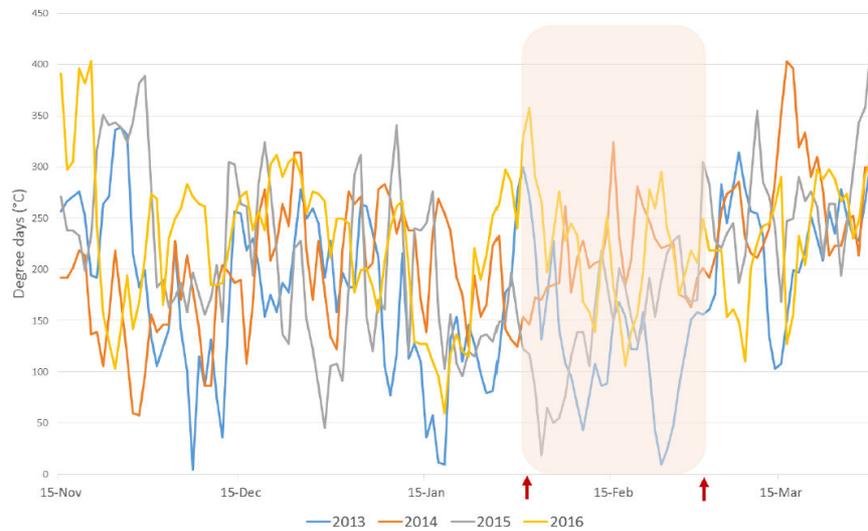


**Annexe I : Figures et tables additionnelles de l'article GWAS de la partie III.**

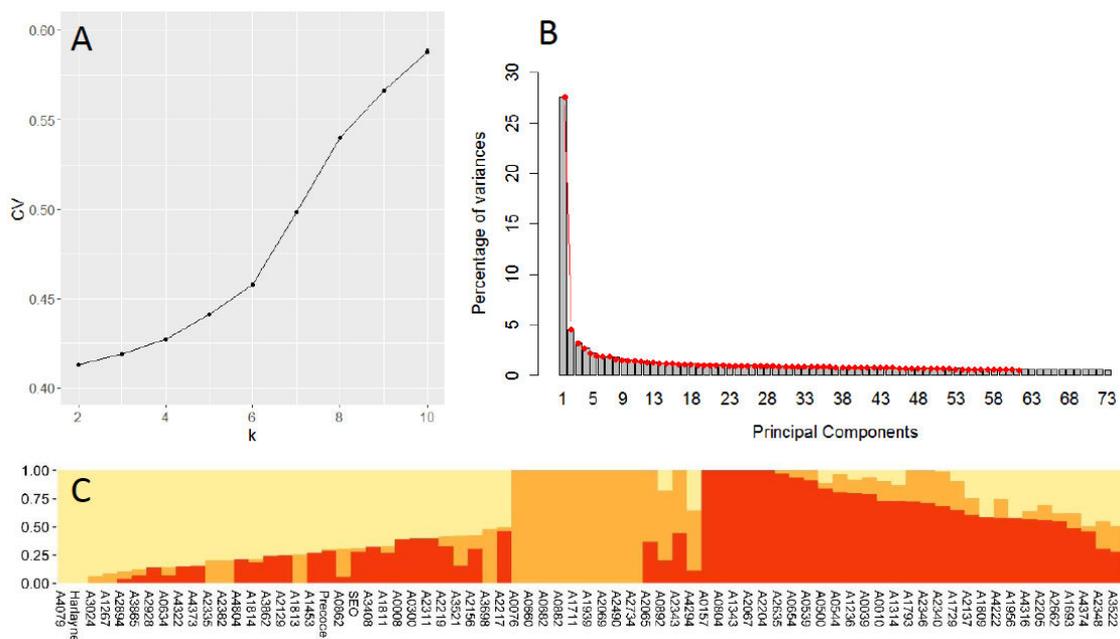
**Additional file 1 (.xls): Information of plant material used in the study: geographical origin, name and number of year repetitions through 2013 to 2016 phenotyping period.**

Accession code	Name	Origin	Number of year repetitions	Accession code	Name	Origin	Number of year repetitions
A0008	Colomer	France	4	A2156	Veecot	Canada	2
A0010	Sucre de Holub	Hungary	1	A2204	Bebeco	Greece	4
A0039	Précoce d'Ampuis	France	4	A2205	Andswee	USA	2
A0076	Pêche de Nancy	France	2	A2217	Sunglo	Canada	2
A0157	Rouge du Roussillon	France	4	A2219	Vivagold	Canada	4
A0300	Mei Hwang	China	4	A2311	Lambertin 1	France	4
A0500	Moniqui	Spain	4	A2335	Mamaia	Romania	1
A0539	Poizat	France	4	A2340	Saturn	Romania	4
A0544	Cafona	Italy	4	A2343	Olimp	Romania	4
A0634	Henderson	USA	4	A2346	Timpurii de Kisinou	Romania	2
A0654	Paviot	France	4	A2348	Erevani	Armenia	1
A0660	Bergeron	France	4	A2382	Stella	USA	3
A0682	Morden 604	Canada	1	A2490	Tardif de Tain	France	4
A0804	Screara	France	4	A2635	Moorpark	USA	2
A0862	Pseudo Royal	USA	4	A2662	Alfred	USA	4
A0882	Madarska Najlepsia	Slovakia	4	A2734	Boucherand Boutard	France	3
A0892	Silistra x Ananas	Romania	2	A2894	Orangered® Bhart	USA	4
A1236	Manicot	France	4	A2928	Early Blush® Rutbhart	USA	4
A1267	Badami	Iran	4	A3024	Kec Psar	Czech Republic	2
A1314	Arrogante	Spain	4	A3408	Ajouc	France	4
A1343	Canino	Spain	3	A3521	A3521	Iran	4
A1453	Perfection	USA	4	A3522	A3522	Iran	3
A1693	Nikitski	Ukraine	4	A3698	Vestar	Slovakia	3
A1711	Avikaline	France	3	A3862	Robada	USA	4
A1729	Priana	Tunisia	4	A3865	IS13-91	USA	4
A1793	Tardif de Bordaneil	France	4	A4079	Mangocot® Rutsey	USA	4
A1809	Precoce de Tyrinthe	Greece	4	A4222	Larclyd	New-Zealand	1
A1811	Harcot	Canada	4	A4294	Incomparable de Malissard	France	3
A1813	Laycot	Canada	3	A4316	Augusta 3	Italy	4
A1814	Hargrand	Canada	4	A4322	Priboto	France	4
A1939	Fantasme	France	4	A4373	A4373	Italy	4
A1956	Palsteyn	South Africa	3	A4374	Murciana	Spain	4
A2065	Marouch 4	Morocco	1	A4804	A4804	Italy	3
A2067	Marouch 14	Morocco	1	Harlayne	Harlayne	Canada	4
A2069	Caid Agdz 2	Morocco	1	Precoce	Precoce d'Imola	Italy	3
A2129	Rouge de Fournes	France	4	SEO	Stark Early Orange	USA	4
A2137	Bakour	Tunisia	3				

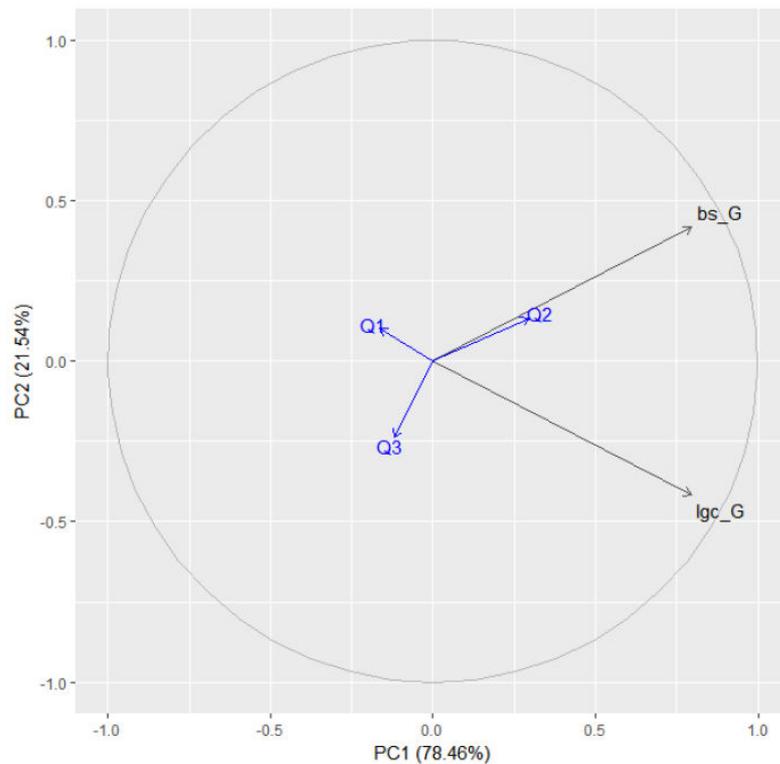
**Additional file 2 (.pdf): Time series of degree-days over winter period in l’Amarine.** 2013 to 2016 annual data from November 15<sup>th</sup> to March 31<sup>st</sup>. February month is delimited by the red arrows.



**Additional file 3 (.pdf): Population Structure (Q) of the apricot core-collection based on data of 21,942 independent SNPs.** Genotypic data pruned in order to keep markers with pairwise  $r^2$  values  $< 0.2$ . (A) Cross-validation estimation from  $k=2$  to  $k=10$  groups. Calculations performed using Admixture program [45]. (B) Barplot of Principal Component Analysis (PCA) eigenvalues from PC1 to PC73. The three first components explain 35.23% of the total variability. (C) Distribution of the ancestral fractions for each accession according to a structure divided into 3 groups: Central and Eastern Asia (yellow), Continental Europe (orange), Irano-Caucasian and Mediterranean Basin area (red).



**Additional file 4 (.pdf): Relation between susceptibility phenotypes *lgc* and *bs G* BLUPs and ancestral fractions in the apricot core-collection.** Ancestral fractions Q1 to Q3 are displayed as supplementary variables (in blue) in the PCA. The contribution of the first component reached 78.46% of the total variability.

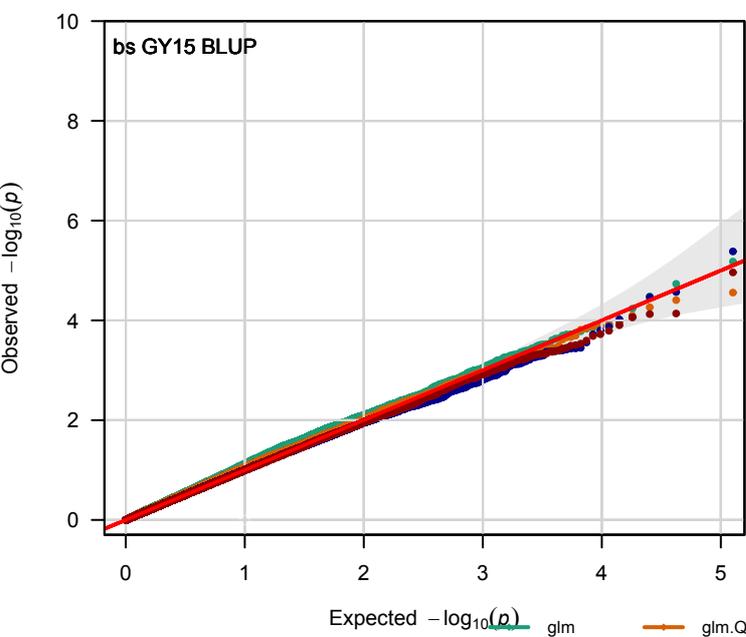
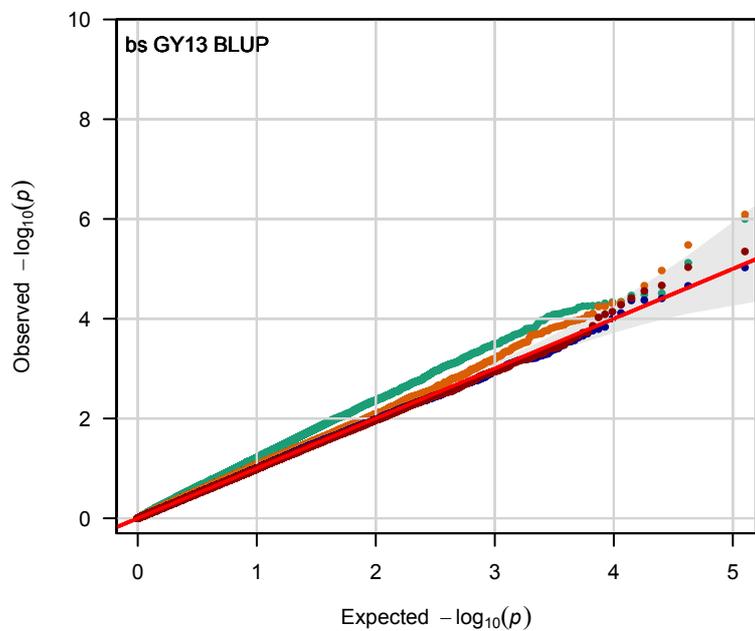
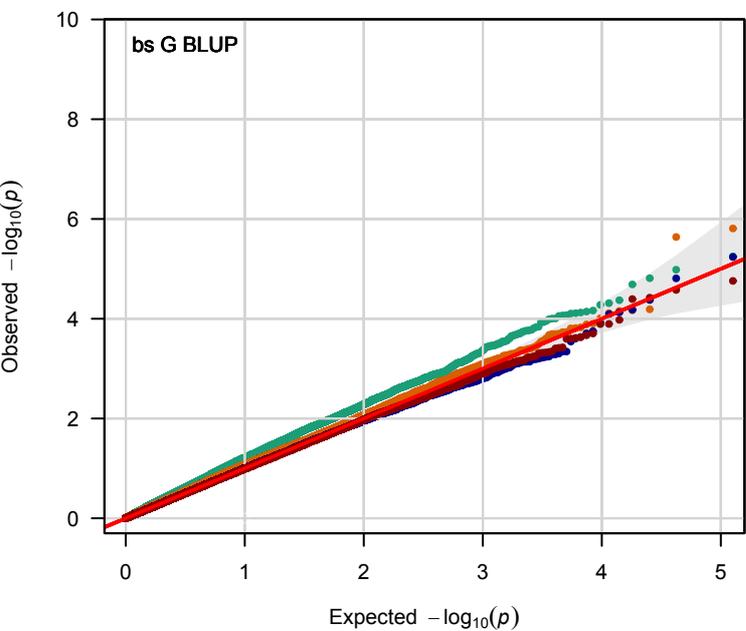
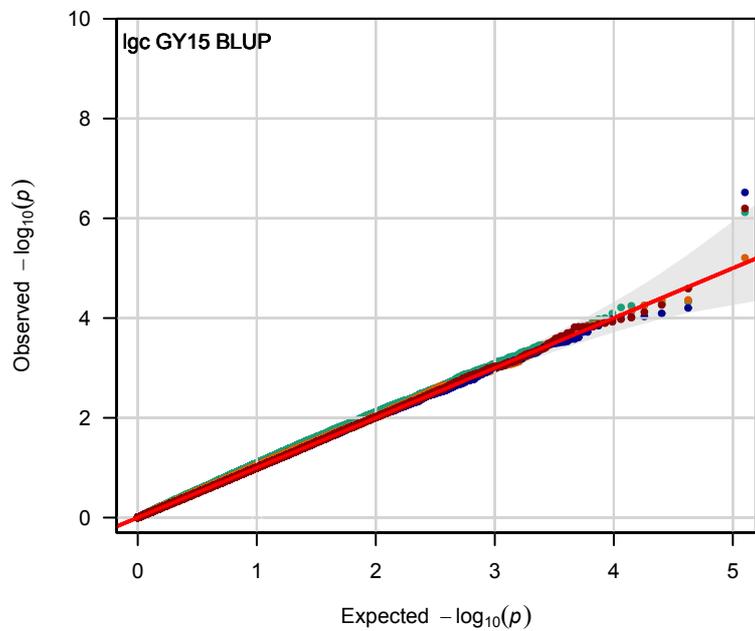
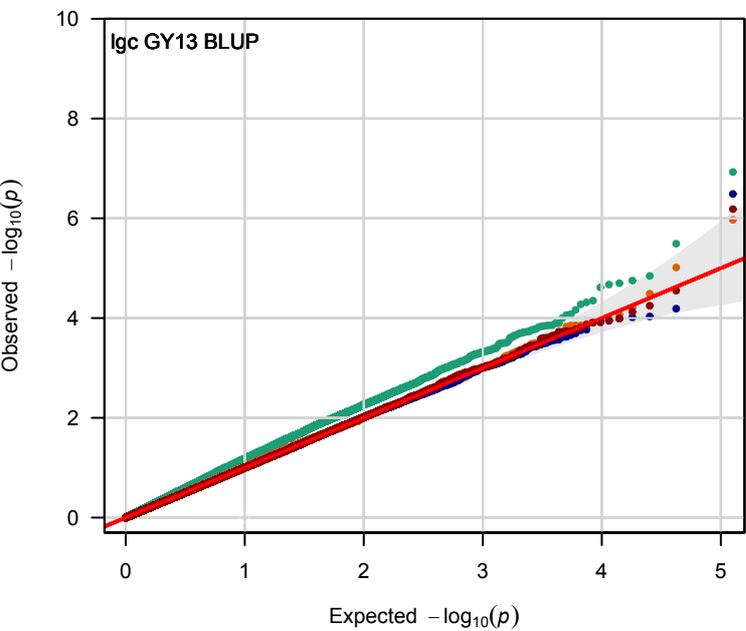


*Pages suivantes*

**Additional file 5 (.pdf): Linkage disequilibrium (LD) decay over physical position in the apricot core-collection.** Pairwise  $r^2$  values were computed considering markers in a 1 kb-window. Estimated LD decay according to a non-linear regression is represented for each chromosome. Black curves show non-corrected estimates. Curves of corrected  $r^2$  estimates with relatedness K (red), structure Q (orange) and both K and Q (blue) are displayed.  $r^2$  mean and standard deviation values are indicated.  $r^2=0.10$  threshold are represented with dashed horizontal lines.

**Additional file 6 (.pdf): QQ (quantile-quantile) plots of the observed p-values distribution for *lgc* and *bs G* and  $G \times Y$  BLUPs.** QQ-plot for *lgc G* BLUP is shown in Figure 4. Non-corrected generalized linear model (glm – green), glm accounting for structure Q (glm.Q – orange), linear mixed model correcting for relatedness K (emmax.K – blue) and both for K and Q (emmax.KQ – brown) output p-values are compared.

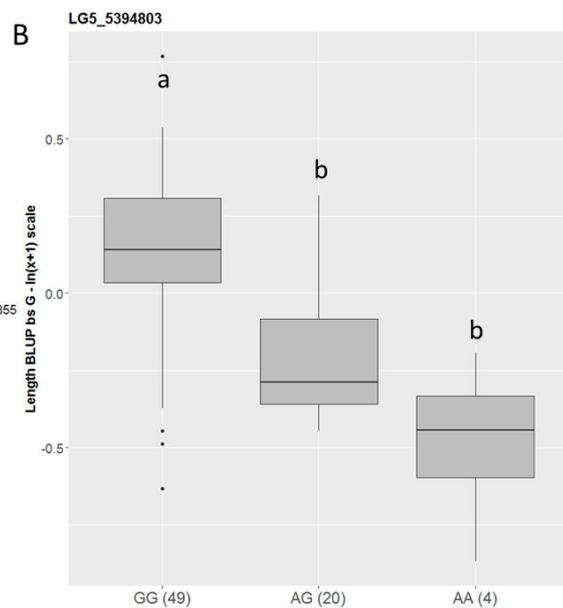
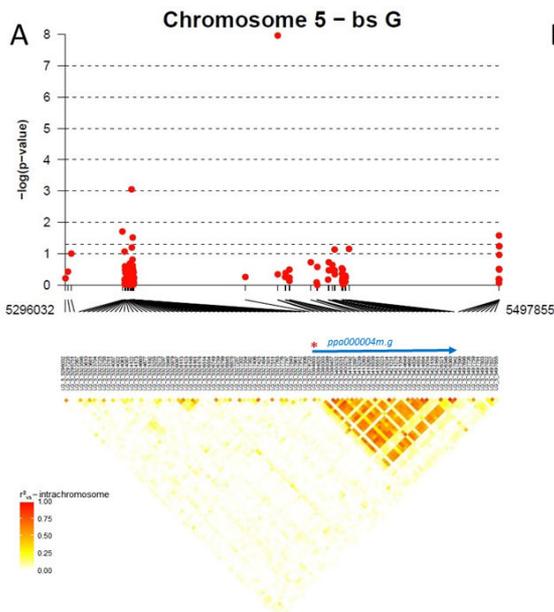
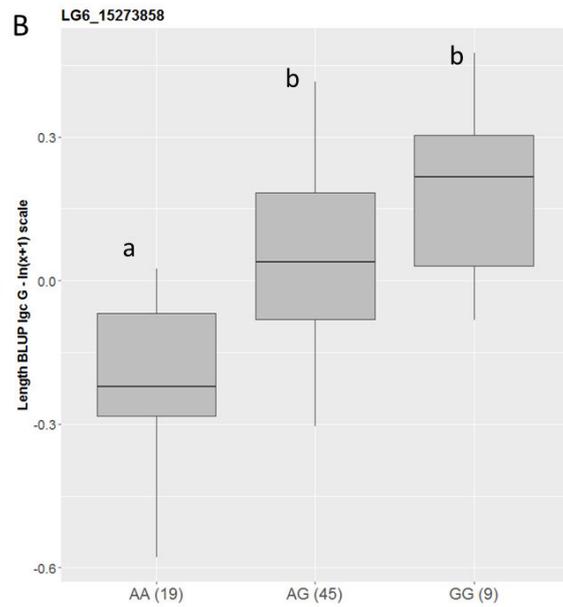
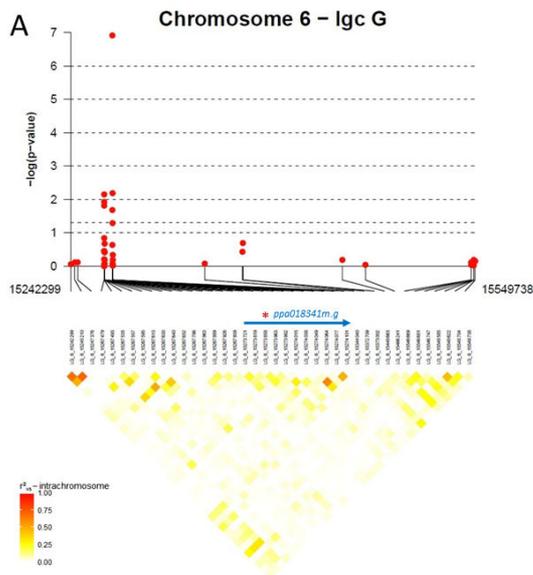




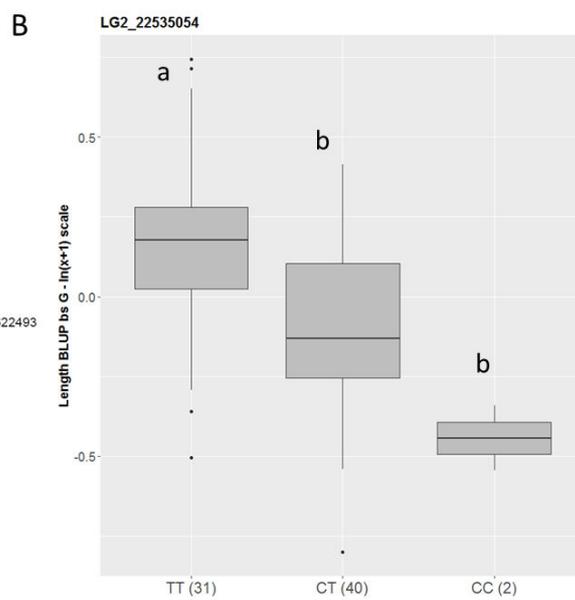
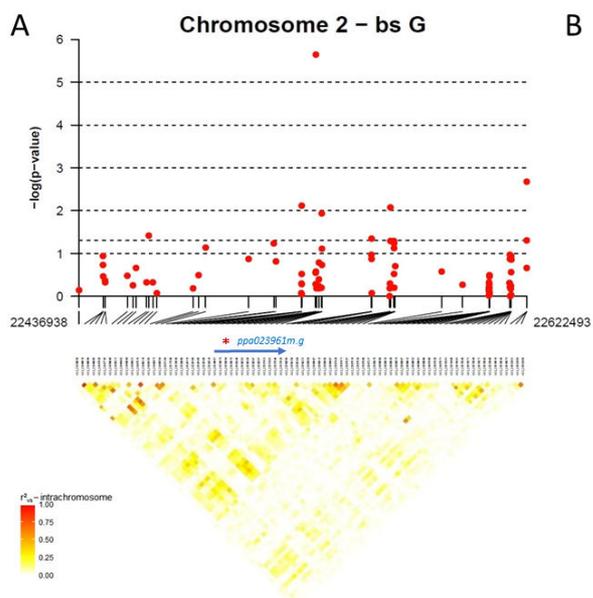
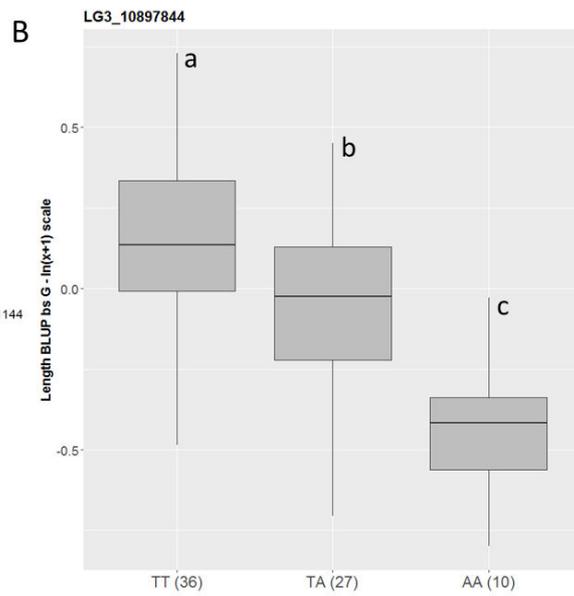
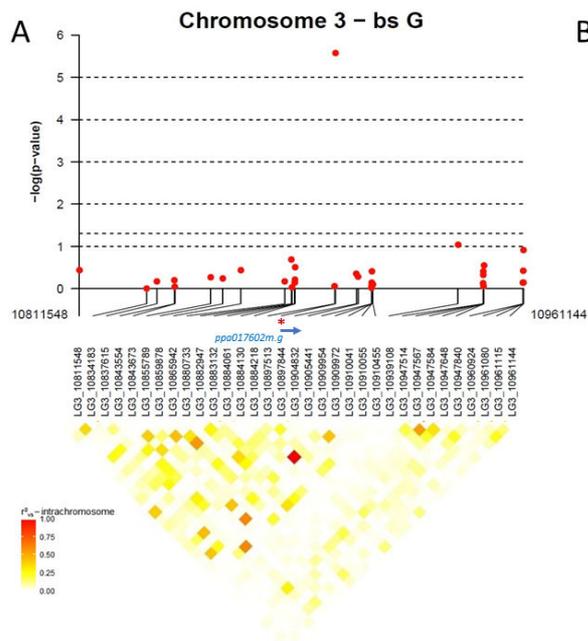
Expected  $-\log_{10}(p)$     glm    glm.Q    emmax.K    emmax.KQ

**Additional file 7 (.pdf): Effect of all detected loci using genome-wide multi-locus and multi-variate (mvLMM) mixed models on *lgc* and *bs G* and  $G \times Y$  BLUPs.**

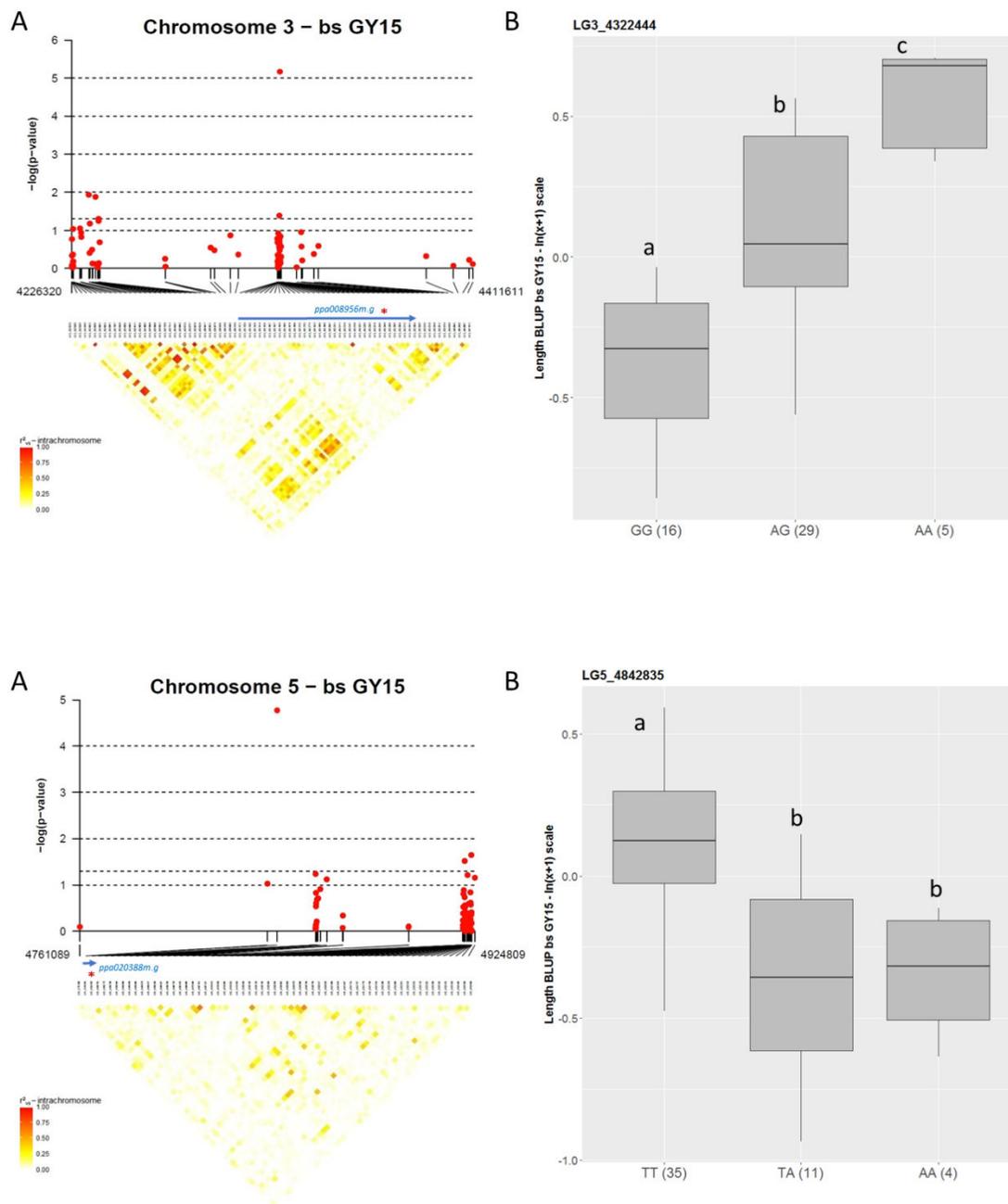
(A) Manhattan Plots displaying the  $-\log_{10}$  (p-values) over physical positions in approximate 100 to 300 kb windows surrounding associations. Significance level determined with eBIC criterion. Detected SNPs were shown with red stars and overlapping candidate genes indicated with blue arrows. As an indication, Bonferroni threshold  $= -\log_{10}(\text{p-value-thr}) = 6.10$ . Pairwise LD heatmaps ( $r_{ij}^2$ ) were drawn within the genomic window around each candidate. (B) Boxplot displaying the allelic effect for the associated marker. For the loci detected with mvLMM on chromosome 7, the allelic effect is represented for both *lgc* and *bs G* BLUPs.



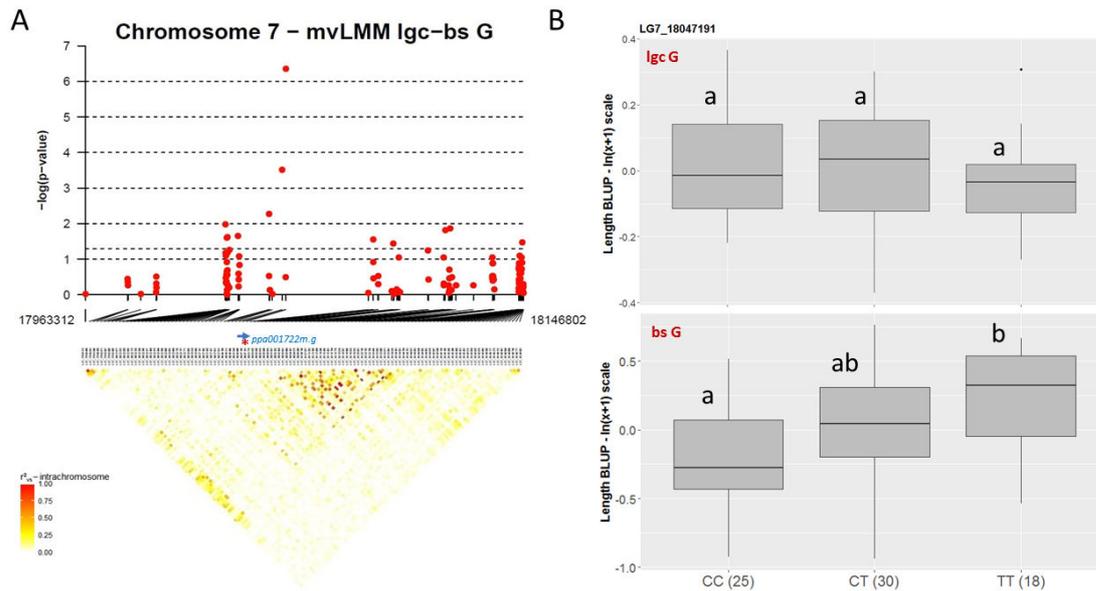
Suite additional file 7



Suite additional file 7



Suite additional file 7



*Page suivante*

**Additional file 8 (.xls): Individual/multi-locus ANOVA and post-hoc mean comparison tests of all associated markers on *G* BLUPs from the within-year model.** \*\*\* shows p-value below 0.001, \*\* between 0.001 and 0.01, \* between 0.01 and 0.05, and 'ns' indicates non-significant p-values. Significant results are indicated in red. Results with p-values slightly over the 5% significance level (noted as (.)) are indicated in italic. SS = ratio sum of squares (factor/total variation). For 'multi-locus' ANOVA, strict-sense heritability estimates  $h^2$  are indicated. Comparison of means between groups were performed for individual locus-ANOVA only on phenotypes associated to a unique locus. AA/BB= homozygotes major/minor allele. AB= heterozygous. Letters indicate means differing significantly between groups (HSD post-hoc test  $\alpha=5\%$ ). Group sizes are indicated between brackets for HSD post-hoc tests.

ANOVA (1 factor)

trait	2_22535054			3_4322444			3_10897844			5_4842835			5_5394803			6_15273858			7_18047191		
	p-value	SS (%)	groups	p-value	SS (%)	groups	p-value	SS (%)	groups	p-value	SS (%)	groups	p-value	SS (%)	groups	p-value	SS (%)	groups	p-value	SS (%)	groups
lgc13	ns	2.66		ns	6.06		ns	2.22		ns	5.37		0.05(.)	8.74		1.9 <sup>E-3</sup> **	17.58	a/b/b	ns	6.71	
lgc14	0.04*	10.85	a/b/b	ns	4.69		ns	3.18		ns	0.15		ns	2.44		ns	1.09		ns	3.41	
lgc15	ns			0.04*	12.56		0.08(.)	10.26		0.02*	16.10		0.05(.)	0.12		9.0 <sup>E-5</sup> ***	32.71		ns	0.49	
lgc16	ns	0.41		ns	1.29		ns	3.58		ns	5.51		ns	0.73		0.02*	14.27	a/b/ab	ns	2.03	
bs13	0.06(.)	8.28		ns	3.42		ns	6.13		ns	5.39		0.02*	11.15		0.06(.)	8.31		2.3 <sup>E-3</sup> **	17.00	
bs15	0.02*	16.20		1.4 <sup>E-4</sup> ***	31.56		0.07(.)	10.78		5.6 <sup>E-3</sup> **	19.81		2.0 <sup>E-3</sup> **	23.26		6.9 <sup>E-3</sup> ***	26.64		7.1 <sup>E-3</sup> **	19.00	
bs16	ns	5.57		ns	1.48		ns	0.80		ns	1.05		ns	7.23		ns	4.85		ns	6.38	

ANOVA (multi-locus) and mean comparison (HSD post hoc test α=5%)

trait	h <sup>2</sup> (%)																				
lgc13	17.58															1.9 <sup>E-3</sup> **	17.58	a/b/b			
lgc14	10.85	0.04*	10.85	a/b/b																	
lgc15	32.71															3.5 <sup>E-5</sup> ***	32.71	a/b/b			
lgc16	14.27															0.02*	14.27	a/b/ab			
bs13	39.21						0.05(.)	6.20	-				8.0 <sup>E-4</sup> ***	16.01	a/b/b				5.4 <sup>E-4</sup> ***	17.00	a/a/b
bs15	64.46				2.2e-6***	31.56	a/b/c			1.1 <sup>E-2</sup> *	8.81	a/b/b	1.8 <sup>E-4</sup> ***	18.65					0.05(.)	5.44	a/b/b

Group means

trait	h <sup>2</sup> (%)	AA	AB	BB	AA	AB	BB	AA	AB	BB	AA	AB	BB	AA	AB	BB	AA	AB	BB	AA	AB	BB
lgc13	17.58																-0.10 ± 0.16a(19)	0.02 ± 0.16b(41)	0.14 ± 0.22b (8)			
lgc14	10.85	0.02 ± 0.05a (24)	-0.01 ± 0.06b (33)	-0.03 b (1)																		
lgc15	32.71																-0.29 ± 0.28a (13)	0.08 ± 0.25b (31)	0.19 ± 0.23b (6)			
lgc16	14.27																0.12 ± 0.20a (15)	-0.06 ± 0.20b (36)	0.06 ± 0.08ab (6)			
bs13	39.21				0.11 ± 0.54 (34)	-0.07 ± 0.41 (24)	-0.21 ± 0.29 (10)				0.11 ± 0.46a (46)	-0.21 ± 0.45b (18)	-0.30 ± 0.52b (4)							-0.18 ± 0.44a (23)	-0.05 ± 0.44a (28)	0.33 ± 0.46b (17)
bs15	64.46				-0.34 ± 0.39a (16)	0.10 ± 0.43b (29)	0.51 ± 0.24c (5)			0.13 ± 0.38a (35)	-0.25 ± 0.56b (11)	-0.47 ± 0.54b (4)	0.17 ± 0.36a (32)	-0.28 ± 0.55b (14)	-0.37 ± 0.42b (4)					-0.28 ± 0.48a (17)	0.17 ± 0.41b (24)	0.08 ± 0.44b (9)

**Additional file 9 (.xls): Marker haplotype composition of the apricot core-collection considering all detected loci.** Groups of markers were constructed by considering all detected loci for each phenotype according to *G* and *G* × *Y* BLUPs. For markers, 0,1 and 2 are allelic codes according to minor allele dose. (i) Global *G* BLUPs from multi-year and within-year models associated with the haplotype. (ii) P-value of association between haplotype and trait. (iii) Number of accessions carrying the haplotype. (iv) Mean ( $\pm$  standard deviation) of associated trait per haplotype. (v) Post-hoc group mean comparison test (HSD  $\alpha=5\%$ ). HSD post-hoc tests were performed on *G* BLUPs from the multi-year and within-year models considering all haplotypes classes with  $n>2$ . Green and red cases: extreme group values related to partially resistant and susceptible accessions for each phenotype. For haplotypes carried by only one individual, assessment was performed in comparison with significantly different groups from HSD post-hoc tests. Bold values show outliers in individual haplotypes from the extreme groups (values over the mean  $\pm$  standard deviation of the extreme significant groups issued from HSD post-hoc tests).

lgc_G	Haplotype	Associated marker [chromosome number_physical position in bp] - SNP Polymorphism major/minor allele			ID (n=73)	lgc_G <sup>(i)</sup>			lgc2013			lgc2014			lgc2015			lgc2016		
		6_15273858 A/G	5_5394803 G/A	7_18047191 C/T		(iii)	(iv)	(v)	5.84e-03**			NS(0.81)			NS (0.1322)			NS (0.2279)		
1	0	0	0		A2204,A2217	2	-0.01 $\pm$ 0.05	abc	2	0.02 $\pm$ 0.22	abc									
2	0	0	1		A1314, A2129,A2894,A3862,A4374	5	-0.14 $\pm$ 0.20	abc	5	0 $\pm$ 0.22	abc									
3	0	1	0		<b>A1814, Harlayne</b>	2	-0.27 $\pm$ 0.03	bc	2	-0.15 $\pm$ 0.07	bc									
4	0	1	1		<b>A0008, A1793, SEO</b>	3	-0.34 $\pm$ 0.22	c	3	-0.22 $\pm$ 0.12	c									
5	0	1	2		A2156,A2219,A2346,A2348	4	-0.22 $\pm$ 0.10	bc	4	-0.13 $\pm$ 0.06	bc									
6	0	2	0		A1236	1	-0.22		1	-0.27										
7	0	2	1		A0804	1	-0.22		1	-0.20										
8	0	2	2		A0862	1	-0.35		1	-0.04										
9	1	0	0		A0010,A0544,A0654,A1453,A2311,A3408,A3865,A4294,Precoce	9	0.01 $\pm$ 0.16	abc	9	-0.02 $\pm$ 0.12	abc									
10	1	0	1		A0039,A0500,A0892,A1693,A1809,A1811,A2343,A2734,A3521,A3698,A4079,A4222,A4316,A4373,A	15	0.07 $\pm$ 0.18	ab	14	0.04 $\pm$ 0.11	abc									
11	1	0	2		A0634,A0682,A0882,A1267,A2069,A2382,A2662,A2928,A3522	9	0.05 $\pm$ 0.15	ab	8	-0.02 $\pm$ 0.18	abc									
12	1	1	0		A0300,A1343,A1956,A2067,A2340,A2635,A4322	7	0.08 $\pm$ 0.19	ab	5	0.17 $\pm$ 0.23	ab									
13	1	1	1		A1729	1	0.30		1	0.17										
14	1	1	2		A1813,A2137,A2335	3	-0.04 $\pm$ 0.14	abc	3	-0.08 $\pm$ 0.11	abc									
15	1	2	1		A0157	1	-0.11		1	-0.16										
16	2	0	0		A0539,A1711,A1939,A2205	4	0.24 $\pm$ 0.18	a	4	0.27 $\pm$ 0.21	a									
17	2	0	1		A0076,A0660,A2490,A3024	4	0.20 $\pm$ 0.15	a	3	0.07 $\pm$ 0.12	abc									
18	2	0	2		A2065	1	-0.08		1	-0.17										
TOT						73			68											

bs_G						bs_G	bs2013			bs2015			bs2016		
	A/G	G/A	T/A	T/C	C/T		0.02602*	0.07755.	2.05e-05 ***	NS (0.8859)					
1	0	0	1	0	1	A2894,A4374	2	-0.01 ± 0.00	ab	2	-0.25 ± 0.11	ab	0	NA	
2	0	0	1	1	0	A2217	1	0.23		1	0.26		0	NA	
3	0	0	1	1	1	A2129	1	-0.21		1	-0.43		1	-0.04	
4	0	0	2	0	0	A2204	1	-0.14		1	-0.41		1	0.05	
5	0	0	2	1	1	A1314,A3862	2	-0.23 ± 0.32	b	2	-0.03 ± 0.55	ab	2	-0.24 ± 0.42	bcd
6	0	1	0	0	2	A2346	1	0.63		1	0.85		0	NA	
7	0	1	0	1	1	A1793	1	-0.43		1	-0.53		1	-0.13	
8	0	1	0	1	2	A2348	1	-0.33		1	-0.31		0	NA	
9	0	1	1	1	1	SEO	1	-0.41		1	-0.60		1	0.33	
10	0	1	1	1	2	A2219	1	-0.53		1	0.10		1	-0.72	
11	0	1	1	2	1	A0008	1	-0.93		1	-0.66		1	-1.05	
12	0	1	2	0	0	A1814	1	-0.51		1	-0.52		1	-0.46	
13	0	1	2	0	2	A2156	1	-0.30		1	-0.26		0	NA	
14	0	1	2	1	0	Harlayne	1	-0.92		1	-0.36		1	-1.26	
15	0	2	0	0	1	A0804	1	-0.20		1	-0.35		1	-0.10	
16	0	2	0	0	2	A0862	1	-0.07		1	0.43		1	-0.53	
17	0	2	0	1	0	A1236	1	-0.87		1	-0.74		1	-0.89	
18	1	0	0	0	1	A4316	1	0.36		1	0.31		1	0.28	
19	1	0	0	0	2	A0682,A0882,A2069,A2382,A2662,A3522	6	0.10 ± 0.17	ab	5	0.80 ± 0.09	a	3	0.38 ± 0.21	ab
20	1	0	0	1	0	A0544,A4294,Precoce	3	0.22 ± 0.04	ab	3	-0.15 ± 0.41	ab	2	0.66 ± 0.16	a
21	1	0	0	1	1	A0892,A1693,A2343,A2734	4	0.10 ± 0.11	ab	4	0.08 ± 0.62	ab	2	0.60 ± 0.02	a
22	1	0	1	0	0	A0654,A3865	2	-0.12 ± 0.07	ab	2	-0.25 ± 0.18	ab	2	-0.19 ± 0.14	bcd
23	1	0	1	0	1	A0039,A0500,A3521,A4222,A4373	5	0.23 ± 0.11	a	5	0.05 ± 0.53	ab	4	0.56 ± 0.29	a
24	1	0	1	0	2	A0634	1	0.24		1	0.15		1	0.22	
25	1	0	1	1	0	A1453	1	0.14		1	0.77		1	-0.47	
26	1	0	1	1	1	A1809,A1811,A3698,A4804	4	-0.07 ± 0.08	ab	3	0.09 ± 0.31	ab	3	0.18 ± 0.06	abc
27	1	0	1	1	2	A2928	1	0.31		1	0.38		0	NA	
28	1	0	1	2	0	A0010	1	-0.27		1	-0.16		0	NA	
29	1	0	2	0	0	A3408	1	-0.53		1	-0.29		1	-0.50	
30	1	0	2	0	1	A4079	1	-0.08		1	0.19		1	-0.05	
31	1	0	2	1	0	A2311	1	-0.39		1	-0.37		0	NA	
32	1	0	2	1	2	A1267	1	0		1	0		1	-0.05	
33	1	1	0	0	1	A1729	1	0.60		1	0.30		1	0.68	
34	1	1	0	0	2	A1813	1	0.34		1	0.07		1	0.44	
35	1	1	0	1	0	A0300,A1343,A2067,A2340,A2635	5	0.08 ± 0.23	ab	4	-0.32 ± 0.53	b	3	-0.47 ± 0.07	d
36	1	1	0	1	2	A2137,A2335	2	-0.08 ± 0.19	ab	2	-0.17 ± 0.65	ab	1	0.16	
37	1	1	1	1	0	A1956,A4322	2	0.08 ± 0.04	ab	1	NA/-0.28		2	-0.26 ± 0.01	cd
38	1	2	0	1	1	A0157	1	-0.20		1	-0.54		1	0.04	
39	2	0	0	0	0	A1711	1	0.51		1	0.68		0	NA	
40	2	0	0	0	2	A2065	1	0.38		1	0.49		0	NA	
41	2	0	0	1	0	A1939	1	-0.01		1	0.26		1	-0.27	
42	2	0	0	1	1	A0076,A0660,A2490	3	0.17 ± 0.17	ab	3	0.16 ± 0.39	ab	3	0.13 ± 0.07	abc
43	2	0	1	0	0	A0539	1	-0.29		1	-0.68		1	0	
44	2	0	1	0	1	A3024	1	0.22		0	NA		1	0.28	
45	2	0	1	1	0	A2205	1	0.07		1	0.04		0	NA	
						TOT	73			68			50		



Les annexes ci-dessous sont indiquées à titre d'information mais n'apparaissent pas dans le manuscrit.

**Additional file 10 (.xls): Q ancestral covariates,  $G$ ,  $G \times Y$  BLUPs from the multi-year and  $G$  BLUPs from the within-year model of phenotypic data in the apricot core-collection.**

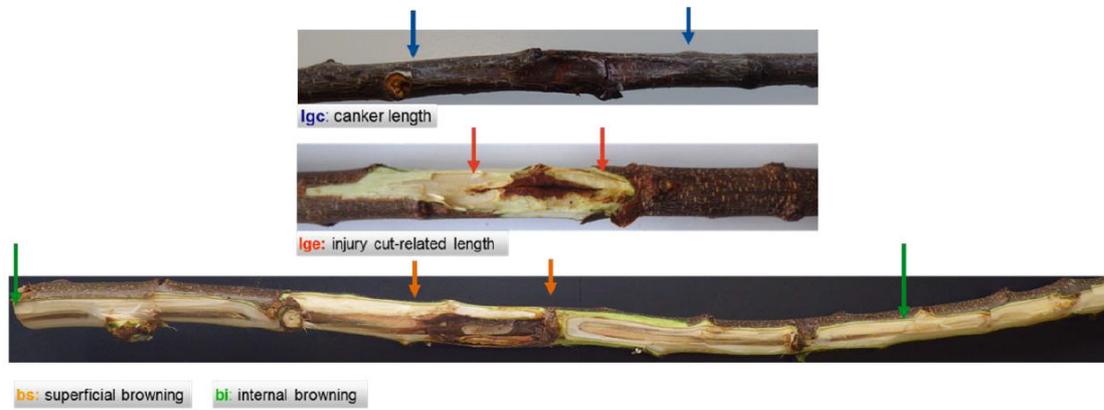
NA: missing values. Q1, Q2 and Q3 covariates are relative to ancestral fractions from (i) Central & Eastern Asia, (ii) Continental Europe and (iii) Irano-Caucasia & Mediterranean Basin, respectively.

**Additional file 11 (.xls): Genotypic data of the 63,236 SNPs used for GWAS.**

MAF: Minor Allele Frequency. Genotypes are coded according to minor allele dose. 0/2=homozygous major/minor allele. 1=heterozygous. Markers names are indicated as LG[chromosome number]\_[physical position in bp].

Annexe II : Figures et tables additionnelles de l'article de l'analyse de liaison sur la Berba et les trois populations interconnectées de la partie IV.

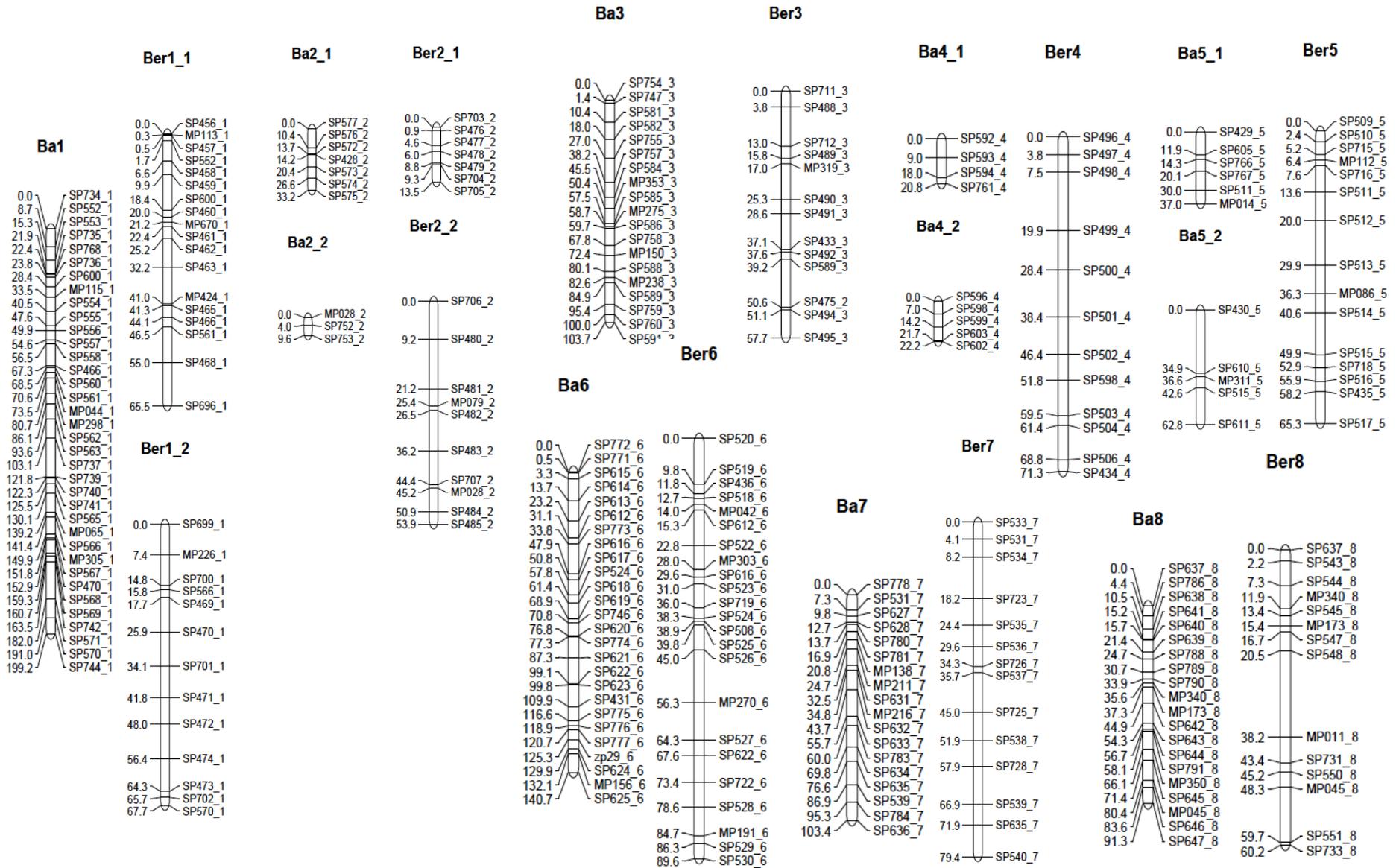
Additional file 1 : Photography of the symptoms issued from controlled inoculations in the orchard.



**Additional file 2 : Detailed statistics on the parental linkage maps of ‘Bergeron’ and ‘Bakour’ and the consensus map over the parental hybrids C092, D099 and C109.**

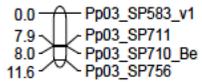
LG	length (cM)	density	STD	number of markers	gap max	coverage <i>Pp</i> <i>v2.0</i> physical	<i>Pp v2.0</i> length	% coverage
<b>Bergeron</b>								
1_1	65.5	3.64	3.47	18	10.5	27742082	47849303	<b>76.81%</b>
1_2	67.7	5.21	3.06	13	8.4	9009810	-	-
2_1	13.5	1.93	1.54	7	4.2	14457815	30405870	<b>73.55%</b>
2_2	53.9	5.39	4.00	10	12	7905105	-	-
3	57.7	4.44	3.81	13	11.4	23134557	27368013	<b>84.53%</b>
4	71.3	5.94	3.30	12	12.4	24567921	25843236	<b>95.07%</b>
5	65.3	4.35	2.86	15	9.9	18081048	18496696	<b>97.75%</b>
6	89.6	3.90	3.08	23	11.3	28778682	30767194	<b>93.54%</b>
7	79.4	5.67	2.42	14	10	21861418	22388614	<b>97.65%</b>
8	60.2	4.30	4.82	14	17.7	21543582	22573980	<b>95.44%</b>
<b>TOT</b>	<b>624.1</b>	<b>4.49</b>	<b>3.424023</b>	<b>139</b>		<b>197082020</b>	<b>225692906</b>	<b>87.32%</b>
<b>Bakour</b>								
1	199.2	5.53	4.39	36.0	18.7	47587188	47849303	<b>99.45%</b>
2_1	33.2	4.74	3.35	7.0	10.4	4780681	30405870	<b>17.01%</b>
2_2	9.6	3.20	1.13	3.0	5.6	390022	-	-
3	103.7	5.46	3.28	19.0	11.2	25340797	27368013	<b>92.59%</b>
4_1	20.8	5.20	3.58	4.0	9.0	779707	25843236	<b>57.12%</b>
4_2	22.2	4.44	3.37	5.0	7.5	13981810	-	-
5_1	37	6.17	3.68	6.0	11.9	4744941	18496696	<b>87.68%</b>
5_2	62.8	12.56	15.04	5.0	34.9	11472600	-	-
6	140.7	5.41	3.90	26.0	14.1	30207094	30767194	<b>98.18%</b>
7	103.4	5.74	3.29	18.0	12.0	22120109	22388614	<b>98.80%</b>
8	91.3	4.57	2.71	20.0	9.4	21970801	22573980	<b>97.33%</b>
<b>TOT</b>	<b>823.9</b>	<b>5.53</b>	<b>4.51</b>	<b>149</b>		<b>183375750</b>	<b>225692906</b>	<b>81.25%</b>
<b>Consensus map on the interconnected progenies</b>								
1	233.32	6.31	4.46	37	19.2	47587188	47849303	99.45%
2_1	54.21	6.78	5.32	8	15.8	22479251	30405870	<b>87.02%</b>
2_2	41.66	8.33	9.65	5	20.7	3979265	-	-
3	99.43	5.23	4.98	19	16.3	27156516	27368013	<b>99.23%</b>
4_1	21.09	2.64	3.33	8	7.9	12021313	25843236	<b>83.21%</b>
4_2	58.45	5.85	7.78	10	23.1	9482898	-	-
5	103.37	6.46	5.56	16	16.5	18081048	18496696	<b>97.75%</b>
6	147.29	6.14	5.92	24	15.9	30207094	30767194	<b>98.18%</b>
7	133.03	6.33	5.40	21	18.5	22116899	22388614	<b>98.79%</b>
8	78.25	5.59	6.56	14	22.2	21808392	22573980	<b>96.61%</b>
<b>TOT</b>	<b>970.1</b>	<b>5.99</b>	<b>5.50</b>	<b>162</b>	-	<b>214919864</b>	<b>225692906</b>	<b>95.23%</b>

Additional file 3 : Parental linkage maps of ‘Bergeron’ (‘Ber’) and ‘Bakour’ (‘Ba’)

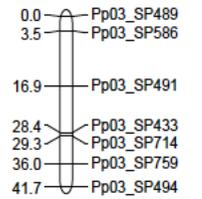




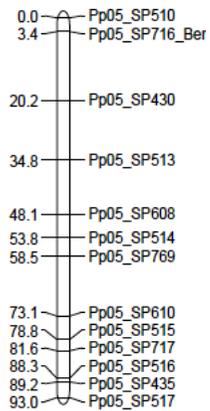
C092\_3\_1



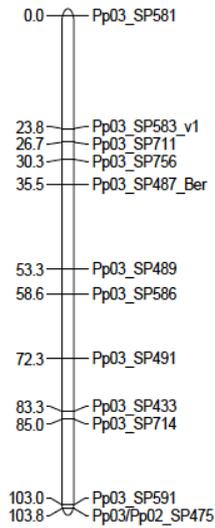
C092\_3\_2



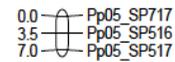
C092\_5



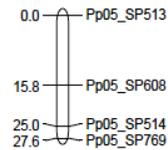
C109\_3



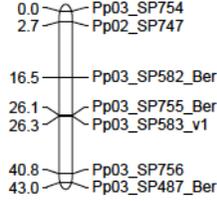
C109\_5\_1



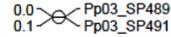
C109\_5\_2



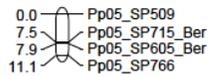
D099\_3\_1



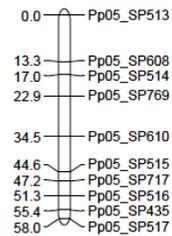
D099\_3\_2



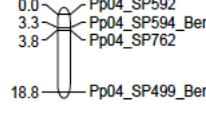
D099\_5\_1



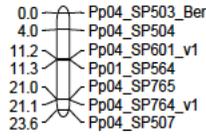
D099\_5\_2



C092\_4\_1



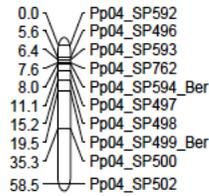
C092\_4\_2



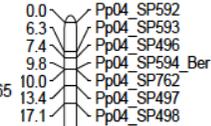
cons4\_1



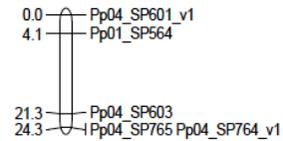
cons4\_2



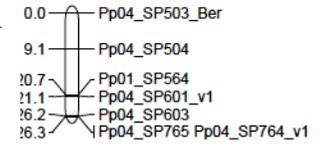
C109\_4\_1



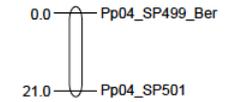
C109\_4\_2



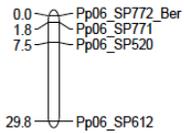
D099\_4\_1



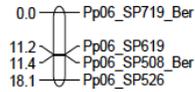
D099\_4\_2



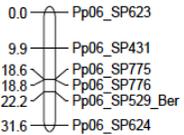
C092\_6\_1



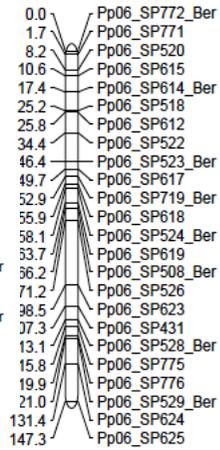
C092\_6\_2



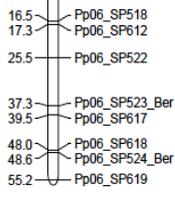
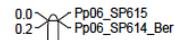
C092\_6\_3



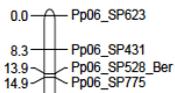
cons6



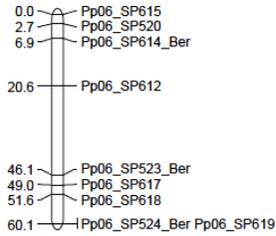
C109\_6\_1



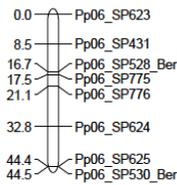
C109\_6\_2



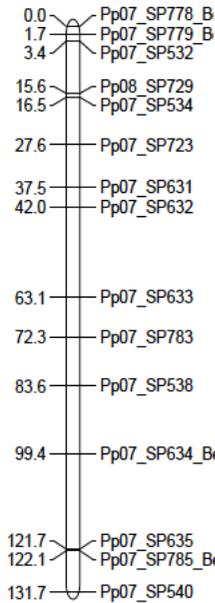
D099\_6\_1



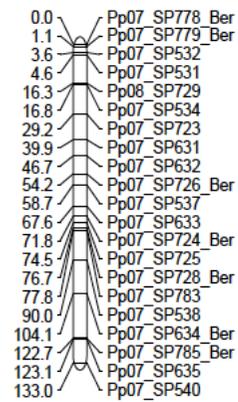
D099\_6\_2



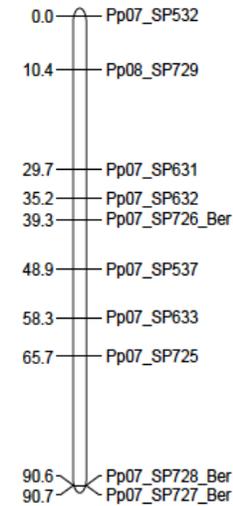
C092\_7



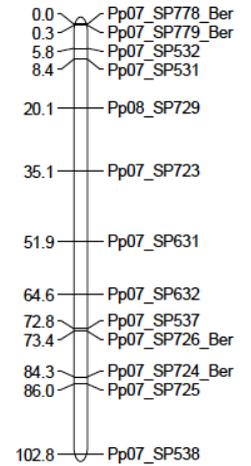
cons7



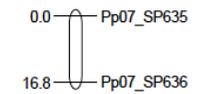
C109\_7



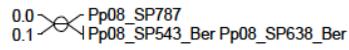
D099\_7\_1



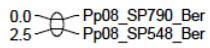
D099\_7\_2



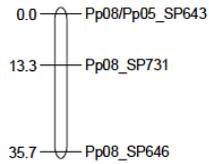
**C092\_8\_1**



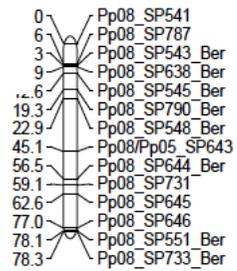
**C092\_8\_2**



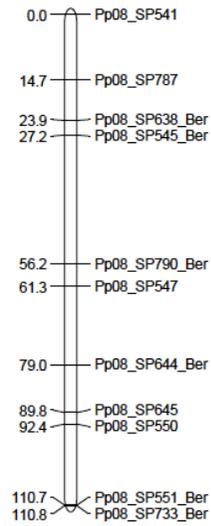
**C092\_8\_3**



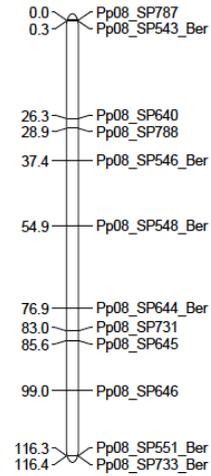
**cons8**



**C109\_8**

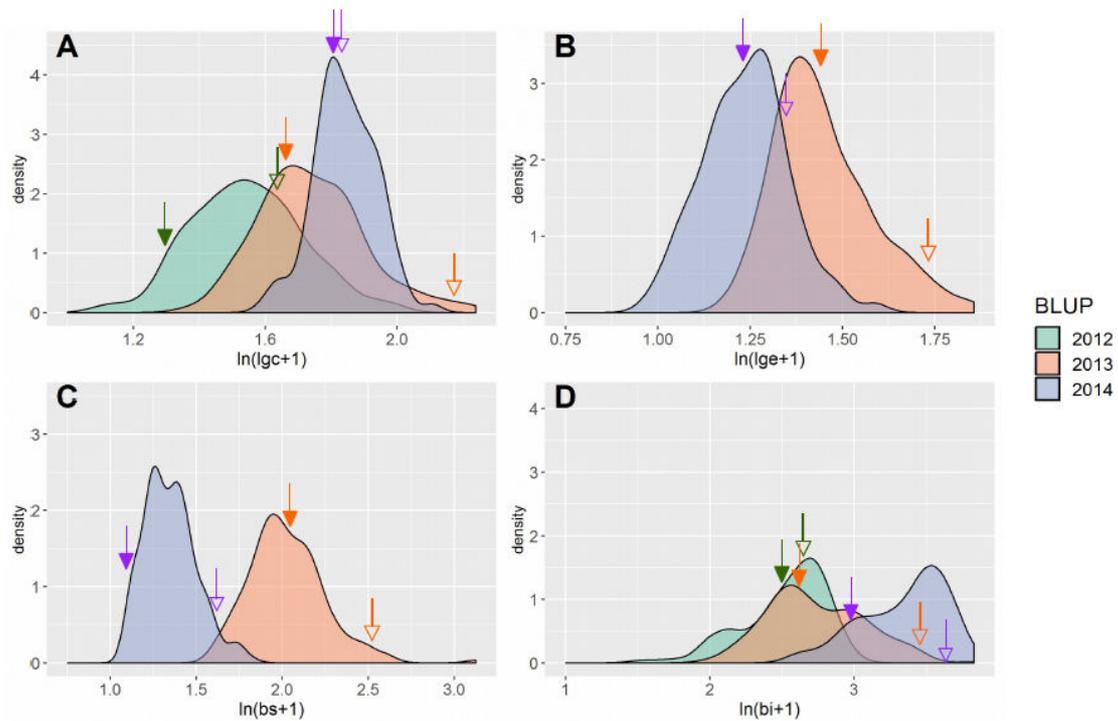


**D099\_8**

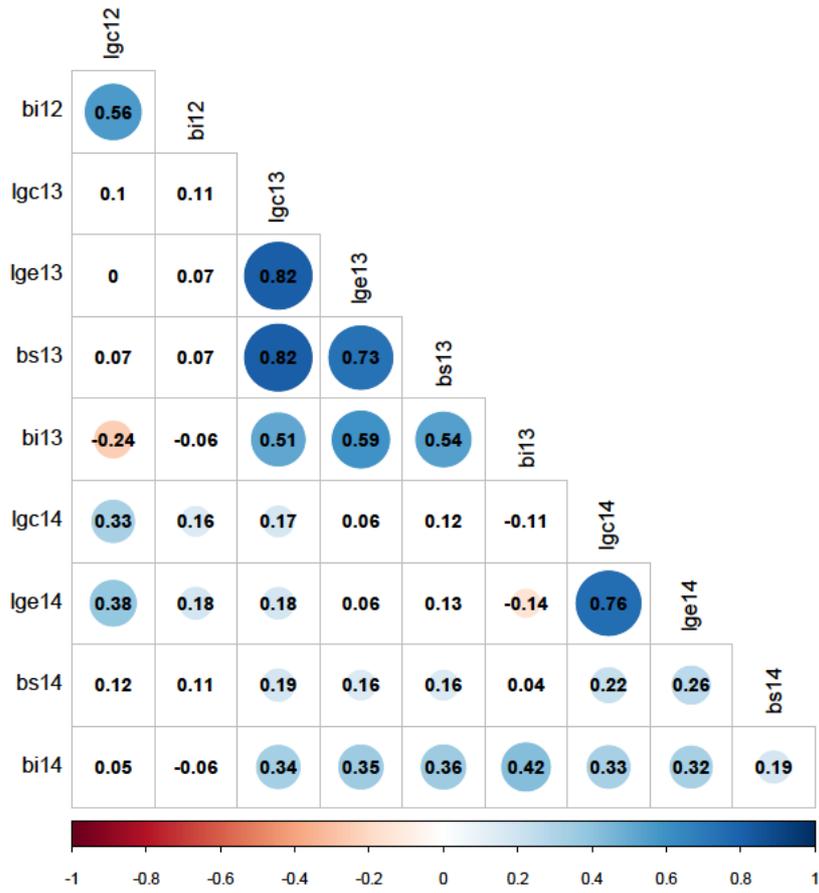


**Additional file 5: Annual genetic distributions of *lgc*, *lge*, *bs* and *bi* BLUPs in the Berba progeny evaluated between 2012 and 2014.**

Density plots of symptoms genetic BLUPs represented in 2012 (green), 2013 (orange) and 2014 (purple). All BLUP values are represented on a  $\ln(x+1)$  scale. Arrows on the x-axis are positioned on the average value of the parents 'Bakour' (filled arrow) and 'Bergeron' (empty arrow).



Additional file 6: Heatmap of within-year and between-years correlations on *lgc*, *lge*, *bs* and *bi* in Berba progeny over years 2012 to 2014



**Additional file 7.** QTLs detected in the three interconnected progenies using within ('disconnected') and across ('connected') family analyses. Genetic effects followed by ., \*, \*\* and \*\*\* are significant at P-value < 0.1, 0.05, 0.01 and 0.001, respectively by Student t-tests. For traits with multiple detected QTLs, R2 of the model are indicated in addition to individual contribution of QTLs (PVE).

Model		Marker										Additive Effects			Dominance Effect		
Trait	category	Year	-logPvalue	R <sup>2</sup>	-logPvalue	PVE	LG	Closest locus	Position into the consensus map (cM)	Support interval into the consensus map (cM)	Pop	Berba hybrid	A4049				
mortality	binary	2017-2018	3.9	0.25	3.9	0.25	3	Pp03_SP491	58.91	54.91-67.28	C092	-0.39 ± 0.21	.				
											C109	0.98 ± 0.33	**				
											D099	-1.46 ± 0.64	*				
lgc	continuous	2017	4.7	0.27	4.7	0.27	4_2	Pp04_SP503_Ber	0	exact position	C092	-0.08 ± 0.02	***	0.04 ± 0.01	**	-0.06 ± 0.01	***
											C109	-0.02 ± 0.02	ns				
											D099	-0.08 ± 0.02	***				
bs	continous	2017	4.4	0.20	4.4	0.20	4_2	Pp04_SP503_Ber	0	exact position	C092	-0.26 ± 0.08	**	0.20 ± 0.06	***	-0.23 ± 0.06	***
											C109	-0.10 ± 0.07	ns				
											D099	-0.37 ± 0.09	***				
bs	continuous	2016	7.5	0.17	5.0	0.10	4_2	Pp04_SP503_Ber	0	exact position	C092	-0.07 ± 0.02	**	0.03 ± 0.02	*	-0.07 ± 0.02	***
											C109	-0.05 ± 0.02	*				
					3.7	0.07	1	Pp01_SP562_Ber	116.55	26.34 - 144.15	C092	0.09 ± 0.04	**	-0.05 ± 0.01	***		
											C109	0.00 ± 0.01	ns				
											D099	-0.03 ± 0.02	ns				

		Marker										Additive Effects			Dominance Effect	
Trait	category	Year	-logPvalue	PVE	LG	Closest locus	Position into the consensus map (cM)	Support interval into the consensus map (cM)	Pop	Berba hybrid	A4049					
mortality	binary	2017-2018	3.18	0.13	8	Pp08_SP790_Ber	16.62	14.62-21.27	C092	-1.05 ± 0.32	**					
mortality	binary	2017-2018	3.40	0.16	3	Pp03_SP433	64.91	50.91-99.4	C109	1.01 ± 0.34	**					
mortality	binary	2017-2018	3.23	0.15	6	Pp06_SP625	147.25	87.17-147.25	D099	0.93 ± 0.41	*					
lgc	continuous	2017	3.80	0.17	4_2	Pp04_SP507	21.09	4.7-21.09	C092	-0.04 ± 0.01	**	-0.04 ± 0.01	**			
lgc	continuous	2017	4.6	0.20	4_2	Pp01_SP564	4.70	0-11.28	C109	-0.02 ± 0.01	ns	0.05 ± 0.01	**		-0.05 ± 0.02	***
bs	continous	2017	3.28	0.13	4_2	Pp04_SP764_v1	17.45	7.99-21.09	C092	0.01 ± 0.07	ns	-0.28 ± 0.07	***			
bs	continous	2017	5.6	0.20	4_2	Pp04_SP504	0.70	0-21.09	C109	-0.13 ± 0.06	*	0.30 ± 0.06	***			
bs	continous	2017	3.4	0.16	6	Pp06_SP518	25.25	21.41-38.36	D099	-0.26 ± 0.07	***					
bs	continuous	2016	3.5	0.16	1	Pp01_SP738	140.15	118.55-154.53	C092	0.06 ± 0.02	**	-0.07 ± 0.02	**	-0.07 ± 0.03	**	
bs	continuous	2016	4.9	0.21	4_2	Pp04_SP601_v1	9.28	0-15.28	C109	-0.06 ± 0.02	**	0.05 ± 0.02	**	-0.04 ± 0.02	*	

## Annexe III : Données additionnelles de l'étude de l'interaction GxG de la partie V.

### Annexe III-1 : Liste du matériel végétal exploité au niveau du dispositif factoriel 20 accessions abricotier x 9 souches *Psy*.

ID	Name	Origin
A0157	Rouge du Roussillon	France
A0660	Bergeron	France
A1145	SEO	USA
A1352	Polonais	France
A1814	Hargrand	Canada
A1939	Fantasme	France
A1963	Orangered Bhart	USA
A2137	Bakour	Tunisia
A2218	Goldrich	Canada
A2382	Stella	USA
A2914	Bergarouge	France
A2928	Early Blush	USA
A3325	2121xG1	France
A3749	Robada	USA
A3950	Bergeval Aviclo	France
A4025	Candide	France
A4034		
A4049		
A4445	Farbaly	France
A4481	Anegat	France

### Annexe III-2 : Protocole d'inoculations contrôlées sur boutures.

Rameaux dormants coupés à une taille de 18cm, conservant la partie centrale.

Paraffinage de la coupe supérieure.

Désinfection des boutures : 1 mn dans bain hypochlorite de sodium à 4°chlorométrique, puis, après rinçage à l'eau stérile, passage rapide dans l'alcool à 95°.

Rinçage et séchage sous la hotte à flux laminaire.

Placer les tiges dans des boîtes transparentes (6 tiges/boîte) avec du papier filtre stérile imbibé d'eau à la base. Avant l'inoculation, les boîtes ont été désinfectées en les plaçant sous les UV dans la hotte (temps de fonctionnement des UV: ¼ heure).

Inoculation (Une souche/boîte) : par une petite fente (1 cm environ) dans l'écorce avec suspension bactérienne *Psy* préparée à 10<sup>8</sup> CFU/ml, 20 µl. Entourer l'incision d'une bande de parafilm. Humidifier les papiers filtre avec 5 ml dans chaque boîte (1 ml au pied de chaque tige).