



HAL
open science

Rôle de l'architecture racinaire dans le contrôle génétique de la diminution des symptômes de pourriture racinaire dus à *Aphanomyces euteiches* chez le pois (*Pisum sativum*)

Aurore Desgroux

► **To cite this version:**

Aurore Desgroux. Rôle de l'architecture racinaire dans le contrôle génétique de la diminution des symptômes de pourriture racinaire dus à *Aphanomyces euteiches* chez le pois (*Pisum sativum*). Biologie végétale. Agrocampus Ouest, 2016. Français. NNT : 2016NSARA076 . tel-01615191v3

HAL Id: tel-01615191

<https://theses.hal.science/tel-01615191v3>

Submitted on 17 Oct 2017

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.

AGRO CAMPUS

OUEST

Aurore DESGROUX • 31 mars 2016

Thèse AGROCAMPUS OUEST
sous le label de l'Université Bretagne Loire
pour obtenir le grade de
DOCTEUR D'AGROCAMPUS OUEST
Spécialité Biologie et Agronomie

ÉCOLE DOCTORALE • Vie-Agro-Santé (VAS)
LABORATOIRE D'ACCUEIL • Institut de génétique,
environnement et protection des plantes (IGEPP)

Rôle de l'architecture racinaire dans le contrôle génétique de la diminution des symptômes de pourriture racinaire dus à *Aphanomyces euteiches* chez le pois (*Pisum sativum*)

Maria MANZANARES-DAULEUX
Professeure, AGROCAMPUS OUEST, UMR INRA-AO-UR1 IGEPP /
présidente

Xavier DRAYE
Professeur, Université Louvain, Belgique / *rapporteur*

Ivan SACHE
Professeur, AgroParisTech / *rapporteur*

Mathilde CAUSSE
Directrice de recherche, INRA Avignon / *examinatrice*

Maxime BONHOMME
Maître de conférences, Université Toulouse / *examinateur*

Marie Laure PILET-NAYEL
Chargée de recherche, HDR, INRA Rennes / *directrice de thèse*

Virginie BOURION
Ingénieur de recherche, INRA Dijon / *co-directrice de thèse*



N° ordre : 2016-11
N° Série : A-76

THESE / AGROCAMPUS OUEST

Sous le label de l'Université Européenne de Bretagne
pour obtenir le diplôme de :

**DOCTEUR DE L'INSTITUT SUPERIEUR DES SCIENCES AGRONOMIQUES,
AGRO-ALIMENTAIRES, HORTICOLES ET DU PAYSAGE**

Spécialité : Biologie et Agronomie

Ecole Doctorale : Vie-Agro-Santé

présentée par :

Aurore DESGROUX

**RÔLE DE L'ARCHITECTURE RACINAIRE DANS LE CONTRÔLE GÉNÉTIQUE DE
LA DIMINUTION DES SYMPTÔMES DE POURRITURE RACINAIRE DUS À
APHANOMYCES EUTEICHES CHEZ LE POIS (*PISUM SATIVUM*)**

soutenue le 31 mars 2016 devant la commission d'Examen

Composition du jury :

Maria MANZANARES- DAULEUX	Professeure, AGROCAMPUS OUEST, UMR INRA-AO-UR1 IGEPP	Présidente
Xavier DRAYE	Professeur, HDR, Université Louvain, Belgique	Rapporteur
Ivan SACHE	Professeur, HDR, AgroParisTech	Rapporteur
Mathilde CAUSSE	Directrice de Recherche, HDR, INRA Avignon	Examinatrice
Maxime BONHOMME	Maître de conférences, Université Toulouse	Examinateur
Marie- Laure PILET-NAYEL	Chargée de recherche, HDR, INRA Rennes, UMR INRA-AO-UR1 IGEPP	Directrice de thèse
Virginie BOURION	Ingénieur de recherche, INRA Dijon	Co-directrice de thèse



Remerciements

Ma thèse a été une grande aventure, pleine de rebondissements, mais surtout de rencontres et de découvertes, qu'elles soient scientifiques ou humaines.

Dans un premier temps, je souhaite remercier du fond du cœur Philippe Declerck, sélectionneur de pois protéagineux. Philippe, lors de notre rencontre à la ferme expérimentale de Milly-la-Forêt, je découvrais alors le monde de la sélection. J'ai intégré ton équipe en tant que saisonnière pour les récoltes, et voilà que je suis repartie de cette saison 2007 avec pour but d'intégrer une école d'ingénieur pour travailler dans l'amélioration des plantes ! Sélectionneur passionné, tu as pris le temps de m'expliquer les bases de la sélection et les enjeux de la culture du pois protéagineux, tout en me transmettant ta passion. Par la suite, lorsque j'ai voulu faire une thèse, tu as su me diriger vers les bonnes personnes, afin de monter un projet. Sans cette rencontre, je ne serais probablement pas ici aujourd'hui. Pour tout cela, merci !

Merci au Ministère de l'agriculture, de l'alimentation et de la forêt pour m'avoir permis de bénéficier du dispositif de formation complémentaire par la recherche, qui permet aux agents du MAAF de suivre une thèse dans le cadre de leur fonction par un détachement. Ce dispositif est riche d'enseignement et permet aux agents d'acquérir rigueur et connaissances scientifiques, afin de mieux dialoguer avec l'ensemble des interlocuteurs du MAAF. Plus particulièrement, merci à Cyril Kao et Alexandre Péry d'avoir suivi mes travaux et de m'avoir épaulée pour la suite de ma FCPR.

Merci également à Terre Univia, les projets ANR Investissement d'avenir PeaMust et Genoepa, ainsi qu'au projet INRA-CropDL (métaprogramme Selgen) d'avoir financé mes travaux de recherche, mes déplacements, mes formations, mes publications et ma participation à des réunions ou conférences.

Merci également aux membres de mon comité de thèse, Maxime Bonhomme, Loïc Pages, Alain Baranger, Anne Moussart, Gérard Duc, Mathieu Siol et Philippe Leterme, pour avoir évalué mon travail et pour vos conseils avisés sur l'évolution de mes recherches.

Un énorme MERCI à ma directrice de thèse et à mes co-encadrantes : Marie-Laure, Virginie et Maria. Marie-Laure, merci pour ton encadrement quotidien, ta rigueur, tes conseils avisés, ton écoute et ta disponibilité sans failles, mais également pour ta bonne humeur et tes encouragements ! Virginie, merci également pour ton encadrement, tes conseils avisés et tes encouragements, pour ton accueil toujours chaleureux et ta disponibilité lors de mes visites à Dijon, ainsi que ton suivi régulier malgré la distance ! Grâce à toi j'ai bien compris que l'observation des racines n'est pas une mince affaire, et surtout que c'est un travail d'équipe !! Maria, merci pour ta disponibilité, tes relectures et tes encouragements tout au long de ces 3 années ! A toutes les trois, cela a été un vrai plaisir de travailler avec vous, et de découvrir ainsi le monde de la recherche ! J'ai énormément appris avec vous et je tiens à souligner la qualité de l'encadrement et de la formation que vous m'avez fournis ! Je vous souhaite bonheur et réussite dans vos projets respectifs.

Merci à l'équipe Résistance et adaptation (Rennes), ainsi qu'à l'équipe Agroécologie (Dijon) que j'ai intégrées pour ces 3 années. Merci à tous pour votre accueil et votre gentillesse. Une pensée particulière à ceux qui ont travaillé de très près ou d'un peu plus loin sur les données de génétique et de phénotypage que j'ai acquises et/ou utilisées. Pour le génotypage : Grégoire Aubert, Nadim Tayeh, Françoise Jacquin, Christophe Piriou, Virginie L'Anthoëne, Gilles Boutet. Pour le phénotypage au champ : Jean-Philippe Rivière, Pierrick Vetel, Pierre Mangin, Norbert Blanc, Henri de Larambergue, Rodolphe. Pour le phénotypage et conditions contrôlées : Gwenola Leroy, Valentin Baudais (stagiaire

M1), Henri Miteul, Véronique Aubert, Marianne Martinello. Pour l'ensemble du phénotypage, merci aux équipes des unités expérimentales de la Motte et d'Époisses, ainsi que les saisonniers qui ont participé aux différents chantiers de lavage et/ou d'étalement des racines, et sans qui ce travail n'aurait pas été possible.

Je remercie également tous ceux qui ont permis mon accueil dans l'UMR IGEPP et le pôle GEAPSI. Merci à Denis Tagu, Régine Delourme et Alain Baranger de m'avoir accueillie sur le site du Rheu. Merci à Christophe Salon, Judith Burstin et Gérard Duc de m'avoir accueillie sur le site de Dijon. Merci également à Pascal Le Neve et Maryline Robert de m'avoir appuyé pour les aspects administratifs tout au long de ma thèse.

Je tiens également à remercier à l'équipe du Service régional de l'alimentation de la DRAAF Midi-Pyrénées à Toulouse qui m'a fait un très bon accueil en septembre 2015, soit six mois avant ma soutenance. Plus spécialement, un énorme merci Guillaume Girard et Catherine Pavé, pour votre patience, votre gentillesse, vos encouragements et votre compréhension lors de cette période de « double-emploi » ! Guillaume, tu as toujours été vigilant et attentif à ma situation, ce qui m'a permis de prendre mes fonctions à la DRAAF tout en finissant la rédaction de mon manuscrit de façon sereine ! Merci pour tout, je souhaite à tout le monde d'avoir un chef au moins aussi génial ! Merci à tout le 7^{ème} étage de la tour infernale pour votre accueil et votre soutien moral !

Mon expérience dans la recherche n'a pas débuté avec ma thèse, mais bel et bien lors de mon stage de fin d'école d'ingénieur, au sein de l'UMR Génétique et amélioration des plantes de l'INRA d'Avignon. Un grand merci à Mathilde Causse pour son encadrement, et plus particulièrement à Laura Pascual qui m'a appris les ficelles du codage sous R, ainsi que la rigueur lors des analyses statistiques et des expérimentations. Laura, en plus d'être une très bonne formatrice, tu es également une très bonne amie, je viendrai bientôt te voir à Madrid ☺ !! Merci également à Justine Gricourt, rencontrée à l'occasion de ce stage et qui m'a énormément soutenue tout au long de ma thèse ! Justine, nos *road trip* et *rando sac à dos* bretons resteront à jamais gravés dans ma mémoire ! Bientôt les Pyrénées et la Corse, une fois que bébé aura grandi un peu !



Que seraient des remerciements sans dument remercier le lieu incontournable de discussions improbables en tous genres (scientifiques ou non), de découvertes culinaires, de débats passionnés sur les derniers films et séries, d'accueil de tous les scientifiques en herbes, stagiaires, techniciens, doctorants, post-doctorants et autres ? J'ai nommée, la Pause BM ! Merci à tous pour tous ces bons moments, ces fous rires et d'avoir supporté mes multiples personnalité parfois déjantées (« *Vous êtes au moins 10 là-haut !?!* »). Spéciale dédicace à Anne Levrel, Laure, Laurie, Jocelyne, Clément L, Angélique, Christophe, Cécile, Françoise, Patrick, Philippe, Gwenaëlle, Émilie, Hélène, Marie-Odile, Sophie R, Sophie P, Solenn, Vanessa, Virginie, Maryse, et ceux que j'ai oublié... car malheureusement il doit y en avoir ! Merci pour ces moments magiques et votre soutien !

Pas de remerciements non plus sans la magnifique équipe des doctorants !!! Pour la plupart regroupés dans le BDF (Bureau du fond), les doctorants forment une équipe soudées qui se serre les coudes, que ce soit pour s'entraider au travail, ou pour boire une bière à l'Artiste !

Merci Carole pour les moments passés ensemble malgré ta fin de thèse lors de mon arrivée ! Je n'ai pas suivi le pas sur la course à pied, mais j'ai retenu tout de même une leçon qui m'a permis de tenir sur la fin : conserver une activité physique, même en période d'*over booking* est super important pour le moral !!

Merci à la génération 2011-2014 ! Pascal, Clément S, Berline, les trois compères de choc ! Pascal, toujours d'un calme olympien, toujours un pull sur les épaule même avec 35°C, toujours un truc à bricoler (« *On pourra faire un diner à la maison, quand j'aurai enlevé le moteur de ma voiture du milieu du salon !* »), ou en train de tester un nouveau « truc » pour fabriquer tes bières ! Clément, pour ton cynisme et ton attitude râleuse qui cache en fait un mec super, plein de second degré et de sympathie ! Berline, vraiment beaucoup plus qu'une collègue, j'ai découvert une véritable amie ! Pour ta bonne humeur, ton dynamisme, ta franchise, ton soutien, tes chansons et tant d'autres choses ! À tous les trois, MERCI !!! On se revoit très bientôt de toute façon ☺🙏 !!

Merci également à mes compères de promo Anne-So, Clément L et Séverine! Clément, pour ta bonne humeur et ton énergie, même si parfois tu ne sais pas t'arrêter ;). Anne-So, pour ta gentillesse, ton écoute, ta bonne humeur, ta patience, nos *tea times* et les sorties à l'Artiste !

Merci à la relève, Sylvain, Yoann et Alexandre ! Courage ça vaut le coup ! Et encore merci pour votre photo soutien pour m'encourager, même à 800km de distance ! ☺

D'autres doctorantes, même si elles ne sont pas de l'équipe, Morgane Gillard et Auriane De tonnac ! Notre passage à l'Enita de Clermont nous aura permis de nous rencontrer et de continuer à nous voir sur Rennes, chacune sur notre thème de recherche, mais toujours avec des discussions hyper intéressantes et des soirées au top ! Courage à toutes les deux qui êtes maintenant sur la dernière ligne droite ! Et merci pour votre soutien, surtout sur la fin ! La Jussie et les Oméga 3 dans le porc n'ont qu'à bien se tenir ☺

Un ENORME MERCI à mes coloc du 28 Ter ! Enora, Quentin, Cédric, Thomas, Malik, Jimmy, merci de m'avoir supporté pendant ces deux années ! J'ai passé de super moments avec vous et votre compagnie m'a bien aidée !! Enora, pour ton optimisme et ton lâché prise. Quentin, nous n'avons pas vécu longtemps ensemble, mais j'ai passé de super moments sur la terrasse ! Bon courage pour ta thèse de médecine !

Les copains de l'Élabo, merci pour les super moments que j'ai passé parmi vous et qui m'ont permis de décompresser ! Une spéciale dédicace à Romu, frangin, ne lâche rien ! Une autre à Yo, Jo et leur p'tit Balou ! Ça a été un plaisir de garder votre petit bout ! Greg, Sly, Judith, Aude, Melenn et tous les autres, merci pour les soirées spectacles et les autres ! Peut-être se croiera-t-on à Aurillac ou ailleurs !

Laurent, merci pour ton soutien, tes encouragements et ta patience ! Merci également à toute ta famille !

Doumi, malgré la distance, ton soutien et ton amitié ont été indispensables pour le moral ! Nos petits week-ends morlaisiens me laissent des souvenirs de trop peu ! Merci pour tout !

Chers coloc et amis toulousains, un grand merci pour votre accueil et votre soutien qui ont été indispensables à la réussite de cette thèse ! Seb, Maggy, Nico et Mélina, Merci de me supporter au quotidien et de m'avoir encouragé sur la fin, malgré mon « *sacré caractère* » =D. Ketty, Boris, Philippe, Marion, Géraldine, Emmanuel, Fabrice, Eve, Etienne, merci à tous ceux que j'ai rencontré et qui mon fait un accueil incroyable dans cette toute nouvelle ville. Merci pour votre soutien sans faille !

Enfin, je finirai ces remerciements par mes parents ! Merci à vous de m'avoir soutenue, accompagnée et d'avoir toujours été présents lors de mon parcours ! Votre soutien et votre amour me sont précieux !

Une page se tourne, mais le livre reste ouvert

Chapitres passés, présents ou futurs, vous faites tous partis de mon livre à jamais

∞ Pour que tous vous vous sentiez heureux comme des petits pois dans leur gousse ∞

Table des matières

Introduction générale	1
Chapitre I : Introduction bibliographique	3
1. La résistance aux maladies et l'architecture des plantes	3
1.1. Quelques notions sur la résistance des plantes aux maladies	3
1.1.1. Résistances qualitative et quantitative.....	3
1.1.2. Les résistances partielles sous contrôle polygénique.....	4
1.1.3. Mécanismes de résistance des plantes aux maladies	5
1.2. L'architecture des plantes peut-elle être utilisée comme un levier pour lutter contre les épidémies ?.....	11
1.2.1. Etude de l'architecture des plantes.....	11
1.2.1.1. Variabilité et plasticité de l'architecture des plantes	11
1.2.1.2. L'architecture racinaire : quels moyens de mesure pour un compartiment difficile d'accès ?	13
1.2.1.3. Le contrôle génétique de l'architecture des plantes	14
1.2.2. Architecture et maladies des plantes	15
1.2.2.1. Cycle épidémique d'un agent pathogène et interaction avec la plante hôte	15
1.2.2.2. Liens entre architecture de la plante et réduction des épidémies.....	15
1.2.2.3. Lien entre architecture racinaire et maladies telluriques	19
1.2.2.4. Interdépendance des déterminismes génétiques de résistance et d'architecture des plantes et construction d'idéotypes	21
2. Dissection génétique des caractères complexes chez les végétaux.....	23
2.1. Les marqueurs moléculaires	23
2.2. Analyse de liaison : approche QTL	23
2.3. Analyse de déséquilibre de liaison : approche génétique d'association	27
2.4. LD-LA : approche mixte	29
3. Le pathosystème Pois/Aphanomyces euteiches	31
3.1. Le pois protéagineux, une culture d'avenir pour la production de protéines végétales en Europe	31
3.1.1. Enjeux et intérêts de la culture de pois	31
3.1.2. Production et rendement.....	33
3.1.3. Types de culture	33
3.1.4. Ravageurs et maladies du pois protéagineux	35
3.1.5. Caractéristiques biologiques du pois protéagineux	37
3.1.6. Ressources biologiques et génomiques pour l'amélioration génétique du pois	39
3.1.6.1. Ressources génétiques	39

3.1.6.2. Ressources génomiques	39
3.2. <i>A. euteiches</i> , un agent pathogène majeur du pois	41
3.2.1. Symptômes et importance économique	41
3.2.2. Caractéristiques de l'agent pathogène	43
3.2.2.1. Classification	43
3.2.2.2. Cycle biologique	43
3.2.2.3. Spécificité d'hôte et pathotypes	44
3.2.3. Méthodes de lutte	44
3.3. Leviers génétiques chez le pois pour lutter contre <i>A. euteiches</i>	47
3.3.1. Sources de résistance et sélection	47
3.3.2. Diversité des locus de résistance à <i>A. euteiches</i> chez le pois	48
3.3.3. Mécanismes sous-jacents aux QTL de résistance	49
3.3.4. Relation entre résistance à <i>A. euteiches</i> et architecture racinaire	51
4. Objectifs et stratégie du travail de thèse	53

Chapitre II : Dissection du déterminisme génétique de la résistance partielle du pois à

A. *euteiches* par génétique d'association..... 57

1. INTRODUCTION DU CHAPITRE	57
2. ARTICLE ASSOCIÉ	57
1. Abstract	61
2. Background	62
3. Methods	65
3.1. Plant material	65
3.2. Phenotyping	66
3.3. Genotyping	67
3.4. Statistical analysis of phenotypic data	68
3.5. Genetic analysis	69
3.5.1. Missing data imputation	69
3.5.2. Linkage disequilibrium analysis	69
3.5.3. Population structure and individual relatedness	70
3.5.4. Association mapping	70
3.5.5. Haplotype analysis	71
3.5.6. Comparative mapping	72
3.5.7. Identification of putative candidate genes	72
4. Results	73
4.1. Analysis of phenotypic data	73
4.1.1. Resistance to <i>A. euteiches</i>	73

4.1.2. Earliness and height	73
4.1.3. Correlations among variables.....	73
4.1.4. Multiple Factors Analysis	75
4.2. Genetic analyses.....	77
4.2.1. Linkage disequilibrium, structure, and kinship in the pea-Aphanomyces collection.....	77
4.2.2. GWA markers and confidence intervals	77
4.2.2.1. Disease resistance	77
4.2.2.2. Linkage between resistance and morphological or developmental traits.....	79
4.2.3. Marker Haplotypes	83
4.2.4. Putative candidate genes	84
5. Discussion	85
5.1. GWA study validated most of the previously detected QTL and identified new resistance loci	86
5.2. GWA study detected resistance loci with a higher resolution than in previous linkage mapping studies	88
5.3. GWA study provides new tools for pyramiding Aphanomyces resistance alleles in breeding programs.....	89
5.4. GWA study identified relevant putative candidate genes underlying Aphanomyces resistance QTL	91
6. Conclusion	92

Chapitre III : Analyse génétique comparative de l'architecture constitutive de la plante et de la résistance à *Aphanomyces euteiches* chez le pois.....103

1. INTRODUCTION DU CHAPITRE	103
2. ARTICLE ASSOCIÉ.....	103
1. Introduction.....	105
2. Material and method	108
2.1. Plant material.....	108
2.2. Phenotyping	108
2.3. Statistical analysis of phenotypic data	110
2.4. Genetic analysis.....	110
2.4.1. Genotyping and consensus map	110
2.4.2. Population structure and individual relatedness and genome wide association	112
2.4.3. Comparative mapping.....	112
3. Results	113
3.1. Analysis of phenotypic data	113
3.2. Genetic analysis.....	119
3.2.1. Linkage disequilibrium, structure, and kinship in the 266-pea-line collection	119

3.2.2. Genome wide association results and comparative mapping	123
4. Discussion	131
4.1. A wide genetic diversity for plant architecture and resistance to <i>A. euteiches</i> in the 266-pea-line collection	131
4.2. Root growth and architecture traits correlated with resistance to <i>A. euteiches</i> in pea.....	132
4.3. Multiple genetic determinants of plant architecture poorly closely-linked to <i>Aphanomyces</i> resistance loci	133
4.4. A major locus associated with resistance and intrinsic root system architecture	133
5. Conclusion	134

Chapitre IV : Architecture induite en réponse à l’infection et résistance à *Aphanomyces euteiches* chez le pois: analyse génétique comparative et rôle dans l’amélioration de la tolérance 135

1. INTRODUCTION DU CHAPITRE	135
2. ARTICLE ASSOCIÉ.....	136
1. Introduction.....	137
2. Material and methods.....	139
2.1. Plant material.....	139
2.2. Controlled conditions experiment	139
2.2.1. Phenotyping in controlled conditions.....	139
2.2.2. Statistical analysis of controlled condition data	140
2.3. Genetic analysis and comparative mapping	141
2.4. Field experiment	141
2.4.1. Field assay	141
2.4.2. Statistical analysis of field experiment	142
3. Results	143
3.1. Controlled conditions experiment	143
3.1.1. Analysis of phenotypic data	143
3.1.2. Correlations among variables.....	149
3.1.3. Genome wide association study for plant response to <i>A. euteiches</i> infection	151
3.1.4. Comparative mapping with previous studies for plant architecture and <i>Aphanomyces</i> resistance	153
3.2. Field experiment of the Apharchi pea line-subset.....	155
3.2.1. Subset selection for field experiment.....	155
3.2.2. Level of soil infection by <i>A. euteiches</i>	157
3.2.3. Disease impact on pea root and shoot architecture	159
3.2.4. Disease impact on pea yield and its components	161

4. Discussion	161
4.1. Plant architecture induced by infection is highly correlated with resistance and less related to intrinsic architecture	161
4.2. Most of genomic regions controlling disease-induced plant architecture were common to those controlling resistance but different from those implied in intrinsic plant architecture.....	163
4.3. Resistance and disease-induced plant architecture were related to a lower impact of <i>Aphanomyces</i> root rot on yield in pea.....	165
5. Conclusion	165
Chapitre V : Synthèse, discussion générale et perspectives	167
1. Synthèse des résultats de la thèse	168
1.1. Variabilité génétique et corrélations entre variables de résistance à <i>A. euteiches</i> et d'architecture racinaire de la plante chez le pois.....	169
1.2. Cartographie comparée des locus de résistance à <i>A. euteiches</i> et d'architecture racinaire de la plante	171
1.3. Exploitation de la variabilité génétique pour diminuer les pertes de rendement	172
2. Apports des résultats de la thèse en recherche et en sélection.....	172
2.1. Relation entre architecture racinaire et résistance chez les pathosystèmes racinaires	173
2.2. Hypothèses sur les mécanismes biologiques sous-jacents aux QTL de résistance aux maladies chez les plantes.....	175
2.3. Combinaison d'allèles de résistance et d'architecture pour la sélection	176
3. Perspectives de recherche.....	178
3.1. Améliorer la qualité la précision et le débit du phénotypage.....	179
3.2. Préciser les liens génétiques entre locus de résistance et d'architecture	180
3.3. Explorer les mécanismes moléculaires associés à la résistance et à l'architecture	181
4. Conclusion générale	181
Références bibliographiques	183
Annexes	195

Liste des abréviations

(CC-)RRI ou INR	Controlled conditions root rot index
2D	Deux dimensions
3D	Trois dimensions
ACP ou PCA	Analyse en composante principale
ADN ou DNA	Acide désoxyribonucléique
AFLP	Amplified fragment length polymorphism
BC	Backcross
Br:TProjArea	Percentage of brown root projected area
CC- ou TIR-NBS-LRR	Coiled coil ou Toll/interleukine nucleotid bindin site leucine riche repeat
CI	Confidence interval
CIM	Composite interval mapping
DI	Dijon-Epoisses
DIA	Disease-induced architecture
DL ou LD	Déséquilibre de liaison
DS	Disease severity
EMMA	Efficient mixed-model association
ET	Ethylène
ETI	Effector triggered immunity
FDR	False discovery rate
GAPIT	Genome association and prediction integrated tool
GBS	Genotyping-by-sequencing
GEAPSI	Déterminismes génétiques et environnementaux de l'adaptation des plantes à des systèmes de culture innovants
GSP	Groupement des sélectionneurs de protéagineux
GWA(S)	Genome wide association (study)
HD	Haploïde doublé
IDA ou ADI	Index de dépérissement aérien
IGEPP	Institut de génétique, environnement et protection des plantes
INRA	Institut national de recherche agronomique
JA	Acide jasmonique
JLAM	Joint linkage association mapping
LAI	Leaf area index
LatRootL	Average lateral root length
LD-LA	Linkage desequilibrium-Linkage analysis
LG	Linkage group
LM	Linear model
LOD score	Logarithm of Odds score
LSMeans	Least square means
MAAF	Ministère de l'alimentation, de l'agriculture et de la forêt
MAF	Minor allele frequency
MAGIC	Nested association mapping
MFA	Multiple factor analysis
MLMM	Multi-locus mixed model
MQTL	Méta-QTL
NAM	Multi-parent advanced generation inter-cross

NGS	Next generation sequencing
NIL	Near isogenic line
NLatRoot	Number of lateral roots
Nnode	Number of nodes
Npod	Number of pods
Npod.PodNod	Number of pods per podded node
NPodNod	Number of podded nodes
Nseed	Number of seeds
Nseed.Pod	Number of seeds per pod
PAC	Politique agricole commune
PAMP	Pathogen associated molecular pattern
PI	Potentiel infectieux
PTI	PAMP triggered immunity
qRT-PCR	Quantitative reverse transcription polymerase chain reaction
QTL	Quantitative trait locus
REN	Rennes-Le Rheu
RFLP	Restriction fragment length polymorphism
RGA	Resistance gene analogue
RI	Riec-sur-Belon
RIL	Recombinant inbred line
RLK	Receptor like kinase
RLP	Receptor like protein
RootB	Root biomass
RootB:TB	Root to total biomass ratio
RootDia	Average root diameter
RSA	Root system architecture
SA	Acide salicylique
SAM ou MAS	Sélection assistée par marqueurs
SeedW	Seed weight
SeedYield	Pooled seeds harvested on a plot
ShootB	Shoot biomass
ShootL	Plant height
SNP	Single nucleotide polymorphism
SSR	Single sequence repeat
TapRootL	Tap root length
TB	Total biomass
TILLING	Target induced local lesions in genomes
TM-LRR	Transmembrane leucine rich repeat
TProjArea	Total root projected area
TRootL	Total root length
TSW	Thousand seeds weight
UMR	Unité mixte de recherche
UNIP	Union nationale de l'interprofession des plantes riches en protéines
USDA	United States department of agriculture
VIGS	Virus induced gene silencing

Introduction générale

Depuis le Grenelle de l'environnement (loi n° 2009-967 du 3 août 2009, chapitre III, article 31) et la Directive 2009/128/CE du Parlement européen, la réduction de l'usage d'intrants et la généralisation des pratiques agricoles durables constituent des enjeux majeurs pour l'agriculture en Europe. En France, le plan Écophyto (septembre 2008 puis octobre 2015) prône notamment une réduction de 50 % de l'usage des produits phytosanitaires d'ici 2025 et l'emploi de pratiques alternatives. Le contexte de changement climatique favorisant l'émergence et le développement des maladies, les alternatives à la lutte chimique contre les bio-agresseurs des cultures présentent un intérêt croissant. L'utilisation des résistances génétiques aux bio-agresseurs constitue un levier important dans la maîtrise des ravageurs des cultures. Toutefois, l'adaptation des bio-agresseurs aux résistances génétiques, notamment celles déterminées par des gènes majeurs, a fréquemment mené à l'observation de contournements de résistances. Les résistances à déterminisme polygénique seraient a priori plus durables mais la connaissance de leurs mécanismes sous-jacents et leur exploitation en sélection restent difficiles. La création de variétés, cumulant résistance intrinsèque et caractères d'architecture de la plante permettant de tolérer ou d'échapper à la maladie, constituerait une voie complémentaire de protection durable des cultures. Depuis une dizaine d'années, plusieurs études ont montré que l'architecture aérienne de la plante ou du couvert pouvait constituer un levier efficace pour le contrôle du développement des épidémies aériennes. Au niveau racinaire, très peu d'études ont été réalisées concernant le rôle de l'architecture de la plante pour limiter les épidémies telluriques. Ceci est principalement dû au fait que le système racinaire est plus difficilement observable *in situ*.

Concernant plus particulièrement les légumineuses, le plan Protéines végétales (2014-2020) vise à s'engager durablement dans le développement de ces cultures pour améliorer la performance environnementale et économique de l'agriculture française, s'inscrivant pleinement dans le projet agro-écologique pour la France (2012). La pourriture racinaire précoce, due à l'agent pathogène *Aphanomyces euteiches*, est l'un des principaux freins à la culture du pois dans de nombreux pays. Cette maladie occasionne de fortes pertes de rendement sur le pois protéagineux mais également sur d'autres espèces de légumineuses cultivées comme la lentille, la luzerne, le haricot, la vesce ou encore la féverole. La combinaison de différentes méthodes de lutte, incluant l'utilisation de variétés résistantes comme levier principal, apparaît la stratégie la plus efficace et durable pour mieux gérer la

maladie. Il n'existe aucune variété commercialisée résistante à ce jour mais des sources de résistance partielle, à déterminisme polygénique, ont été identifiées chez le pois aux États-Unis et en France ces vingt dernières années. Un programme de sélection récurrente phénotypique a été mis en place par le GSP (Groupement des sélectionneurs de protéagineux) en 1995 afin de créer des géniteurs de pois agronomiques présentant des niveaux de résistance accrus, utilisables en création variétale. Récemment, l'architecture racinaire a été identifiée comme un caractère de la plante pouvant présenter un intérêt pour limiter la sévérité des maladies racinaires, par croissance rapide des racines et production de larges masses racinaires. Cette thèse a visé à explorer le rôle de l'architecture racinaire dans la réduction des symptômes de pourriture racinaire chez le pois, en précisant les bases génétiques de l'architecture de la plante et de la résistance partielle polygénique à *A. euteiches* chez celle-ci.

La thèse a été réalisée à l'Institut de génétique, environnement et protection des plantes (IGEPP), Unité mixte de recherche (UMR) Institut national de recherche agronomique (INRA) / Agrocampus Ouest / Université Rennes 1, au sein de l'équipe Résistance et adaptation, et à l'UMR Agroécologie INRA / AgroSup Dijon / Université de Bourgogne, au sein du pôle Déterminismes génétiques et environnementaux de l'adaptation des plantes à des systèmes de culture innovants (GEAPSI). Elle a été financée par le ministère de l'Agriculture, de l'Agroalimentaire et de la Forêt (MAAF), via une mise à disposition, dans le cadre d'une formation complémentaire par la recherche (FCPR), proposée par la DGER (Direction générale de l'enseignement et de la recherche) aux agents du MAAF. Son fonctionnement a été cofinancé par Terres Univia (fiches annuelles, convention INRA-UNIP) pour les expérimentations au champ et en conditions contrôlées ; les projets INRA-CropDL (2013-2015) (métaprogramme SELGEN) et ANR-GENOPEA (2010-2014) pour les travaux de génotypage à l'aide d'une puce SNP Illumina Infinium ; le projet Investissement d'avenir ANR-PeaMUST (2012-2019) pour les déplacements, formations, publications, participations à réunions ou conférences et autres frais.

Chapitre I

Introduction bibliographique

Chapitre I : Introduction bibliographique

1. La résistance aux maladies et l'architecture des plantes

1.1. Quelques notions sur la résistance des plantes aux maladies

1.1.1. Résistances qualitative et quantitative

Les végétaux sont généralement soumis à un grand nombre de stress abiotiques (climat, composition des sols, nutriments, etc.) et biotiques (bactéries, oomycètes, champignons, nématodes, virus, insectes et autres ravageurs). La résistance des plantes aux agents pathogènes se distingue suivant la pathogénicité de l'agent pathogène au sein d'une espèce de plantes et le type d'interaction plante hôte-pathogène.

Lorsque l'ensemble des géotypes d'une même espèce est résistant à l'ensemble des isolats d'un agent pathogène, la résistance est dite **non-hôte**. *A contrario*, lorsque certains géotypes sont résistants et d'autres sensibles, la résistance est dite **hôte**.

On distingue deux types de résistance au sein des interactions hôte-pathogène.

Lorsque la colonisation ou la croissance de l'agent pathogène est stoppée à un stade précoce du développement de la maladie, on parle d'**interactions incompatibles** ; l'agent pathogène est dit avirulent. La plante présente alors une **résistance totale**, qui se traduit souvent par des symptômes d'hypersensibilité (HR) résultant de la mort programmée des cellules avoisinant le point d'infection (Poland et al. 2009). Les symptômes sont en général peu visibles et ne provoquent pas de dégâts sur la plante. Cette résistance à caractère binaire (résistant/sensible) est également qualifiée de **résistance qualitative**. D'un point de vue génétique, elle est également appelée résistance monogénique car souvent contrôlée par un seul gène de résistance à effet fort, appelé gène « R ».

Lorsque l'agent pathogène colonise efficacement la plante hôte et se développe activement dans ses tissus, on parle d'**interactions compatibles** ; l'agent pathogène est dit virulent. Certains géotypes peuvent présenter des symptômes réduits. On dit qu'ils présentent une **résistance partielle**, car leurs défenses n'empêchent pas totalement le développement de la maladie. Cette résistance, quantifiable par la quantité de symptômes observables, est également appelée **résistance quantitative**. D'un point de vue génétique, elle est souvent sous le contrôle de plusieurs gènes à effets souvent faibles (résistance polygénique), identifiés

au sein de régions génomiques contrôlant la variation quantitative ou QTL (*Quantitative Trait Loci*).

La distinction entre résistance quantitative et qualitative n'est pas toujours aisée car des gènes R peuvent parfois avoir des effets quantitatifs, tandis que des QTL de résistance partielle peuvent avoir des effets forts. Poland et al. (2009) suggèrent plutôt qu'il existerait un continuum entre résistance totale et sensibilité, sur lequel les gènes R se situeraient préférentiellement à l'une des extrémités et les QTL de résistance tout au long du continuum.

Les résistances monogéniques exercent généralement une pression plus forte sur les populations pathogènes et sélectionnent plus efficacement les individus virulents (*i. e.* qui colonisent efficacement la plante hôte) apparus par mutation. À l'inverse, les résistances polygéniques sont plus difficilement contournables par les populations d'agents pathogènes. Avec une mutation effective vis-à-vis d'un QTL de résistance, l'agent pathogène ne gagnera qu'un avantage marginal sur sa capacité d'infection. En présence d'une résistance polygénique, les populations pathogènes devront acquérir et sélectionner des mutations à plusieurs endroits de leur génome pour pouvoir coloniser la plante hôte. Les résistances polygéniques, bien que souvent moins efficaces à un instant donné, présentent donc l'avantage d'être *a priori* plus durables dans le temps et l'espace par rapport aux résistances monogéniques.

1.1.2. Les résistances partielles sous contrôle polygénique

En 2001, Kover and Caicedo (2001) ont effectué une synthèse des QTL de résistance aux maladies détectés dans 85 expériences. Les auteurs ont montré que la résistance partielle était contrôlée en moyenne par 4 à 6 QTL, avec un maximum de 18 QTL, généralement répartis sur l'ensemble du génome. Le pourcentage de variation phénotypique expliquée par les QTL individuellement était généralement faible (inférieur à 20 %) et variait fortement suivant les pathosystèmes et les QTL considérés. Le pourcentage de variation phénotypique expliqué par l'ensemble des QTL d'une même étude variait également selon l'héritabilité du caractère dans le dispositif de test utilisé.

Les QTL de résistance partielle ont souvent été montrés non spécifiques d'un groupe de virulence (ou pathotype ou race) de l'agent pathogène considéré. Cependant, des études ont montré que certains QTL pouvaient être spécifiques d'une ou plusieurs races. Par exemple, Li et al. (2006) ont étudié la résistance du riz à *Xanthomonas oryzae* à partir de deux populations de riz et 10 races de *X. oryzae*. Ils ont montré que certains QTL étaient spécifiques d'un sous-

ensemble de ces 10 races alors que d'autres QTL étaient détectés avec toutes les races étudiées. Si les QTL de résistance peuvent être à large spectre ou spécifiques d'une race au sein de l'espèce d'un agent pathogène, ils peuvent également être spécifiques ou non d'un agent pathogène. Certaines études ont montré que les QTL pouvaient parfois conférer une résistance à plusieurs agents pathogènes. Par exemple, Manosalva et al. (2009) ont mis en évidence qu'un QTL majeur de résistance du riz à *Magnaporthe oryzae* conférait également une résistance à d'autres agents pathogènes.

La détection des QTL de résistance a souvent été décrite comme dépendante de l'environnement. Un environnement correspondant à la fois à un lieu et une année d'expérimentation, les conditions climatiques peuvent varier fortement d'un environnement à un autre, influençant l'expression du caractère mesuré. L'environnement peut donc influencer sur le nombre et l'effet des QTL détectés. C'est pourquoi la résistance est souvent évaluée dans plusieurs environnements (3 au minimum) afin d'évaluer la stabilité des QTL ainsi détectés. L'identification des QTL a également été montrée comme dépendant du fonds génétique et des croisements utilisés pour la construction de la population de détection. Ainsi, Pilet et al. (2001) ont démontré que la détection des QTL de résistance à *Leptosphaeria maculans* chez le colza dépendait à la fois du type de croisement (haploïdes doublés ou F2-3) et des génotypes parentaux de la population de détection (Darmor-*bzh* x Yudal ou Darmor x Samourai). De même, une étude récente (Lavaud et al. 2015) a démontré, grâce à la construction de lignées quasi-isogéniques (NILs), que l'efficacité des QTL introgressés dans différents fonds génétiques pouvait être variable.

1.1.3. Mécanismes de résistance des plantes aux maladies

Les mécanismes associés aux gènes majeurs de résistance (gènes R) sont relativement bien documentés. On compte à ce jour 87 gènes R clonés, à partir de plusieurs espèces (blé, orge, tomate, pomme de terre, laitue, poivre noir, riz, maïs, *Arabidopsis thaliana*, lin, betterave et tabac) et pour une large diversité d'agents pathogènes (Sekhwal et al. 2015). À partir des séquences de protéines des gènes R, un grand nombre d'analogues de gènes de résistance (RGA) a pu être identifié chez 26 espèces. Ces RGA contiennent des séquences conservées des domaines associés à la résistance chez les gènes R. Les domaines protéiques associés à la résistance peuvent être classés en plusieurs groupes et sous-groupes (Sekhwal et al. 2015) (Figure I.1) :

- les sites de liaison nucléotidique à répétitions riches en leucine (*Nucleotide binding*

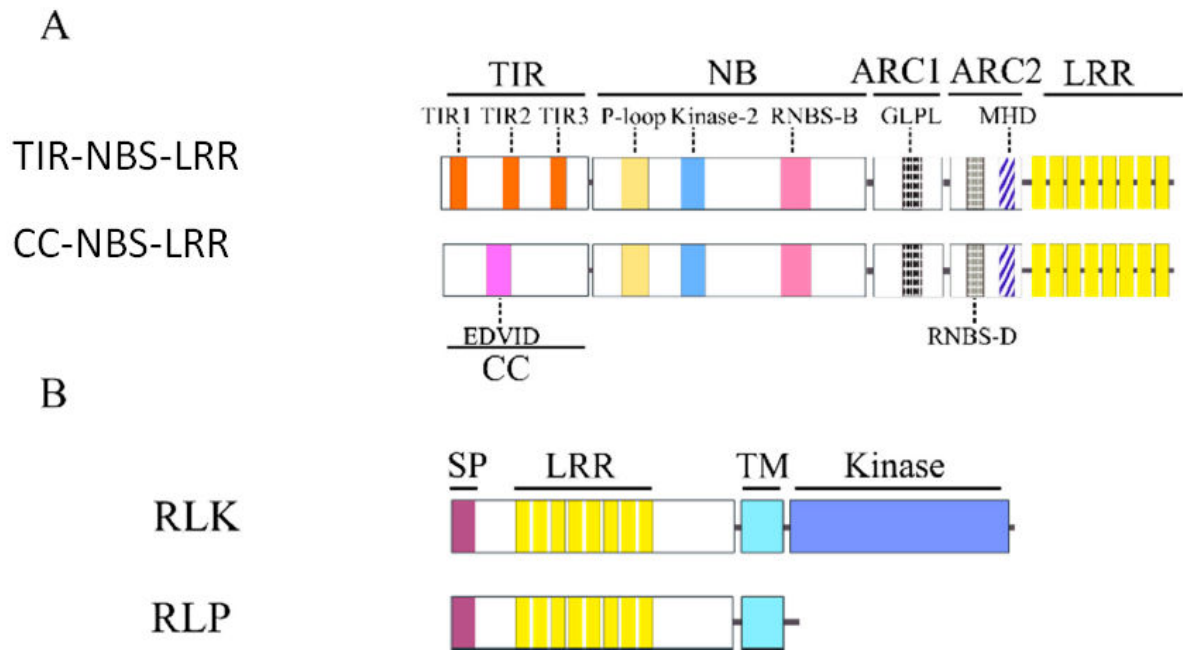


Figure I.1 : Représentation schématique des structures communes des quatre protéines majeures de résistances chez les plantes (d’après *Sehwal et al. (2015)*)

Les motifs sont représentés par des bandes colorées et identifiés par le nom des domaines protéiques. La longueur des domaines n’est pas représentée à l’échelle pour faciliter leur visualisation.

A : Représentation des domaines typiques des protéines TIR-NBS-LRR et CC-NBS-LRR. Seuls les motifs hautement conservés sont représentés.

B : Représentation des domaines des protéines RLK et RLP.

TIR : *Toll/interleukine-1 receptor* ; CC : *Coiled-coil* ; SP : *Signal peptide* ; TM : *transmembrane* ; LRR : *Leucine-rich repeat*

site leucine rich repeat : NBS-LRR). Les NBS-LRR sont répartis en deux sous-groupes : les récepteurs *toll/interleukine* (TIR-NBS-LRR) et les superhélices (*coiled coil* : CC-NBS-LRR). Ces domaines NBS-LRR reconnaissent des protéines cibles de l'agent pathogène (effecteurs) directement adressés dans la cellule hôte et induisent une réponse de type ETI (*effector triggered immunity*) ;

- les domaines transmembranaires à répétitions riches en leucine (*transmembrane leucine rich repeat* : TM-LRR). Les TM-LRR sont également répartis en deux sous-groupes : les analogues aux récepteurs kinases (*Receptor like kinase* : RLK) et les protéines analogues à d'autres récepteurs (*receptor like protein* : RLP). Ces domaines sont des récepteurs reconnaissant certains motifs. Ils permettent la médiation de l'immunité déclenchée par la reconnaissance d'une molécule pathogène (*pathogen associated molecular pattern* : PAMP) (*PAMP triggered immunity* : PTI). Les PAMP sont des motifs conservés sur les protéines de virulence de la plupart des agents pathogènes comme la chitine, les flagelles ou les lipopolysaccharides. La PTI permet la reconnaissance d'un large spectre de pathogènes et confère donc une résistance à large spectre ;
- d'autres types de domaines, tels que les répétitions de pentatricopeptides (PPR) ou les peroxydases. Ces autres types de protéines sont impliqués dans des mécanismes plus variés comme la détoxification de phytotoxines, la modulation de la réponse de défense et la régulation de la transcription d'autres gènes (Poland et al. 2009).

Toutefois, si les mécanismes sous-jacents aux gènes R et RGA sont bien connus, ceux sous-jacents aux QTL de résistance le sont beaucoup moins. Poland et al. (2009) ont suggéré différentes hypothèses sur les gènes sous-jacents au QTL de résistance. On trouve parmi ces hypothèses :

- des gènes impliqués dans le **développement ou la morphologie de la plante hôte**. Différentes études ont montrée l'association de certains caractères morphologiques ou de développement à la résistance aux maladies. Par exemple, l'ouverture du couvert, la hauteur de la plante, l'angle entre les feuilles ou encore l'espace entre les nœuds foliaires ont souvent été cités comme des caractères influençant la résistance aux pathogènes aériens (voir paragraphe 1.2). À une autre échelle, l'ouverture ou la densité des stomates ont également été montrées déterminantes dans la diminution de la colonisation d'une plante hôte par des agents pathogènes ;
- des gènes impliqués dans la **reconnaissance basale des agents pathogènes via les**

Tableau I.1 : Gènes clonés et domaines protéiques identifiés dans des QTL de résistance aux maladies chez les plantes

Domaines protéiques	Locus	Pathosystème	Référence
ABC (ATP-Binding Cassette) transporter	<i>Lr34</i>	Blé / rouille	Krattinger et al. (2009)
Kinase-START (Steroidogenic Acute Regulatory protein-related lipid Transfer)	<i>Yr36</i>	Blé / rouille	Fu et al. (2009)
Hexose transporter	<i>Lr67</i>	Blé / Rouille, oïdium	Moore et al. (2015) White and Frommer (2015)
Wall-associated receptor-like kinase	<i>Htn1</i>	Maïs / dessèchement	Hurni et al. (2015)
Heavy metal-transport detoxification	<i>Pi21</i>	Riz / pyriculariose	Fukuoka et al. (2009)
Serine hydroxymethyltransferase	<i>Rhg4</i>	Soja / nématode	Liu et al. (2012)
NBS-LRR	<i>Pi35</i>	Riz / pyriculariose	Fukuoka et al. (2014)
sulphotransferase	<i>STV11</i>	Riz / Virus	Wang et al. (2014)

-

PAMP (principalement chitine ou flagelline) pourraient avoir subi des mutations, altérant la reconnaissance mais conférant malgré tout une résistance partielle. Par exemple, il a été montré qu'une mutation dans le gène *FLS2* (*FLAGELLIN-SENSITIVE 2*, gène R permettant la reconnaissance de la flagelline des bactéries et entraînant une résistance totale) chez *Arabidopsis thaliana* conférait une résistance partielle aux bactéries (Poland et al. 2009) ;

- des gènes impliqués dans la **détoxification des phytotoxines** produites par l'agent pathogène ou dans la **production d'antibiotiques**. Les agents pathogènes nécrotrophes sont connus pour produire des phytotoxines qui vont permettre la dégradation des cellules de la plante hôte afin d'en tirer des nutriments. Ainsi, l'oxalate réductase, produite par certaines plantes hôtes partiellement résistantes, permet de réduire le *burst* oxydatif provoqué par la phytotoxine acide oxalique. Les phytoalexines, comme la camalexine ou le glutathion, seraient impliquées dans des fonctions antibiotiques ou cyto-protectrices ;
- des gènes impliqués dans la **transduction des signaux** de défense des plantes. Un stress est parfois transmis de façon systémique chez les plantes par la production de phytohormones telles que l'acide salicylique (SA), l'acide jasmonique (JA) ou l'éthylène (ET). Lors d'une colonisation par un agent pathogène, s'il y a production de phytohormone, il s'agira le plus souvent d'ET dans le cas d'un pathogène nécrotrophe et de SA ou JA dans le cas d'un pathogène biotrophe. Chez *Arabidopsis thaliana*, des mutants de facteurs de transcription (WRKY) impliqués dans les voies de signalisation régulées par l'acide salicylique et l'acide jasmonique, ont permis d'augmenter la sensibilité au pathogène nécrotrophe *Botrytis cinerea* (Qiu et al. 2007; Zheng et al. 2006).
- des **gènes R à effets faibles** ou contournés à effets résiduels. L'effet partiel d'un QTL pourrait provenir de la modification des sites de reconnaissance et/ou des sites effecteurs, conduisant à des réponses de défense de la plante moins importantes et/ou moins rapides. Chez le riz, des substitutions dans le domaine LRR du gène *Xa21* ont conféré une résistance partielle à *Xanthomonas oryzae* pv. *oryzae* (Wang et al. 1998). Chez le maïs, le QTL *RCG1* a été identifié comme étant un gène de type NB-LRR impliqué dans la résistance partielle à l'antracnose (Broglie et al. 2006).

Quelques QTL de résistance ont été clonés à ce jour, révélant des fonctions diverses des gènes sous-jacents codant pour des protéines à domaines variés (Tableau I.1). Cook et al. (2012) ont également démontré que trois gènes multiples répétés dans un segment de 31 kb, codant pour

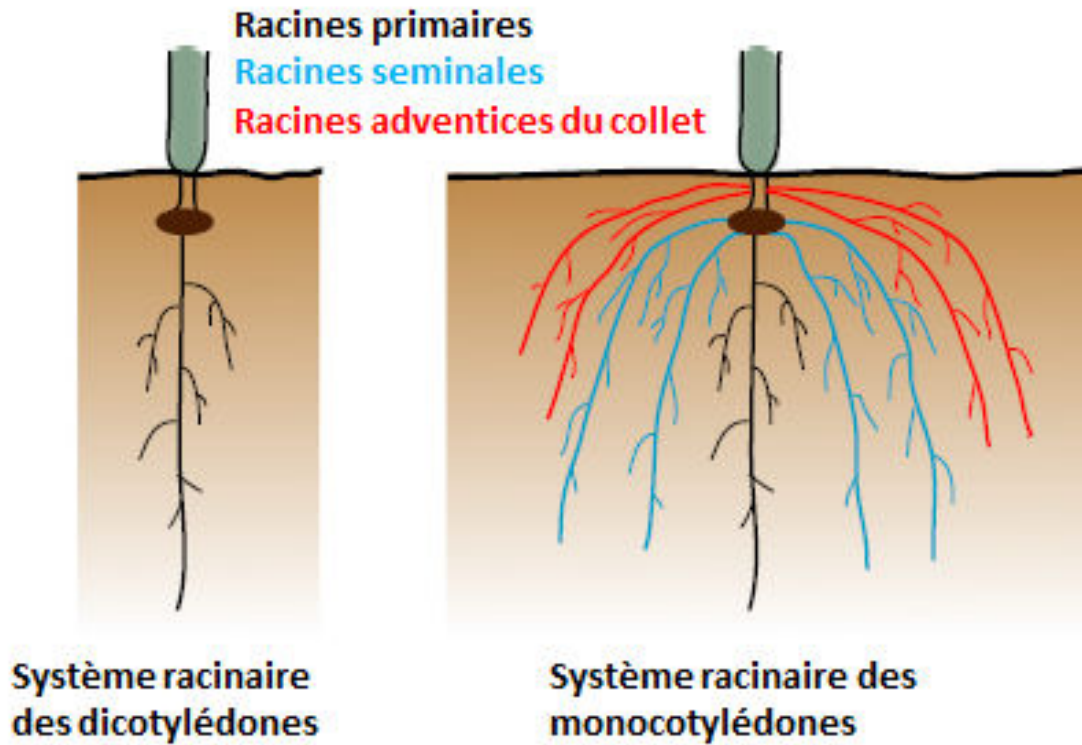


Figure I.2 : Systèmes racinaires des dicotylédones et des monocotylédones

Les racines primaires sont les premières à sortir, initiées dès le stade embryonnaire. Les racines séminales et adventives sont initiées au stade post-embryonnaire, au niveau de la graine ou du collet, respectivement. (D'après (Wachsman et al. 2015))

un transporteur d'acide aminé, une protéine α -SNAP (*soluble N-ethylmaleimide-sensitive attachment protein*) et une protéine WI12 à domaine inductible par blessure, étaient sous-jacents au QTL *Rhg1* de résistance au nématode à kyste chez le soja.

1.2. L'architecture des plantes peut-elle être utilisée comme un levier pour lutter contre les épidémies ?

1.2.1. Etude de l'architecture des plantes

1.2.1.1. Variabilité et plasticité de l'architecture des plantes

L'architecture aérienne des plantes est déterminée à la fois par le génotype de la plante (espèce et variété), l'environnement et le stade de la plante (Calonnec et al. 2012; Costes et al. 2013). L'environnement influence l'architecture aérienne par la distribution de la lumière, la quantité d'eau ou de nutriments disponibles, la température ou le vent. Dans le cas de plantes cultivées, l'agriculteur peut également influencer l'architecture des plantes par des pratiques culturales comme la coupe chez les plantes pérennes, la fertilisation chez les plantes annuelles ou encore la densité du couvert déterminé lors du semis ou de l'implantation.

Cinq critères sont généralement utilisés pour décrire l'architecture aérienne d'une plante : la croissance, la ramification, la différenciation morphologique, l'orientation des axes de ramification et la localisation des organes reproducteurs (Costes et al. 2013). Ces caractéristiques traduisent un potentiel génétique qui peut être partiellement modulé par les conditions environnementales.

La morphologie des racines ainsi que leur croissance sont intimement liées au génotype de la plante, ainsi qu'aux propriétés du sol ou du milieu dans lequel elles poussent. Les racines ont la capacité de s'adapter de façon importante à leur environnement local (Downie et al. 2015). Par exemple, la morphologie des racines peut être impactée par la compaction du sol, la disponibilité en nutriments ou en eau, mais également par des facteurs biotiques comme les organismes pathogènes ou symbiotiques (*Rhizobium*, mycorhizes, etc.).

Chez les monocotylédones, l'architecture du système racinaire est complexe. Plusieurs racines primaires sont initiées au stade embryonnaire et des ramifications se développent au stade post-embryonnaire, à partir du collet, de la graine et des racines primaires (Wachsman et al. 2015) (**Figure I.2**). Chez les dicotylédones, l'architecture racinaire est généralement plus simple, car organisée en une racine principale, appelée pivot, et une série de ramifications appelées racines latérales, bien que des racines adventives peuvent également apparaître au

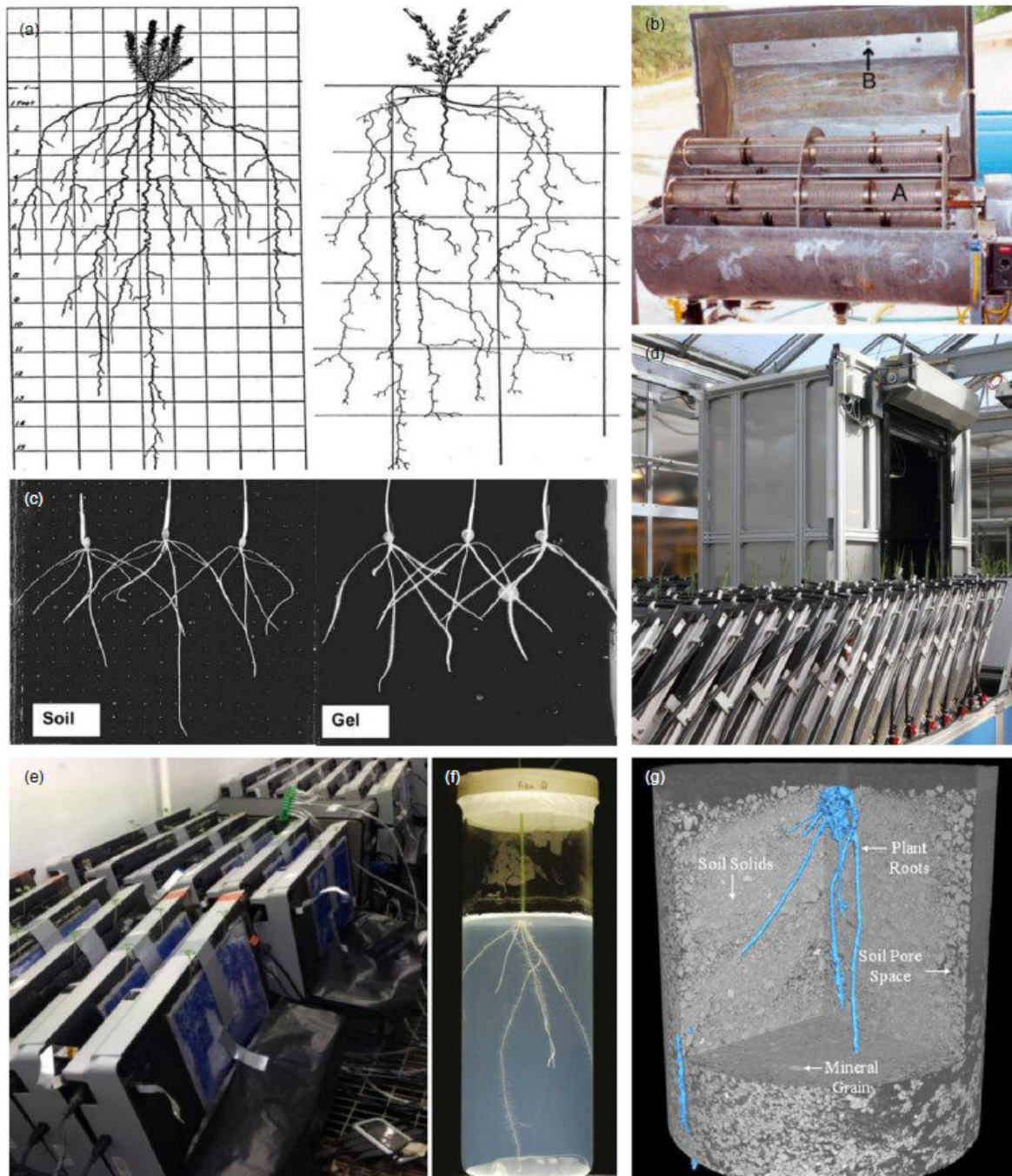


Figure I.3 : Évolution des méthodes de mesures du système racinaire, du dessin manuel à la tomographie par rayon X, (d'après *Downie et al. (2015)*)

(a) Tracé manuel du système racinaire après prélèvement destructeur (Weaver 1919). (b) Extraction automatisée des racines au champ (Benjamin and Nielsen 2004). (c) Jeunes plantules d'orge cultivées dans des systèmes 2D de sol ou de gel et scannées pour illustrer le type de croissance racinaire (Bengough et al. 2004). (d) Système de phénotypage racinaire automatisé, GROWSCREEN-Rhizo, Jülich Plant Phenotyping Center, Germany (Nagel et al. 2012). (e) Banque de scanner à prise d'image automatisée pour l'évaluation de la dynamique de croissance des racines sur papier filtre (Adu et al. 2014). (f) Racines poussant dans un système à base de gel pour une prise d'image par tomographie optique en 3D (Clark et al. 2011). (g) Racines *in situ* poussant dans du sol, prise d'image par rayons X (Zappala et al. 2013).

niveau du collet (Sorin et al. 2005). La description de l'architecture des systèmes aériens et racinaires est basée sur le même concept : le nombre d'éléments (feuilles, branches, racines), leur taille (longueur, diamètre) et leur organisation spatiale (distance entre éléments, angle...).

1.2.1.2. L'architecture racinaire : quels moyens de mesure pour un compartiment difficile d'accès ?

En plus de constituer un organe très plastique à son environnement, le système racinaire est difficile à observer *in situ* du fait de sa croissance sous la surface du sol. Afin de pouvoir observer la variabilité morphologique des racines, celles-ci sont généralement extraites de leur substrat et lavées, puis des comptages et mesures sont faits. Il s'agit d'une technique délicate, fastidieuse et chronophage. Elle est de surcroît destructive et ne permet donc pas d'observer une dynamique de croissance sur les racines extraites (Downie et al. 2015). Depuis 25 ans, se sont développés des systèmes d'acquisition d'images du système racinaire, à l'aide de caméras ou de scanners. Ces techniques permettent une mesure non manuelle, notamment des longueurs et diamètre des racines (Kirchhof 1992; Lebowitz 1988). Les images peuvent être prises *ex situ* après lavage et étalement des racines sur une surface plane, ou *in situ* dans des dispositifs adaptés. Ces dispositifs en 2D permettent une observation non destructive et sont appelés rhizotrons. Il s'agit en général de deux plaques (ou d'une poche) au milieu desquelles est disposé un substrat et entre lesquelles le système racinaire va se développer. L'observation ou la prise d'image des racines se fait alors par transparence du support. Conjointement à la baisse des coûts d'acquisition d'image, beaucoup d'efforts ont été déployés afin de développer des outils informatiques permettant de faciliter et automatiser la collecte de données à partir d'images en 2D. À ce jour, 31 logiciels d'analyse d'image de systèmes racinaires sont disponibles (<http://www.plant-image-analysis.org/>), certains étant libres d'accès - par exemple SmartRoot, basé sur ImageJ (Lobet et al. 2011) ou archiDART un package dans R (Delory et al. 2015) - et d'autres sous licence commerciale - par exemple WinRHIZO® ; (Arsenault et al. 1995). Les techniques d'observation et d'analyse du système racinaire continuent d'évoluer (**Figure I.3**). En effet, les rhizotrons classiques ne permettent pas une croissance ou une observation en trois dimensions (3D) du système racinaire, plus représentatives d'une croissance en milieu naturel. C'est pourquoi, d'autres systèmes ont été développés plus récemment pour une observation du système racinaire en 3D, utilisant des scanners, des caméras, la fluorescence ou les radiations (imagerie par rayons X, par exemple) (**Figure I.3**). Enfin, pour faciliter l'automatisation des prises d'images sur les compartiments racinaire et aérien, des plates-formes de phénotypage haut-débit ont vu le jour récemment.

Ces plates-formes possèdent généralement des rhizotrons (plats ou cylindriques) positionnés sur des supports mobiles qui permettent le déplacement automatisé des plantes pour la prise d'image (par exemple la plate-forme GROWSCREEN Rhizo de Jülich http://www.fz-juelich.de/ibg/ibg-2/EN/methods_jppc/GROWSCREEN_rhizo/_node.html et la plate-forme *Plant Phenotyping Platform for Plant and Micro-organism Interactions* - 4PMI de l'INRA de Dijon). Au sein de la Plateforme de phénotypage haut-débit 4PMI mise en place à l'UMR Agroécologie (INRA Dijon, France), un focus particulier est apporté au phénotypage racinaire et aux interactions entre plantes et micro-organismes, en particulier telluriques. L'objectif est, grâce à un projet novateur d'observation automatisée des racines, de répondre aux besoins de caractérisation d'un nombre élevé d'unités biologiques nécessaire, en particulier, dans des approches de génétique quantitative ou d'association. Les RhizoTubes de la 4PMI (1 200 prévus d'ici fin 2016) sont des rhizotrons cylindriques permettant de visualiser le développement et la croissance du système racinaire, de manière non-destructive et dynamique - Brevet Salon et al. (2014). Ces RhizoTubes sont déplacés automatiquement dans les plateformes équipées de convoyeurs, de la zone de croissance en serre vers la zone d'alimentation en solution nutritive et les cabines de phénotypage.

Toutes ces dernières technologies restent chères, du fait de l'acquisition de matériel et de logiciels d'analyse, et demandent un espace de stockage informatique important.

1.2.1.3. Le contrôle génétique de l'architecture des plantes

L'architecture aérienne est le résultat de plusieurs caractères morphologiques et organisationnels des éléments qui la composent. Ces caractères (nombres de tiges et de feuille, forme, taille et organisation spatiale des éléments) sont le plus souvent des caractères quantitatifs qui sont contrôlés par plusieurs gènes (Perez-Perez et al. 2010). Par exemple, Ding et al. (2015) ont détectés 14 QTL contrôlant l'angle des feuilles chez le maïs, dont 7 communs avec les études précédemment publiées. Les QTL d'architecture aérienne de la plante sont impliqués dans de fortes interactions génotype x environnement et dépendent de la population d'étude.

L'architecture racinaire est le plus souvent contrôlée par plusieurs QTL à effets faibles (Burton et al. 2014; de Dorlodot et al. 2007; Orman-Ligeza et al. 2014; Rogers and Benfey 2015). Dans la plupart des cas, les caractères de l'architecture racinaire ont une faible héritabilité. Les QTL détectés présentent souvent de fortes interactions avec l'environnement, en lien avec la plasticité élevée du système racinaire. Leur validation au champ dans plusieurs

environnements est donc cruciale mais également coûteuse et longue. Le clonage de QTL d'architecture racinaire n'a jusqu'à présent pas été réalisé avec succès chez les plantes cultivées (Orman-Ligeza et al. 2014). Chez le riz, espèce cultivée la plus étudiée pour son système racinaire, plus de 80 études réalisées depuis les années 1990 ont permis de détecter plusieurs centaines de QTL d'architecture racinaire. Or les résultats ont été difficiles à comparer du fait de l'absence de standardisation des méthodes d'observation et des conditions de culture, de l'usage de plantes d'âges différents et de populations biparentales d'étude diversifiées. Malgré tout, une recherche de QTL a permis de mettre en évidence un gène contrôlant l'architecture globale des racines (Wachsman et al. 2015), le gène *DEEPER ROOTING 1 (DRO1)*. Ce gène contrôle l'angle des racines en réponse au gravitropisme (orientation de la croissance en fonction de la gravité), permettant un ancrage plus profond du système racinaire et une meilleure tolérance à la sécheresse chez le riz.

1.2.2. Architecture et maladies des plantes

1.2.2.1. Cycle épidémique d'un agent pathogène et interaction avec la plante hôte

Le cycle de vie d'un agent pathogène consiste en plusieurs phases : l'interception et la reconnaissance de l'agent pathogène par la plante, l'infection (germination et pénétration), la colonisation de la plante, la reproduction sexuée ou asexuée de l'agent pathogène, et la dispersion à d'autres organes ou d'autres plantes (Tivoli et al. 2012b). Le succès du développement de l'agent pathogène sur une plante hôte peut donc dépendre de la distribution, de la forme et de la composition des organes sensibles.

Une **épidémie** est un phénomène dynamique dans le temps et dans l'espace qui traduit la colonisation d'une **population d'hôtes** par un agent pathogène. Le succès d'une épidémie dépend de la proportion de tissus sensibles au sein de la population, du taux de transmission de la maladie et de la durée de la période infectieuse (Calonnec et al. 2012). Une épidémie peut donc être fonction des modifications structurales de la population hôte (distribution des individus ou des organes sensibles).

1.2.2.2. Liens entre architecture de la plante et réduction des épidémies

La conférence « Plant and Canopy Architecture Impact on Disease Epidemiology and Pest Development » qui s'est déroulée à Rennes en 2012, a rassemblé la communauté scientifique internationale autour de la question de l'impact de l'architecture des plantes et des couverts

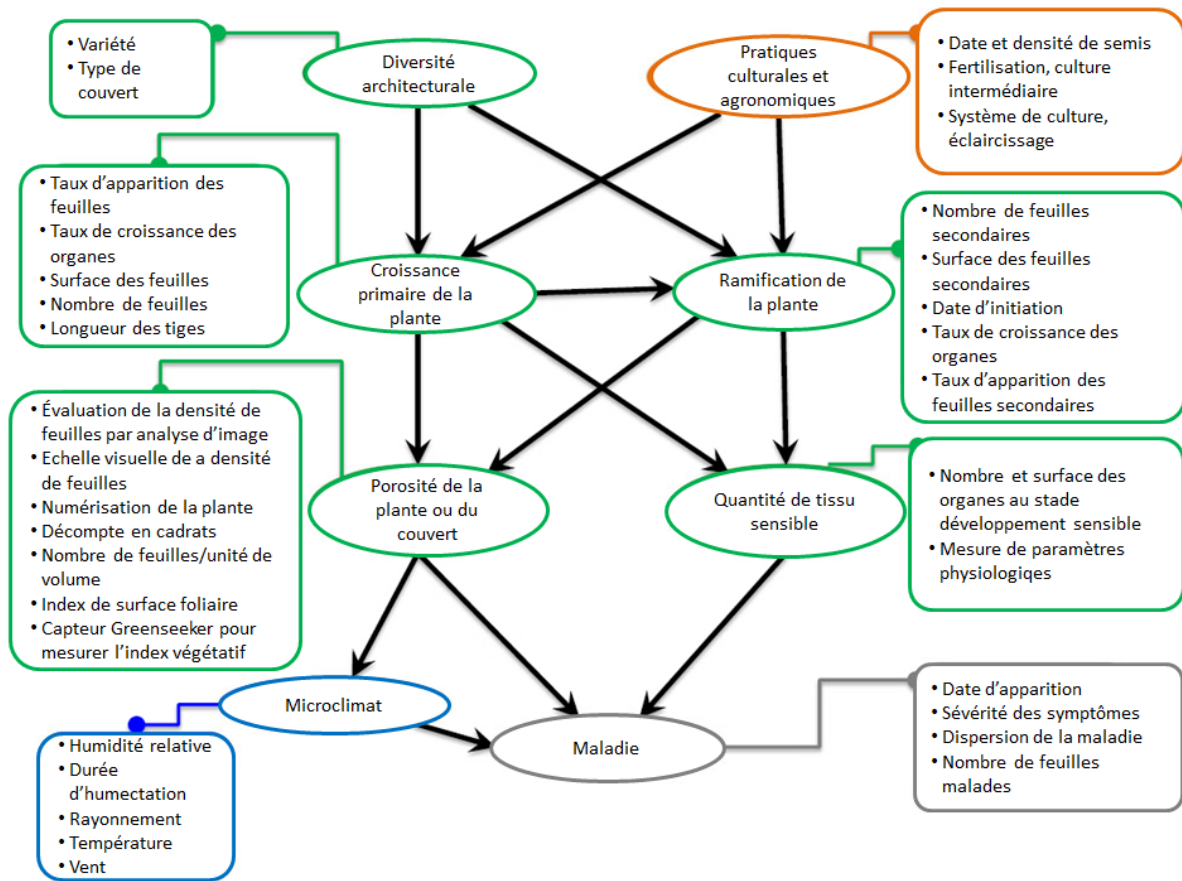


Figure I.4 : Interactions potentielles entre les variables d'architecture du couvert, le microclimat et la maladie

Les caractères « Diversité architecturale », « Pratiques culturales », « Croissance primaire », « Ramification », « Porosité », « Quantité de tissu sensible », « Microclimat » et « Maladie » sont caractérisés par un certain nombre de variables simples mesurables. (d'après Calonnec et al. (2012)).

sur l'épidémiologie des maladies (Tivoli et al. 2012a). Différents concepts y ont été abordés incluant :

- la **définition de l'architecture de la plante et du couvert**. L'impact des caractères architecturaux sur le développement des maladies se distingue à différentes échelles (Costes et al. 2013). L'**organe** influence les interactions hôtes/pathogènes au niveau tissulaire par sa nature, sa surface et sa forme. La **plante** influence la propagation des agents pathogènes en son sein, ainsi que le microclimat, favorable ou défavorable pour le développement de l'agent pathogène, par sa forme et sa topologie. La **taille du couvert** influence la propagation plante à plante et le microclimat du couvert par sa densité et son organisation spatiale. L'architecture aérienne des végétaux varie fortement avec le type de plante (monocotylédone/dicotylédone ; annuelle/pérenne ; herbacées/ligneux) et le stade de la plante. L'importance relative de chaque échelle et de chaque stade dans la défense des plantes dépend de l'espèce végétale et de l'agent pathogène considérés ;
- l'analyse des **principales caractéristiques architecturales influençant les épidémies** et le développement des agents pathogènes. L'interception des agents pathogènes aériens par les plantes est rendue plus probable par une plus grande surface foliaire d'une plante ou d'un couvert (Tivoli et al. 2012b). La plupart des épidémies dues à des agents pathogènes fongiques ou bactériens sont donc favorisées par une forte densité de feuillage. De plus, la germination des spores et le développement de ces agents pathogènes sont généralement favorisés par un fort taux d'humidité dans le couvert, corrélé à sa densité. Ainsi, Richard et al. (2013) ont montré que les génotypes de pois ayant une architecture aérienne plus dense sont plus favorables aux infections par *Dydimella pinodes*. La densité du couvert favorise également la propagation des spores d'organe à organe (dispersion verticale) ou de plante à plante (dispersion horizontale). Plus les organes sensibles sont proches, plus la probabilité de transmission de l'infection est forte par transport aérien (vent) ou par *splashing* (gouttes de pluie) (Calonnec et al. 2012; Tivoli et al. 2012b) ;
- la **description des caractères architecturaux** et leur impact sur les épidémies. Les caractères architecturaux qui impactent potentiellement la résistance des plantes aux maladies peuvent être mesurés grâce à différentes variables simples (Calonnec et al. 2012) (**Figure I.4**). Une fois ces caractères quantifiés, les mesures peuvent être comparées aux variables caractérisant la maladie pour mettre en évidence les

corrélations ou les impacts les plus forts. Par exemple, Richard et al. (2013) ont mesuré la durée d'humectation, la vitesse du vent et la température au sein du couvert de 3 géotypes de pois contrastés pour la hauteur et la fermeture du couvert, l'index de surface foliaire (*Leaf Area Index* : LAI) et la densité foliaire. Ils ont ainsi mis en évidence que la durée d'humectation était négativement corrélée à la hauteur et au LAI, suggérant ainsi que les couverts denses entraînaient un microclimat favorable au développement de *Dydimella pinodes*. Chez le haricot, Miklas et al. (2013) ont montré qu'un couvert végétal plus ouvert et haut, ainsi qu'une bonne tenue de tige étaient corrélés à un moindre niveau de sclérotiniose au champ ;

- **l'impact de la modification de l'architecture sur l'application de fongicides et le rendement.** La modification de l'architecture du couvert permet, selon les pathosystèmes (McDonald et al. 2013), de réduire les symptômes, l'application de fongicides (Miklas et al. 2013) et les pertes de rendement, par échappement ou par tolérance à la maladie. L'échappement d'une plante à la maladie se définit par sa capacité à éviter la rencontre ou les interactions compatibles avec l'agent pathogène ; il peut être spatial (ex. : répartition des organes rendant difficile la dissémination de l'agent pathogène entre eux), temporel (ex. : précocité ou tardiveté de la plante permettant d'éviter la rencontre agent pathogène/plante hôte au stade de développement de la plante le plus sensible) ou microclimatique (ex. : architecture de la plante permettant un microclimat défavorable au développement et à la dissémination de l'agent pathogène). L'échappement permet généralement d'éviter les symptômes mais pas de les supprimer complètement. La tolérance d'une plante se traduit par sa capacité à être moins impactée par la maladie (croissance, rendement) qu'une plante sensible, malgré un niveau de symptômes similaire. Les changements d'architecture des plantes sont obtenus par des pratiques culturales (taille, palissage, densité de semis) ou par l'utilisation de géotypes ayant une architecture défavorable au développement des agents pathogènes. Ainsi, les agriculteurs peuvent agir sur la densité du couvert (semis ou implantation), le choix des variétés en fonction de leur port aérien, ou encore sur la taille des arbres en cultures pérennes afin de limiter les interactions hôte/pathogène ainsi que les conditions favorables (microclimat dans le couvert) au développement de l'agent pathogène.

1.2.2.3. Lien entre architecture racinaire et maladies telluriques

Si le compartiment aérien est relativement bien documenté, un petit nombre d'études, souvent conduites sur un faible nombre de génotypes, s'est intéressé à **l'impact de l'architecture racinaire sur le contrôle des maladies** telluriques. Snapp et al. (2003) ont étudié les corrélations entre différents caractères d'architecture racinaire et la résistance à la pourriture racinaire due à *Fusarium solani* chez le haricot commun. En utilisant 17 lignées recombinantes de haricot commun issues d'un croisement entre un génotype sensible et un génotype résistant, ils ont montré que le nombre de racines latérales à l'interface entre les racines et la tige, ainsi que le diamètre des racines, étaient positivement corrélés à la résistance. À partir de 10 génotypes de haricot commun, Román-Avilés et al. (2004) ont montré que la résistance à *F. solani* était améliorée par un plus grand nombre de racines secondaires et une plus forte biomasse racinaire. De la même manière, Dias et al. (2004) ont associé à la résistance à *Monosporascus cannonbalus* un système racinaire plus vigoureux et plus ramifié, dans une population biparentale de melon. Sur deux génotypes de haricot commun, respectivement résistant et sensible à *F. solani*, greffés de façon réciproque, Cichy et al. (2007) ont montré un contrôle de la résistance par le génotype racinaire. En 2015, (Hagerty et al.) ont mis en évidence une corrélation négative entre le diamètre du pivot et la résistance à *Aphanomyces euteiches* et *F. solani*, dans une population de 177 lignées recombinantes de haricot commun en conditions infectées. Chez le framboisier, Graham et al. (2011) ont mis en évidence une corrélation négative entre la pourriture racinaire due à *Phytophthora fragariae* et le diamètre d'extension des racines latérales par rapport au pied-mère et le nombre de racines partant du pied-mère. L'étude de la **réponse du système racinaire de la plante à une maladie** a également été très peu documentée. La vigueur racinaire en réponse à une infection peut représenter une caractéristique d'architecture majeure pour limiter les épidémies. En effet, les racines constituent une réserve dynamique utilisable pour produire de nouveaux organes. Certaines espèces comme le maïs, le raygrass et le peuplier sont même capables de créer des organes racinaires de réserve en réponse à l'attaque d'agents pathogènes (Erb et al. 2009). Ainsi, des génotypes de tomate résistants à *Rhizoctonia solani* ont montré un système racinaire plus ramifié en présence de l'agent pathogène (Berta et al. 2005; Sampo et al. 2007). Cette augmentation de la ramification en présence de l'agent pathogène a également été montrée par Trotta et al. (1996) sur des génotypes de tomates résistants à *Phytophthora nicotianae* var. *parasitica*. Dans les pathosystèmes tomate/*R. solani* (Larkin and Fravel 1998) et *Medicago sativa*/*Pithium* sp. (Sampo et al. 2007), l'infection induit un plus faible nombre de ramifications par rapport aux

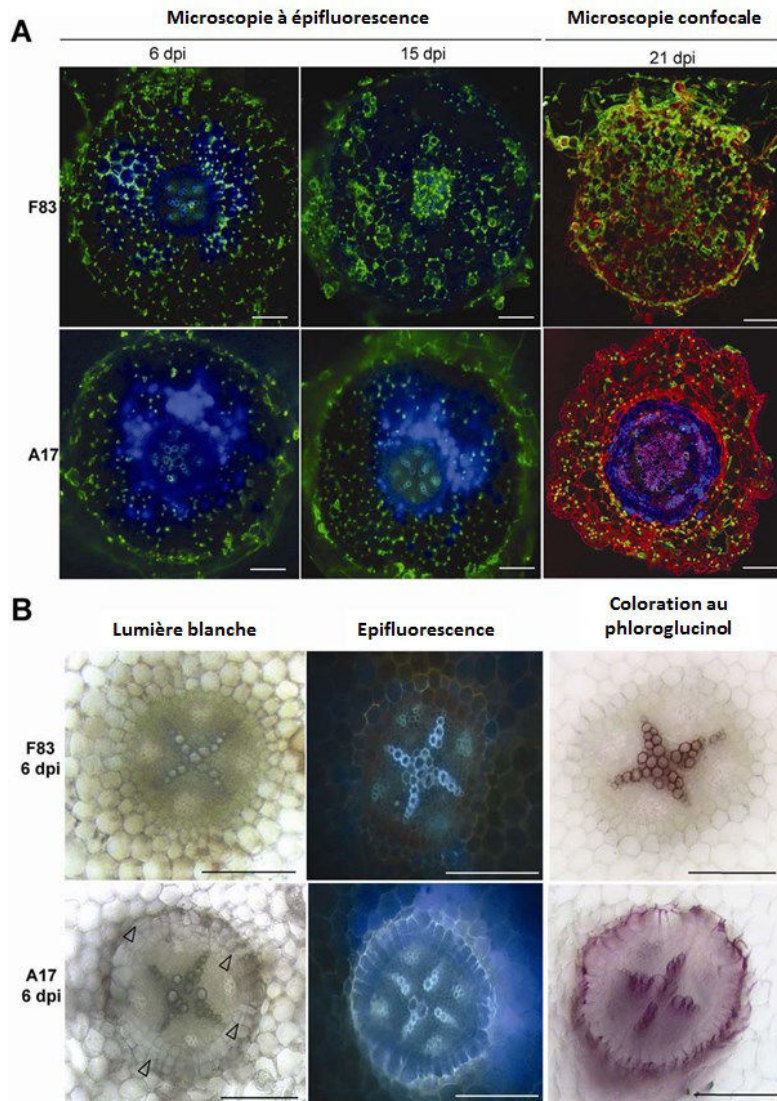


Figure I.5 : Analyses cytologiques comparant les réponses de deux génotypes de *Medicago truncatula* à une infection par *Aphanomyces euteiches* (d'après Djébali *et al.* (2009))

Le génotype A17 est considéré comme résistant et F83005.5 (F83) comme sensible ; Barre = 200 μm ; dpi : nombre de jours post-inoculation

A : Observation du développement d'*A. euteiches* (vert) et de l'accumulation de composés phénoliques (bleu) dans les deux génotypes. Les coupes transversales de racines inoculées colorées avec de l'isothiocyanate de fluorescéine sont observées au microscope à épi-fluorescence (à 6 et 15 dpi) ou par microscopie confocale (à 21 dpi). La concentration de mycélium dans les tissus racinaires est proportionnelle au niveau de fluorescence de la plante. Le mycélium n'est pas observé dans le cylindre central d'A17. À 21 dpi, les parois cellulaires (rouges) sont finement dessinées dans les racines d'A17 et renforcées dans les couches de cellules environnant la stèle racinaire. L'auto-fluorescence bleue est principalement contenue dans la stèle, où aucun mycélium n'est présent. Au contraire, tous les tissus racinaires de F83 sont entièrement colonisés par le mycélium d'*A. euteiches*, et de nombreuses parois cellulaires sont interrompues, y compris celles autour de la stèle, provoquant l'invasion de l'agent pathogène.

B : Observation des changements cytologiques à 6 dpi dans les racines inoculées d'A17 et F83 et les racines non inoculées d'A17. Des couches de cellules supplémentaires sont observées tout autour de la stèle d'A17, suite à des divisions cellulaires plus importantes du péricycle (flèches). Les parois cellulaires extérieures du péricycle sont renforcées avec des composés fluorescents sous lumière UV chez A17. Par coloration au phloroglucinol, ces composés ont été identifiés comme étant proches de la lignine. Aucune de ces caractéristiques n'est observée dans les racines infestées de F83.

contrôles sains.

À une échelle beaucoup plus fine, Djébali et al. (2009) ont montré que la résistance de *Medicago truncatula* à *A. euteiches* était associée à des mécanismes de protection du cortex racinaire par (i) une division fréquente des cellules du péricycle, (ii) un dépôt de lignine autour du péricycle, et (iii) une accumulation de composés phénoliques solubles (**Figure I.5**). Ces modifications s'accompagnent d'une plus grande production de racines latérales chez les lignées partiellement résistantes que chez la lignée sensible, en réponse à l'infection par *A. euteiches*. Plus récemment, une étude de génétique d'association a mis en évidence un QTL impliqué à la fois dans une diminution des symptômes de pourriture racinaire et une induction de racines latérales (Bonhomme et al. 2014). Enfin, une corrélation entre augmentation du nombre de racines latérales et tolérance accrue à *A. euteiches* a été observée chez un mutant du gène récepteur de cytokinine MtCRE1 (Laffont et al. 2015). Ainsi, toutes les études menées sur le lien entre architecture racinaire et maladies telluriques s'accordent à dire qu'un système racinaire vigoureux ou ramifié contribuerait à un niveau de résistance supérieur aux agents pathogènes racinaires.

1.2.2.4. Interdépendance des déterminismes génétiques de résistance et d'architecture des plantes et construction d'idéotypes

Les liens observés entre les caractères d'architecture des plantes et la réduction des épidémies ou des symptômes de maladie posent la question de l'interdépendance du contrôle génétique des caractères d'architecture et de résistance des plantes. Dans certains pathosystèmes, des colocalisations ont été mises en évidence entre QTL d'architecture et QTL de résistance. Chez le pois, Prioul et al. (2004) ont montré que des QTL de résistance à *D. pinodes* colocalisaient avec des QTL de date de floraison et de hauteur. De plus, Giorgetti (2013) a identifié 2 zones génomiques de colocalisation de QTL pour la résistance à *D. pinodes* et la taille des stipules. Dans le pathosystème haricot commun/*S. sclerotinium*, Miklas et al. (2013) ont mis en évidence des colocalisations entre QTL de résistance à la pourriture blanche et QTL d'architecture aérienne permettant l'échappement de la plante à la maladie (tenue de tige, hauteur, longueur des entre-nœuds, porosité du couvert...). Chez le framboisier, Graham et al. (2011) ont montré une colocalisation entre des QTL de résistance à la pourriture racinaire et des QTL de distance d'expansion des racines par rapport au pied-mère.

Ces résultats suggèrent que les caractères de résistance aux maladies et d'architecture peuvent être sélectionnés conjointement. Cependant, aucune étude à ce jour n'a permis de préciser si

ces colocalisations étaient dues à des gènes d'architecture à effets pléiotropes, à des gènes identiques présentant d'autres fonctions et contrôlant la résistance et l'architecture (colocalisation vraie), ou à des gènes étroitement liés contrôlant respectivement la résistance et l'architecture (chevauchement d'intervalles de confiance de QTL différents).

La sélection d'idéotypes variétaux combinant des locus de résistance et d'architecture défavorables à une maladie, qu'ils soient liés ou indépendants génétiquement, constitue une stratégie d'intérêt pour limiter les épidémies de façon *a priori* plus durable. Toutefois, le développement de telles variétés n'est pas si simple car les caractères architecturaux défavorables au développement de maladies ne sont pas forcément favorables à leur exploitation agronomique. Par exemple, une plante avec un feuillage peu dense et de longs entre-nœuds sera moins susceptible d'intercepter un agent pathogène et d'entretenir un microclimat favorable à son développement, mais cette plante sera également plus sensible à la verse, ce qui causera des pertes de rendement. Il est donc indispensable de définir des compromis pour limiter le développement des maladies, l'usage des pesticides et les pertes de rendement conjointement (Andrison et al. 2013). Miklas et al. (2013) ont prouvé l'efficacité de cette stratégie dans le pathosystème haricot commun/ *Sclerotinia sclerotinium*.

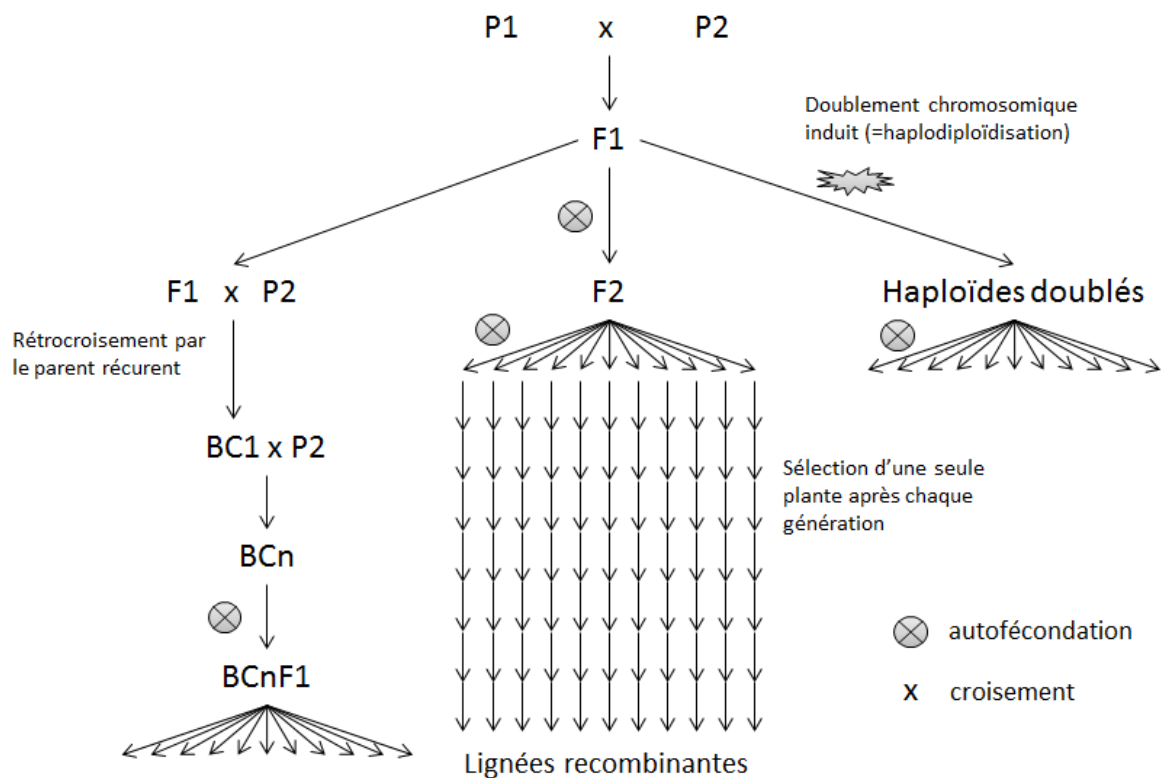


Figure I.6 : Principaux types de populations biparentales pour la cartographie de liaison

BC : backcross ; P : parent

2. Dissection génétique des caractères complexes chez les végétaux

La détection de QTL met en regard la variation quantitative observée d'un caractère avec la variabilité génétique d'une population, représentée par des marqueurs moléculaires ou morphologiques balisant une carte génétique ou un génome. Deux approches principales sont utilisées pour élucider le contrôle génétique de caractères complexes : l'analyse de liaison, le plus souvent basée sur des populations biparentales, et l'analyse d'association basée sur des populations non apparentées.

2.1. Les marqueurs moléculaires

Une grande variété de marqueurs moléculaires a été utilisée durant les vingt dernières années pour les analyses génétiques des caractères complexes. Ces marqueurs présentent un polymorphisme au niveau de l'ADN (variabilité du code génétique dans une séquence d'ADN) basé sur :

- la présence ou l'absence de sites de restriction (*Restriction Fragment Length Polymorphism* : RFLP) ;
- la longueur de fragments amplifiés (*Amplified Fragment Length Polymorphism* : AFLP) ;
- le nombre de séquences répétées en tandem (*Single Sequence Repeat* : SSR) ;
- un nucléotide unique (*Single Nucleotide Polymorphism* : SNP).

Alors que les trois premiers types de marqueurs présentent souvent plus de deux allèles, les marqueurs SNP sont en général bialléliques. Les marqueurs SNP sont donc moins informatifs, mais ils sont présents en très grand nombre dans le génome. De plus, avec l'évolution rapide des techniques de séquençage, les marqueurs SNP sont maintenant peu coûteux à développer et à analyser sur un grand nombre d'individus. Ils représentent donc aujourd'hui les marqueurs moléculaires les plus couramment utilisés en génétique.

2.2. Analyse de liaison : approche QTL

Depuis une vingtaine d'années, la détection de QTL a été largement menée à l'aide de populations biparentales en ségrégation. Ces populations, issues de croisement entre deux parents contrastés pour un ou plusieurs caractères d'intérêt, comprennent des populations F₂, d'haploïdes doublés (HD), de lignées recombinantes (*Recombinant Inbred Lines* : RIL) et de rétrocroisement avec un parent récurrent (*backcross* : BC) (**Figure I.6**). Initialement, lorsqu'un petit nombre de marqueurs était disponible pour couvrir le génome, la détection de QTL était réalisée **marqueur par marqueur**. Dans ce cas, il s'agissait de tester si les sous-classes de la population pour un caractère donné, étaient liées à la variation d'allèle du

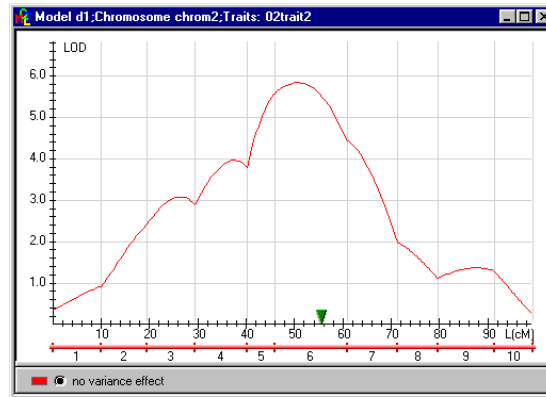


Figure I.7 : Représentation graphique de la statistique LOD score le long d'un chromosome par cartographie d'intervalle, logiciel MultiQTL (extrait de *Charmet (2011)*)

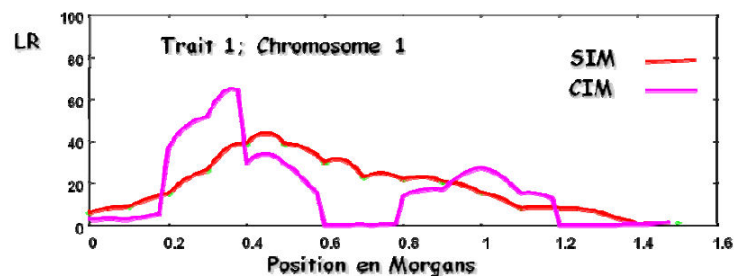


Figure I.8 : Comparaison des méthodes de cartographie d'intervalles simple (SIM) et composite (CIM) : tracé du rapport de vraisemblance (Likelihood Ratio) en fonction de la position le long du chromosome (extrait de *Charmet (2011)*)

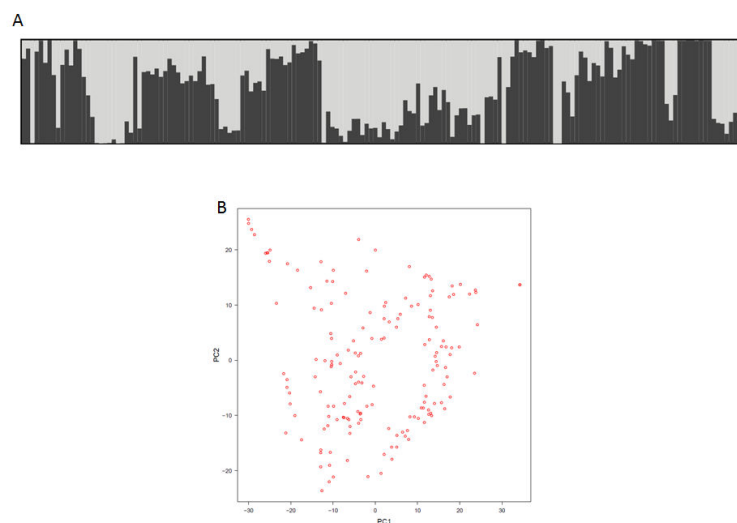


Figure I.9 : Exemple de représentation graphique de la structure d'une même population avec la méthode (A) logiciel STRUCTURE et (B) ACP sur les marqueurs (package R GAPIT)

L'analyse avec le logiciel STRUCTURE indique deux sous-populations alors que l'ACP ne montre pas de structuration claire de la population étudiée.

marqueur (Sax 1923). Une analyse de variance à un facteur (le marqueur) était alors réalisée. Or, si le marqueur n'était pas localisé exactement au QTL, la fréquence de recombinaison entre le marqueur et le vrai QTL n'était pas prise en compte, conduisant à des biais de détection.

Par la suite, des méthodes basées sur des marqueurs flanquants d'intervalles ont été développées. Les sous-classes de la population ont ainsi été comparées aux allèles des deux marqueurs flanquants, prenant ainsi en compte le taux de recombinaison entre ces deux marqueurs. Il s'agit de la **cartographie d'intervalles**, utilisant principalement la méthode du maximum de vraisemblance pour estimer la position et l'effet du QTL (Lander and Botstein 1989). La position du QTL est alors fixée et l'espace de variation est exploré pas à pas de manière exhaustive. À chaque pas, le rapport entre la vraisemblance de la présence du QTL et celle de l'hypothèse nulle sans QTL est calculé. Généralement, on représente graphiquement une fonction logarithmique de ce rapport, le LOD score (*Logarithm of Odds*), en fonction de la position (**Figure I.7**). La position la plus probable du QTL est celle qui maximise la vraisemblance du modèle avec le QTL, et donc celle où le LOD score est à son maximum. L'intervalle de confiance est souvent déterminé autour du maximum lorsque le LOD score est abaissé de 1 par rapport à sa valeur la plus élevée. Toutefois, en utilisant cette méthode, la variance inter-classe comprend toutes les sources de variation non prises en charge par le QTL, y compris les QTL localisés en dehors de l'intervalle exploré.

Des méthodes de **composite interval mapping** (CIM) se sont alors développées. Elles permettent de mettre en covariables les QTL détectés, pour la suite du scan du génome et se réalisent en deux étapes : (i) une régression multiple de l'ensemble des marqueurs sur le phénotype pour le choix des covariables et (ii) une recherche de QTL pas à pas sur un intervalle par régression linéaire ou par maximum de vraisemblance, après avoir fixé les covariables (Zeng 1994). Ces méthodes permettent à la fois un gain de puissance (détection des QTL ayant des effets plus faibles) et une meilleure précision (intervalles de confiance plus étroits) (**Figure I.8**).

Cependant ces intervalles restent généralement larges et ne permettent pas le clonage positionnel et l'identification de gènes candidats. De plus, les effets sont basés uniquement sur le contraste entre les deux allèles parentaux.

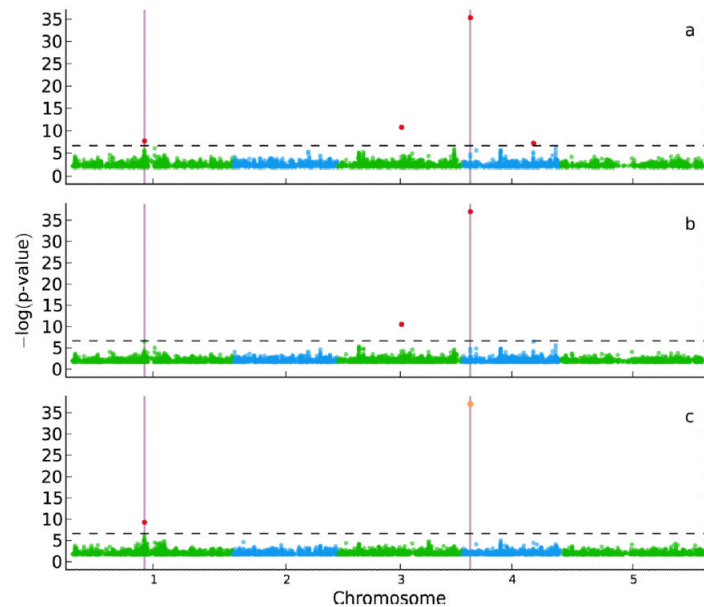


Figure I.10 : Représentation graphique des résultats d’analyse MLMM de génétique d’association pour deux SNP causaux simulés (lignes verticale) (extrait de Segura et al. 2012)

Caractère simulé avec une héritabilité de 25 % pour un set de données génotypiques réelles d’*Arabidopsis thaliana*.

(a) Un scan du génome avec une régression linéaire à un seul marqueur détecte 4 marqueurs significatifs (seuil de détection représenté par la ligne verticale en pointillés) représentés en rouge. La moitié des SNP détectés sont des faux positifs et l’autre moitié des vrais positifs. Le taux de faux positifs (FDR) est donc de 50 % et la puissance de détection de 100 %.

(b) Un scan du génome avec un modèle mixte (prise en compte de l’apparentement) élimine un faux positif mais aussi un vrai positif. Le FDR est donc similaire (50 %) et la puissance de détection diminue (50 %).

(c) L’ajout du SNP le plus significatif en cofacteur dans le modèle mixte (orange) permet de retrouver le deuxième SNP causal et d’éliminer le deuxième faux positif. Le FDR est donc à 0 % et le pouvoir de détection à 100 %.

Tableau I.2 : Comparaison des avantages et inconvénients des différents types de populations pour l’analyse de liaison et/ou d’association (d’après Pascual et al. 2015a)

DL : déséquilibre de liaison

SAM : sélection assistée par marqueurs

	Populations biparentales	Populations multi-parentales	Population de diversité
Temps de développement	Intermédiaire	Long	Court
Précision de cartographie des allèles fréquents	+	++	+++
Précision de cartographie des allèles rares	+++	+	-
Evènements de recombinaison lors de la construction	+	++	+++
Nombre de marqueurs nécessaires	+	++	+++
Structure de la population à prendre en compte	Non	Non	Oui
Principaux avantages	Cartographie des allèles rares Facilité d’analyse	Plusieurs allèles en ségrégation Effet des allèles parentaux pour la SAM ou l’identification de QTL	Précision de détection due aux recombinaisons historiques
Principaux inconvénients	Variation des allèles limités aux deux parents Intervales de confiance des QTL larges	Temps de construction très long Nécessité d’une population large	Un fort DL limite la précision La structure de la population est responsable de faux positifs

Des méthodes de **méta-analyse** ont été également développées (Goffinet and Gerber 2000), permettant de réduire les intervalles en comparant les résultats de plusieurs études. Les QTL issus d'études indépendantes (populations et environnements indépendants) sont comparés grâce à des marqueurs communs entre les différentes cartes, permettant ainsi une réduction de l'intervalle de confiance et une estimation des positions des QTL indépendants des populations et des environnements étudiés.

2.3. Analyse de déséquilibre de liaison : approche génétique d'association

La construction de populations biparentales pour l'analyse de liaison est souvent longue et coûteuse. De plus les variations génétiques sont limitées aux différences entre les deux parents de la population d'étude et les événements de recombinaison limités lors de la construction de la population. Contrairement à la cartographie par intervalle, la génétique d'association est basée sur le déséquilibre de liaison (DL) entre individus non apparentés d'une population (Gupta et al. 2014). Cette méthode tire parti des événements de recombinaison historiques au sein d'une population naturelle pour identifier des marqueurs associés aux variations phénotypiques de caractères quantitatifs. Du fait des événements de recombinaison plus nombreux, un plus grand nombre de marqueurs est nécessaire pour les représenter. L'essor de la génétique d'association bénéficie donc de la diminution du coût de génotypage à haute densité et du développement des nouvelles méthodes de séquençage (*Next-Generation Sequencing* : NGS), permettant le génotypage par séquençage (*Genotyping-By-Sequencing* : GBS).

Si la génétique d'association a été développée sur des populations naturelles (individus indépendants), de telles populations ne sont pas fréquentes chez les espèces cultivées. En effet, les événements de domestication et de sélection ont fait que les individus étudiés ont souvent des ancêtres communs. La variation d'un caractère peut donc parfois être liée à la présence de sous-populations ou d'apparentements entre les individus. Afin de s'affranchir de ces biais, les méthodes de génétique d'association prennent généralement en compte une matrice de structure de la population (Q) et une matrice d'apparentement entre les individus (K) comme cofacteurs dans l'analyse (Gupta et al. 2014). D'ordinaire, la structure de la population (matrice Q) peut être décrite par deux types de paramètres : (i) le nombre de sous-populations et la probabilité d'appartenance des individus à une sous-population, estimés avec le logiciel STRUCTURE ou (ii) les coordonnées des individus sur un certain nombre de composantes principales issues d'une ACP (Analyse en composantes principales). Les deux méthodes d'analyse de la structure sont réalisées à partir de données de génotypage des

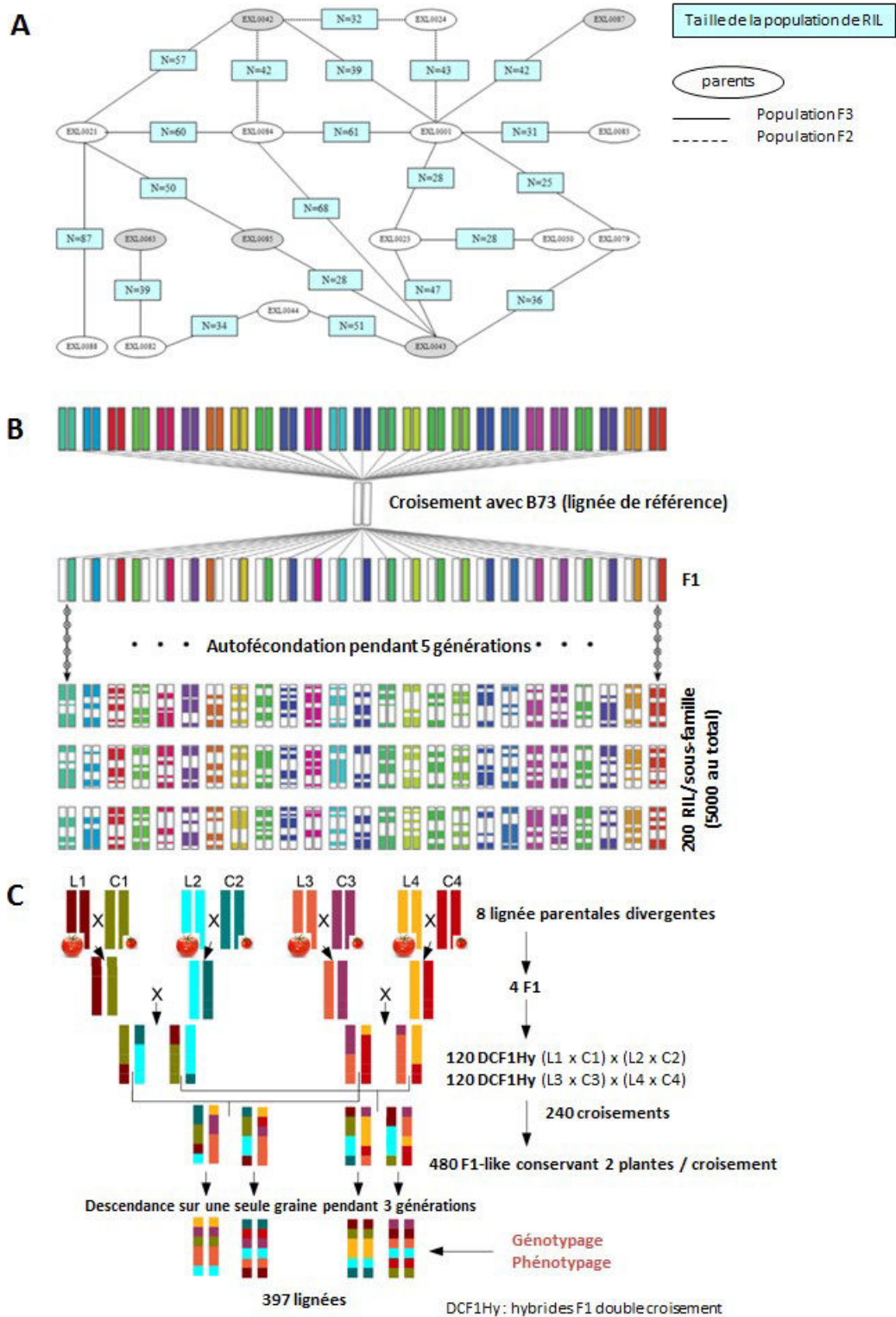


Figure I.11 : Exemples de construction de populations multi-parentales

- (A) population biparentales connectées, d'après Bardol et al. (2013) ;
 (B) population NAM chez le maïs à partir de 25 lignées parentales et une lignée de référence (B73), d'après Wallace et al. (2014) ;
 (C) population MAGIC chez la tomate à partir de 8 lignées parentales inter-croisées, d'après Pascual et al. (2015b).

individus pour des marqueurs pris au hasard dans la matrice de génotypage et répartis sur l'ensemble du génome. Si la méthode STRUCTURE (Pritchard et al. 2000) a été couramment utilisée, elle a également souvent été remise en question car elle surestimerait le nombre de sous-populations sous certaines conditions (**Figure I.9**). Quoi qu'il en soit, ces deux méthodes ne sont pas efficaces pour capturer la structure des populations complexes. Il est donc courant d'associer à la matrice Q, une matrice d'apparentement entre les individus, ou matrice de Kinship (K). L'analyse de l'association marqueur-phénotype est alors réalisée en utilisant un modèle linéaire mixte (MLM, (Yu et al. 2006)) : $Y = X\alpha + Q\beta + Ku + \varepsilon$, avec Y le phénotype, X les allèles au marqueur, α l'effet fixe du locus, Q, la matrice de structure, β l'effet fixe de la population, K la matrice d'apparentement, u l'effet aléatoire de l'apparentement et ε la résiduelle.

Sur le même principe que le CIM en cartographie par intervalle, Segura et al. (2012) ont récemment développé un modèle multilocus (*Multilocus Linear Mixed Model* : MLMM) qui prend en compte la structure et l'apparentement de la population mais inclus également une méthode de sélection de marqueurs significatifs en tant que co-facteurs par régression *forward* et *backward*. Cette méthode prend en compte l'effet de confusion potentiel des locus du fonds génétique en raison du DL. Elle permet à la fois de supprimer des marqueurs faux positifs (dus à un déséquilibre de liaison entre les marqueurs) et faux négatifs (effet faible masqué par un autre marqueur à effet fort), et d'augmenter la puissance de détection (**Figure I.10**). Certains modèles multi-caractères ont également été développés (Gupta et al. 2014) mais aucun permettant une analyse à la fois multi-locus et multi-caractères simultanément.

2.4. LD-LA : approche mixte

Récemment, des méthodes d'analyse mixte de liaison et d'association, nommées *joint linkage association mapping* (JLAM) ou LD-LA (*Linkage Disequilibrium-Linkage Analysis*), se sont développées à partir de populations multi-parentales. Ces méthodes présentent l'intérêt de tirer parti des avantages des approches de liaison et d'association, incluant notamment la connaissance de la structuration des populations, la prise en compte d'événements de recombinaison anciens et récents et la détection des allèles rares (Tableau I.2). Les populations multi-parentales utilisées par cette approche peuvent inclure (i) des populations biparentales connectées, partageant des parents communs, (ii) des populations NAM (*Nested Association Mapping*) composées de RIL issues de croisements entre une lignée parentale commune et plusieurs lignées parentales d'origines diversifiées et (iii) des populations MAGIC (*Multi-parent Advanced Generation Inter-Cross*), dans lesquelles des croisements sont réalisés entre parents deux à deux, puis entre descendants, afin de mélanger au maximum

les génomes de tous les parents (**Figure I.11**).

3. Le pathosystème Pois/Aphanomyces euteiches

3.1. Le pois protéagineux, une culture d'avenir pour la production de protéines végétales en Europe

3.1.1. Enjeux et intérêts de la culture de pois

Originnaire du Moyen-Orient, le pois protéagineux (*Pisum sativum* L.) est une culture récente en Europe. Son essor débute dans les années quatre-vingt suite à l'embargo des États-Unis sur le tourteau de soja. Cet embargo rend indispensable le développement de cultures riches en protéines pour se substituer l'apport protéique du tourteau de soja dans les rations animales en Europe. Le pois protéagineux est alors une bonne alternative, présentant une teneur en protéines intermédiaire (24 % de la matière sèche) entre les céréales (13 %) et le soja (40 %). En Europe, le pois protéagineux a un fort débouché en alimentation animale (principalement monogastriques et oisellerie), mais également à l'export pour l'alimentation humaine (pois verts de casserie et pois jaunes de bonne qualité visuelle).

La culture de pois possède plusieurs atouts économiques au niveau de l'exploitation agricole et pour l'environnement. L'introduction du pois dans les rotations céréalières permet de diminuer l'utilisation d'intrants, du fait de sa capacité à fixer l'azote atmosphérique (symbiose avec *Rhizobium*), et d'augmenter le rendement des cultures suivantes. Ainsi, l'enfouissement des pailles de pois permet de diminuer la fertilisation azotée des cultures de céréales ou de colza suivantes de 20 à 60 kgN/ha en moyenne (UNIP and ARVALIS 2014). Un précédent pois permet d'augmenter le rendement d'une culture de blé de 6 à 10 q/ha. Le pois a une période de semis décalée ainsi que des ravageurs et des substances actives de désherbage différents par rapport aux principales cultures. Il permet donc de rompre le cycle des ravageurs et des adventices, facilitant leur contrôle sur les cultures suivantes. Par la limitation de l'utilisation d'engrais azotés et de produits phytosanitaires sur les cultures suivantes, le pois participe à une diminution de l'impact environnemental de l'agriculture. Du fait de son cycle court et précoce en semis de printemps, il permet également une économie de l'irrigation, participant à la préservation des nappes phréatiques et limitant le lessivage des reliquats.

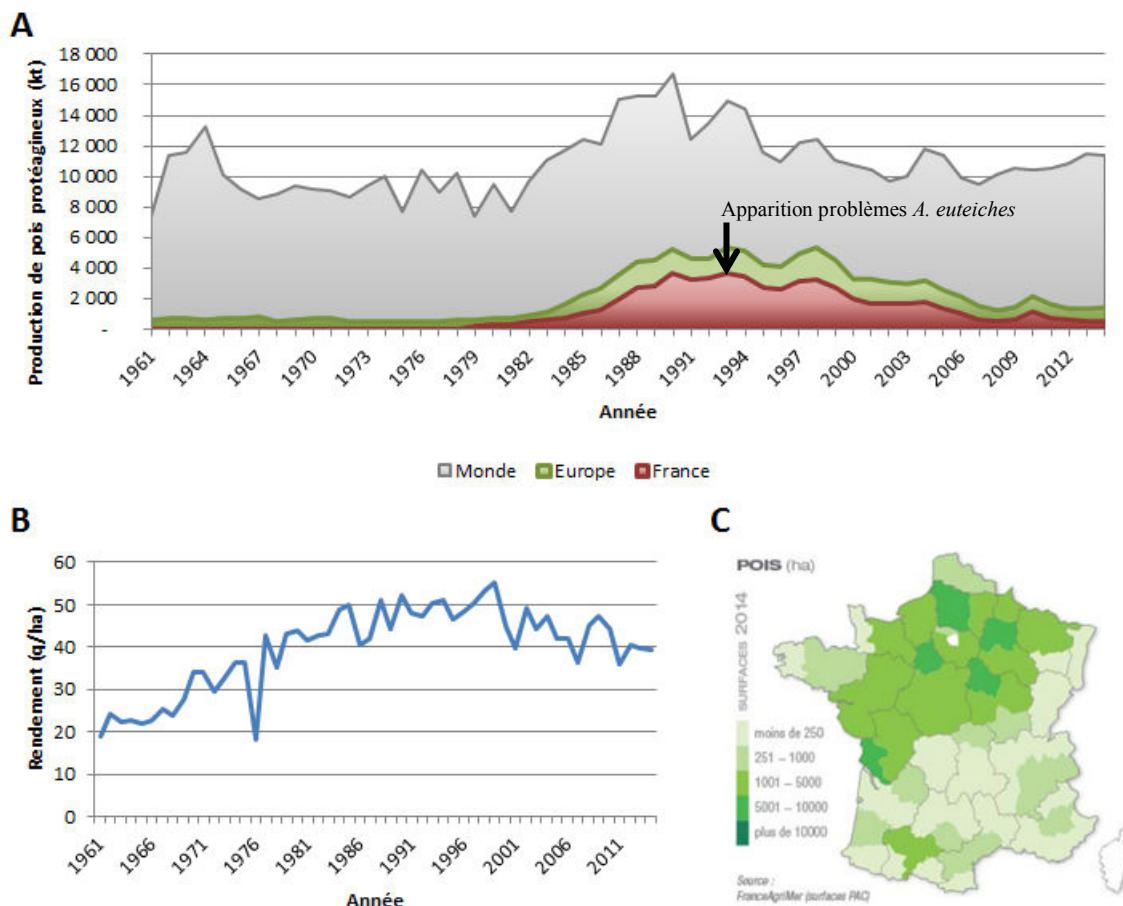


Figure I.12 : Évolution de la production de pois en France

(A) Évolution de la production de pois protéagineux dans le monde, en Europe et en France de 1961 à 2014

(B) Évolution du rendement du pois protéagineux en France de 1961 à 2014

(C) Surfaces cultivées en pois protéagineux en France en 2014

(A) et (B) source : (FAOSTAT 2015)

(C) source : (Terres Univia 2014)

3.1.2. Production et rendement

Le pois est cultivé pour trois usages principaux pour lesquels les variétés utilisées sont spécifiques :

- les pois fourragers sont récoltés pour leur qualité fourragère à destination animale ;
- les pois potagers, récoltés non matures pour leurs graines, sont soit consommés frais soit transformés sous forme de conserves ou surgelés ;
- les pois protéagineux, récoltés à maturité des graines, sont riches en protéines et utilisés aussi bien en alimentation humaine qu'animale, entiers ou sous forme de farine. Ils constituent la majeure partie de la production de pois dans le monde.

La France est aujourd'hui le sixième pays producteur de pois protéagineux dans le monde (environ 500 Mt en 2013), et le premier producteur en Europe (FAOSTAT 2015). Les principales régions productrices sont celles du Nord de la France (**Figure I.12**). Après un essor de la culture dans les années quatre-vingt, la production de pois en France a marqué un palier dans les années quatre-vingt-dix. Les années deux-mille ont vu une très nette diminution de la production, notamment due à la première réforme de la PAC (Politique agricole commune) en 1992 puis l'Agenda 2000 défavorables aux prix du pois à l'hectare. En 2010, le plan « Protéines » redonne de l'intérêt à la culture de pois en favorisant la culture des légumineuses. En 2014, le plan « Protéines végétales pour la France » fait suite à celui de 2010, confirmant la nécessité de développer la filière des légumineuses à graines.

3.1.3. Types de culture

On distingue deux grands types cultureux de pois protéagineux, l'un semé au printemps, l'autre à la fin de l'automne. Le pois de printemps est actuellement prédominant en France. Contrairement aux céréales pour lesquels les types « hiver » sont définis par leur besoin en hivernage, les pois d'hiver se caractérisent par leur résistance au gel. Ils sont particulièrement intéressants pour leur potentiel de rendement, ainsi que leur cycle décalé permettant d'échapper à certains stress, notamment pendant la floraison et en fin de cycle. Cependant, ils sont souvent victimes des périodes de gels tardives après la reprise de la végétation. On distingue deux types de pois d'hiver : le type « *Hr* » (*Highly responsive to photoperiod for floral initiation*) et le type « *hr* ». Contrairement au type « *hr* », le type « *Hr* » est sensible à la photopériode (durée du jour) pour l'initiation de la floraison (Lejeune-Hénaut et al. 2008). Il est donc moins sensible aux gelées tardives si un redoux a lieu en fin d'hiver.

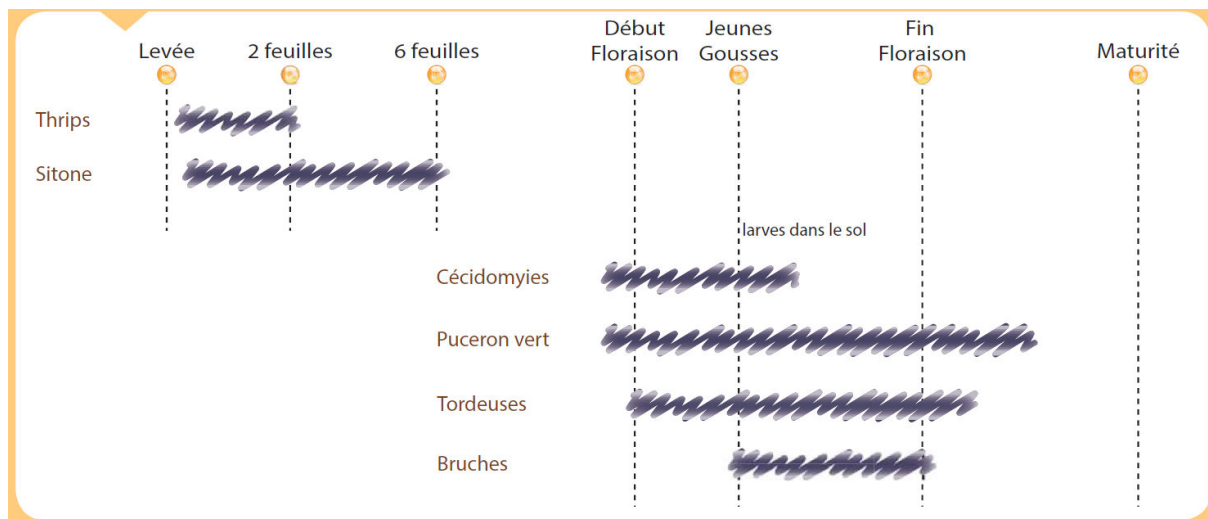


Figure I.13 : Ravageurs du pois de printemps - Période de surveillance et de lutte (source : UNIP and ARVALIS (2014))

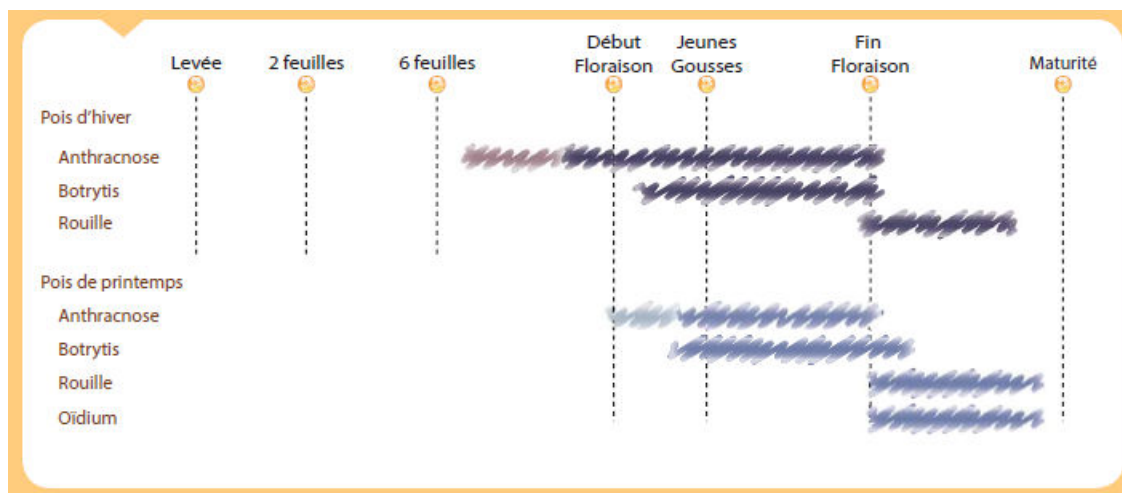


Figure I.14 : Maladies aériennes du pois - périodes de surveillance et de lutte (source : UNIP and ARVALIS (2014))

La culture du pois en France a été impactée par de fortes fluctuations de rendement (**Figure I.12**) dues à sa sensibilité aux stress abiotiques et biotiques. Concernant les stress abiotiques, le pois est notamment sensible aux gelées tardives, aux coups de chaleur et aux stress hydriques pendant la floraison.

3.1.4. Ravageurs et maladies du pois protéagineux

Un certain nombre de ravageurs et de maladies peuvent également être préjudiciable pour le rendement. Parmi les ravageurs, les thrips, les sitones, les pucerons verts du pois et les cécidomyies du pois peuvent avoir une forte influence sur le rendement alors que les tordeuses et les bruches affectent principalement la qualité de la production mais peu le rendement (**Figure I.13**). La maladie aérienne la plus dommageable est l'ascochytose (jusqu'à - 20 q/ha), mais il existe également des risques de mildiou (jusqu'à - 20 q/ha), de botrytis (jusqu'à - 10 à - 15 q/ha), de rouille (jusqu'à - 5 à - 12 q/ha), de sclérotiniose (jusqu'à - 5 q/ha), d'oïdium et de bactériose suivant les régions (UNIP and ARVALIS 2014) (**Figure I.14**). L'agent pathogène le plus préoccupant en culture de pois, *Aphanomyces euteiches*, est d'origine tellurique et peut provoquer des pertes de rendement considérables voire totales dans les parcelles fortement infestées. D'autres maladies racinaires existent mais sont moins préoccupantes, telles que celles causées par *Fusarium solani*, *Fuarium oxysporum* et *Verticillium* ou, de façon plus rare, par des nématodes. Le développement de moyens de lutte contre *A. euteiches*, notamment de type génétique, est donc devenu une préoccupation majeure de la filière pois en France, impliquant les instituts techniques, les sélectionneurs ainsi que les chercheurs, en collaboration.

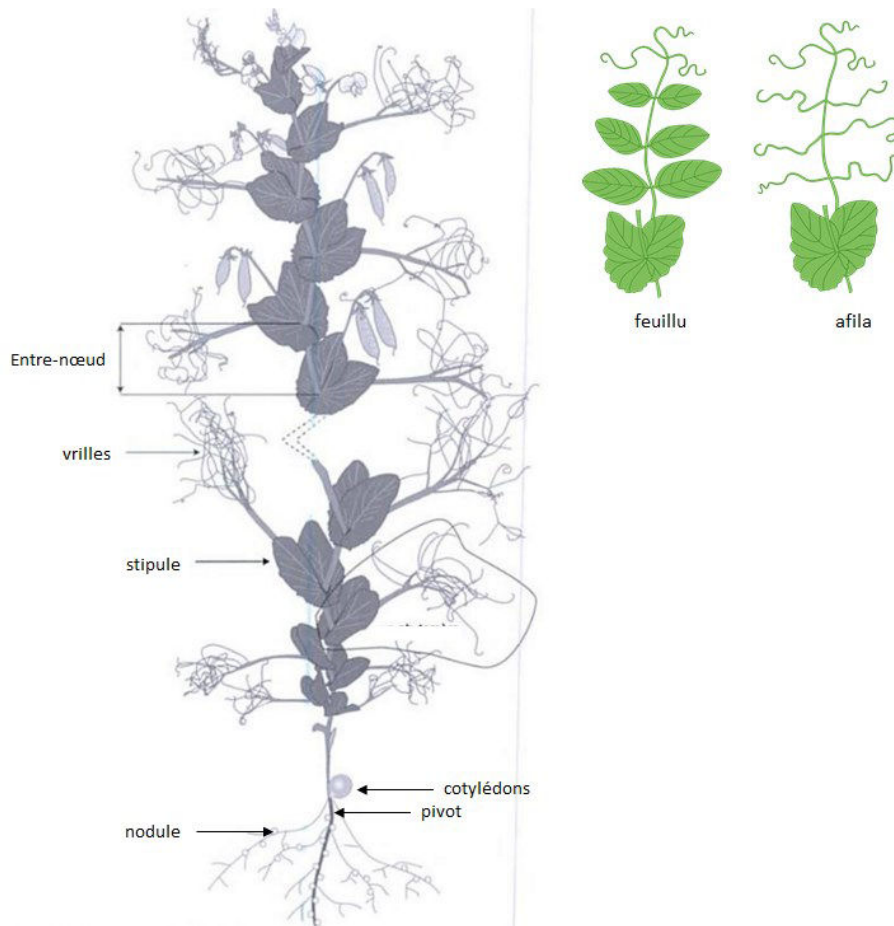


Figure I.15 : Organisation d'une plante de pois (d'après *Lecoœur (2005)*)

3.1.5. Caractéristiques biologiques du pois protéagineux

Le pois (*Pisum sativum*, L.) est une légumineuse annuelle et cléistogame (*i. e.* autopollinisation stricte des fleurs qui ne s'ouvrent pas). L'apex de la tige différencie des nœuds alternes successifs (**Figure I.15**). Les feuilles normales sont composées et munies à leur base de larges stipules à forme arrondie. Une variation particulière de forme de feuilles, nommée *afila*, présente un intérêt agronomique. Cette variation, remplaçant des folioles par des vrilles (**Figure I.15**), permet de diminuer la verse (par accrochage entre les plantes) et donc de faciliter la récolte tout en diminuant la densité de feuillage (diminution potentielle de maladies aériennes). Toutes les variétés nouvellement inscrites en pois protéagineux sont maintenant de type *afila*. Appartenant à la famille des Légumineuses, le pois présente des nodosités (ou nodules) sur ses racines. Ces nodosités sont le résultat d'une symbiose avec des bactéries du genre *Rhizobium*, fixatrices de l'azote atmosphérique. Depuis les travaux de Mendel (1822-1884), le pois est reconnu pour sa richesse en variations morphologiques. Les grains présentent une grande variabilité d'une variété à l'autre, pour leur couleur (jaune, vert, brun, marbré), leur forme (lisse ou ridé) ou leur taille. La large variabilité de son architecture racinaire a également été démontrée au sein de collections d'accessions de pois dans différentes conditions contrôlées et/ou au champ (Bourion et al. 2007; Bourion et al. 2010; McPhee 2005). Cette variabilité touche la plupart des caractères de l'architecture globale comme la longueur de pivot, la longueur totale de racines, la surface racinaire ou le poids sec. Une classification des différents types d'architecture racinaire basée sur la longueur de pivot, la densité, la distribution et la longueur des racines secondaires a été proposée (McPhee 2005). Bourion et al. (2007) et Bourion et al. (2010) ont montré que la variabilité de l'architecture racinaire était observable à des stades avancés et qu'il existait une variabilité de dynamique de croissance des racines. Les auteurs ont mis en évidence une relation positive entre la croissance du système racinaire et l'établissement de nodules permettant la nutrition azotée. Des liens génétiques ont aussi été identifiés entre caractères développement racinaire, nodulaire et d'acquisition d'azote, démontrant l'impact des développements racinaire et nodulaire sur l'acquisition d'azote et le rendement (**Figure I.16**).

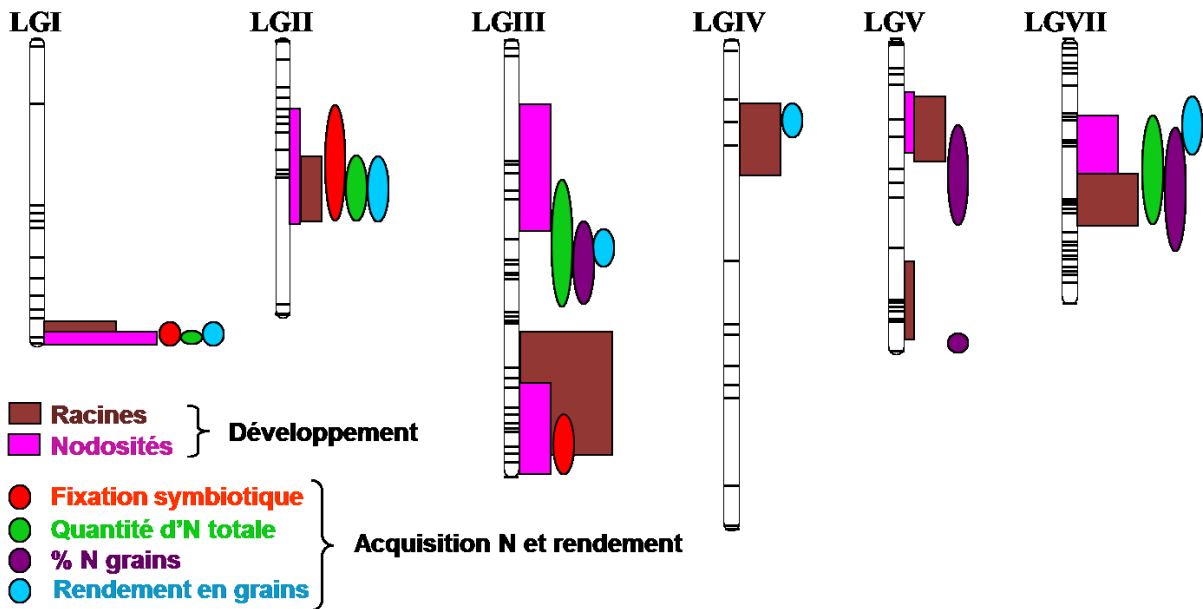


Figure I.16 : QTL d'architecture racinaire et de nutrition azoté (d'après(Bourion et al. 2010))

Approche QTL : Des zones du génome communes aux caractères de développement racinaire, nodulaire et d'acquisition d'azote montrent l'impact des développements racinaire et nodulaire sur l'acquisition d'azote et le rendement

3.1.6. Ressources biologiques et génomiques pour l'amélioration génétique du pois

3.1.6.1. Ressources génétiques

Près de 74 000 accessions de pois sont disponibles dans plus de 28 collections nationales ou internationales conservant entre 200 et 9 700 accessions (Warkentin 2015). Environ 2 % sont des accessions sauvages, apportant de la diversité par rapport aux accessions cultivées. Il n'existe pas de *core collection* mondiale représentant la variabilité de l'espèce *Pisum sativum*, mais plusieurs collections de référence constituées à partir de données passeport ou phénotypiques pour des caractères d'intérêt (origine géographique, qualité des graines, morphologie, architecture de la plante, résistance aux maladies...) (Warkentin 2015), telle que la collection de référence de 372 accessions développée à l'INRA (Burstin et al. 2015).

Par ailleurs, des collections de mutants naturels ou induits ont également été développées, en particulier au John Innes Center, Royaume-Uni [1 400 mutants « Fast Neutron » ; (Domoney et al. 2013)], à l'USDA, États-Unis (Marx *Pisum* Genetic Stock collection extraite de 80 000 accessions ; www.ars.usda.gov) et à l'INRA, France [deux populations de 5 000 mutants TILLING (*Targeting Induced Local Lesions IN Genomes*) produites et maintenues à Dijon ; www.urgv.evry.inra.fr/UTILLdb ; (Dalmais et al. 2008)].

Diverses populations de RIL ont été également construites pour étudier des caractères d'intérêt tels que la teneur en protéines des graines (Burstin et al. 2007; Tar'an et al. 2004), la nutrition azotée (Bourion et al. 2010), la résistance aux maladies (Fondevilla et al. 2008; Hamon et al. 2013; Prioul et al. 2004) ou à la tolérance au gel (Lejeune-Hénaut et al. 2008). Des lignées quasi-isogéniques (*Near-Isogenic-Lines* : NIL) comportant des QTL de résistance aux stress ont récemment été développées pour la validation de leurs effets dans différents fonds génétiques (Hascoët et al. 2014; Lavaud et al. 2015).

3.1.6.2. Ressources génomiques

Le pois est une espèce diploïde ($2n = 14$) dont le génome présente une taille élevée de 4 450 Mpb (Tayeh et al. 2015b).

Avec l'essor des NGS, plusieurs programmes nationaux et internationaux ont récemment développé des ressources génomiques de grand intérêt. Actuellement, un consortium international a été développé pour générer la séquence complète du génome du pois (Madoui et al. 2015). Alors qu'en 1925, Wellensiek développait la première carte génétique du pois,

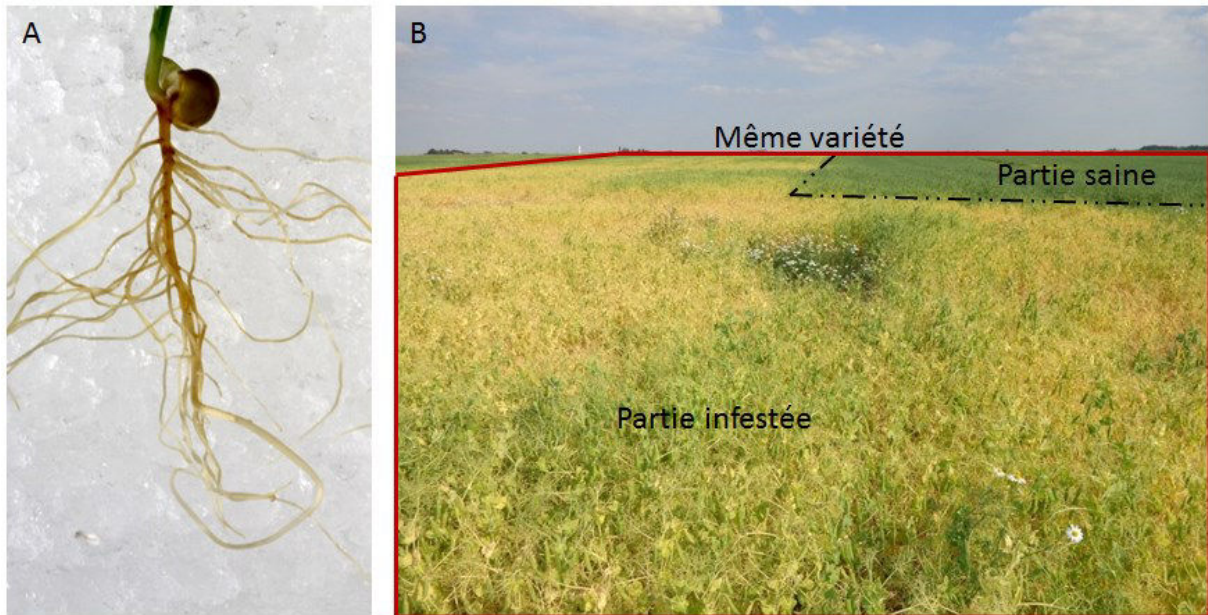


Figure I.17 : Symptômes provoqués par *Aphanomyces euteiches* sur pois

A : Symptômes sur racine (source : INRA) ; (B) Parcelle de pois (même variété) avec deux historiques différents (partie saine/partie infestée). Potentiel infectieux de 2,5 sur la partie gauche.

basée sur une dizaine de marqueurs morphologiques, des progrès considérables ont été fait avec l'apparition des marqueurs SSR et SNP. Aujourd'hui, au minimum 52 cartes génétiques existent, basées sur différentes populations biparentales de F2 ou de RIL. À partir de ces cartes génétiques, des cartes consensus ont été construites comprenant jusqu'à plus de 60 000 marqueurs (Boutet et al. 2016; Tayeh et al. 2015a; Tayeh et al. 2015b). D'autres ressources génomiques intéressantes ont récemment été développées comme l'atlas de gènes du génome du pois (Alves-Carvalho et al. 2015) ou des banques d'ADN génomique de type BAC (Bacterial Artificial Chromosome) (<http://cnrgv.toulouse.inra.fr/fr>), permettant aux chercheurs d'accéder à l'organisation physique du génome du pois.

3.2. *A. euteiches*, un agent pathogène majeur du pois

3.2.1. *Symptômes et importance économique*

Décrit pour la première fois dans les années 1920 (Jones and Drechsler 1925), l'oomycète tellurique *A. euteiches* est l'un des agents pathogènes principaux du pois. Provoquant une pourriture précoce des racines, il est devenu un problème majeur dans les zones de culture intensive de pois en Amérique du Nord, Europe, Australie, Nouvelle-Zélande et Japon. En France, bien que la maladie ait été observée dès les années 1930, elle n'est devenue problématique qu'à partir des années 1990, dans le Bassin parisien et le Nord de la France, suite à l'intensification de la production.

Lorsque les conditions climatiques sont favorables, les symptômes sont visibles dès le stade 3-4 feuilles (Wicker et al. 2001) et apparaissent en deux phases. (i) Les premiers symptômes sont visibles sur les **parties racinaires**. On observe la présence de lésions molles et translucides qui apparaissent sur les radicelles 3 à 4 jours après l'infection. Ces lésions évoluent ensuite en une pourriture brun-miel qui s'étend sur tout le système racinaire (Figure I.17.A). Cette pourriture peut également atteindre l'épicotyle, qui présente alors un étranglement caractéristique. Chez les génotypes très sensibles, ces symptômes entraînent la mort de la plante. Plus tard dans la saison, les racines et les nodosités se dessèchent. (ii) À partir du stade de la floraison, les symptômes sont également visibles sur les **parties aériennes** et se traduisent par un jaunissement, un nanisme puis un dessèchement précoce des plantes. À l'échelle de la parcelle, les symptômes se manifestent sous forme de foyers, d'abord vert pâle puis jaunes (**Figure I.17.B**). Les foyers se développent en fonction de l'historique (travail du sol, rotation...) et de l'hygrométrie (points plus humides) de la parcelle. Lorsque l'infestation est particulièrement importante, les plantes touchées produisent

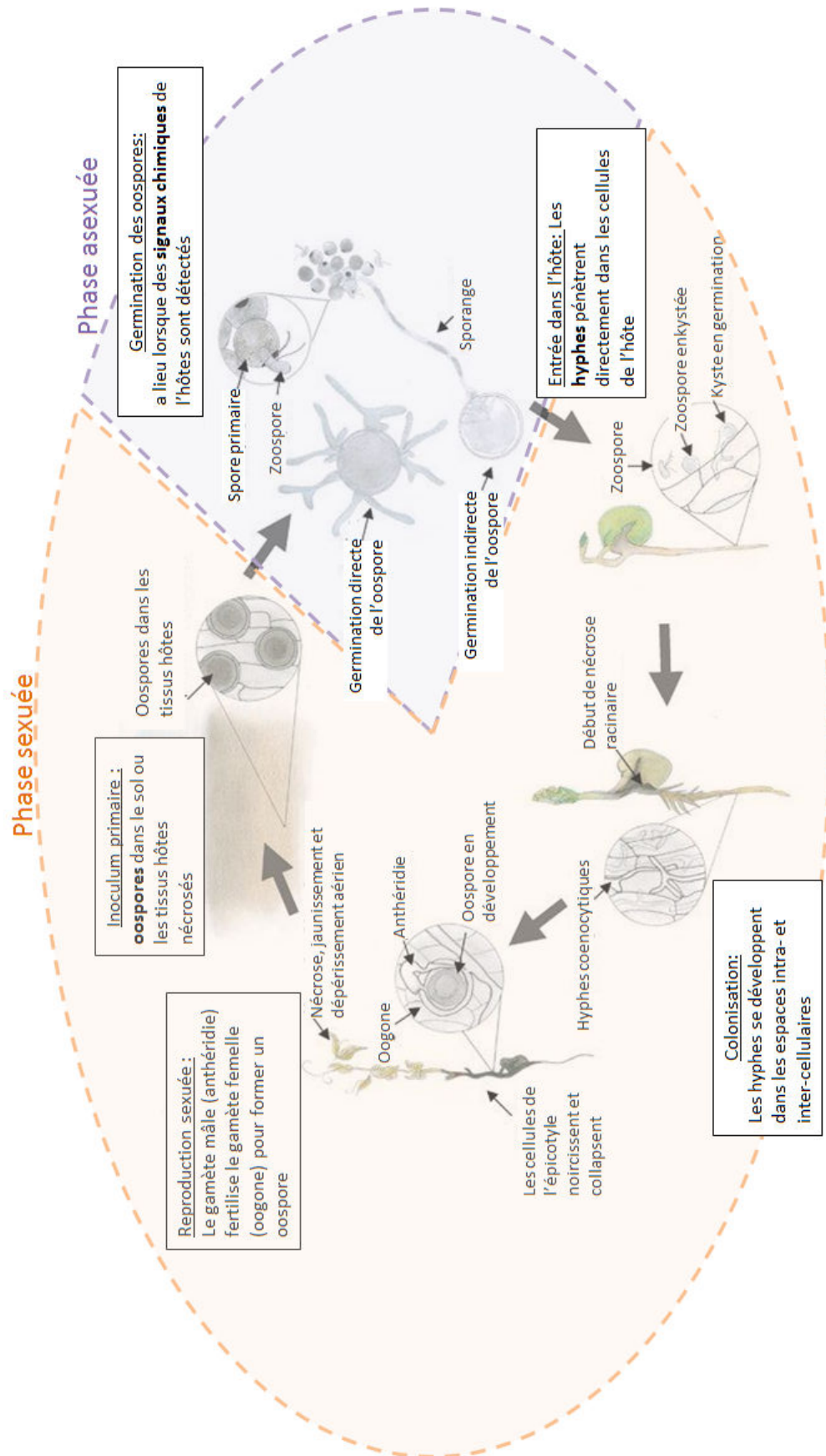


Figure I.18 : Cycle infectieux d'*Aphanomyces euteiches* (d'après Hughes and Grau (2007))

très peu d'étages de gousses et la maladie peut engendrer des pertes de rendements de 20 à 50 q/ha.

3.2.2. Caractéristiques de l'agent pathogène

3.2.2.1. Classification

A. euteiches appartient à la classe des Oomycètes, l'ordre des Saprolegniales et la famille des Saprolegniaceae (Gaulin et al. 2007). Le genre *Aphanomyces* regroupe plusieurs agents phytopathogènes, dont les deux plus importants sont *A. cochlioides* (sur betterave et épinard) et *A. euteiches* (sur de nombreuses espèces de légumineuses). *A. euteiches* est considéré comme un agent pathogène homothallic (gamètes des deux sexes produits par un seul gamétophyte) à mode de vie davantage hémibiotrophe (colonisation des tissus vivants, formes sexuées identifiées dans les tissus morts).

3.2.2.2. Cycle biologique

Le cycle infectieux d'*A. euteiches* est constitué d'une phase sexuée et d'une courte phase asexuée (**Figure I.18**). L'inoculum primaire correspond aux oospores (issues de la reproduction sexuée), présentes sur les résidus de plantes hôtes ou à l'état libre dans le sol. Les oospores, formes de survie qui peuvent résister à des situations extrêmes, ont une durée de vie particulièrement longue (10 à 20 ans) (Papavizas and Ayers 1974). Lorsqu'elles rencontrent des conditions favorables (humidité, température, signaux chimiques émis par la plante hôte), les oospores germent de manière directe ou indirecte. La germination directe permet la production d'hyphes directement issues de l'oospore, tandis que la germination indirecte conduit à la production de sporanges qui libèrent ensuite des zoospores biflagellées. Les zoospores libérées peuvent se déplacer dans l'eau libre du sol grâce à leurs flagelles sur environ 10-15 cm pour atteindre des racines à infester (Willoquet et al. 2007). Elles sont ensuite attirées par chimiotactisme vers les racines avant de s'enkyster à leur contact (Cannesan et al. 2012; Deacon and Saxena 1998). La germination des zoospores conduit alors à la production d'hyphes qui progressent dans les espaces inter- et intra- cellulaires de la racine. Environ 4 jours après l'infection, les premières oogones (organes produisant les gamètes femelles) sont formées, la reproduction sexuée peut donc commencer. Les anthéridies (gamètes mâles) pénètrent les oogones grâce à un tube de fertilisation, conduisant à la production d'oospores diploïdes.

La dissémination d'*A. euteiches* peut se faire (i) à courte distance par mobilité des zoospores dans le sol ou par contact entre les racines ; (ii) à moyenne distance par le travail du sol, le

transfert de sol contaminé sur le matériel agricole, ou les ruissellements au sein des parcelles ; (iii) à longue distance par le transport de sol ou de plantes contaminés.

3.2.2.3. Spécificité d'hôte et pathotypes

Le genre *Aphanomyces* rassemble plusieurs espèces phytopathogènes capables d'infecter plusieurs familles botaniques en conditions contrôlées incluant les légumineuses, les graminées, les brassicacées et les solanacées (Papavizas and Ayers 1974). Au sein des légumineuses, des souches d'*A. euteiches* ont été isolées en milieu naturel à partir de pois, haricot, luzerne, trèfle, lentille et féverole. Les souches isolées du pois présentent un large spectre d'hôtes au sein des légumineuses. Certaines espèces de légumineuses hôtes sont très sensibles et présentent peu de variabilité intra-spécifique pour la résistance (pois, lentille, luzerne et haricot), d'autres présentent une forte variabilité intra-spécifique (féverole, vesce et trèfle) ou sont très résistantes (lupin et pois chiche) (Moussart et al. 2008).

Sur pois, Wicker and Rouxel (2001) ont montré qu'il existait une variabilité de pouvoir pathogène selon les souches d'*A. euteiches*. Deux pathotypes principaux ont ainsi été identifiés parmi 109 souches d'*A. euteiches*, se distinguant par leurs pouvoirs pathogènes différents sur une gamme d'hôtes de six génotypes (Wicker et al. 2003). Le pathotype I a inclus une grande majorité de souches, dont toutes les souches françaises testées, tandis que le pathotype III est apparu spécifique de souches américaines. Deux souches de référence, RB84 (France) et Ae109 (États-Unis), ont été choisies dans plusieurs études pour représenter respectivement les pathotypes I et III (Hamon et al. 2010; Hamon et al. 2013; Lavaud et al. 2015; Moussart et al. 2008).

3.2.3. Méthodes de lutte

Actuellement, il n'existe aucune méthode de lutte curative pour réduire la dissémination et l'impact de la maladie. Certains produits phytosanitaires ont été rapportés comme ayant des effets inhibiteurs sur la maladie, mais à des doses très élevées, non autorisés sur pois ou comportant des substances actives interdites en France.

Aujourd'hui, la gestion du risque repose essentiellement sur deux mesures prophylactiques (démarches préventives par des leviers agronomiques) :

- le **test *Aphanomyces*** (Crosson and Moussart 2011) : ce test prédictif permet d'évaluer le potentiel infectieux (PI) d'une parcelle. Il s'agit d'un outil d'aide à la décision pour l'implantation d'une parcelle de pois. Si le PI est inférieur à 1, la culture

de pois est possible ; si le PI est entre 1 et 2,5, la culture de pois peut être envisagée en semis d'hiver qui échappe partiellement à la maladie ; et si le PI est supérieur à 2,5, la culture de pois est déconseillée ;

- la **rotation des cultures** : il est conseillé d'espacer deux cultures de pois de 5-6 ans minimum et d'alterner avec de la féverole ou du lupin, dont les variétés cultivées sont résistantes à la maladie, afin d'éviter l'augmentation du risque. Certains précédents culturaux, tels que les Brassicacées, pourraient avoir un effet répresseur sur la maladie, par la libération dans le sol de composés volatils tels que les glucosinolates (Hossain et al. 2012).

D'autres mesures prophylactiques semblent également avoir un effet pour limiter le développement épidémique de la maladie sur pois :

- le **semis d'hiver** : un semis d'hiver permettrait de diminuer l'impact de la maladie, par échappement partiel résultant en un âge plus avancé des tissus racinaires lors de la période favorable au développement d'*A. euteiches* (résistance ontogénique). Des travaux sont en cours à l'INRA Le Rheu afin d'étudier la nuisibilité de la maladie en fonction du stade de développement de la plante ;
- le **travail du sol** : le développement de la maladie étant favorisé par la présence d'eau libre dans le sol, l'amélioration du drainage et de la structure du sol par le travail du sol et l'incorporation de matière organique permettraient de limiter les symptômes (Allmaras et al. 2003) ;
- la **lutte biologique** : l'utilisation de bactéries ou champignons antagonistes, ainsi que les mycorhizes, auraient un effet néanmoins faible sur le développement d'*A. euteiches* (Wicker 2001) ;
- la **résistance génétique** : aucune variété de pois résistante à *A. euteiches* n'existe à ce jour, bien que le levier génétique paraisse le plus encourageant. En effet, plusieurs sources de résistance partielle ont été mises en évidence aux États-Unis et en France (Hamon 2010; Pilet-Nayel et al. 2007).

La combinaison de plusieurs moyens de lutte prophylactique semble aujourd'hui la stratégie la plus prometteuse pour gérer la pourriture racinaire précoce due à *A. euteiches* chez le pois et limiter durablement les pertes de rendement.

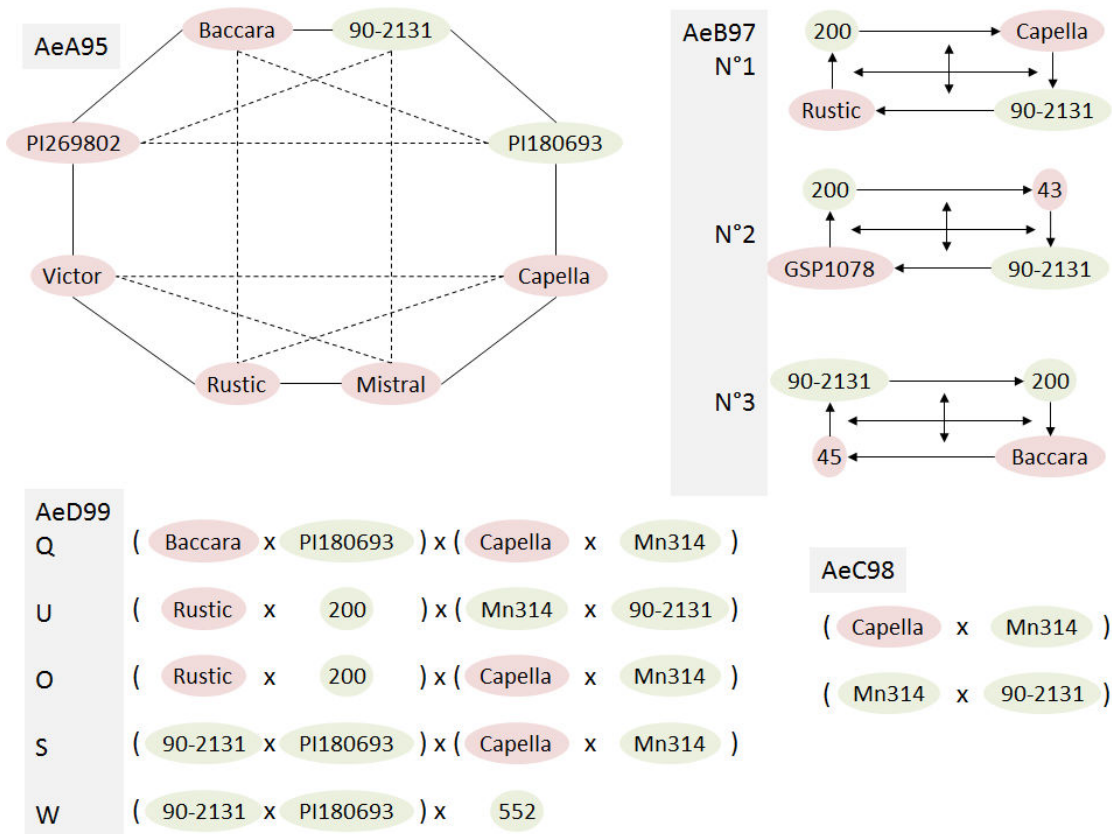


Figure I.19 : Programmes de croisements du schéma de sélection récurrente du GSP pour la résistance à *A.euteiches* (Moussart et al. 2007; Roux-Duparque and Boitel 2004)

Les lignées rouges sont sensibles et les lignées vertes sont partiellement résistantes. Les meilleures F1 des programmes de croisements AeA95, AeB97 et AeC98 ont été sélectionnées pour le programme de croisement AeD99. Les sous-programmes AeD99 Q et U ont été réalisés dans l'objectif de la création de géniteurs pour la sélection, tandis que les sous-programmes AeD99 O, S et W ont été créés pour la validation des QTL de résistance.

3.3. Leviers génétiques chez le pois pour lutter contre *A. euteiches*

La résistance et l'architecture de la plante au niveau racinaire, constituent des leviers génétiques potentiels pour lutter contre la pourriture racinaire due à *A. euteiches* chez le pois. Le premier a été largement documenté et des outils ont été générés pour aider la sélection de variétés résistantes. Le deuxième n'a été que suggéré dans la littérature et son exploitation potentielle reste à explorer.

3.3.1. Sources de résistance et sélection

Les travaux de recherche et de sélection de lignées de pois résistantes à *A. euteiches* ont débuté dans les années 1930 au Nord des États-Unis et se sont intensifiés jusqu'à aujourd'hui. En 1960, l'évaluation de 800 lignées a mis en évidence 18 lignées partiellement résistantes, dont la lignée *PII80693* (Lockwood and Ballard 1960). Entre 1990 et 2000, trois équipes américaines ont publié des résultats de travaux de sélection pour la résistance à *A. euteiches* chez le pois (Davis 1995; Gritton 1990; Kraft 1992). Ils ont ainsi identifié des lignées partiellement résistantes comme, respectivement, les lignées *200* et *552*, *90-2079* et *90-2131* ainsi que *MN313* et *Mn314*. À la fin des années 1990, un large programme de criblage de ressources génétiques a également conduit à l'identification de sources de résistance efficaces vis-à-vis de souches américaines (Malvick and Percich 1999). De la même manière, en France, Pilet-Nayel et al. (2007) ont sélectionné 20 sources de résistance partielle vis-à-vis de souches françaises à partir du criblage de 1 900 accessions, mettant ainsi en évidence la rareté des sources de résistance au sein de l'espèce *Pisum*.

En 1995, le Groupement des sélectionneurs de protéagineux (GSP, France) a développé un programme de sélection récurrente pour la création de géniteurs partiellement résistants de type agronomique (Roux-Duparque and Boitel 2004). Plusieurs programmes de croisements ont été mis en place à partir de sources de résistance identifiées aux États-Unis et confirmées en France (*PII80693*, *200*, *552*, *90-2131*) et de lignées agronomiques sensibles (Figure I.19). La sélection phénotypique à l'aide de tests au champ et en conditions contrôlées a permis l'identification de lignées présentant des niveaux accrus de résistance partielle. Néanmoins, le potentiel de rendement des lignées résistantes n'a pas pu être suffisamment amélioré.

Bien que ces travaux aient abouti à des résultats intéressants, la sélection phénotypique de lignées de pois résistantes à *A. euteiches* est restée difficile du fait du déterminisme fortement polygénique de la résistance, de la variabilité de l'agent pathogène (Wicker et al. 2001) ainsi que des interactions avec d'autres agents pathogènes du complexe racinaire comme *Fusarium*

solani (Papavizas and Ayers 1974). Depuis le début des années 2000, des études de détection de QTL de résistance à *A. euteiches* ont été menées afin d'explorer la diversité des locus contrôlant la résistance et d'identifier des marqueurs génétiques utilisables en sélection assistée par marqueurs.

3.3.2. Diversité des locus de résistance à *A. euteiches* chez le pois

La première étude de détection de QTL de résistance à *A. euteiches*, réalisée sur 45 RIL issues du croisement entre *OSU1026* (sensible) et *Mn313* (partiellement résistant), a permis d'identifier un QTL majeur sur le groupe de liaison (LG) IV, associé à la résistance au champ sur un lieu et une année (Weeden et al. 2000). Dans les années 2000, des travaux collaboratifs entre l'INRA et l'USDA (United States Department of Agriculture) ont permis de développer un large programme de détection de QTL dans quatre populations de RIL générées à partir de croisements incluant les meilleures sources de résistance identifiées précédemment (*90-2131*, *90-2079*, *PI180693* et *552*). Ainsi, Pilet-Nayel et al. (2002) ont identifié 7 régions génomiques associées à la résistance partielle au champ sur 2 lieux et 2 années, à partir d'une population de 127 RIL issues du croisement *Puget* x *90-2079*. Trois d'entre elles ont été identifiées de manière stable sur plusieurs environnements ou pour plusieurs variables, incluant *Aph1*, un QTL majeur expliquant jusqu'à 47 % de la variation phénotypique, ainsi qu'*Aph2* et *Aph3*. Ces trois QTL ont également été validés par des tests en conditions contrôlées avec une souche américaine (*SP7*) et une souche française (*Ae106*) (Pilet-Nayel et al. 2005). *Aph1* a ensuite été détecté comme un gène majeur vis-à-vis d'une souche de référence du pathotype III (*Ae109*) et comme un QTL à effet faible vis-à-vis d'une souche de référence française du pathotype I (*RB84*) (Pilet-Nayel et al., unpublished). La lignée *90-2079*, identifiée comme résistante dans les régions des États-Unis comportant le pathotype III, s'est avérée ainsi sensible en France. Par la suite, Hamon et al. (2011) ont mené une étude sur 111 RIL « *DSP* x *90-2131* », 178 RIL « *Baccara* x *PI180693* » et 178 RIL « *Baccara* x *552* ». Chaque population a été génotypée à l'aide de 126 à 168 marqueurs génétiques, et phénotypée en conditions contrôlées (2 à 6 souches par population) et au champ sur un réseau franco-américain de parcelles infestées (4 à 11 environnements par population). Une méta-analyse des QTL détectés à partir des trois populations et de celle précédemment étudiée (Pilet-Nayel et al. 2002), a permis d'identifier 27 méta-QTL de résistance par regroupement (cluster) des 115 QTL initiaux. Parmi ces 27 méta-QTL, 7 régions majeures ont été détectées de manière consistante dans plusieurs populations de RIL, à partir de plusieurs environnements ou

plusieurs souches. Ces régions ont été nommées *Ae-Psx.y*, *x* étant le nom du LG et *y* le numéro du QTL sur le LG (**Figure I.20**).

Les allèles de résistance de ces principaux QTL ont été introgressés, seuls ou en combinaison, dans des lignées agronomiques par *Back-cross* assistée par marqueurs (Lavaud et al. 2015). Les NIL ainsi créées ont permis de valider l'effet des QTL *Ae-Ps4.5*, *Ae-Ps7.6*, *Ae-Ps2.2* et *Ae-Ps5.1*, individuellement ou en combinaison en conditions contrôlées. Cependant, des interactions entre QTL et fonds génétiques modulant l'effet des QTL ont également été mis en évidence.

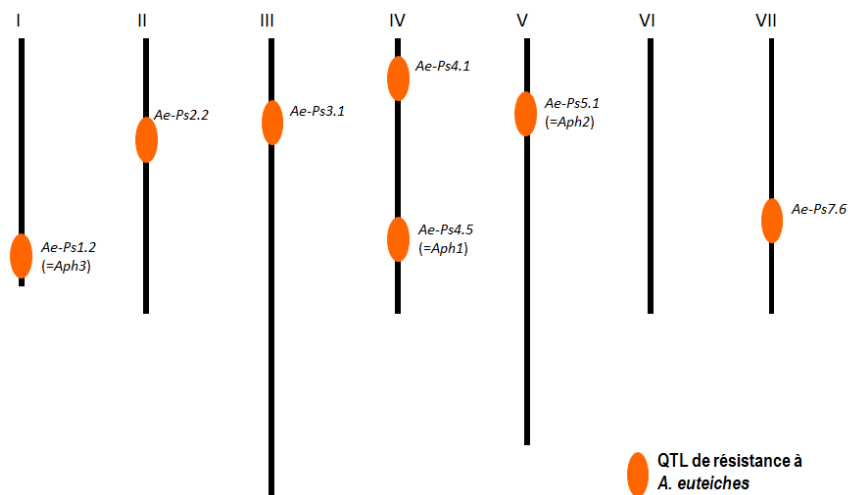


Figure I.20 : Les sept principaux QTL de résistance à *A. euteiches* identifiés chez le pois par analyse de liaison

Les QTL ont été détectés à partir de quatre populations de RIL issues de croisement biparentaux incluant les sources de résistance 90-2079, 90-2131, PII80693 et 552 (Hamon et al. 2011; Hamon et al. 2013; Pilet-Nayel et al. 2002). Les sept principaux QTL représentés sont ceux qui ont été détectés de manière consistante entre populations et environnements.

3.3.3. Mécanismes sous-jacents aux QTL de résistance

Les gènes causaux sous-jacents aux QTL de résistance à *A. euteiches* ne sont pas connus à ce jour. Des gènes candidats ont été identifiés chez le pois par analyses transcriptomiques dans des lignées résistantes et sensibles. Morin et al. (2008) ont sélectionné par « Suppression Subtractive Hybridation » un total de 372 gènes différentiellement exprimés chez les génotypes de pois 552 (partiellement résistant) et *Baccara* (sensible) en conditions inoculées ou non par *A. euteiches*. Plus récemment, Hosseini et al. (2015) ont identifiés respectivement 574 et 817 gènes exprimés différentiellement 6 heures et 20 heures après inoculation par *A. euteiches* chez le pois. Ces gènes sont associés à des réponses immunitaires chez les plantes, comme le renforcement des parois cellulaires, la signalisation hormonale ou le

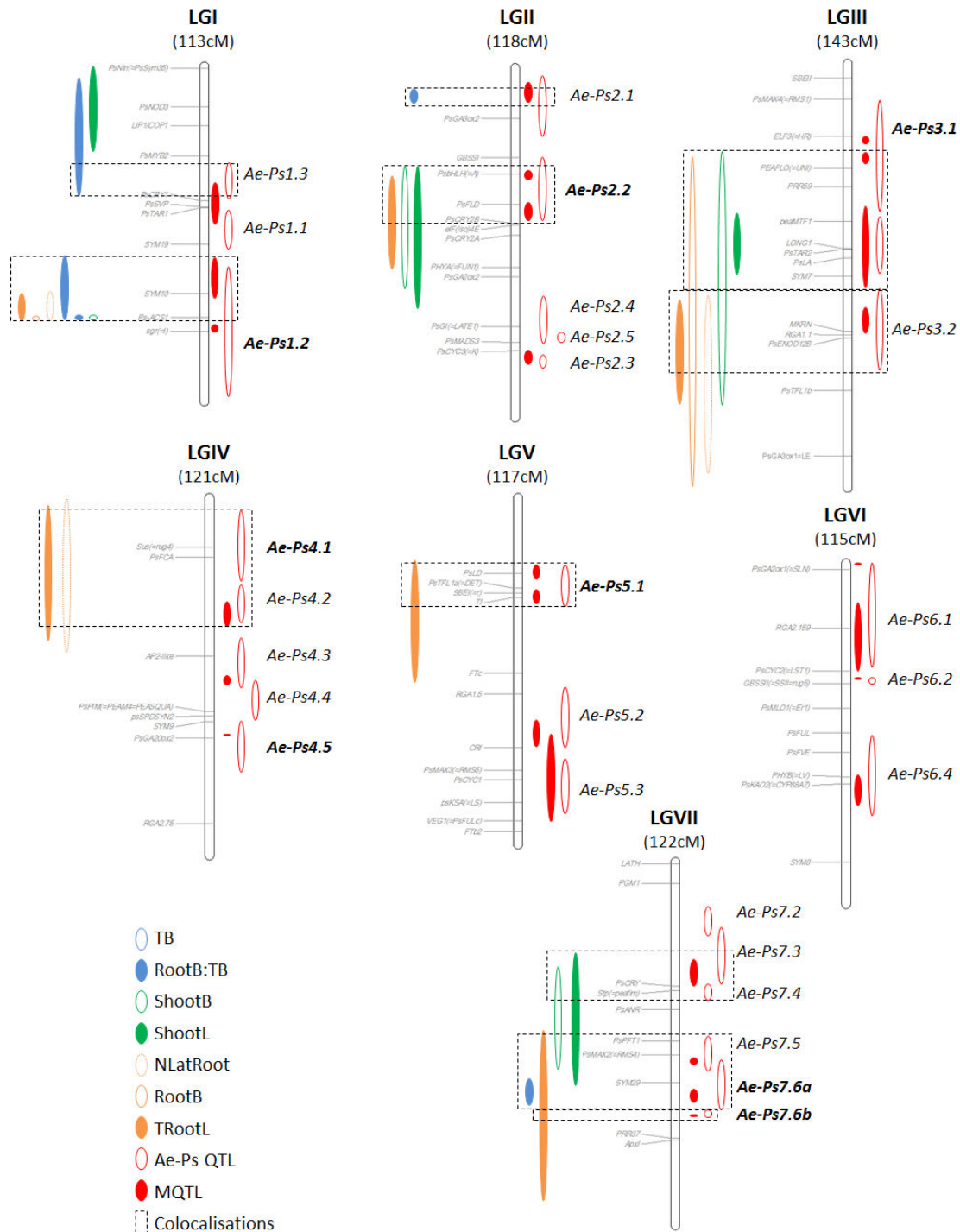


Figure I.21 : Chevauchement des QTL de résistance à *A. euteiches* et des QTL d'architecture de la plante chez le pois

Les QTL de résistance à *A. euteiches* sont extraits de Hamon et al. (2013) : les QTL *Ae-Ps* sont les régions génomiques principales détectées à partir de quatre populations de RIL (*Puget* x *90-2079*, *DSP* x *90-2131*, *Baccara* x *PI180693* et *Baccara* x *552*), tandis que les MQTL sont les méta-QTL issus de la méta-analyse des QTL *Ae-Ps*.

Les QTL d'architecture de la plante de pois sont issus de l'étude de Bourion et al. (2010) sur une population de RIL (*Cameor* x *Ballet*).

Les QTL ont été projetés sur une carte consensus grâce à des marqueurs moléculaires communs. Des chevauchements entre QTL de résistance et QTL d'architecture sont visibles sur 6 des 7 groupes de liaison.

métabolisme de phénylpropanoïdes.

Chez la légumineuse modèle *M. truncatula*, Djébali et al. (2009) et Badis et al. (2015) ont montré par analyse cytologique que la résistance était liée à (i) une accumulation de composés phénoliques solubles et une accélération de la fréquence de divisions dans les cellules du péricycle ; et (ii) à un dépôt de lignine dans les parois des cellules du péricycle, protégeant ainsi le cylindre central de l'invasion par *A. euteiches*. Badis et al. (2015) ont également identifié plus de 700 gènes sur-exprimés dans une NIL porteuse d'un allèle de résistance à *A. euteiches* au locus majeur *prAe1* (Djébali et al. 2009) par rapport à la NIL sensible correspondante. Le rôle de gènes impliqués dans les voies de biosynthèse des flavonoïdes a été particulièrement validé par qRT-PCR (*quantitative Reverse Transcription Polymerase Chain Reaction*) et dosages de métabolites.

3.3.4. Relation entre résistance à *A. euteiches* et architecture racinaire

Kraft and Boge (2001) et McPhee (2005) ont suggéré que le déterminisme génétique de l'architecture racinaire pourrait jouer un rôle dans l'échappement ou la tolérance à la pourriture racinaire précoce et la réduction des symptômes chez le pois, par croissance rapide des racines et production de larges masses racinaires. En particulier, le génotype *PII80693*, partiellement résistant à *A. euteiches*, a été identifié comme présentant le plus grand rapport poids sec racinaire/ poids sec aérien parmi 330 accessions testées (McPhee 2005). De la même manière, chez *Medicago truncatula*, Djébali et al. (2009) et Bonhomme et al. (2014) ont identifié certains génotypes résistants présentant une croissance accrue de racines latérales en réponse à l'infection par *A. euteiches*. D'après Bourion et al. (2010) et Hamon et al. (2013), des QTL majeurs d'architecture racinaire et de résistance à *A. euteiches* colocaliseraient dans dix régions du génome du pois, dont cinq correspondraient au principaux QTL de résistance à *A. euteiches* détectés sur les LG I (*Ae-Ps1.2*, région du gène *af*), II (*Ae-Ps2.2*), IV (*Ae-Ps4.1*), V (*Ae-Ps5.1*) et VII (*Ae-Ps7.6*) (**Figure I.21**).

4. Objectifs et stratégie du travail de thèse

L'amélioration variétale des espèces cultivées est un enjeu majeur pour l'agriculture, notamment vis-à-vis des résistances aux bioagresseurs dans un contexte de diminution des produits phytosanitaires. Pour cela, une bonne compréhension du contrôle génétique de ces résistances est indispensable. Afin d'améliorer la durabilité des résistances totales et/ou d'améliorer l'efficacité des résistances partielles, l'utilisation de leviers architecturaux des plantes semble être une stratégie complémentaire intéressante. Le lien entre l'architecture de la plante et le développement des épidémies aériennes a été relativement bien exploré ces dernières années. Cependant, peu de travaux portent sur les maladies racinaires d'origine tellurique.

Dans un contexte où la volonté est de développer l'usage des légumineuses dans les rotations agricoles, le développement de solutions pour gérer la pourriture racinaire due à *A. euteiches* est un enjeu majeur. En effet, cet agent pathogène est capable d'infester plusieurs espèces de légumineuses cultivées, constituant un frein à leur culture dans les régions touchées. Du fait des connaissances déjà acquises sur le pathosystème pois/*A. euteiches* et de l'enjeu économique et environnemental du re-développement de la culture de pois, ce pathosystème constitue un modèle de choix, à la fois pour l'étude de l'impact de l'architecture racinaire sur la résistance aux maladies racinaires et pour l'étude de la résistance polygénique à *A. euteiches* dans le but de développer des variétés résistantes.

L'objectif de ce travail de thèse a été de préciser les déterminants génétiques de la résistance du pois à *A. euteiches* et de déterminer le rôle de l'architecture racinaire dans le contrôle génétique de cette maladie. Ces objectifs s'appuient sur trois questions de recherche principales : (i) Existe-t-il une diversité génétique de facteurs impliqués dans la résistance à *A. euteiches* ? (ii) L'architecture racinaire intrinsèque de la plante, qualifiée dans ce manuscrit « d'architecture constitutive », est-elle corrélée à la résistance du pois à *A. euteiches* via un contrôle génétique commun ? (iii) L'architecture racinaire générée en réponse à la maladie, qualifiée dans ce manuscrit « d'architecture induite », est-elle corrélée à la résistance et partage-t-elle des déterminants génétiques communs ? peut-elle être impliquée dans la tolérance de la plante à la maladie ?

Pour répondre à ces questions, la démarche de mon travail de thèse s'est organisée en trois étapes principales, correspondant aux chapitres suivants :

- **Préciser le déterminisme génétique de la résistance à *A. euteiches* chez le pois, par génétique d'association génome entier (GWA)** dans un panel d'association comportant 175 lignées de pois. Le panel d'association a été en partie constitué à partir des lignées du programme de sélection récurrente du GSP (voir section 3.3.1) permettant ainsi d'augmenter la fréquence des allèles de résistance. Des données phénotypiques ont été collectées au champ sur plusieurs années et plusieurs lieux (9 environnements au total), ainsi qu'en conditions contrôlées avec deux souches (RB84 et Ae109). Le génotypage de la collection a été réalisé à l'aide d'une puce Illumina Infinium de 15 000 SNP développée récemment (Tayeh et al. 2015a). L'exploitation de ces différentes données a permis (i) d'identifier avec précision des locus associés à la résistance, et (ii) d'identifier des haplotypes de marqueurs et gènes candidats d'intérêt pour la sélection et la compréhension des mécanismes de résistance. L'originalité de ces travaux par rapport aux précédents réside dans l'approche de **génétique d'association** utilisant des **données de génotypage denses** et un **modèle multi-locus** pour augmenter la puissance de détection et de précision (chapitre II).
- Explorer le **déterminisme génétique de l'architecture racinaire constitutive** du pois à un stade précoce et mener une **étude comparative de la variabilité phénotypique et des locus pour l'architecture racinaire et la résistance à *A. euteiches***. Le panel d'association étudié dans le chapitre II a été élargi avec 92 lignées supplémentaires utilisées à l'INRA de Dijon pour étudier la relation entre architecture racinaire et nutrition azotée. Ce panel élargi a été phénotypé pour l'architecture racinaire à 8 et 14 jours en condition contrôlées sans maladie et pour la résistance à *A. euteiches* en conditions contrôlées avec une souche (RB84). Les données phénotypiques ont été comparées par une approche de corrélations entre les variables d'architecture racinaire et de résistance. L'analyse de génétique d'association sur le panel a permis de déterminer et comparer les déterminants génétiques de l'architecture constitutive et de la résistance (chapitre III).
- Evaluer la variabilité d'architecture induite, traduite par la **vigueur racinaire, de la plante en réponse à l'infection** dans le panel élargi et réaliser une analyse comparative des déterminants génétiques de l'architecture induite et de la résistance. Puis, dans un deuxième temps, analyser la relation entre la résistance et/ou la vigueur racinaire et les moindres pertes de rendements au champ dans un sous-set de lignées contrastées. Le panel d'association élargi utilisé dans le

chapitre III a été phénotypé pour l'architecture à 14 jours en conditions contrôlées avec et sans inoculation par *A. euteiches* (souche RB84). Des calculs de ratios ont été calculés à partir de ces données afin de déterminer la variabilité de réponse du système racinaire à l'infection dans le panel. L'analyse de génétique d'association a permis de détecter des locus d'architecture induite et de les comparer avec les locus d'architecture constitutive et de résistance détectés dans les chapitres précédents. Un sous-ensemble de 15 génotypes a été sélectionné parmi les lignées du panel sur la base de leurs niveaux de résistance et d'architecture constitutive et induite contrastés. Ces lignées ont été expérimentées en parcelles saines et infestées au champ sur 4 environnements afin de mettre en relation leurs caractéristiques de résistance et d'architecture avec les pertes de rendement et de déterminer notamment si la vigueur racinaire permettait une meilleure tolérance de la plante à la maladie (chapitre IV).

Les résultats détaillés de chacune de ces étapes sont présentés et discutés sous forme d'articles scientifiques en anglais, dont un accepté dans une revue internationale à comité de lecture. Le dernier chapitre présente une synthèse de l'ensemble des résultats, une discussion sur leurs apports pour la sélection et la recherche, ainsi que des perspectives de recherche.

Chapitre II

Dissection du déterminisme
génétique de la résistance
partielle du pois à
Aphanomyces euteiches par
génétique d'association

Chapitre II : Dissection du déterminisme génétique de la résistance partielle du pois à *A. euteiches* par génétique d'association

1. INTRODUCTION DU CHAPITRE

Dans ce chapitre, nous avons étudié, par génétique d'association, l'organisation structurale des régions génomiques impliquées dans la résistance partielle à *A. euteiches* au sein d'une collection de lignées de pois. Nous avons ensuite mené une étude de la composition en haplotypes de marqueurs favorables des différentes lignées de la collection aux locus les plus consistants détectés. Enfin, nous avons recherché des fonctions putatives de gènes candidats sous-jacents aux locus de résistance, par une approche génomique comparative avec les génomes de *Medicago truncatula*, *Glycine max* and *Arabidopsis thaliana*.

Grâce à un modèle multi-locus, une collection de 175 lignées (collection « *Aphanomyces* ») et un génotypage dense (13.2K SNPs), nous avons pu détecter avec précision des locus associés à la résistance en conditions contrôlées et au champ (9 environnements de 2010 à 2013), confirmant six des sept principaux QTL détectés dans des études précédentes à partir de populations de RILs. Nous avons ainsi pu réduire les intervalles de confiance de ces six principaux QTL. Nous avons également mis en évidence que le cumul d'haplotypes de résistance favorables permettait d'augmenter les niveaux de résistance partielle. Enfin, nous avons identifié, au niveau des régions génomiques détectées, une diversité de gènes candidats putatifs impliqués dans la reconnaissance des plantes aux agents pathogènes, la modification des parois cellulaires, la signalisation hormonale, la régulation de la transcription ou la détoxification.

Ce chapitre est présenté sous la forme d'un article publié dans *BMC Genomics*.

2. ARTICLE ASSOCIÉ

Les références citées sont situées à la fin du chapitre (p.94-101)

Les fichiers additionnels sont présentés en annexe 1

Genome-wide association mapping of partial resistance to *Aphanomyces euteiches* in pea

Aurore Desgroux ^{1,2}, Virginie L'Anthoëne ^{1,10}, Martine Roux-Duparque ^{3,11}, Jean-Philippe Rivière ^{1,4}, Grégoire Aubert ², Nadim Tayeh ², Anne Moussart ^{4,5}, Pierre Mangin ⁶, Pierrick Vetel ^{1,4}, Christophe Piriou ^{1,4}, Rebecca J. McGee ⁷, Clarice J. Coyne ⁸, Judith Burstin ², Alain Baranger ^{1,4}, Maria Manzanares-Dauleux ^{1,9}, Virginie Bourion ², Marie-Laure Pilet-Nayel ^{1,4*}

¹ INRA, UMR IGEPP 1349, Institut de Génétique et Protection des Plantes, Domaine de la Motte au Vicomte, BP 35327, 35653 Le Rheu Cedex, France

² INRA, UMR 1347 Agroécologie, 17 rue de Sully, 21065 Dijon Cedex, France

³ GSP, Domaine Brunehaut, 80200 Estrées-Mons Cedex, France

⁴ PISOM, UMT INRA/Terres Inovia, UMR IGEPP 1349, Domaine de la Motte au Vicomte, BP 35327, 35653 Le Rheu Cedex, France

⁵ Terres Inovia, 11 rue de Monceau, CS 60003, 75378 Paris Cedex, France

⁶ INRA, Domaine Expérimental d'Époisses, UE0115, 21110 Bretenières Cedex, France

⁷ USDA, ARS, Grain Legume Genetics and Physiology Research Unit, Pullman, WA 99164-6434, USA

⁸ USDA, ARS, Western Regional Plant Introduction Station, Washington State University, Pullman, WA 99164-6402, USA

⁹ AgroCampus Ouest, UMR IGEPP 1349 IGEPP, 65 rue de Saint Briec, 35042 Rennes Cedex, France

¹⁰ Current address: Nestlé R&D Center Tours, 101 Avenue Gustave Eiffel, 37097 Tours Cedex 2, France

¹¹ Current address: Chambre d'Agriculture de l'Aisne, 1 rue René Blondelle, 02007 Laon Cedex, France

Corresponding author

* MLPN: Marie-Laure.Pilet@rennes.inra.fr

1. Abstract

Background

Genome-wide association (GWA) mapping has recently emerged as a valuable approach for refining the genetic basis of polygenic resistance to plant diseases, which are increasingly used in integrated strategies for durable crop protection. *Aphanomyces euteiches* is a soil-borne pathogen of pea and other legumes worldwide, which causes yield-damaging root rot. Linkage mapping studies reported quantitative trait loci (QTL) controlling resistance to *A. euteiches* in pea. However the confidence intervals (CIs) of these QTL remained large and were often linked to undesirable alleles, which limited their application in breeding. The aim of this study was to use a GWA approach to validate and refine CIs of the previously reported *Aphanomyces* resistance QTL, as well as identifying new resistance loci.

Methods

A pea-*Aphanomyces* collection of 175 pea lines, enriched in germplasm derived from previously studied resistant sources, was evaluated for resistance to *A. euteiches* in field infested nurseries in nine environments and with two strains in climatic chambers. The pea-*Aphanomyces* collection was genotyped using 13,204 SNPs from the recently developed GenoPea Infinium® BeadChip.

Results

GWA analysis detected a total of 52 QTL of small size-intervals associated with resistance to *A. euteiches*, using the recently developed Multi-Locus Mixed Model. The analysis validated six of the seven previously reported main *Aphanomyces* resistance QTL and detected novel resistance loci. It also provided marker haplotypes at 14 consistent QTL regions associated with increased resistance and highlighted accumulation of favourable haplotypes in the most resistant lines. Previous linkages between resistance alleles and undesired late-flowering alleles for dry pea breeding, were mostly confirmed but the linkage between loci controlling resistance and coloured flowers was broken due to the high resolution of the analysis. A high proportion of the putative candidate genes underlying resistance loci encoded stress-related proteins and others suggested that the QTL are involved in diverse functions.

Conclusion

This study provides valuable markers, marker haplotypes and germplasm lines to increase levels of partial resistance to *A. euteiches* in pea breeding.

Keywords

Root rot, plant disease resistance, GWAS, pea (*Pisum sativum*), Quantitative trait loci, marker haplotype, candidate genes

2. Background

Research interest in polygenic resistance to plant diseases has increased worldwide in the past ten years to meet the challenge of sustainable agriculture. Recommendations to reduce chemical inputs and the frequent breakdown of major resistance genes in plants have encouraged the integration of polygenic resistance into cultivars of many crops. However, breeding strategies for polygenic resistance, which is controlled by many genes, have not been as widely developed as for monogenic or oligogenic resistance (controlled by one or few genes, respectively) [1]. Breeding schemes for polygenic traits are costly and time-consuming. The effects of Quantitative Trait Loci (QTL) controlling resistance are not always conserved in different genetic backgrounds and environments and markers tightly linked to resistance loci have also often been lacking. Further research is needed to validate QTL effects, reduce their confidence intervals (CIs) and identify their underlying causal genes [2, 3], to encourage and optimize QTL use in breeding [4].

The identification of plant resistance QTL has broadly been explored using linkage mapping populations derived from crosses between two parental genotypes [1]. With the decrease in genotyping costs and the massive development of markers in the recent past years, genome-wide association (GWA) studies, are becoming common approaches to detect natural variation underlying complex traits, especially polygenic resistance to major diseases, in a large range of crop species [5-7], including legumes [8, 9]. The advantages of GWA studies compared to bi-parental linkage mapping include access to wider genetic diversity, higher recombination rates due to the evolutionary history of the species, and thus substantially refined genomic regions associated with trait variations [5, 10]. Accurate marker density for GWA studies depends on the rate of linkage disequilibrium (LD) decay and should be higher in species with a rapid LD decay (a few kb, such as in maize) than in those with a slow LD decay (~ 100 kb, such as in rice) [5]. Segura et al. [11] recently proposed a Multi-Locus Mixed Model (MLMM) approach, in order to improve GWA studies precision and power of detection, and it was successfully applied [12]. The reduction of CIs of genomic regions associated with traits of interest, opened the possibility of identifying haplotypes for marker

assisted selection (MAS) [8, 13] and pinpointing interesting candidate genes underlying QTL [14, 15]. However, the GWA approach has also been reported to have poor power to detect rare alleles associated with the trait of interest, leading to missing heritability [5, 16], and complementarity between linkage and GWA approaches has been underlined [7, 17]. Multi-parental designs, including Nested Association Mapping (NAM) [18, 19], Multi-parent Advanced Generation Inter-Cross (MAGIC) [20, 21] and breeding line populations [22, 23] were proven to efficiently increase power of GWA studies to detect rare variants, for which rates are increased by selection of rare-allele-carrier parental lines [7, 16].

Dry pea (*Pisum sativum*) is the third most important pulse crop worldwide [24], for which yield has been unstable for the past decades, mainly due to biotic and abiotic stresses. One of the most damaging biotic stresses of peas is *Aphanomyces* root rot due to *Aphanomyces euteiches* [25]. The soil-borne root pathogen, first described in 1925 [26], has been mainly reported as a yield limiting factor in the United States of America (USA) and Europe for more than twenty years [27-29], and more recently in Canada [30]. Two main pathotypes of *A. euteiches* were described by Wicker and Rouxel [28], including pathotype I predominant in France and pathotype III detected in some regions of the USA (Onfroy et al., personal communication). Both pathotypes cause honey brown necrosis symptoms on pea roots and epicotyls, resulting in dwarfism, foliage yellowing and then death of plants in highly infested fields. Increasing yield loss due to *A. euteiches* in dry and green pea production has been noted in Western Europe due to short crop rotations of susceptible pea varieties and the long lifespan of oospores [27]. The development of resistant cultivars has been considered as a major objective for the past two decades in France, as only prophylactic and cropping methods are available to manage the disease. Pea lines partially resistant to *A. euteiches* were identified from germplasm screening and breeding programs conducted in the USA [31-35], and more recently, from a French germplasm screening program of approximately 1,900 *Pisum* lines [36]. The most resistant lines were integrated into crossing programs to develop breeding lines [37, 38], recombinant inbred lines (RILs) [39-43] and near-isogenic lines (NILs) [44]. Breeding lines with increased levels of resistance to *A. euteiches* were selected in a phenotypic recurrent selection-based breeding program developed by GSP (Groupement des Sélectionneurs de Pois Protéagineux, France) [37, 38]. RILs have also been used for discovery of *Aphanomyces* resistance QTL [39-41, 43]. A total of 27 meta-QTL were identified on a consensus genetic map from four RIL populations [43]. Eleven of them, corresponding to seven genomic regions, were detected on six of the seven pea linkage groups (LGs), with high consistency over locations, years, isolates and populations [43]. Marker

assisted back-crossing was used to introgress each of the seven consistent genomic regions into one of the susceptible RIL parents and two main spring and winter pea varieties. The resulting NILs were used to validate individual or combined major resistance QTL effects [44]. Lavaud et al. [44] considered large QTL intervals for NIL creation, which brought undesirable morphological (coloured flowers, normal leaves) or developmental (late flowering) alleles linked to resistance alleles at several QTL.

Massive numbers of Single-Nucleotide Polymorphism (SNP) markers were recently developed from whole genome cDNA (coding deoxyribonucleic acid) [45-47] or genomic sequencing of pea lines [48, 49]. A GenoPea Infinium® BeadChip was developed by Tayeh et al. [49], containing 13,204 SNPs, all located in gene-context sequences. Pea diversity panels, especially the USDA (United States Department of Agriculture) core collection and the INRA (Institut National de la Recherche Agronomique) reference collection, were used to determine associations between low to medium density genetic markers (137-1233) and traits of interest [9, 50-52]. However, only a few sources of resistance to *A. euteiches* were identified in these collections (Pilet-Nayel et al., unpublished data), as was also found in larger *Pisum* screening programs for *Aphanomyces* resistance [34, 36].

The aim of this study was to validate and refine the CIs of previously reported *Aphanomyces* resistance QTL, as well as identifying new resistance loci, using a GWA approach. A novel panel, enriched in pea lines partially resistant to *A. euteiches* from gene pools previously studied, was designed including breeding and germplasm lines [37, 38, 43]. The recent GenoPea Infinium® BeadChip was used for high-density SNP genotyping of the collection [49]. The GWA study detected SNPs and LD blocks associated with *Aphanomyces* resistance from data collected in nine field environments and two strains under controlled conditions. The genomic positions and CIs of resistance loci detected by GWA study were compared to those previously identified by linkage analysis [41, 43]. The GWA study also identified loci associated with morphological and developmental traits, in order to analyse their linkages with *Aphanomyces* resistance loci. The GWA study allowed marker haplotypes and putative candidate genes to be identified, for future pyramiding of resistance alleles in breeding and to investigate the molecular basis of polygenic resistance.

3. Methods

3.1. Plant material

A collection of 175 *Pisum sativum* lines, referred to as the “pea-Aphanomyces collection, was used for association mapping. It includes cultivars, breeding lines and germplasm lines, supplied by public and private programs in legume genetics research and breeding (Additional file II.1):

(i) One hundred and one breeding lines, partially resistant or susceptible to *A. euteiches*, originating from the Aphanomyces recurrent selection-based breeding program conducted by GSP [37, 38]. The breeding scheme included four crossing programs (namely AeA95, AeB97, AeC98 and AeD99), established from 1995 to 1999. Parental lines of the crosses included partially resistant germplasm lines confirmed in French field conditions (*PII80693*, *90-2131* and *552*) [31-34] and dry spring or winter pea cultivars with good agronomic values but susceptible to the pathogen [37, 38]. Simple or double crosses between multiple parents were designed in the AeA95, AeB97 and AeC98 programs.. The best F₁ hybrids from the three programs were selected based on resistance levels of their inbred progenies (F₅-F₆). The selected F₁ hybrids were then inter-crossed in the AeD99 program which was divided into two parts. The first part (AeD99QU) was generated from two crosses and aimed at developing pea lines with improved resistance levels for breeding purposes. The second part (AeD99OSW) was generated from three other crosses and was conducted as a divergent selection program of resistant and susceptible lines for QTL validation studies. In each of the four crossing programs, selection for resistance was conducted from F₃ or F₄ generations, alternatively in growth chamber and in French field nurseries. Resistance levels and agronomic values of the best F₇ and F₈ lines were then assessed in infested field trials and in healthy field conditions.

(ii) Twenty RILs from INRA and USDA research programs selected for Aphanomyces resistance and for morphological and developmental traits. These RILs included: (ii_a) four lines from the *Baccara* x *PII80693* RIL population [41], (ii_b) eight lines from the *Baccara* x *552* RIL population [41], and (ii_c) eight lines from the *Dark Skin Perfection (DSP)* x *90-2131* RIL population [53, 43].

(iii) Eighteen parental lines of RIL populations used for QTL analysis of Aphanomyces resistance [39-41, 43], and/or of the Aphanomyces recurrent selection-based breeding program conducted by the GSP. These lines included the six pea differential genotypes defined by Wicker et al. [29] to verify disease severity and strain pathotypes in pathogenicity tests.

(iv) Twenty lines representing new sources of resistance, selected from the INRA screening program of *Aphanomyces* resistance sources previously conducted [36]. The new sources of resistance selected included pea breeding and germplasm lines from *Aphanomyces* resistance screening programs in the USA and from collections of genetic resource centers in France, the Netherlands, Russia and the USA (Pilet-Nayel et al., unpublished).

(v) Sixteen spring or winter pea varieties currently or formerly grown in Europe.

The pea-*Aphanomyces* collection thus combined different genetic sources of resistance, especially originating from three major resistant pea lines effective in France (*90-2131*, *552* and *PI180693*) and analysed in previous QTL studies [43]. The collection included plant material described both for different end-uses (food, feed or fodder peas) and for different sowing times (spring and winter peas). Morphological descriptions (seed type, foliage type and flower colour) and available pedigree information of the lines are presented in Additional file II.1.

3.2. Phenotyping

The collection was evaluated for *Aphanomyces* resistance in inoculated controlled conditions (CC) and infested fields. In CC, resistance tests were performed in 2009 on all the lines but the 20 RILs with two pure-culture reference strains of *A. euteiches* described in [41], *i.e.* the French RB84 strain from pathotype I and the American Ae109 strain from pathotype III. CC resistance tests were also conducted in 2013 on all the lines with the RB84 strain only. The experiments used a modified version of the standardized test developed by Moussart et al. [54], as described in [41]. The pea lines were evaluated for resistance to each strain, using a randomised complete block design with four and three blocks in 2009 and 2013 tests, respectively, and five seeds per block sown in a pot. Resistance tests were conducted in a growth chamber (25°C for 16h-day, 23°C for 8h-night), as described in [41]. Seven-day old seedlings were inoculated with a solution of 10^3 zoospores per plant, produced as described by Moussart et al. [54]. Disease severity was assessed on each plant seven days after inoculation, using a 0 (no symptoms) to 5 (dead plant) scoring scale proportional to the percentage of browning symptoms on roots and epicotyls [41]. A CC root rot index (CC_RRI) was calculated as the mean disease score on all plants in a pot.

In the field, the collection was evaluated for *Aphanomyces* resistance in infested nurseries over four years (from 2010 to 2013) and three locations in France (Riec-sur-Belon, Finistère (RI); Dijon-Epoisses, Côtes d'Or (DI) described in [41]) and in the USA (Kendrick, Idaho (KEN)). The collection was phenotyped each year at the two French locations and in 2012 at

KEN, USA, for a total of nine environments. Field assays were carried out using randomised complete block designs with three replicates and an adjacent susceptible control (Solara) every two lines, as described in [41]. Two disease criteria were used to assess resistance in each plot, as reported in [41]: (i) a field root rot index (Field_RRI), using the 0–5 CC scoring scale, evaluated on ten plants per plot for each year in French nurseries and (ii) a field aerial decline index (Field_ADI), measuring yellowing symptoms on a plot, evaluated once or twice (Field_ADI1 and Field_ADI2) in all the USA and French disease nurseries, using a 1 (green plant) to 5 or 8 (dead plant) disease scoring scale depending on the nursery. The RRI and ADI scores on each plot were adjusted for local disease variation measured by scores on the adjacent susceptible control, as described in [41].

In the field, the collection was also phenotyped for developmental traits in a healthy nursery at Rennes-Le Rheu (Ille-et-vilaine, France (REN)) for two years (2010 and 2012), using a randomised complete block design with three replicates, as described in [43]. Earliness was scored on each plot as the number of calendar days to 50 % bloom (FLO1), to 100 % bloom (FLO2) and to ripening defined as 100 % of dried plants (RIPE) from the first day of the year. Average plant height (HT) was evaluated on five plants per plot at 100 % bloom.

3.3. Genotyping

The collection was genotyped using three morphological genes (*Af*, afile/normal leaves; *R*, smooth/wrinkled seeds; *A*, anthocyanin production related to white/coloured flowers), specific-PCR markers for two known-function genes (Clpser and SugTrans, [55]), 45 simple sequence repeat (SSR) markers from [56] and 13,204 SNP markers from [49]. DNA was extracted from each pea line from approximately 1 g of young leaves collected on plants grown in a greenhouse, as described by Doyle and Doyle [57]. Concentrations were adjusted to 10 ng/μl for SSR genotyping and to 50 ng/μl for SNP genotyping.

SSR fragments were amplified by polymerase chain reaction (PCR) and analysed using GeneMapper® software v.4.0 (Applied Biosystems®, USA), as described in [44], except for the 20 RILs of the collection which were genotyped for SSRs as in [41]. Out of the 45 SSRs, 28 and 17 were located within and outside the main genomic regions previously associated to *Aphanomyces* resistance, respectively (Additional File II.2) [41, 43]. Each SSR revealed between two and 16 alleles in the collection. Each line of the collection was coded for each SSR allele as homozygous for the considered allele (“BB”), heterozygous (“AB”) or homozygous for another allele (“AA”). A total of 441 SSR coded alleles were included in the genotyping matrix.

The 13,204 SNPs were part of a large pea SNP resource of 248,617 robust filtered SNPs discovered from whole-genome Illumina HiSeq2000 sequencing of 16 diverse pea genotypes [49]. These 13,204 SNPs were all located in gene-context sequences, each originating from a separate transcript [58], and were included in the Illumina Infinium® BeadChip developed by Tayeh et al. [49]. Infinium genotyping and data analysis were conducted as described in Tayeh et al [49]. A total of 12,067 SNPs were selected for clearly being bi-allelic in the collection. Each line was coded “AA” or “BB” when homozygous for the first or second allele and “AB” when heterozygous.

3.4. Statistical analysis of phenotypic data

Phenotypic data obtained on the collection for resistance to *A. euteiches*, earliness and height were analysed, for each variable in each field environment or CC experiment (individual analysis), and then for all year x location field environments (global analysis), using the R 3.1.1 program [59]. In the individual analysis, phenotypic variables were analysed using a linear model (LM) [R function lm] including G (genotype) and R (replicate) as fixed factors. In the global analyses, E (environment) and GxE (genotype x environment) interaction were added as fixed factors. Normality of residuals and homogeneity of their variances were checked using Skewness, Kurtosis and Shapiro-Wilk statistics ($P \geq 0.05$), as well as Bartlett test ($P \geq 0.05$), respectively [R functions skewness and kurtosis of package agricolae, [60]; plotresid of package RVAideMemoire, [61]; bartlett.test]. Mean-based heritability (h^2) was calculated for each variable from variance estimates in the individual and global LM analysis, using the formulas $h^2 = \sigma_G^2 / [\sigma_G^2 + (\sigma_E^2 / r)]$ and $h^2 = \sigma_G^2 / [\sigma_G^2 + (\sigma_{GE}^2 / E) + (\sigma_E^2 / rE)]$, respectively, where σ_G^2 is the genetic variance, σ_{GE}^2 the GxE interaction variance, σ_e^2 the residual variance, E the number of environments and r the number of replicates per line. Least Square Means (LSMeans) were calculated from each LM analysis (R function lsmeans of package lsmeans [62]). Histograms of LSMeans frequency distributions were drawn using the R function hist.

Pearson correlation analysis was carried out between LSMeans obtained from the individual and global analysis (R function corr.test of package psych, [63]). The significance of the Pearson correlations was evaluated with a false discovery rate correction for multiple testing (corrected p -value < 0.05 ; [64]). The heatmap of the Pearson coefficients (r) was drawn using the R function heatmap.2 (package gplots) [65].

A Multiple factor analysis (MFA) was performed for the different resistance variable categories (CC_RRI; Field_RRI; Field_AD11 and Field_AD12), using LSMeans from

individual and global LM analysis (R function MFA of package FactoMineR, [66]). For MFA, LSMeans missing values were handled with the R package missMDA [67].

3.5. Genetic analysis

The genotyping dataset of the collection, composed of three morphological genes, two specific-PCR markers for known-function genes, 441 SSR alleles and 12,067 SNP markers was filtered using PLINK 1.9 software [68-70]. Six of the 175 pea lines with missing data for more than 10 % of SNP markers were not included in the GWA analysis. Markers with missing data that exceeded 10 % or with a minor allele frequency (MAF) that did not exceed 5 % in the 169 remaining lines, respectively, were also removed for the GWA analysis. A total of 9,980 markers, including three morphological genes, two specific-PCR markers of known-function genes, 189 SSR alleles and 9,786 SNP markers were thus retained in the genetic analysis.

3.5.1. Missing data imputation

The raw dataset of 9,786 SNP markers was imputed using the R function `knncatimputeLarge` (package `scime`, [71]). This function imputed missing values, which corresponded to 0.45 % of the total dataset, regarding the *k* nearest neighbour SNPs without missing values. Imputation parameters were tested with 10 replicates, using a subset of 5,001 SNP of the dataset with no missing values. For each replicate, 0.45 % of missing values were randomly simulated. Parameters tested were one to 10 nearest neighbours and four different methods to determine distances between SNPs [72]. Distance calculation methods were based on corrected Pearson's contingency coefficient, scaled Manhattan distance, simple matching coefficient or Cohen's kappa. Imputed values of simulated missing data were then compared to real values. The lowest error rate mean (9.86 %) over the 10 replicates was found with the Cohen's kappa distance calculation method and six nearest neighbours. These parameters were applied to the 9,786-SNP-markers and 189-SSR-allele dataset to impute missing values.

3.5.2. Linkage disequilibrium analysis

Pairwise LD between markers was explored within LGs from imputed genotypic data using PLINK 1.9 software. Obtained square correlation coefficient (r^2) values were then plotted against genetic distances (cM) according to the consensus map from Tayeh et al. [49], namely TMap in this study, to estimate the LD decay. LD decay regression curves were fitted to the observed LD decay plots (R 3.1.1 program; [59]), following Sved [73] method with $r_{\text{expected}}^2 = 1/(1 + 4N_e \times c)$, N_e effective population size and *c* recombination rate between

two markers. A nonlinear model was fitted on the pairwise LD data, then least-squares estimates were computed (R function `nls`) and N_e was predicted from this fitted nonlinear model (R function `predict`). The LD decay rate of the population was measured as the genetic distance (cM) where the average r^2 dropped to half its maximum value ($r^2 = 0.5$) [74].

3.5.3. Population structure and individual relatedness

A principal Component Analysis (PCA) and a Kinship relatedness matrix were used to estimate the structure of the collection from genotypic data, using the EMMA (efficient mixed-model association) method in the GAPIT (Genome Association and Prediction Integrated Tool) R package ([75], see the GAPIT R script at [76]). PCA and Kinship matrices were calculated based on a subset of 2,937 SNP markers, among the 9,980 imputed filtered markers on the TMap, corresponding to a single randomly-chosen marker per genetic position. Regarding to eigenvalues, the three first principal components (PCs) of the PCA were selected to take into account structure population in the GWA analysis. Clustering of individuals was considered based on their Kinship coordinates and was implemented in the GAPIT package with the UPGMA (Unweighted Pair Group Method with Arithmetic mean) method.

3.5.4. Association mapping

GWA analyses were performed using a modified version of the multi-locus mixed model (MLMM) R package [11]. Briefly, the PCA matrix of population structure and the Kinship matrix obtained from GAPIT were defined as cofactors in the MLMM (see the `mlmm_cof.r` R script at [77]). Significant SNP markers were also used as cofactors in a forward/backward approach.

The initial script of `mlmm_cof` was modified to define the multiple-Bonferroni (mBonf) threshold using the formula: $mBonf = [-\log(\alpha/m)]$, with $\alpha = 0.10$, the overall false positive threshold and $m = 2,937$, the number of markers selected at non-redundant genetic positions on the TMap. Thus, the mBonf threshold was set at 4.47, which corresponded to a p -value of $2.5E-5$. GWA study was performed from LSMeans scores of all the variables described in the phenotypic data analyses section, as well as from coordinates of lines on the two first PCs of MFA resistance variables analysis namely MFA.Dim.1 and MFA.Dim.2. In each GWA analysis, the optimal MLMM step was determined as the largest stepwise mixed model regression in which all cofactors have $-\log(p\text{-value})$ above the mBonf threshold defined. At the optimal MLMM step, reliability of the analysis was evaluated based on plot of mBonf criteria among forward and backward steps, `qqplot` at the optimal step with mBonf criteria,

and plot of partition of variance among steps (Additional File II.3). The MLM outputs used in this study were the partition of variance (percentage of variance explained by PCA, Kinship, markers in cofactors, and unexplained variance), the number and names of markers as cofactors at the optimal step and the *p-value* and allelic effect of each significant marker.

Local LD analysis was used to define the CIs around significant associated markers detected by GWA study using Plink 1.9 software. A LD block was determined as the interval containing all markers in LD ($r^2 > 0.2$) with the significant associated marker, as described in [17]. LD blocks were named with the LG name (Roman numeral) and then an Arabic numeral in genetic position order. A LD block was considered as consistent for each trait (resistance to *A. euteiches*, or earliness or height), when it contained either one marker associated with at least two variables of the trait, or two or more linked markers ($r^2 > 0.2$), each associated with at least one variable of the trait.

3.5.5. Haplotype analysis

At each consistent LD block associated with *Aphanomyces* resistance, marker haplotypes, *i.e.* genotypic patterns at the given LD block, were identified among all the lines of the collection based on non-imputed raw genotyping data. Haplotypes were named with the LD block name and an Arabic numeral. For each trait significantly associated with marker(s) in a given consistent LD block, mean phenotypic scores of pea line groups comprising more than eight lines (5 % of the total number of lines) and carrying different haplotypes were compared, using the Tukey-HSD test ($\alpha = 5 \%$; R package multcomp, [78]). Favourable and unfavourable haplotypes were defined as those meeting the three following criteria: (i) carrying favourable and unfavourable allele(s), respectively, at the disease trait-associated marker(s), (ii) without missing or heterozygous genotypic data at the other markers in the LD block and (iii) showing a significantly lower or higher disease mean score (breaking of group means considered, $P < 0.05$) than the other favourable or unfavourable haplotypes, respectively, for the highest number of disease traits among those associated with the LD block. According to these criteria, more than one favourable or unfavourable haplotype per LD block could be defined. Missing haplotypes were defined in lines showing missing genotypic data or heterozygosity for at least one marker in the considered LD block. Each line of the collection was described for its number of favourable haplotypes at all the consistent LD blocks. A Tukey-HSD test ($\alpha = 5 \%$) was then used to compare the mean numbers of favourable haplotypes in three groups of pea lines without missing haplotypes. The three groups were defined based on their MFA.Dim1 scores and corresponded to lines with high

(25 % lowest scores), intermediate (50 % interquartile scores) or low (25 % highest scores) levels of resistance, respectively.

3.5.6. Comparative mapping

The consensus map from Hamon et al. [43], namely HMap in this study, summarized individual- and meta-QTL previously mapped for *Aphanomyces* resistance and developmental traits [41, 43], while the consensus TMap of Tayeh et al. [49] contained all the markers used in the present study. Thus, for comparative mapping, markers and QTL of the HMap were projected onto the TMap, using Biomercator V4.2 software [79]. The level of connectivity between the two maps was estimated using the “InfoMap” tool of the software. Maps were compiled using the “iterative map projection” tool from Biomercator V4.2 based on loci position data. TMap was used as reference map and HMap was projected based on common loci. Inversions of common loci were automatically resolved. Visualization of loci detected by association and linkage mapping on the resulted consensus map (namely THMap) was computed using MapChart 2.1 software [80].

3.5.7. Identification of putative candidate genes

Using annotation data for SNP anchored sequences from [49], a search was carried out for the putative gene and protein functions contained in each LD block for resistance to *A. euteiches*, earliness and height. The annotation data described predicted protein functions for each transcript sequence. The annotation was obtained following Blastx searches against *P. sativum*, *M. truncatula*, *G. max* and *A. thaliana* protein sequences. If at least two annotations were consistent, others missing, and at least one e-value was lower than 1E-50, the corresponding putative protein function was assigned to the sequence underlying the SNP. The putative annotation of the SNP-anchored sequence was scored with a ‘disagreeing hits’ comment when annotations disagreed, and with a ‘not assigned’ comment when at least three annotations were missing and/or all e-values were higher than 1E-50. When possible, putative protein functions were attributed to known mechanisms of plant development and responses to biotic stress, based on the literature. SNPs which were anchored to previously cloned pea genes and were located close to LD blocks associated with resistance or developmental variables were identified based on information provided in [49].

4. Results

4.1. Analysis of phenotypic data

4.1.1. Resistance to *A. euteiches*

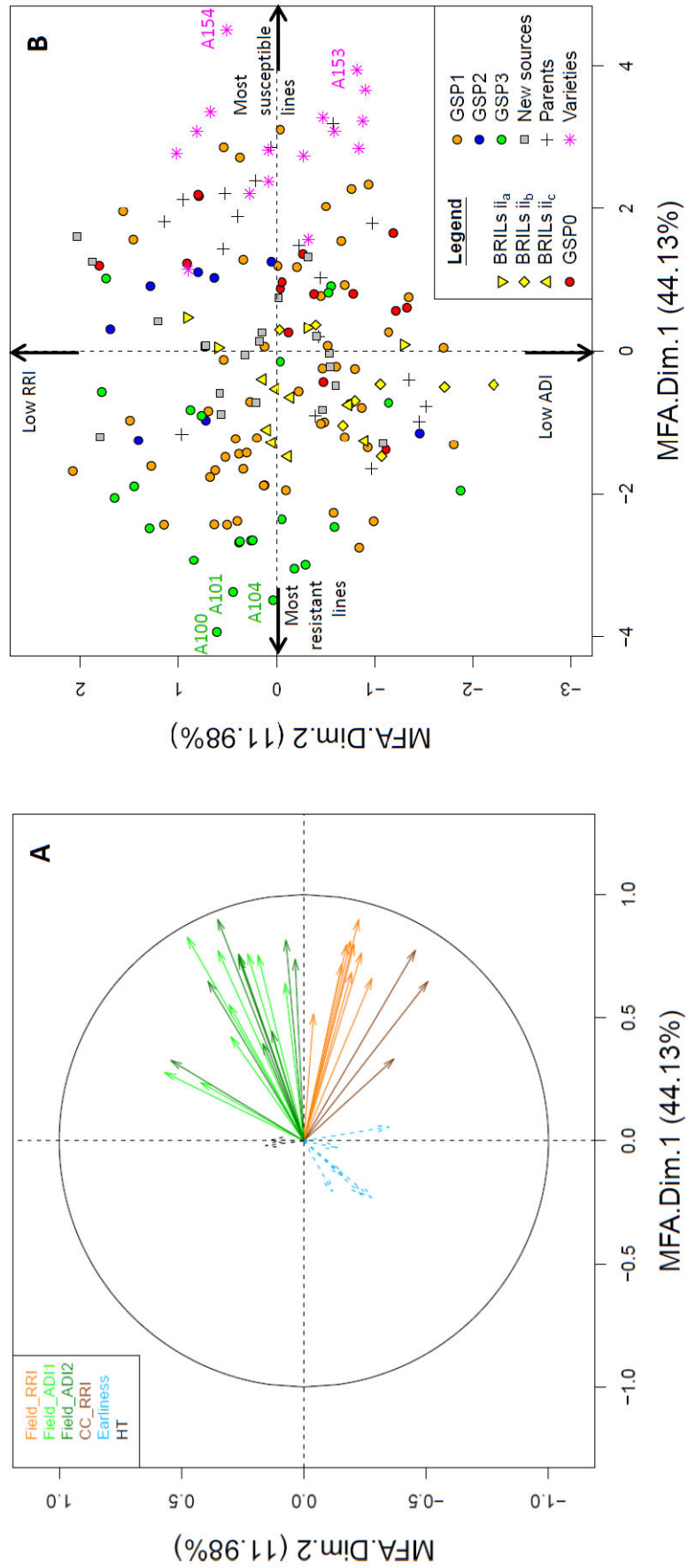
Global statistical analysis of RRI and ADI disease scores, obtained on the pea-*Aphanomyces* collection in field infested nurseries over the nine environments studied, showed highly significant GxE interactions ($P < 0.001$) (Additional File II.4). These results confirmed the relevance of data analysis in each environment. Individual analysis of disease scores in each field environment, as well as in each CC experiment, revealed highly significant genotypic effects ($P < 0.001$) for all the disease variables but for Field_RRI at RI in 2013 ($P < 0.01$) and Field_ADI2 at RI in 2010 ($P < 0.05$) (Additional File II.4). Heritability of resistance ranged from 0.28 (Field_RRI, RI 2013) to 0.96 (Field_ADI2, DI 2012), depending on the variable, and was high for most of the resistance variables ($h^2 > 0.60$, except for Field_RRI at RI in 2013 and Field_ADI at RI in 2010, 2011, 2013 and KEN in 2012). Heritability values were especially high for CC_RRI scorings ($h^2 > 0.78$), especially with the Ae109 strain ($h^2 = 0.91$). Frequency distributions of LSMeans values for each individual resistance variable tended to fit normal curves (Additional File II.5), except for CC_RRI_09_Ae109 which showed a skewed distribution with some highly resistant lines.

4.1.2. Earliness and height

Global statistical analysis of earliness and height scores, obtained on the collection in the field healthy nursery at REN over two years (2010, 2012), showed highly significant GxE interactions ($P < 0.001$) (Additional File II.4). Individual analysis of the scores in each year revealed a highly significant genotypic effect ($P < 0.001$). Heritabilities of earliness and height traits in each environment were very high ($h^2 > 0.87$). Frequency distributions of LSMeans values for each individual developmental variable tended to fit normal curves and were consistent between the two years (Additional File II.5).

4.1.3. Correlations among variables

All Field_RRI and CC_RRI scoring data were significantly and positively correlated between each other (corrected $P < 0.001$, $0.30 < r < 0.79$), as well as closely clustered, except that obtained with the Ae109 strain (Additional File II.6). Most of the Field_ADI data were slightly and positively correlated to CC_RRI data (corrected $P < 0.01$, $0.25 < r < 0.57$) with the RB84 strain but not with the Ae109 strain. Most of the Field_ADI1 and Field_ADI2



Legend on next page

Figure on previous page

Figure II.1: Graphical representation of Multiple Factors Analysis (MFA) of the phenotypic data

Correlation circle of variables (coloured according to groups) (A) and genotype factor map (coloured according to origin) (B), on the first two principal components of the MFA of disease variables. Earliness and height variables are supplemental variables and thus are projected onto the dimensions but were not included in the analysis. All the variables are abbreviated as described in Table 1. GSP0-3: lines from the GSP breeding program containing in their pedigree zero to three of the previously described major sources of resistance (*P1180693*, *90-213* and *552*) [43] (Additional File II.1), respectively. Parents: Parental lines of GSP breeding programs and/or RIL populations. BRILs ii_a, ii_b, ii_c: Selected lines from the *Baccara* x *P1180693*, *Baccara* x *552* and *DSP* x *90-2131* RIL population, respectively; New sources: Selected accessions from the large INRA screening program; Varieties: Registered dry pea varieties. Extreme genotypes are labelled on the two main principal components.

scoring data (68 %) were also significantly and positively correlated between each other (corrected $P < 0.01$, $0.22 < r < 0.79$). Field_ADI scoring data were significantly and positively correlated with Field_RRI scoring data (corrected $P < 0.03$, $0.19 < r < 0.79$), except those from 2011 (RI and DI).

All earliness and height data were significantly and positively correlated between each other (corrected $P < 0.005$, $0.26 < r < 0.92$) (Additional File II.6). Earliness data were slightly and negatively correlated with most of the Field_ADI data (corrected $P < 0.05$, $-0.5 < r < -0.2$).

4.1.4. Multiple Factors Analysis

The two first PCs of the MFA analysis of disease resistance variables explained a total of 56 % of the inertia (MFA.Dim.1: 44.13 % and MFA.Dim.2: 11.89 %; Figure 1A). Three groups of variables could be distinguished, including CC_RRI, Field_RRI and Field_ADI. A total of 77 % of the variables, especially the CC_RRI_RB84 variables, were highly correlated with MFA.Dim.1 ($r^2 > 0.5$) and well represented on the first axis ($\cos^2 > 0.5$) (Additional file II.7).

Pea lines with extreme negative or positive coordinates on MFA.Dim.1 were considered as the most resistant or susceptible lines, respectively. MFA.Dim.2 was mainly constructed from RB84 CC_RRI variables ($r^2 \leq -0.5$) in contrast to Field_ADI variables ($r^2 > 0.5$). Pea lines with extreme negative or positive coordinates on MFA.Dim.2 had low levels of aerial symptoms or root symptoms, respectively. Dispersion of the pea lines on the two principal MFA components showed that the frequency of resistant and susceptible lines was homogeneous among the collection and within some groups of lines (Figure 1B), which is

Table II.1: Genome-wide association analysis results using the multi-locus mixed model (MLMM) method in the pea *Aphanomyces* collection

Variable (a)	Number of markers as cofactors (b)	Range of <i>p</i> -value (c)	% of variance explained by			Unexplained (g)
			PCA (d)	Kinship (e)	Markers as cofactors (f)	
Field_RRI_10.13	2	7.86E-11 - 8.10E-06	48%	20%	19%	14%
Field_RRI_10_RI	2	1.81E-05 - 2.37E-05	13%	28%	25%	34%
Field_RRI_11_RI	2	9.92E-08 - 6.58E-06	42%	9%	20%	29%
Field_RRI_12_RI	5	3.14E-07 - 1.50E-05	20%	18%	39%	23%
Field_RRI_13_RI	0	-	26%	25%	0%	48%
Field_RRI_10_DI	1	5.78E-06	24%	28%	10%	38%
Field_RRI_11_DI	2	4.87E-07 - 5.90E-07	54%	25%	9%	12%
Field_RRI_12_DI	3	4.76E-06 - 7.46E-06	36%	14%	23%	28%
Field_RRI_13_DI	3	6.00E-08 - 2.17E-05	28%	13%	23%	37%
CC_RRI_09_RB84	1	5.40E-12	43%	8%	19%	30%
CC_RRI_09_Ae109	2	8.82E-28 - 1.60E-06	8%	10%	68%	13%
CC_RRI_13_RB84	3	7.34E-11 - 3.89E-06	47%	6%	22%	25%
Field_ADI1_10.13	4	7.86E-11 - 2.43E-06	32%	25%	27%	16%
Field_ADI1_10_RI	4	3.12E-08 - 2.25E-05	10%	4%	39%	48%
Field_ADI1_11_RI	0	-	5%	50%	0%	45%
Field_ADI1_12_RI	3	3.99E-07 - 8.98E-06	23%	50%	15%	13%
Field_ADI1_13_RI	5	3.85E-08 - 3.32E-05	31%	0.001%	37%	32%
Field_ADI1_10_DI	0	-	37%	45%	0%	18%
Field_ADI1_11_DI	0	-	2%	17%	0%	81%
Field_ADI1_12_DI	7	1.13E-10 - 2.53E-05	29%	21%	39%	12%
Field_ADI1_13_DI	2	2.53E-10 - 1.16E-08	35%	15%	26%	23%
Field_ADI1_12_KEN	1	2.64E-05	32%	25%	8%	34%
Field_ADI2_10.13	3	1.15E-07 - 8.27E-06	39%	47%	8%	6%
Field_ADI2_10_RI	0	-	15%	0.004%	0%	85%
Field_ADI2_11_RI	2	7.77E-07 - 3.61E-06	0%	43%	23%	34%
Field_ADI2_12_RI	0	-	39%	47%	0%	14%
Field_ADI2_13_RI	0	-	39%	45%	0%	17%
Field_ADI2_10_DI	6	7.56E-08 - 1.27E-05	20%	0.002%	38%	42%
Field_ADI2_12_DI	0	-	24%	50%	0%	26%
Field_ADI2_13_DI	9	3.25E-12 - 2.34E-5	36%	5%	45%	14%
Field_ADI2_12_KEN	1	2.55E-05	9%	48%	11%	32%
MFA.Dim.1	4	1.51E-08 - 1.26E-05	56%	15%	21%	8%
MFA.Dim.2	2	1.01E-05 - 1.48E-05	9%	78%	1%	13%
FLO1_10_REN	5	1.24E-17 - 1.97E-05	9%	17%	64%	10%
FLO1_12_REN	7	6.29E-16 - 3.31E-06	8%	23%	63%	6%
FLO2_10_REN	8	8.50E-12 - 2.21E-06	5%	0.001%	71%	24%
FLO2_12_REN	3	2.68E-15 - 3.69E-08	5%	19%	52%	24%
RIPE_10_REN	7	5.02E-13 - 7.74E-06	-2%	3%	72%	27%
RIPE_12_REN	2	2.69E-17 - 4.63E-06	-1%	25%	49%	28%
HT_10_REN	3	3.76E-26 - 5.00E-06	-1%	37%	53%	12%
HT_12_REN	2	1.62E-23 - 2.55E-05	-3%	23%	76%	3%

(a) Variables are abbreviated as follows: CC: Controlled conditions experiments, Field: Infested field experiments, RRI: Root rot index, IDA1: First aerial decline index, IDA2: Second aerial decline index, FLO1: date of 50 % bloom, FLO2: date of 100 % bloom, RIPE: date of 100 % dried plants, HT: height of plants, 09 to 13: year of field evaluation, All: Global variables over field environments, RI: Riec-sur-Belon, France, DI: Dijon-Epoisses, France, REN: Rennes-Le Rheu, France, KEN: Kendrick (ID), USA, RB84 and Ae109: two pure-cultured strains; (b) number of markers used as cofactors at the optimal step of the multi-locus mixed model (MLMM) analysis; (c) Range of *p*-values of the significant markers, significance threshold is *p*-value < 3.4E-05 as described in Methods section; (d) Range of allelic effects of the significant markers in absolute values; Percentage of phenotypic variance explained by: (e) the principal component analysis (PCA) matrix of the collection, (f) the Kinship relatedness matrix among lines of the collection, (g) all cofactor markers and (h) the unexplained variance qualified as "missing heritability".

optimal for GWA analysis of the resistance. The three most resistant lines were GSP breeding lines, namely *AeD99OSW -45-8-7 (A100)*, *AeD99OSW -50-2-5 (A104)* and *AeD99OSW-47-6-1 (A101)*. The two most susceptible lines were *Safranos (A154)* and *Marignan (A153)*, registered as French varieties.

4.2. Genetic analyses

4.2.1. Linkage disequilibrium, structure, and kinship in the pea-Aphanomyces collection

From the imputed genotypic data of the collection obtained with the filtered 9,980 markers, the LD decay was estimated to range from 0.08 to 0.16 cM, depending on the LG, and averaged 0.12 cM over all the pea LGs (Additional File II.8). Based on non-redundant marker positions on the consensus genetic map of Tayeh et al. [49] (TMap), the average distance between two markers used in the analysis was 0.27 cM.

From the defined subset of 2,937 SNPs located at non-redundant genetic positions on the TMap, PCA analysis of the collection identified three first PCs that explained a total of 20.91 % of genetic variation in the collection (Additional file II.9 D). The first PC contributed to 9.28 % of the variation and the second and third PCs to 6.34 % and 5.29 %, respectively. Pea lines were clustered slightly on PCs 1 to 3 depending on their origin, mainly varieties, distinct RIL populations, and groups of GSP breeding lines with any, two or three partially resistant parents in their pedigree (Additional File II.9 A to C). The first three PCs were added to the GWA model as cofactors to take into account this slight population structure. From the subset of 2,937 SNPs, the Kinship matrix of genetic similarities revealed a moderate relatedness among lines of the collection ($0.5 < r^2 < 0.65$ for the major part of the lines), including sub-groups with higher relatedness ($r^2 > 0.8$) according to pedigree, end use or sowing type (Additional File II.10). Thus, the Kinship relatedness matrix was also added to the GWA model as a cofactor.

4.2.2. GWA markers and confidence intervals

4.2.2.1. Disease resistance

GWA analysis identified a total of 79 markers, located on the seven LGs, associated with 33 global or individual *Aphanomyces* resistance variables, including (i) 28 Field_RRI or _ADI variables from nine field environments, (ii) the three CC_RRI variables and (iii) the two MFA variables (Table II.1 and Additional File II.11). Zero to nine markers were significantly associated with each variable, with a *p-value* that ranged from 8.82E-28 to 3.32E-05,

depending on the marker. Zero to nine cofactors were thus retained in the MLM for each disease resistance variable, explaining a total of zero to 68 % of the phenotypic variation depending on the variable and in accordance with the heritability of the trait. Missing heritability (unexplained variance) ranged from 6 to 85 % and PCA and Kinship individually captured between zero and 78 % of the variance, depending on the variable. Allelic effects of markers associated with Field_ADI variables fluctuated widely, depending on the marker (absolute values from 0.13 to 0.61) (Table II.1 and Additional File II.11). Markers associated with Field_RRI had similar allelic effects as those associated with CC_RRI with the RB84 strain (absolute values from 0.13 to 0.29), but lower than those associated with CC_RRI with the Ae109 strain (absolute values from 0.27 to 0.81).

A total of 52 CIs, ranging from 0 to 5.4 cM, were defined around all the significant disease-trait-associated markers, which included markers in LD ($r^2 > 0.2$) with the targeted marker(s). Three kinds of LD blocks were identified, based on their significance and consistency (Figure 2): (i) Fourteen LD blocks were considered as consistent since these included two to six disease trait-associated markers (double red stars on Figure 2; $3.25E-12 < P < 2.64E-5$). (ii) Four disease LD blocks included a single highly significant disease trait-associated marker (single red stars on Figure 2; $8.82E-28 < P < 1.13E-10$). Three of these markers were associated with CC_RRI variables. (iii) Most of the disease LD blocks (67 %) included a single moderately significant disease-trait-associated marker ($3E-08 < P < 3.3E-05$).

Out of the 14 consistent LD blocks, seven were Field_ADI specific, three were Field_RRI specific and four were common to Field_ADI and _RRI. The seven markers associated with global Field_ADI variables were all located in common LD blocks as individual Field_ADI variables. Out of the two markers associated to the global Field_RRI variable, one was located in the same LD block as one containing individual Field_RRI variables. Even if CC_RRI and Field_RRI data were correlated ($r^2 > 0.5$), no common LD block was detected. Among the four markers associated with MFA.Dim.1, two were included in Field_RRI specific LD blocks and one in a LD block that contained Field_ADI associated markers. The two markers associated with MFA.Dim.2 variable were included in Field_ADI specific LD blocks.

The THMap, which resulted from the projection of the HMap onto the TMap, was based on 144 common markers (13 to 29 markers per LG; Additional File II.12), most of which were located in previous QTL regions. The projection led to a total of 41 inversions of marker positions, which did not likely affect the QTL projection.

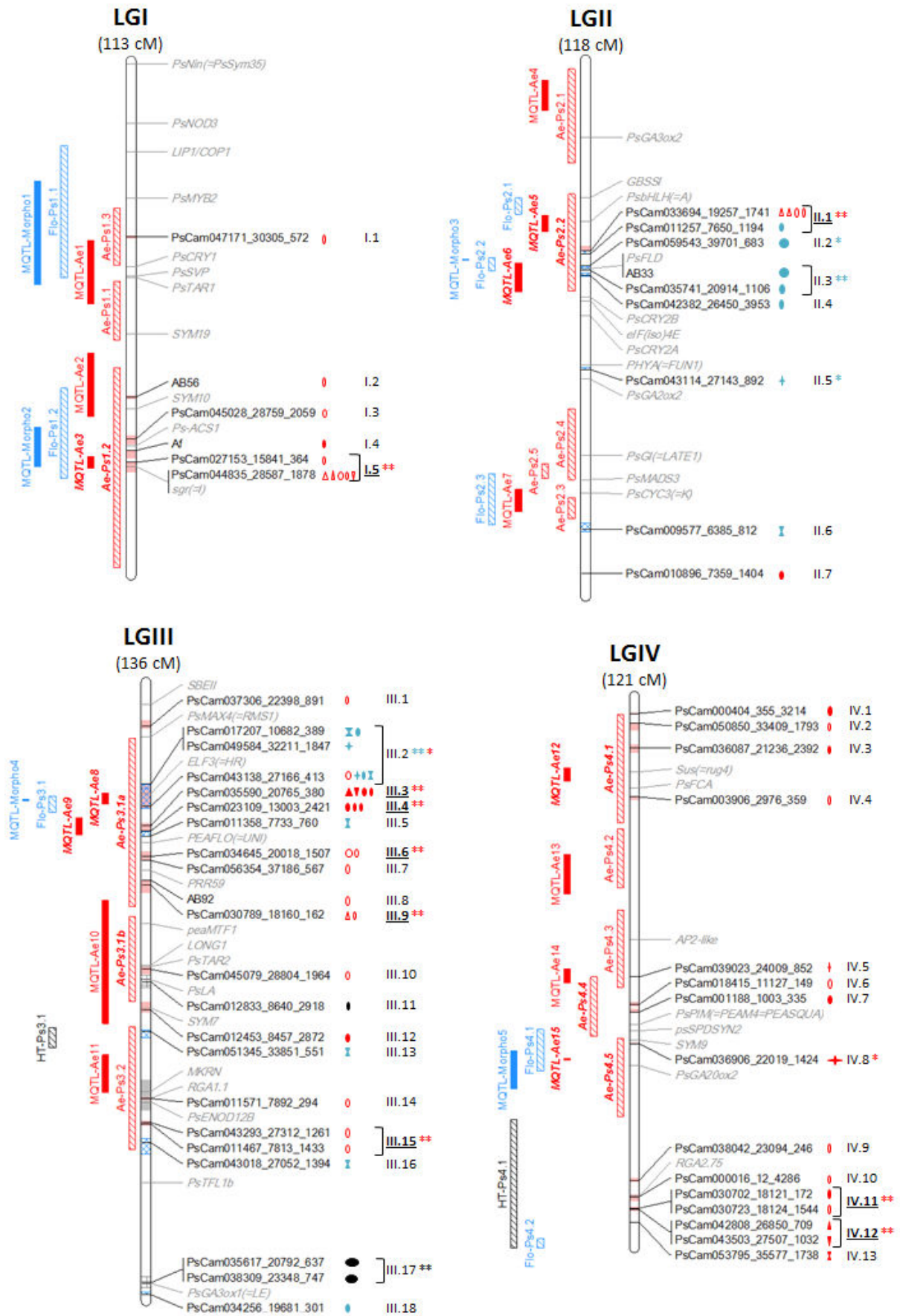
Comparison of linkage and association mapping results revealed that 19 of the 27 genomic regions including individual *Ae-Ps* QTL and 12 of the 27 *MQTL-Ae* meta-QTL previously reported [41, 43] were also detected by association mapping (Figure 2 and Additional File II.11). (i) Eight out of the 14 consistent disease LD blocks (*I.5*, *II.1*, *III.3*, *III.4*, *III.6*, *III.9*, *VII.13* and *VII.14*) co-localized with four of the seven main *Ae-Ps* QTLs previously identified and used for NILs construction by Lavaud et al [44], *i.e.* *Ae-Ps1.2*, *Ae-Ps2.2*, *Ae-Ps3.1a-b* and *Ae-Ps7.6a-b*. Moreover, the highly consistent LD block *VII.16* was located just 0.45 cM away from *Ae-Ps7.6b*. Two other consistent LD blocks co-localized with two less consistent *Ae-Ps* QTL (*Ae-Ps3.2* and *Ae-Ps7.3*). (ii) One of the four disease LD blocks including single highly significant disease-trait-associated markers (*IV.8*) co-localized with the main previously detected QTL *Ae-Ps4.5*. (iii) Among the 40 remaining disease LD blocks identified in this study, three co-localized with the previous main QTL *Ae-Ps4.1* and 11 were located in regions not previously reported for resistance to *A. euteiches*. The main QTL *Ae-Ps5.1* previously identified on LGV was not detected in this study.

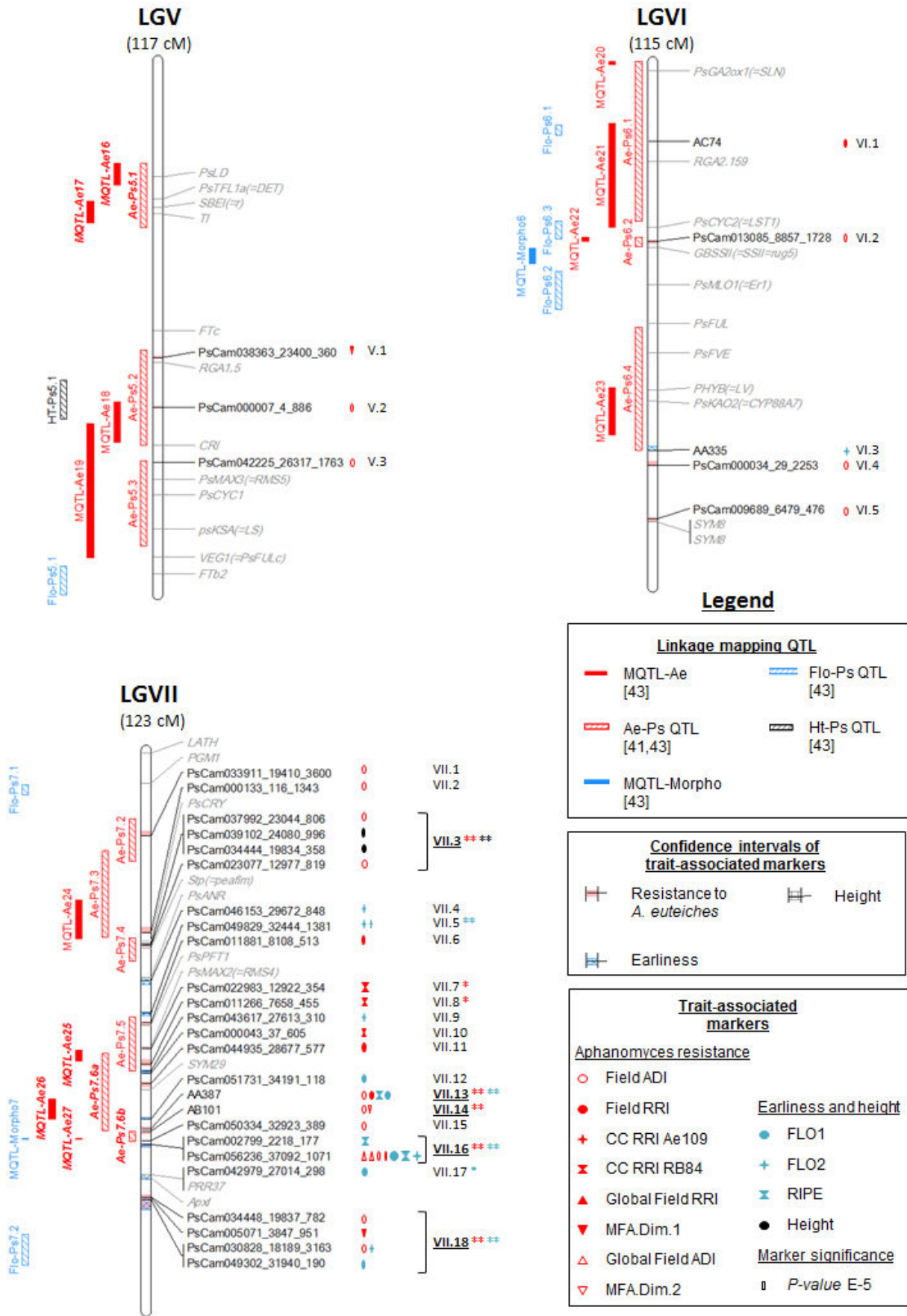
4.2.2.2. Linkage between resistance and morphological or developmental traits

A total of 25 markers, distributed over four LGs, were significantly associated ($1.24E-17 < P < 1.97E-05$) with the six earliness variables scored in the pea-*Aphanomyces* collection in the REN healthy field nursery over two years (Table II.1 and Additional File II.11). Twenty LD blocks were defined around earliness-associated markers, covering 0.3 to 5.4 cM. Six were consistent, since they were associated with two to six variables, and had different allelic effects ranging from one to 8.4 days (double blue stars on Figure 2; $1.24E-17 < P < 1.97E-05$). Three LD blocks were not consistent but included highly significant single trait-associated markers and had high allelic effects from 2.5 to 12.8 days (single blue stars on Figure 2; $6.29E-16 < P < 4.30E-10$). Three of the 14 QTL previously identified in Hamon et al (2013) for earliness, *i.e.* *Flo-Ps2.2*, *Flo-Ps2.3* and *Flo-Ps3.1*, were also detected by association mapping.

A total of five markers, corresponding to three LD blocks distributed over two LGs, were significantly associated with the two height variables scored ($3.76E-26 < P < 2.55E-05$). Two LD blocks (*III.7* and *VII.3*) on LGIII and LGVII, ranging from 0.5 to 2.4 cM, were consistent, since they each included two height-associated markers. In particular, LD block *III.17* was highly significant ($3.76E-26 < P < 1.62E-23$) and presented an allelic effect which

Chapitre II : Dissection du déterminisme génétique de la résistance partielle du pois à *Aphanomyces euteiches* par génétique d'association





Legend on next page

Figure on previous page

Figure II.2: Comparative genetic map of genome-wide association (GWA) and previously detected linkage quantitative trait loci (QTL)

The comparative genetic map was constructed from the projection of the consensus map from Hamon et al [43] onto the consensus map from Tayeh et al [49]. Linkage groups (LG) are named from I to VII and their size is indicated in cM Haldane. Resistance-, earliness- and height- associated markers, LD blocks and QTL are indicated in red, blue and black, respectively. **To the right of each LG:** Names of significant trait-associated markers identified by GWA study are indicated. The shading in the LG bar represents the confidence intervals around the significant trait-associated markers, based on linkage disequilibrium (LD) value of $r^2 > 0.2$. Symbols shown on the right of each marker and described in the legend indicate trait(s) to which the marker was significantly associated. Sign width is proportional to the significance level (p -value) of the marker-trait association. Names of LD blocks are indicated to the extreme right of each LG. The brackets indicate markers that were attributed to a same LD block ($r^2 > 0.2$). LD blocks labelled by double and single asterisks correspond to consistent (at least two significant markers with $P < 5E-10$) and highly significant but not consistent (one significant marker with $P < 5E-10$) blocks, respectively. Genomic positions of cloned pea genes are indicated in grey. The resistance genes cluster identified by Tayeh et al [49] is represented by light grey shading on the bottom of LGIII. **To the left of each LG:** Projected QTL and Meta-QTL described in [41, 43] are represented with the same colour trait codes as for markers detected by GWA study. Solid bars represent Meta-QTLs [43], while hatched bars represent initial QTL before meta-analysis [41, 43]. The main *Aphanomyces* resistance QTL and Meta-QTL [43] names are in bold italic.

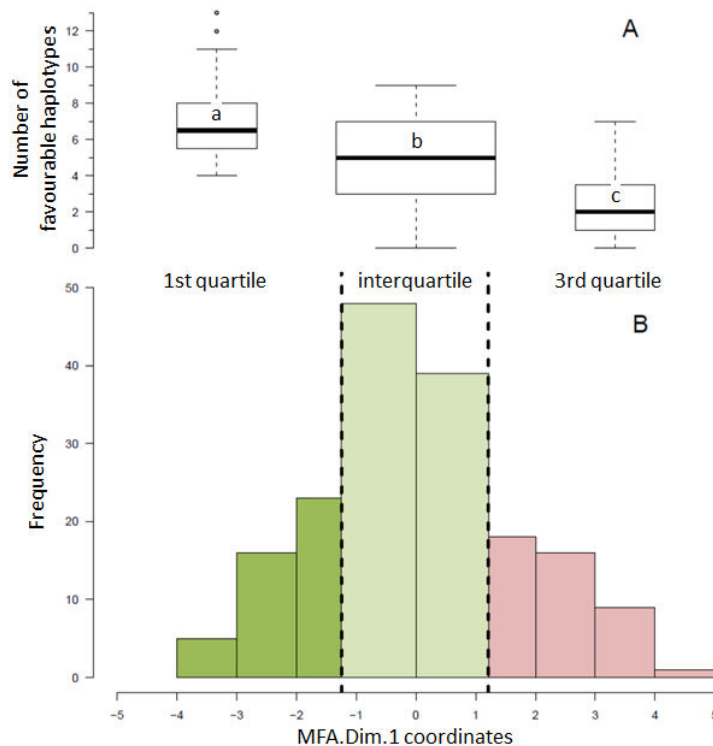


Figure II.3: Mean comparison of favourable haplotype numbers in resistant, intermediate and susceptible lines

(A): Box plot of the number of favourable linkage disequilibrium (LD) block haplotypes carried by pea lines (without missing haplotypes) from the pea-*Aphanomyces* collection belonging to the first quartile (25 % lowest values), interquartile (50 % intermediate values) and third quartile (25 % highest values) of coordinates on the first principal component of the multiple factors analysis (MFA.Dim.1). The three groups include pea lines with a high (first quartile), intermediate (interquartile) and low (third quartile) level of resistance. Letters indicate significantly different groups based on mean comparison (Tukey-HSD, $\alpha = 0.05$).

(B): Distribution histogram of MFA.Dim.1 for the pea lines (without missing haplotypes) of the collection. The first quartile (25 % lowest coordinates; dark green) carries the highest number of favourable haplotypes, the interquartile (intermediate values; light green) carries an intermediate number of favourable haplotypes and the third quartile (25 % highest values; light red) carries the lowest number of favourable haplotypes.

contributed to height differences of more than 30 cm. None of the three LD blocks co-localized with previous QTL detected for height in [43] (*HT-Ps*).

Out of the total 75 LD blocks identified for resistance to *A. euteiches*, earliness and height, only five (*II.1*, *III.2*, *VII.13*, *VII.16* and *VII.18*) were detected for both resistance and earliness, and one (*VII.3*) for both resistance and height (Figure II.2 and Additional File II.11). In the five resistance and earliness common LD blocks, three SNP markers and one SSR marker (*AA387*) were associated with both resistance and earliness. Allelic effects at these markers were opposing for resistance and earliness, *i.e.* the resistance-enhancing alleles conferred later bloom.

One LD block identified for resistance to *A. euteiches* also co-localized with the *Af* morphological gene which controls leaf type on LGI. No LD block co-localized with the *A* morphological gene (*i.e.* *PsbHLH* gene) which control anthocyanin production on LGII. Linkage between *Aphanomyces* resistance and normal leaves at the *Ae-Ps1.2* QTL reported in [41, 43] was thus confirmed, while linkage between resistance and coloured flowers at the *Ae-Ps2.2* QTL was broken.

4.2.3. Marker Haplotypes

At each of the 14 consistent disease LD blocks, three to 26 haplotypes were identified, depending on the LD block (Additional File II.13). Mean comparison of phenotypic LSMeans between marker haplotype groups of a LD block, allowed one or two favourable haplotypes per LD block to be selected, except at LD block *IV.12* for which four favourable haplotypes were identified. A total of 22 favourable haplotypes were identified over the 14 consistent disease resistance LD blocks. A total of 112 haplotypes, carrying at least one favourable allele at the disease-trait-associated marker(s) of the blocks, were also identified among the 14 consistent disease resistance LD blocks. In a subset of 84 lines without missing haplotypes, the 25 % most resistant lines, according to MFA.Dim.1 coordinates, showed a mean frequency of favourable haplotypes which was significantly higher (eight favourable haplotypes on average for the 14 LD blocks) than that of the lines classified in the intermediate or susceptible groups (six and four favourable haplotypes on average, respectively) (Figure II.3 and Additional File II.14). Lines *AeD99OSW-49-5-7* (*A103*), *AeD99OSW-45-8-7* (*A100*) and *AeD99OSW-37-3-4* (*A092*) showed the highest number of favourable haplotypes defined (13, 12 and 11 respectively) at the 14 consistent disease LD blocks in the subset of 84 lines. The group of susceptible lines carried the highest number of unfavourable haplotypes (three unfavourable haplotypes on average, versus two and one for

intermediate and resistant groups, respectively). Two of the favourable haplotypes for resistance to *A. euteiches* (II.1.02 and VII.16.07) were the worst haplotypes for earliness as it gave the latest flowering time. Another one (II.1.06) was unfavourable for earliness as it gave intermediate flowering time. One of the favourable haplotypes for resistance to *A. euteiches* (VII.3.04) was unfavourable for plant height as it gave higher plants. Early bloom and small height are two important selection criteria in dry pea breeding programs.

4.2.4. Putative candidate genes

Out of the 550 markers in the 52 disease-related LD blocks, 211 were located in gene sequences and could be assigned to putative protein functions (Additional File II.15). Out of these 211 annotated SNPs, 135 SNPs corresponded to genes with putative functions reported to be involved in plant response to biotic stresses. Among these genes, 21 encode for protein domains reported as involved in plant recognition of pathogens (including Leucine Rich Repeat-LRR-domains), 65 for domains reported to contribute to signal transduction (including Serine-Threonine-Tyrosine Protein Kinases and proteins involved in Ethylene biosynthesis), 12 for proteins playing a role in transcription regulation (including a WRKY transcription factor, a VQ motif and a basic helix-loop-helix domain) and 47 for proteins involved in plant defence responses such as cell wall modifications (including homeobox leucine zipper proteins), oxidative burst (including Ras GTPases), detoxification (including an ABC transporter C and cytochrome P450 enzymes) or cell death (including chaperone Dnaj proteins). A subset of 24 putative proteins were also reported to be involved in plant growth or development, out of which 14 were not reported to be related to pathogen resistance. Four disease-trait-associated markers in LD blocks III.14, IV.10, V.1 and VI.1, were close (0.1-4.4 cM) but not in LD with pea resistance genes analogues (RGA) coding for nucleotide binding site Leucine Rich Repeat (NBS-LRR) proteins (*RGAI.1*, *RGA2.75*, *RGAI.5* and *RGA2.159*, respectively) [81, 82]. Three RRI-associated markers in LD blocks III.12, IV.8 and VII.11, and one Field_ADI-associated marker in LD block VI.5, were close (0.5-0.8 cM) to cloned genes involved in pea-rhizobia symbiosis (*SYM7*, *SYM9*, *SYM29* and *SYM8*, respectively). A Field_ADI-associated marker in LD block I.5 co-localized with the *SGR* gene (*Stay Green*) which controls chlorophyll catabolism during foliage senescence [83]. Among the 18 consistent or significant disease-related LD blocks, 56 SNP markers were attributed to putative sequences coding for interesting candidate proteins involved in plant resistance to pathogens, including Leucine-rich repeats (LRR) in LD blocks I.5, IV.8 and

VII.3, Serine/threonine protein kinases in six LD blocks [84] and an ABC transporter [85] in LD block *III.6*.

Among the 15 additional LD blocks exclusively related to earliness, 55 of the 110 SNP markers could be assigned to a putative function, 24 of which are involved in plant growth or development. Three earliness-associated markers in LD blocks *II.3*, *III.2* and *III.5*, were close (0-2.1 cM) to cloned genes involved in flowering time or indeterminacy (*PsFLD*, *Hr* and *PEAFLO* respectively) [86-88]. Among the three height-related LD blocks, 26 of the 44 SNP markers could be assigned to a function. Eight of them were reported to be involved in plant growth and development, including two SNPs in LD blocks *III.11* and *III.17* designed in sequences of pea cloned genes controlling plant height, *i.e.* *La* (*PsLA*) [89] and *Le* (*PsGA3ox1*) [90], respectively. The height-associated markers in the third height-related LD block *VII.3* was located just 0.8 cM from the SNP marker designed in the sequence of the *Cry* gene, involved in plant growth [89].

5. Discussion

This work describes the first GWA approach in pea to decipher polygenic control of partial resistance to *A. euteiches*, using novel plant material as well as high density and quality genomic resources. The study evaluated a pea-*Aphanomyces* collection of 175 lines, enriched in sources of resistance to *A. euteiches*, which were identified in the last 15-years of research and breeding programs. Genetic diversity and recombination events screened in the collection were higher than those previously studied in bi-parental populations. We took advantage of the GenoPea Infinium® BeadChip recently generated [49], containing 13,204 SNPs precisely anchored onto a consensus genetic map generated from genotyping data on more than 1,300 RILs. This novel genomic resource greatly enhanced the resolution of the genetic analysis in the association mapping collection compared to the few hundred genetic markers used in previous linkage studies of *Aphanomyces* resistance in pea [41, 43]. We used the same resistance phenotyping methodologies and *Aphanomyces* field network as in the previous QTL analysis of resistance, which led to high accuracy when comparing the results of these two approaches. Finally, this study used a MLM to perform GWA study. This model corrects for population structure and individual relatedness, and takes into account significantly associated SNPs as cofactors to detect other SNPs. This approach limits the number of false positive and false negative associations [11].

Specifically, the GWA study (i) identified 14 consistent loci out of a total of 52 small sized LD-based CIs detected for resistance to *A. euteiches*; (ii) validated and refined six of the seven major resistance QTL previously identified; (iii) confirmed or broke negative associations with morphological and developmental loci; (vi) pinpointed interesting haplotypes and determined their frequency in the most resistance pea lines, for application in breeding; (v) identified relevant and interesting putative candidate genes underlying main resistance loci.

5.1. GWA study validated most of the previously detected QTL and identified new resistance loci

This study **validated most of the previously detected *Aphanomyces* resistance QTL in pea**. Out of the 52 resistance-associated loci, 41 were estimated to be located in 19 of the 27 projected *Ae-Ps* QTL, including six of the seven main consistent QTL, and 12 of the 26 projected meta-QTL (*MQTL-Ae*) previously identified from bi-parental population studies [41, 43]. Four and two of the main previously consistent resistance QTL (*Ae-Ps1.2*, *Ae-Ps2.2*, *Ae-Ps3.1*, *Ae-Ps7.6a-b* and *Ae-Ps4.1*, *Ae-Ps4.4-4.5*) were re-detected in this study by eight of the 14 consistent resistance loci and seven of the other single-variable specific loci, respectively. Twelve previously detected minor-effect QTL were also confirmed with 18 resistance-associated loci, including two that were consistently detected. GWA validation of previous linkage resistance QTL was expected, as the pea-*Aphanomyces* collection was enriched in genetic pools derived from sources of resistance studied in previous QTL analysis (*90-2131*, *552* and *PII80693*). A total of 87 % of the 121 GSP breeding lines and RILs of the collection were derived from at least one of these three sources of resistance. In crops, GWA studies usually used panels with a good representation of species genetic diversity, such as diversity collections developed in pea [52, 91]. For pea *Aphanomyces* resistance, the GWA study required an adapted panel with a higher resistance allele frequency than to the one previously observed in the pea natural genetic variation (Pilet-Nayel et al., unpublished) [36]. However, Kwon et al. [50] identified a SSR marker (PSARGDECA_F) associated to *Aphanomyces* resistance [34] in the USDA pea core-collection, which is estimated to be located in the *Ae-Ps7.6* QTL region, close to LD block *VII.11*, using comparative genetic maps [43, 45, 49]. Marker and methodology tools were also used to optimize comparative analysis of GWA and previous linkage results. Most of the SSR markers, previously associated with resistance QTL [43], were integrated into the consensus SNP genetic map of Tayeh et al. [49]. Common markers were used as bridges for comparative mapping through

linkage QTL projection. Furthermore, the field and climate chamber conditions were the same in [41, 43] and in this study, which increased the accuracy of comparative mapping.

Eight of the 27 projected *Ae-Ps* QTL **previously associated with the resistance** [43] were **not identified by GWA in this study**, including one main consistent QTL (*Ae-Ps5.1*). Several factors could lead to non-validation of QTL by GWA. These include the low detection power of minor-effect QTL, low allele frequency, GxE interactions, epistatic interactions and gaps in marker coverage [5]. In this study, the detection power of low effect QTL may have not been optimal due to the moderate population size (175 lines). Markers with low MAF (MAF < 5 %) were sometimes clustered in some genomic regions, such as on LGV, and were removed from the analysis, creating gaps in the genome-wide scan of marker-trait associations. GxE interactions for field *Aphanomyces* resistance were observed in this study over the nine environments tested, as previously reported and discussed regarding pedo-climatic conditions, pathogen diversity and the presence of other root rot pathogens [60]. They may have resulted in the detection of QTL specific to the field environments presently studied compared to previous ones, which were submitted to different climatic conditions and pathogen population pressures. Epistatic interactions have also been reported for *Aphanomyces* resistance [41] and some QTL effects were recently shown to increase or be lost when transferred into a new genetic background [44]. Finally, despite the unprecedented marker density used in GWA analysis in pea, marker coverage gaps cannot be excluded, due to the restricted choice of SNPs in the Infinium SNP chip used and the unavailable information on physical distances between markers. The non-detection of the main consistent QTL *Ae-Ps5.1* was particularly unexpected since it was identified from CC experiments with several strains including RB84 and Ae109 used in this study, and contributed up to 38 % of the phenotypic variation in the DSP x 90-2131 RIL population [43]. However, Lavaud et al [44] recently showed that the effect of *Ae-Ps5.1* was lost when transferred into pea lines other than *DSP*, suggesting epistatic interactions between this QTL and genetic background. In addition, a very low allele frequency was observed for the SNP designed in the *Ae-Ps5.1* closely linked marker *PsLD* (MAF = 0.8 %) [86, 41] and for its neighbour SNP in high LD ($r^2 = 0.55$; MAF = 1.4 %), which were both consequently removed from the analysis.

In this study 11 **new *Aphanomyces* resistance loci** were identified, including three consistent ones (*IV.11*, *IV.12* and *VII.18*). Comparative mapping was based on few markers (mainly SSRs) which resulted in some inversions. Thus, some projection inaccuracies may have wrongly pointed out these loci as new ones, especially for the *VII.16* interval. However, the higher resolution of GWA analysis may also have highlighted resistance alleles from the

studied germplasm, which would not have been previously detected due to low marker coverage. Most probably, new favourable alleles are present in the pea-*Aphanomyces* collection, which included new sources of resistance [36] which were not investigated in previous linkage-QTL analyses.

In other pathosystems, GWA mapping often validated linkage mapping QTL and detected new loci. For example, Samayoa et al. [15] and Zhao et al [14] reported similar results when comparing linkage and association mapping of Mediterranean corn borer resistance in maize and *Sclerotinia sclerotinium* resistance in soybean, respectively..

5.2. GWA study detected resistance loci with a higher resolution than in previous linkage mapping studies

In the present GWA study, **QTL resolution increased** compared to previous linkage mapping studies of *Aphanomyces* resistance. The 52 resistance-associated intervals were detected with much smaller sizes (0-5.2 cM, depending on the locus, 0.9 cM in average) than those previously identified in QTL mapping studies (2.1-43.9 cM, depending on the *Ae-Ps* QTL, 17.4 cM in average). The LD decay rate in the pea-*Aphanomyces* collection was low ($r^2 < 0.5$ within a map distance of 0.12 cM) compared to the one observed in a similar study in pea [9] (average $r^2 < 0.17$ within a map distance of 5–10 cM). This low LD decay rate is particularly interesting as LD decay of autogamous (self-pollinated) species have generally been reported slow (~100kb), as in rice, foxtail millet or soybean, compared to outcrossing species (~2kb) [5]. This may be especially due to a high number of recombination events in the breeding lines coming from the GSP recurrent selection program (representing 58 % of the lines in the collection), which were mainly derived from double crosses between multiple parents (Additional File II.1). The average distance between two adjacent marker positions of the TMap considered in this study (0.27 cM) was higher than the LD decay rate. Since marker genetic positions may not have covered all the putative recombination points in the collection, there is still the potential to increase genotyping density of the collection, to improve QTL detection resolution. Previous meta-QTL analysis of *Aphanomyces* resistance attempted to reduce QTL CIs but this reduction was potentially over-estimated because of the only partial independence on populations and environments compared [43]. The MLM model recently proposed by Segura et al. [11], used cofactors that captured background genetic variance during genome scan to improve the precision of cofactor detection as found in Composite Interval Mapping.

Comparative analyses of the QTL detected by linkage and association studies are often not carried out on comparable genetic distance scales in both studies. Consequently, several association mapping loci could correspond to a single QTL interval identified by linkage mapping. Ingvarsson & Street [10] reported frequent splitting of linkage mapping intervals into much smaller association mapping intervals. Split association mapping intervals involved in a same larger linkage mapping interval could correspond to several linked genes controlling the trait or to a single imprecisely located region due to GxE effects on the phenotype. Our findings suggest linkage between multiple genes hypothesis when closely located LD blocks were detected from different variables, as seen for example for the *Ae-Ps7.6* region on LGVII (*VII.7-8* and *VII.13-16* blocks). In other cases the results were consistent with single genes hypothesis when neighbouring blocks were detected for the same type of variable, as in the *Ae-Ps3.1* region on LGIII (*III.3-4* for Field_RRI variables; *III.6-7* or *III.8-9* for Field_ADI variables).

5.3. GWA study provides new tools for pyramiding *Aphanomyces* resistance alleles in breeding programs

This study identified **marker haplotypes** at consistent *Aphanomyces* resistance loci, the **pyramiding of which was associated with increased levels of partial resistance** in pea lines of the collection. At the 14 consistent LD blocks identified, 22 favourable haplotypes carrying the favourable alleles at disease trait-associated markers were significantly associated with enhanced resistance levels. The highest resistant lines of the pea-*Aphanomyces* collection carried a significantly higher number of favourable haplotypes. The five most resistant lines with no missing haplotypes were breeding lines from the AeD99OSW GSP program, all derived from direct crosses between the three most partially resistant germplasm studied in previous QTL analysis [(90-2131 x *PII80693*) x 552]. These five breeding lines had a higher level of partial resistance than their individual parents. They combined between eight and 13 of the favourable haplotypes selected at the 14 consistent resistance loci, whereas each of their single parental lines combined between three to nine favourable haplotypes. From this study, different combinations of haplotypes seemed to be related to enhanced levels of resistance. The best combinations were quite difficult to identify since they could vary depending on the lines. However, the favourable haplotypes defined at several blocks (*I.5*, *II.1*, *III.3*, *III.4*, *III.6* and *III.9*) appeared to be frequently represented in the most resistant lines.

Our findings suggested that the pyramiding of *Aphanomyces* resistance alleles with moderate to low effects is a powerful strategy to develop pea lines with increased levels of partial resistance to *A. euteiches*. In the past 30 years, such a pyramiding strategy, based on phenotypic recurrent selection schemes, was successfully developed for root rot diseases in public USA pea breeding programs [31, 32]. In particular, Lewis and Gritton [92] developed a phenotypic recurrent selection protocol consisting of one cycle per year, including selection of F₂ lines on aerial symptoms of the plants in an infested field nursery, followed by intercrosses and selfing of the best lines. The 552 germplasm line was selected from the eighth cycle of recurrent selection (Roux-Duparque, Pers. Comm.), suggesting that this line accumulated resistance alleles. Accumulation of resistance alleles in the 552 germplasm line was demonstrated in this study (nine favourable haplotypes identified). More recently, Lavaud et al. [44] reported that the combination of two or three of the main consistent *Aphanomyces* resistance QTL identified in [43] could increase the partial resistance level, depending on genetic backgrounds. In several other pathosystems, combinations of resistance alleles were shown to successfully enhance resistance levels, including stem rust resistance in wheat [93, 94], spot leaf diseases resistance in spring wheat [95] and spot blotch resistance in barley [96]. We also identified **negative linkages between resistance and morphological or developmental alleles**, to be considered with caution for application in dry pea breeding. Undesirable linkages were previously reported between *Aphanomyces* resistance and coloured flowers, long internodes, normal leaves and late-flowering [43], and have been avoided in breeding programs [32]. In the present study, associations were highlighted between resistance and late flowering alleles at markers in the consistent disease LD blocks *VII.13*, *VII.16* and *VII.18*. In non-consistent disease LD blocks, undesirable links were also highlighted between resistance and normal leaves (LD block *I.4*) or late flowering (LD block *II.2*). In contrast, high resolution of association mapping in this study also enabled linkage between resistance and coloured flowers (*A* gene) alleles to be broken in the genomic region corresponding to the previous main resistance QTL *Ae-Ps2.1* (LD block *II.1*). Thus, different pea lines of the collection carrying favourable haplotypes in the LD block *II.1* have white flowers. However, in the most resistant lines of the collection, such as *AeD99OSW- 45-8-7*, some negative associations could not be broken, as those lines presented normal leaves and some of them still had coloured flowers. Due to the high resolution of GWA study, we could also hypothesize that loci controlling resistance and delayed leaf senescence (*SGR* gene) [83] in the LD block *I.5* are distinct since they were not in LD ($r^2 > 0.2$). Morphological and developmental-trait-loci which are not associated with *Aphanomyces* resistance could also be

detected, suggesting that resistance could be improved without including undesirable traits. For height and earliness, the height locus *III.17* on the *Le* gene and the earliness loci *II.2*, *II.3*, *II.5*, *VII.5* and *VII.17* were identified and not associated with resistance loci. Genetic associations between plant disease resistance and undesired developmental traits such as late-flowering have been commonly reported in the literature [2]. Association genetic studies could have confirmed such negative linkages [97] but sometimes allowed them to be broken, as observed in our study

5.4. GWA study identified relevant putative candidate genes underlying Aphanomyces resistance QTL

The high resolution of the present association mapping study and the recent availability of pea transcriptome sequences connected to the 13,204 SNPs used in this study, constitute unprecedented advantages for identifying putative candidate genes underlying *Aphanomyces* resistance QTL. Indeed, the detection of plant height QTL in previously cloned genes involved in plant growth, and the proximity of flowering genes to earliness LD blocks suggests that the resolution of the present GWA study was high enough to pinpoint relevant candidate genes.

GWA study identified a **high frequency** (64 %) of **putative candidate genes corresponding to a diversity of stress-related protein functions**, underlying resistance-associated intervals. To date, the low number of resistance QTL cloned suggested that the molecular functions underlying them could be diverse [2, 3, 98]. Hypothetical gene-encoded proteins underlying resistance QTL included genes involved in plant-pathogen recognition, signal transduction, activation or repression of transcription, defence responses and developmental or morphological modifications with pleiotropic effects. Such a diverse range of genes underlying resistance intervals was also identified in this study. Putative candidate genes underlying major effect QTL *Ae-Ps4.4-4.5*, recently validated for pea resistance to the American *A. euteiches* strain Ae109 [44] (LD block *IV.8*), included basic helix-loop-helix (bHLH) and ethylene-responsive transcription factors and a NBS-LRR. A second main effect QTL *Ae-Ps7.6*, validated for pea resistance to the French *A. euteiches* strain RB84 (LD blocks *VII.8* and *VII.10*), included two putative infection-induced proteins. Some different types of putative stress-related genes clustered in the disease LD blocks identified in this study and thus may be strong candidates for underlying one of the known mechanisms. For example, LD blocks *III.9* and *I.5* contained six and four putative candidate genes, respectively, involved

in mechanisms of pathogen recognition, transcription regulation, signal transduction and/or defence response.

Other LD blocks seemed to be interesting as they are very close to NBS-LRR cloned genes (*III.14*, *IV.10*, *V.I* and *VI.1*), however they were detected in single environments and had moderate effects. Genes encoding NBS-LRR protein domains are known to be clustered in crop genomes, such as described in *Medicago truncatula* [99], since they are known to be subjected to rapid evolution through local duplications. In pea, Tayeh et al. [49] recently reported a large cluster of 14 NBS-resistance-like genes at the bottom of LGIII. This genomic region, including SNP markers assigned to these 14 NBS-like genes, co-localized with disease LD block *III.14* and is close to consistent LD block *III.15* identified in this study.

Some disease LD blocks contained putative genes similar to those underlying disease resistance QTL cloned in plants, such as ABC transporters (*III.6*) and NBS-LRR proteins (*I.5*, *IV.8* and *VII.3*). An ABC transporter protein have been shown to underlie the *Lr34* major QTL controlling durable resistance to leaf rust stripe rust and powdery mildew in wheat [101]. Other LD blocks could also be detected close to cloned symbiosis genes (*III.2*, *IV.8*, *VI.5* and *VII.11*), suggesting possible common mechanisms between plant-pathogens and plant-beneficial organisms interactions [100]. Some LD blocks also included plant or root development genes (*IV.3*, *IV.7*, *VII.7* and *VII.10*), suggesting their involvement in disease resistance, tolerance or avoidance as reported for several pathosystems [101].

6. Conclusion

In the present GWA study, most of the previously identified *Aphanomyces* resistance QTL were validated with a finer resolution than before, by taking advantage of the new plant and genomic resources developed in pea. New points of discussion were raised regarding comparison of association and linkage analysis to dissect polygenic disease resistances in plants. This study provides new tools for breeding, including pea germplasm and SNP markers associated with resistance, as well as useful information about marker haplotypes at main resistance loci and undesirable allele linkages with resistance. It demonstrates that pyramiding resistance alleles is a key strategy for increasing levels of partial resistance to *Aphanomyces* root rot in pea. The choice of resistance loci to be pyramided remains to be explored, especially using NILs [44]. This study also identified relevant candidate genes to be confirmed and validated in future studies. These will benefit from further genomic resources

in progress, developed in part from a complete pea reference genome sequence becoming available, for increasing resolution of the GWA study.

Competing interests

The authors declare that they have no competing interests

Authors' contributions

AD participated in phenotypic data acquisition, performed all the phenotypic and GWA analyses and drafted the manuscript. VLA contributed to genotyping data acquisition (SSR), association analysis and to the draft manuscript. MRD managed the GSP breeding program and constructed breeding lines included in the pea-Aphanomyces collection. GA coordinated the SNP genotyping, performed its analysis and contributed to genetic analysis. NT provided the genetic map, gene annotation and contributed to genetic analysis. AM contributed to phenotypic data acquisition and to the design of the pea-Aphanomyces collection. JPR and PV coordinated and participated in field experiments at Riec-sur-Belon and Rennes-Le Rheu. PM coordinated and participated in field experiments at Dijon-Epoisses. CJC and RJMG designed field Aphanomyces assays in the USA and participated in field experiments at Kendrick. CP prepared the DNA for SNP genotyping. JB designed and coordinated the Infinium® BeadChip genotyping of the pea-Aphanomyces collection. AB participated to the design of the GSP breeding scheme and of the pea-Aphanomyces collection. MMD and VB participated in the conception of the study and in drafting the manuscript. MLPN coordinated the overall study and the manuscript draft. All authors read and approved the final manuscript.

Acknowledgements

This work was supported by the pre-doctoral fellowship of Aurore Desgroux, from MAP (Ministère de l'Agriculture, de l'Agroalimentaire et de la Forêt, Paris, France). It was funded by Terres Univia (Paris, France) for phenotyping and SSR genotyping data and for the engineer fellowship of Virginie L'Anthoëne, and by INRA (Métaprogramme SELGEN, France) (CropDL project) and the ANR project (ANR-09-GENM-026) for SNP genotyping data. We greatly acknowledge GSP for having provided breeding lines improved for Aphanomyces resistance. We warmly thank the INRA experimental units of Dijon-Epoisses and Rennes-Le Rheu, as well as UNILET (Union Nationale Interprofessionnelle des Légumes Transformés), Quimperlé, France, for their contribution to the field experiments. We greatly

acknowledge Solène Coëdel from the Genouest genotyping platform and Marie-Christine Le Paslier from the Evry genotyping platform, for their contribution to the genotyping of the pea-Aphanomyces collection. We are very grateful to Mathieu Siol and Anne-Sophie Bouchet, who provided help in genome-wide association analysis and to Gwenola LeRoy who helped in the controlled conditions experiments. We greatly thank Clément Lavaud for his critical review of the manuscript. We thank Leigh Gebbie for revising the English written style of the manuscript.

References

1. St Clair DA. Quantitative disease resistance and quantitative resistance Loci in breeding. *Annu Rev Phytopathol.* 2010;48:247-68. doi:10.1146/annurev-phyto-080508-081904.
2. Poland JA, Balint-Kurti PJ, Wisser RJ, Pratt RC, Nelson RJ. Shades of gray: the world of quantitative disease resistance. *Trends Plant Sci.* 2009;14(1):21-9. doi:10.1016/j.tplants.2008.10.006.
3. Kou Y, Wang S. Broad-spectrum and durability: understanding of quantitative disease resistance. *Curr Opin Plant Biol.* 2010;13(2):181-5. doi:10.1016/j.pbi.2009.12.010.
4. Boyd LA, Ridout C, O'Sullivan DM, Leach JE, Leung H. Plant-pathogen interactions: disease resistance in modern agriculture. *Trends Genet.* 2013;29(4):233-40. doi:10.1016/j.tig.2012.10.011.
5. Huang X, Han B. Natural variations and genome-wide association studies in crop plants. *Annu Rev Plant Biol.* 2014;65:531-51. doi:10.1146/annurev-arplant-050213-035715.
6. Korte A, Farlow A. The advantages and limitations of trait analysis with GWAS: a review. *Plant Methods.* 2013;9:29. doi:10.1186/1746-4811-9-29.
7. Gupta PK, Kulwal PL, Jaiswal V. Association mapping in crop plants: opportunities and challenges. *Adv Genet.* 2014;85:109-47. doi:10.1016/B978-0-12-800271-1.00002-0.
8. Bao Y, Vuong T, Meinhardt C, Tiffin P, Denny R, Chen S et al. Potential of Association Mapping and Genomic Selection to Explore PI 88788 Derived Soybean Cyst Nematode Resistance. *Plant Genome.* 2014;7(3):0. doi:10.3835/plantgenome2013.11.0039.
9. Cheng P, Holdsworth W, Ma Y, Coyne CJ, Mazourek M, Grusak MA et al. Association mapping of agronomic and quality traits in USDA pea single-plant collection. *Mol Breed.* 2015;35(2). doi:10.1007/s11032-015-0277-6.
10. Ingvarsson PK, Street NR. Association genetics of complex traits in plants. *New Phytol.* 2011;189(4):909-22. doi:10.1111/j.1469-8137.2010.03593.x.
11. Segura V, Vilhjalmsson BJ, Platt A, Korte A, Seren U, Long Q et al. An efficient multi-locus mixed-model approach for genome-wide association studies in structured populations. *Nat Genet.* 2012;44(7):825-30. doi:10.1038/ng.2314.
12. Sauvage C, Segura V, Bauchet G, Stevens R, Do PT, Nikoloski Z et al. Genome-Wide Association in Tomato Reveals 44 Candidate Loci for Fruit Metabolic Traits. *Plant Physiol.* 2014;165(3):1120-32. doi:10.1104/pp.114.241521.
13. Begum H, Spindel JE, Lalusin A, Borromeo T, Gregorio G, Hernandez J et al. Genome-wide association mapping for yield and other agronomic traits in an elite breeding population of tropical rice (*Oryza sativa*). *PLoS One.* 2015;10(3):e0119873. doi:10.1371/journal.pone.0119873.
14. Zhao X, Han Y, Li Y, Liu D, Sun M, Zhao Y et al. Loci and candidate gene identification for resistance to *Sclerotinia sclerotiorum* in soybean (*Glycine max* L. Merr.) via association and linkage maps. *Plant J.* 2015;82(2):245-55. doi:10.1111/tpj.12810.

15. Samayoa LF, Malvar RA, Olukolu BA, Holland JB, Butron A. Genome-wide association study reveals a set of genes associated with resistance to the Mediterranean corn borer (*Sesamia nonagrioides* L.) in a maize diversity panel. BMC Plant Biol. 2015;15:35. doi:10.1186/s12870-014-0403-3.
16. Myles S, Peiffer J, Brown PJ, Ersoz ES, Zhang Z, Costich DE et al. Association mapping: critical considerations shift from genotyping to experimental design. Plant Cell. 2009;21(8):2194-202. doi:10.1105/tpc.109.068437.
17. Pascual L, Albert E, Sauvage C, Duangjit J, Bouchet J-P, Bitton F et al. Dissecting quantitative trait variation in the resequencing era: complementarity of bi-parental, multi-parental and association panels. Plant Sci. 2015. doi:10.1016/j.plantsci.2015.06.017.
18. Yu J, Holland JB, McMullen MD, Buckler ES. Genetic Design and Statistical Power of Nested Association Mapping in Maize. Genetics. 2008;178(1):539-51.
19. Benson JM, Poland JA, Benson BM, Stromberg EL, Nelson RJ. Resistance to gray leaf spot of maize: genetic architecture and mechanisms elucidated through nested association mapping and near-isogenic line analysis. PLoS Genet. 2015;11(3):e1005045. doi:10.1371/journal.pgen.1005045.
20. Cavanagh C, Morell M, Mackay I, Powell W. From mutations to MAGIC: resources for gene discovery, validation and delivery in crop plants. Curr Opin Plant Biol. 2008;11(2):215-21. doi:http://dx.doi.org/10.1016/j.pbi.2008.01.002.
21. Pascual L, Desplat N, Huang BE, Desgroux A, Bruguier L, Bouchet JP et al. Potential of a tomato MAGIC population to decipher the genetic control of quantitative traits and detect causal variants in the resequencing era. Plant Biotechnol J. 2015;13(4):565-77. doi:10.1111/pbi.12282.
22. Zila CT, Ogut F, Romay MC, Gardner CA, Buckler ES, Holland JB. Genome-wide association study of Fusarium ear rot disease in the U.S.A. maize inbred line collection. BMC Plant Biol. 2014;14:372. doi:10.1186/s12870-014-0372-6.
23. Hao D, Cheng Y, Chen G, Lu H, Shi M, Zhang Z et al. Identification of significant single nucleotide polymorphisms for resistance to maize rough dwarf disease in elite maize (*Zea mays* L.) inbred lines. Euphytica. 2014;203(1):109-20. doi:10.1007/s10681-014-1277-z.
24. FAOSTAT. FAOSTAT data base. Food and agriculture organisation of the United Nations - statistic division. 2015. <http://faostat3.fao.org/home/E>. Accessed 17 janv 2016.
25. McPhee K. Dry pea production and breeding - A mini-review. Journal of Food, Agriculture & Environment. 2003;1(1):64-9.
26. Jones FR, Drechsler C. Root Rot of Peas Caused by *Aphanomyces euteiches*. J Agric Res. 1925;30(4):293-325.
27. Papavizas GC, Ayers WA. *Aphanomyces* species and their root diseases in pea and sugarbeet: a review. United States Dept of Agriculture. 1974; Technical bulletin n°1485:1-158.

28. Wicker E, Rouxel F. Specific behaviour of French *Aphanomyces euteiches* Drechs. populations for virulence and aggressiveness on pea, related to isolates from Europe, America and New Zealand. Eur J Plant Pathol. 2001;107:919-29.
29. Wicker E, Moussart A, Duparque M, Rouxel F. Further contributions to the development of a differential set of pea cultivars (*Pisum sativum*) to investigate the virulence of isolates of *Aphanomyces euteiches*. Eur J Plant Pathol. 2003;109:47-60.
30. Chatterton S, Bowness R, Harding MW. First Report of Root Rot of Field Pea Caused by *Aphanomyces euteiches* in Alberta, Canada. Plant disease. 2015;99(2):288. doi:10.1094/PDIS09140905PDN.
31. Gritton ET. Registration of five pea root rot resistant germplasm lines of processing pea. Crop Sci. 1990;30:1166-7.
32. Kraft JM. Registration of 90-2079, 90-2131, and 90-2322 pea germplasms. Crop Sci. 1992;32:1076.
33. Davis DW. MN 144, MN 313, and MN 314: Garden Pea Lines Resistant to Root Rot Caused by *Aphanomyces euteiches* Drechs. HortScience. 1995;30(3):639-40.
34. Malvick DK, Percich JA. Identification of *Pisum sativum* Germ Plasm with Resistance to Root Rot Caused by Multiple Strains of *Aphanomyces euteiches*. Plant disease. 1999;83(1):51-4.
35. Kraft JM, Coffman VA. Registration of 96-2052, 96-2058, 96-2068, 96-2198, and 96-2222 Pea Germplasms. Crop Sci. 2000;40:301-2.
36. Pilet-Nayel L, Coyne CJ, Hamon C, Lesne A, Le Goff I, Esnault R et al., editors. Understanding genetics of partial resistance to *Aphanomyces* root rot in pea for new breeding prospects. 3rd International *Aphanomyces* Workshop on Legumes; 2007 07-09 November; Rennes, Brittany, France.
37. Roux-Duparque M, Boitel C, editors. Breeding peas for resistance to *Aphanomyces* root rot: current main outputs of three breeding programmes. 5th Eur Conf Grain Legumes; 2004 7-11th June; Dijon, France.
38. Moussart A, Devaux C, Muel F, Pilet-Nayel M, Baranger A, Tivoli B et al., editors. Improving partial resistance to *Aphanomyces* root rot in GSP breeding program. 3rd International *Aphanomyces* Workshop on Legumes; 2007 07-09 November; Rennes, Brittany, France.
39. Pilet-Nayel ML, Muehlbauer FJ, McGee RJ, Kraft JM, Baranger A, Coyne CJ. Quantitative trait loci for partial resistance to *Aphanomyces* root rot in pea. Theor Appl Genet. 2002;106(1):28-39. doi:10.1007/s00122-002-0985-2.
40. Pilet-Nayel ML, Muehlbauer FJ, McGee RJ, Kraft JM, Baranger A, Coyne CJ. Consistent Quantitative Trait Loci in Pea for Partial Resistance to *Aphanomyces euteiches* Isolates from the United States and France. Phytopathology. 2005;95(11):1287-93. doi:10.1094/PHYTO-95-1287.

41. Hamon C, Baranger A, Coyne CJ, McGee RJ, Le Goff I, L'Anthoene V et al. New consistent QTL in pea associated with partial resistance to *Aphanomyces euteiches* in multiple French and American environments. *Theor Appl Genet.* 2011;123(2):261-81. doi:10.1007/s00122-011-1582-z.
42. McGee RJ, Coyne CJ, Pilet-Nayel ML, Moussart A, Tivoli B, Baranger A et al. Registration of Pea Germplasm Lines Partially Resistant to *Aphanomyces* Root Rot for Breeding Fresh or Freezer Pea and Dry Pea Types. *Journal of Plant Registrations.* 2012;6(2):203-7. doi:10.3198/jpr2011.03.0139crg.
43. Hamon C, Coyne CJ, McGee RJ, Lesné A, Esnault R, Mangin P et al. QTL meta-analysis provides a comprehensive view of the moderately low diversity of loci controlling partial resistance to *Aphanomyces euteiches* in four pea sources of resistance. *BMC Plant Biol.* 2013;13:45.
44. Lavaud C, Lesné A, Piriou C, Le Roy G, Boutet G, Moussart A et al. Validation of QTL for resistance to *Aphanomyces euteiches* in different pea genetic backgrounds using Near Isogenic Lines. *Theor Appl Genet.* 2015. doi:10.1007/s00122-015-2583-0.
45. Duarte J, Riviere N, Baranger A, Aubert G, Burstin J, Cornet L et al. Transcriptome sequencing for high throughput SNP development and genetic mapping in Pea. *BMC Genomics.* 2014;15:126. doi:10.1186/1471-2164-15-126.
46. Sindhu A, Ramsay L, Sanderson L-A, Stonehouse R, Li R, Condie J et al. Gene-based SNP discovery and genetic mapping in pea. *Theor Appl Genet.* 2014;127(10):2225-41. doi:10.1007/s00122-014-2375-y.
47. Leonforte A, Sudheesh S, Cogan NO, Salisbury PA, Nicolas ME, Materne M et al. SNP marker discovery, linkage map construction and identification of QTLs for enhanced salinity tolerance in field pea (*Pisum sativum* L.). *BMC Plant Biol.* 2013;13:161. doi:10.1186/1471-2229-13-161.
48. Boutet G, Alves-Carvalho S, Falque M, Peterlongo P, Lhuillier E, Bouchez O et al. SNP discovery and genetic mapping using Genotyping by Sequencing of whole genome genomic DNA from a pea RIL population. *BMC Genomics.* submitted.
49. Tayeh N, Aluome C, Falque M, Jacquin F, Klein A, Chauveau A et al. Development of two major resources for pea genomics: the GenoPea 13.2K SNP Array and a high density, high resolution consensus genetic map. *Plant J.* 2015. doi:10.1111/tpj.13070.
50. Kwon S-J, Brown AF, Hu J, McGee R, Watt C, Kisha T et al. Genetic diversity, population structure and genome-wide marker-trait association analysis emphasizing seed nutrients of the USDA pea (*Pisum sativum* L.) core collection. *Genes Genomics.* 2012;34(3):305-20. doi:10.1007/s13258-011-0213-z.
51. Diapari M, Sindhu A, Warkentin TD, Bett K, Tar'an B. Population structure and marker-trait association studies of iron, zinc and selenium concentrations in seed of field pea (*Pisum sativum* L.). *Mol Breed.* 2015;35(1). doi:10.1007/s11032-015-0252-2.
52. Burstin J, Salloignon P, Chabert-Martinello M, Magnin-Robert JB, Siol M, Jacquin F et al. Genetic diversity and trait genomic prediction in a pea diversity panel. *BMC Genomics.* 2015;16:105. doi:10.1186/s12864-015-1266-1.

53. McGee RJ, Coyne CJ, Pilet-Nayel M-L, Moussart A, Tivoli B, Baranger A et al. Registration of Pea Germplasm Lines Partially Resistant to *Aphanomyces* Root Rot for Breeding Fresh or Freezer Pea and Dry Pea Types. *J Plant Regist.* 2012;6(2):203-7. doi:10.3198/jpr2011.03.0139crg.
54. Moussart A, Wicker E, Duparque M, Rouxel F, editors. Development of an efficient screening test for pea resistance to *Aphanomyces euteiches*. 4th European conference on grain legumes; 2001 8-12 July; Cracow, Poland.
55. Aubert G, Morin J, Jacquin F, Loridon K, Quillet MC, Petit A et al. Functional mapping in pea, as an aid to the candidate gene selection and for investigating synteny with the model legume *Medicago truncatula*. *Theor Appl Genet.* 2006;112(6):1024-41. doi:10.1007/s00122-005-0205-y.
56. Loridon K, McPhee K, Morin J, Dubreuil P, Pilet-Nayel ML, Aubert G et al. Microsatellite marker polymorphism and mapping in pea (*Pisum sativum* L.). *Theor Appl Genet.* 2005;111:1022-31. doi:10.1007/s00122-0050014-3.
57. Doyle JJ, Doyle JL. Isolation of plant DNA from fresh tissue. *Focus.* 1990;12:13-5.
58. Alves-Carvalho S, Aubert G, Carrère S, Cruaud C, Brochot A-L, Jacquin F et al. Full-length de novo assembly of RNA-seq data in pea (*Pisum sativum* L.) provides a gene expression atlas and gives insights into root nodulation in this species. *Plant J.* 2015;84:1-19. doi:10.1111/tpj.12967.
59. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2014. <http://www.R-project.org/>. Accessed 13 Jul 2015.
60. de Mendiburu F. agricolae: Statistical Procedures for Agricultural Research. R package version 1.2-1. 2014. <http://CRAN.R-project.org/package=agricolae>. Accessed 13 Jul 2015.
61. Hervé M. RVAideMemoire: Diverse basic statistical and graphical functions. R package version 0.9-36. 2014. <http://CRAN.R-project.org/package=RVAideMemoire>. Accessed 13 Jul 2015.
62. Lenth RV, Hervé M. lsmeans: Least-Squares Means. R package version 2.15. 2015. <http://CRAN.R-project.org/package=lsmeans>. Accessed 13 Jul 2015.
63. Revelle W. psych: Procedures for Psychological, Psychometric, and Personality Research. R package version 1.5.1. 2015. <http://CRAN.R-project.org/package=psych>. Accessed 13 Jul 2015.
64. Benjamini Y, Hochberg Y. Controlling the false discovery rate - A practical and powerful approach to multiple testing. *J R Stat Soc series B Stat Methodol.* 1995;57(1):289-300.
65. Warnes GR, Bolker B, Bonebakker L, Gentleman R, Liaw WHA, Lumley T et al. gplots: Various R Programming Tools for Plotting Data. R package version 2.16.0 2015. <http://CRAN.R-project.org/package=gplots>. Accessed 13 Jul 2015.

66. Husson F, Josse J, Le S, Mazet J. FactoMineR: Multivariate Exploratory Data Analysis and Data Mining. R package version 1.29. 2015. <http://CRAN.R-project.org/package=FactoMineR>. Accessed 13 Jul 2015.
67. Husson F, Josse J. missMDA: Handling missing values with/in multivariate data analysis (principal component methods). R package version 1.7.3. 2014. <http://CRAN.R-project.org/package=missMDA>. Accessed 13 Jul 2015.
68. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D et al. PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *Am J Hum Genet.* 2007;81:559-75.
69. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience.* 2015;4:7. doi:10.1186/s13742-015-0047-8.
70. Purcell SM, Chang CC. PLINK 1.9. 2015. <https://www.cog-genomics.org/plink2>. Accessed 13 Jul 2015.
71. Schwender H, Fritsch A. scribe: Analysis of High-Dimensional Categorical Data such as SNP Data. R package version 1.3.3. 2013. <http://CRAN.R-project.org/package=scribe>. Accessed 13 Jul 2015.
72. Schwender H. Imputing Missing Genotypes with Weighted k Nearest Neighbors. *J Toxicol Environ Health.* 2012;75(8-10):438-46. doi:10.1080/15287394.2012.674910.
73. Sved JA. Linkage disequilibrium and homozygosity of chromosome segments in finite populations. *Theor Popul Biol.* 1971;2(2):125-41. doi:[http://dx.doi.org/10.1016/0040-5809\(71\)90011-6](http://dx.doi.org/10.1016/0040-5809(71)90011-6).
74. Huang X, Wei X, Sang T, Zhao Q, Feng Q, Zhao Y et al. Genome-wide association studies of 14 agronomic traits in rice landraces. *Nat Genet.* 2010;42(11):961-7. doi:<http://www.nature.com/ng/journal/v42/n11/abs/ng.695.html#supplementary-information>.
75. Lipka AE, Tian F, Wang Q, Peiffer J, Li M, Bradbury PJ et al. GAPIT: genome association and prediction integrated tool. *Bioinformatics.* 2012;28(18):2397-9. doi:10.1093/bioinformatics/bts444.
76. Buckler Lab for Maize Genetics and Diversity. GAPIT source code and tutorial. 2012. <http://www.maizegenetics.net/#!gapit/cmkv>. Accessed 27 november 2015.
77. Segura V. Multi-Locus Mixed-Model (MLMM) script. 2013. <https://sites.google.com/site/vincentosegura/mlmm>. Accessed 13 Jul 2015.
78. Hothorn T, Bretz F, Westfall P. Simultaneous Inference in General Parametric Models. *Biom J.* 2008;50(3):346-63. doi:10.1002/bimj.200810425.
79. Sosnowski O, Charcosset A, Joets J. BioMercator V3: an upgrade of genetic map compilation and quantitative trait loci meta-analysis algorithms. *Bioinformatics.* 2012;28(15):2082-3.

80. Voorrips RE. MapChart: Software for the Graphical Presentation of Linkage Maps and QTLs. *J Hered.* 2002;93(1):77-8.
81. Timmerman-Vaughan GM, Frew TJ, Weeden NF. Characterization and linkage mapping of R-gene analogous DNA sequences in pea (*Pisum sativum* L.). *Theor Appl Genet.* 2000;101(1-2):241-7. doi:10.1007/s001220051475.
82. Prioul-Gervais S, Deniot G, Receveur EM, Frankewitz A, Fourmann M, Rameau C et al. Candidate genes for quantitative resistance to *Mycosphaerella pinodes* in pea (*Pisum sativum* L.). *Theor Appl Genet.* 2007;114(6):971-84. doi:10.1007/s00122-006-0492-y.
83. Armstead I, Donnison I, Aubry S, Harper J, Hörtensteiner S, James C et al. Cross-Species Identification of Mendel's I Locus. *Science.* 2007;315(5808):73-.
84. Afzal AJ, Wood AJ, Lightfoot DA. Plant Receptor-Like Serine Threonine Kinases: Roles in Signaling and Plant Defense. *Mol Plant Microbe Interact.* 2008;21(5):507-17. doi:10.1094/mpmi-21-5-0507.
85. Walter S, Kahla A, Arunachalam C, Perochon A, Khan MR, Scofield SR et al. A wheat ABC transporter contributes to both grain formation and mycotoxin tolerance. *J Exp Bot.* 2015;66(9):2583-93. doi:10.1093/jxb/erv048.
86. Hecht V, Foucher F, Ferrandiz C, Macknight R, Navarro C, Morin J et al. Conservation of Arabidopsis flowering genes in model legumes. *Plant Physiol.* 2005;137(4):1420-34. doi:10.1104/pp.104.057018.
87. Weller JL, Liew LC, Hecht VFG, Rajandran V, Laurie RE, Ridge S et al. A conserved molecular basis for photoperiod adaptation in two temperate legumes. *Proc Natl Acad Sci U S A.* 2012;109(51):21158-63.
88. Hofer J, Turner L, Hellens R, Ambrose M, Matthews P, Michael A et al. UNIFOLIATA regulates leaf and flower morphogenesis in pea. *Curr Biol.* 1997;7(8):581-7. doi:10.1016/s0960-9822(06)00257-0.
89. Weston DE, Elliott RC, Lester DR, Rameau C, Reid JB, Murfet IC et al. The Pea DELLA Proteins LA and CRY Are Important Regulators of Gibberellin Synthesis and Root Growth. *Plant Physiol.* 2008;147(1):199-205. doi:10.1104/pp.108.115808.
90. Reinecke DM, Wickramaratna AD, Ozga JA, Kurepin LV, Jin AL, Good AG et al. Gibberellin 3-oxidase Gene Expression Patterns Influence Gibberellin Biosynthesis, Growth, and Development in Pea. *Plant Physiol.* 2013;163(2):929-45. doi:10.1104/pp.113.225987.
91. Warkentin TD. Pea. In: De Ron AM, editor. *Grain Legumes. Handbook of Plant Breeding*, vol 10: Springer-Verlag New York; 2015.
92. Lewis ME, Gritton ET. Use of One Cycle of Recurrent Selection per Year for Increasing Resistance to *Aphanomyces* Root Rot in Peas. *Journal of the American Society for Horticultural Science.* 1992;117(4):638-42.
93. Letta T, Maccaferri M, Badebo A, Ammar K, Ricci A, Crossa J et al. Searching for novel sources of field resistance to Ug99 and Ethiopian stem rust races in durum wheat via

association mapping. *Theor Appl Genet.* 2013;126(5):1237-56. doi:10.1007/s00122-013-2050-8.

94. Naruoka Y, Garland-Campbell KA, Carter AH. Genome-wide association mapping for stripe rust (*Puccinia striiformis* F. sp. *tritici*) in US Pacific Northwest winter wheat (*Triticum aestivum* L.). *Theor Appl Genet.* 2015;128(6):1083-101. doi:10.1007/s00122-015-2492-2.

95. Gurung S, Mamidi S, Bonman JM, Xiong M, Brown-Guedira G, Adhikari TB. Genome-wide association study reveals novel quantitative trait Loci associated with resistance to multiple leaf spot diseases of spring wheat. *PLoS One.* 2014;9(9):e108179. doi:10.1371/journal.pone.0108179.

96. Zhou H, Steffenson B. Genome-wide association mapping reveals genetic architecture of durable spot blotch resistance in US barley breeding germplasm. *Mol Breed.* 2013;32(1):139-54. doi:10.1007/s11032-013-9858-4.

97. Goudemand E, Laurent V, Duchalais L, Tabib Ghaffary S, Kema GJ, Lonnet P et al. Association mapping and meta-analysis: two complementary approaches for the detection of reliable *Septoria tritici* blotch quantitative resistance in bread wheat (*Triticum aestivum* L.). *Mol Breed.* 2013;32(3):563-84. doi:10.1007/s11032-013-9890-4.

98. Michelmore RW, Christopoulou M, Caldwell KS. Impacts of Resistance Gene Genetics, Function, and Evolution on a Durable Future. *Annu Rev Phytopathol.* 2013;51(1):291-319. doi:10.1146/annurev-phyto-082712-102334.

99. Ameline-Torregrosa C, Wang B-B, O'Bleness MS, Deshpande S, Zhu H, Roe B et al. Identification and Characterization of Nucleotide-Binding Site-Leucine-Rich Repeat Genes in the Model Plant *Medicago truncatula*. *Plant Physiol.* 2008;146(1):5-21. doi:10.1104/pp.107.104588.

100. Podolich O, Ardanov P, Zaets I, Pirttilä A, Kozyrovska N. Reviving of the endophytic bacterial community as a putative mechanism of plant resistance. *Plant and Soil.* 2015;388(1-2):367-77. doi:10.1007/s11104-014-2235-1.

101. Costes E, Lauri PE, Simon S, Andrieu B. Plant architecture, its diversity and manipulation in agronomic conditions, in relation with pest and pathogen attacks. *Eur J Plant Pathol.* 2013;135(3):455-70. doi:10.1007/s10658-012-0158-3.

Chapitre III

Analyse génétique
comparative de l'architecture
constitutive de la plante et de
la résistance à *Aphanomyces*
euteiches chez le pois

Chapitre III : Analyse génétique comparative de l'architecture constitutive de la plante et de la résistance à *Aphanomyces euteiches* chez le pois

1. INTRODUCTION DU CHAPITRE

Dans ce chapitre, nous avons étudié l'organisation structurale des régions génomiques impliquées dans le contrôle de l'architecture constitutive racinaire, aérienne et globale de la plante, comparativement à celles impliquées dans la résistance intrinsèque à *A. euteiches* chez le pois. Nous avons ensuite mené une étude comparative des régions détectées dans ce chapitre, par rapport à celles identifiées dans le chapitre II et précédemment, afin de mettre en évidence des co-localisations de ces régions.

L'analyse a été réalisée sur une collection de 266 lignées de pois, incluant les 175 lignées de la collection « *Aphanomyces* » étudiée dans le chapitre II et 104 lignées d'une collection « Racines » (Bourion in prep) dont 13 lignées communes à la collection « *Aphanomyces* ». Par rapport au chapitre II, les données de génotypage ont été complétées sur l'ensemble de la collection pour les 13.2K marqueurs SNP utilisés dans le chapitre II, et obtenues pour 953 marqueurs SNP supplémentaires (Boutet et al. 2016) ; l'analyse a été réalisée selon le même modèle multi-locus. La collection a été évaluée en 2013 pour des caractères d'architecture sur jeunes plantes dans deux conditions contrôlées non inoculées à Dijon et à Rennes et pour la résistance intrinsèque en conditions contrôlées inoculées à Rennes.

Les résultats obtenus nous ont permis (i) d'identifier des corrélations entre la résistance à *A. euteiches* et des caractères de développement racinaire et aériens ; et (ii) d'identifier une diversité de régions génomiques impliquées dans le contrôle de ces caractères, dont une région commune à effet majeur associée à la fois la résistance et un caractère d'architecture racinaire constitutive.

Ce chapitre est présenté sous la forme d'un article en préparation pour soumission.

2. ARTICLE ASSOCIÉ

*Les références bibliographiques mentionnées sont situées à la fin du manuscrit
Les fichiers additionnels sont présentés en annexe 2*

Genetic determinism of root growth and architecture traits shares a common locus with *Aphanomyces* root rot resistance QTL in pea

1. Introduction

Plant architecture has often been reported as contributing to modify susceptibility of organs to plant pathogens (or to plant pests) or to favour mechanisms leading to infection escape and to increased tolerance (Ney et al. 2013). Spatial disease avoidance was shown to result from combination of architectural features with unfavourable effects on disease development and severity (Tivoli et al. 2012b). Plant or canopy architecture unfavourable to an aerial epidemic may result in the total avoidance of disease expression, but more frequently reduces disease severity rather than preventing infection completely (Tivoli et al. 2012b). Any change in the position of susceptible plant organs may also minimize epidemic potential and dynamics of one or more pests (Calonnec et al. 2012). For example, variations in number or size of stomata on leaves may decrease disease severity by reducing entry points for pathogens that penetrate through stomata (Calonnec et al. 2012). Correlations between variables describing plant growth or architecture and variables describing disease severity were tested to identify which of them had the highest impact on reducing epidemics (Calonnec et al. 2012). Combining genetic resistance and the most unfavourable architectural traits to diseases would thus be a strategy of interest to reduce epidemics. For instance, breeding cultivars for an upright growth habit, lodging-resistance, and partial intrinsic resistance has improved the management of white mold of bean, soybean, canola, peanut, and potato (McDonald et al. 2013). For the last decades, breeding efficiency has largely benefited from improved phenotyping techniques and the development of molecular markers and of whole genome sequences, which helped in mapping genes controlling architectural traits and partial resistance. In some cases, genes responsible for architectural traits and resistance QTL (Quantitative Trait Loci) were found to overlap. For example, Prioul et al. (2004) found co-localisations of QTL for resistance to *Mycosphaeralla pinodes*, plant height and flowering date in the same genomic region in pea. Similarly, colocalisations were observed between QTL of resistance to white mold in snap bean and QTL controlling canopy height, lodging traits or internode length (Miklas et al. 2013; Tivoli et al. 2012b). Thus, several traits of plant architecture have been shown to improve plant resistance to diseases, especially by reducing the probability of spatial plant-pathogen interactions. Most of these studies concerned aerial

plant architecture and pathogens. Few works have been done on the relationship between root architecture and root pathogens very probably due to the difficulty of studying root pathosystems. However, large number of lateral roots, large root diameter or large root dry weight were shown to improve *Fusarium* root rot resistance in common bean (Cichy et al. 2007; Román-Avilés et al. 2004; Snapp et al. 2003) and pea (based on number of roots, length and surface area of roots) (Kraft and Boge 2001).

Roots are crucial for water and nutrient supply, as well as anchorage into the soil. The spatial configuration of the root system, so-called the root system architecture (RSA) varies greatly according to both intrinsic and environmental determinants (Malamy 2005). Intrinsic determinants are essential for organogenesis and growth, and determine the characteristic of RSA in a given plant species. Environmental determinants are numerous and include the soil nutritional composition, its density and compaction, its salinity and water content, and the presence of micro-organisms. RSA is the result of growth and developmental processes. The root system originates from a primary root that develops during embryogenesis. This primary root produces secondary roots, which in turn produce tertiary roots. All secondary, tertiary, quaternary and further roots are referred as lateral roots. RSA is generally achieved by measuring variables such as secondary lateral root number, root length and average diameter. In contrast, variables describing the variety of components constituting the relationship between the root segments (e.g. type and angle of connexion between roots; root gradients) are considered as relative to the root system topology or structure (Hodge et al. 2009). The intrinsic determinants of RSA are those which are essential for lateral root initiation, developmental patterning of the primordium, and lateral root formation and growth (Malamy 2005). In response to environmental stimuli, Plants can optimize their RSA by initiating more or less lateral root primordia and influencing growth of primary or lateral roots. Thus, RSA results from the expression of a great number of quantitative traits mainly controlled by a large number of loci (Hodge et al. 2009). Few studies managed to show influence of some RSA traits on disease severity. However, large number of lateral root, large root diameter or large root dry weight were shown to improve *Fusarium* root rot resistance in common bean (Cichy et al. 2007; Román-Avilés et al. 2004; Snapp et al. 2003). *Aphanomyces* root rot resistance was correlated with a high number of lateral roots in *Medicago truncatula*, (Bonhomme et al. 2014; Djébali et al. 2013; Laffont et al. 2015). Kraft and Boge (2001) also showed that large-rooted pea lines (based on number of roots, length and surface area of roots) were more resistant to *Fusarium* root rot.

Dry pea (*Pisum sativum*), one of the most cultivated grain legumes throughout the world, is used for animal feeding and human nutrition. Its symbiosis with *Rhizobium* to catch atmospheric dinitrogen makes it valuable to include in crop rotation and to decrease chemical nitrogen use. However, *Aphanomyces euteiches*, a soil borne pathogen infecting roots of a large legume host range, is a limiting factor of pea crop development, since it can cause high yield losses in highly infected fields (Gaulin et al. 2007). The disease causes browning of the roots from the three to four-leaf stage of the plant when climatic conditions are favourable (temperature higher than 10°C and available free water in the soil). Resistance to *A. euteiches* in pea has been well explored in the last decade and has been shown to be partial and under polygenic control (Desgroux et al. 2016; Hamon et al. 2011; Hamon et al. 2013; Pilet-Nayel et al. 2002, 2005). Furthermore, pea shows diverse RSA between genotypes, from seedling stage to mature plant (Bourion et al. 2010; McPhee 2005). Pea root architecture is also under polygenic control as three, eight and 21 QTL have been reported for root dry matter, number of lateral roots and root length, respectively (Bourion et al. 2010). In peas, relationships between resistance to *A. euteiches* and root architecture have not been explored to date.

The aim of the present study is to determine the potential effect of intrinsic RSA traits on plant resistance to soil-borne root pathogens. The pathosystem *P. sativum* / *A. euteiches* was used to (i) **explore genetic diversity** between disease resistance fourteen days post sowing in controlled conditions and RSA in different pathogen free controlled conditions at seven and fourteen days post sowing; (ii) **explore relationships** between disease resistance and RSA; and (iii) **identify and compare genomic regions** controlling disease resistance and RSA through genome wide association (GWA) analysis. A pea collection of 266 pure lines, segregating for intrinsic RSA traits and resistance to *A. euteiches*, was phenotyped for both in controlled conditions. The lines were genotyped using about 15 SNP markers. Phenotypic correlations and one genomic QTL co-localization were found between a trait of RSA and controlled conditions resistance to *A. euteiches* in pea, suggesting that RSA traits could be selected to improve resistance to *A. euteiches*.

2. Material and method

2.1. Plant material

A collection of 266 *P. sativum* lines, including the 175 lines of the pea *Aphanomyces* collection described in Desgroux et al. (2016) and the 104 lines of the pea Root collection presented in Bourion et al. (2010) (Additional file III.1) was used in this study. Thirteen pea lines were common to both collections. The pea *Aphanomyces* collection was composed of: (i) one hundred one breeding lines, partially resistant or susceptible to *Aphanomyces euteiches*, originating from the *Aphanomyces* recurrent selection-based breeding program conducted by GSP (Roux-Duparque and Boitel 2004); (ii) twenty RILs from INRA and USDA research programs selected for *Aphanomyces* resistance and for morphological and developmental traits; (iii) eighteen parental lines of RIL populations used for QTL analysis of *Aphanomyces* resistance, and/or of the *Aphanomyces* recurrent selection-based breeding program conducted by the GSP; (iv) twenty lines representing new sources of resistance, selected from the INRA screening program of *Aphanomyces* resistance sources previously conducted (Pilet-Nayel et al 2007); (v) sixteen spring or winter pea varieties currently or formerly grown in Europe.

2.2. Phenotyping

All lines were assayed in three experiments, including one experiment in a greenhouse at INRA-Dijon (Exp#1) and two experiments in a climatic chamber at INRA-Le Rheu (Exp#2 and Exp#3).

The Exp#1 was carried in a greenhouse (semi-controlled conditions: 20°C for 16h-day, 18°C for 8h-night) located at INRA-Dijon (Burgundy, France), using a randomised complete block design with two replicates. Four surface sterilized seeds per block were sown in a pot (12cm x 12cm x 20cm) for each pea line. Plants were grown in a 1:1 (v/v) mixture of sterilized atapulgite and clay balls (2-6 mm diameter) (Bourion et al. 2010) and watered as needed. In each pot, two eight-day old seedlings were carefully uprooted, washed and stored at 5°C before plant architecture measurements.

The Exp#2 was carried out in a climatic chamber (25°C for 16h-day, 23°C for 8h-night) located at INRA-Le Rheu (Brittany, France). All the lines were assayed using a randomised complete block design with three replicates. Five seeds per replicate were sown in a pot (9 cm x 9 cm x 9.5 cm) for each pea line. Plants were grown in vermiculite substrate and watered as

needed. Fourteen-day old seedlings were uprooted and two seedlings per pot were carefully washed before plant architecture trait measurements.

In Exp#1 and Exp#2 plant traits relative to both growth and architecture were measured as described in Bourion et al. (2010). Tap root length (TapRootL), number of first lateral roots (NLatRoot) and plant height (ShootL) were first measured on each plant. Then, root and aerial parts were separated at the cotyledons insertion point; roots were carefully spread on a transparent sheet and scanned with a blue background at 300 dpi (Exp#1: A3 colour scanner (Epson; Tokyo, Japan); Exp#2: A4 colour scanner (Epson Perfection V37 J232C; Tokyo, Japan)). Images were then analysed using Winrhizo® Software (Regent Instruments, Quebec, Canada) with a home-made colour scale to take into account the maximum root area. Several traits were scored from image analyses: total projected area (TProjArea), total root length (TRootL) and average root diameter (RootDia). Average lateral root length (LatRootL) was computed from measurements and image analyses as following: $LatRootL = (TRootL - TapRootL) / NLatRoot$. Roots and shoot parts were then dried at 80°C during 48 hours and weighted to obtain root biomass (RootB) and shoot biomass (ShootB). Total biomass (TB) and root to total biomass ratio (RootB:TB) were computed as $TB = RootB + ShootB$ and $RootB:TB = RootB / TB$.

The Exp#3, was conducted in the same climatic chamber as Exp#2 (25°C for 16h-day, 23°C for 8h-night), as described in Desgroux et al. (2016). Seven-day old seedlings were inoculated with a pure-culture of the RB84 French reference strain of *A. euteiches* described in (Hamon et al. 2011), referred as from pathotype I (Wicker et al. 2001). Inoculation was performed with a solution of 10^3 zoospores per plant, produced as described by Moussart et al. (2001). Seven days after inoculation, disease severity (DS) was assessed on five plants per plot, using a 0 (no symptoms) to 5 (dead plant) scoring scale proportional to the percentage of browning symptoms on roots and epicotyls (Hamon et al. 2011). For each genotype, a Root Rot Index (RRI) was calculated as the mean of the five scores in a pot.

Among the five plants, two were chosen for further disease assessment. Their roots were carefully spread on a transparent sheet and scanned as described for Exp#2. Image analyses were then performed using the Winrhizo® Software with a home-made colour scale which enabled to discriminate healthy and diseased root projected area. Total projected area (TProjArea) and brown projected area (BrProjArea) were used to compute the percentage of

brown projected area (Br:TProjArea) using the formula $Br:TProjArea = BrProjArea \times 100 / TProjArea$. This variable represented the percentage of root system diseased.

2.3. Statistical analysis of phenotypic data

Phenotypic datasets, obtained from the pea collection for plant architecture traits and resistance to *A. euteiches*, were analysed using R 3.1.1 program (R Core Team 2014). For each variable, a linear model (LM) [R function `lm`] was used, including G (genotype), R (replicate) and P (plant) as fixed factors. Normality of residuals and homogeneity of their variance were checked using Skewness, Kurtosis and Shapiro-Wilk statistics ($p\text{-value} \geq 0.05$), as well as Bartlett test ($p\text{-value} \geq 0.05$) as described in Desgroux et al. (2016), respectively. Mean-based heritability (h^2) was calculated for each variable from variance estimates using the formula $h^2 = \sigma_G^2 / [\sigma_G^2 + (\sigma_E^2/r)]$, where σ_G^2 is the genetic variance, σ_E^2 the residual variance and r the number of replicates per line. Least Square Means (LSMeans) were calculated from each LM analysis (R function `lsmeans` of package `lsmeans`, Lenth and Hervé (2015)). Histograms of LSMeans frequency distributions were drawn using the R function `hist`. Pearson correlation analysis was carried out between LSMeans of the different variables (R function `corr.test` of package `psych`, Revelle (2015)). The significance of the Pearson correlations was tested with a false discovery rate correction for multiple testing (corrected $p\text{-value} < 0.05$; Benjamini and Hochberg (1995)). A regression curve was drawn between the Br:TProjArea and DS variables using the linear model $Br:TProjArea \sim DS$.

2.4. Genetic analysis

2.4.1. Genotyping and consensus map

Pea lines were genotyped using a total of 14,157 SNP markers, including 13,204 SNP markers from the GenoPea 13.2K Infinium SNP Array (Tayeh et al. 2015a) and 953 SNP markers from Boutet et al. (2016). Genotyping data with the 13,204 SNPs were obtained as described in Desgroux et al. (2016). Genotyping data with the 953 SNPs were obtained in this study from the same DNA samples as previously, using KASP™ SNP assays carried out in LGC Genomics service lab, UK (<http://www.lgcgenomics.com>) as described in Boutet et al. (2016). Each line was coded “AA” or “BB” when homozygous for the first or second allele and “AB” when heterozygous.

The genotyping dataset of the pea collection, composed of 14,157 SNP markers was filtered for genotyping quality and was thus reduced to 12,812 SNP markers that were then filtered using PLINK 1.9 software (Chang et al. 2015; Purcell et al. 2007; Purcell and Chang 2015).

Six of the 266 pea lines with missing data for more than 10% of SNP markers were not included in the GWA analysis (the same lines as in Desgroux et al. (2016). Markers with missing data that exceeded 10% or with a minor allele frequency (MAF) that did not exceed 5% in the 260 remaining lines, were also removed from the GWA analysis as described in Desgroux et al (2016). A total of 11,789 SNP markers were thus retained in the genetic analysis.

The genotyping raw data set of 11,789 SNP markers, containing 0.7% of missing values, was imputed using the R function `knncatimputeLarge` (package `scrime`, Schwender and Fritsch (2013)). Imputation parameters were tested with 10 replicates, using a subset of 4,805 SNP markers of the dataset with no missing values. For each replicate, 0.7% of missing values were randomly simulated and parameters of imputation tested for one to 50 nearest neighbours and four different methods to determine distances between SNPs, as described in Desgroux et al. (2016). Parameters with the lower error rate over the 10 replicates (eight nearest neighbours and Cohen's kappa method; error rate: 17.06%) were applied to the 11,789-SNP-marker dataset to impute missing values.

The two sets of markers used in this study, deriving from the GenoPea Infinium SNP Array and the KASPTM assays, respectively, had already been mapped to different genetic maps (Boutet et al. 2016; Desgroux et al. 2016; Tayeh et al. 2015a). For this study, a consensus marker map was obtained by projection of the 953 SNP markers genotyped on the collection, from the 64,263 markers map of Boutet et al. (2016) (namely BP-WGGBS map), onto the consensus map established by Desgroux et al. (2016) (namely THMap), based on 910 SNP markers from Duarte et al. (2014) common to the two maps, using Biomecatator V4.2 software (Sosnowski et al. 2012). The THMap contained individual- and meta-QTL for *Aphanomyces* resistance from linkage (Hamon et al. 2011; Hamon et al. 2013) or association (Desgroux et al. 2016) mapping. Seventy-eight markers were suppressed because they were located within inversions. The level of connectivity between the two maps was estimated with the "InfoMap" tool of the software. Extraction of the BP-WGGBS map was projected onto the THMap with the "iterative map projection" tool of the software, based on loci position data.

Pairwise Linkage Disequilibrium (LD) between markers was explored within Linkage Groups (LGs) from imputed genotypic data using PLINK 1.9 software. Obtained square correlation coefficient (r^2) values were then plotted against genetic distances (cM) according to the consensus map obtained in this study after projection, to estimate the LD decay. LD

decay curve and rate for each LG were estimated as described in Desgroux et al. (2016), based on the consensus genetic map from Tayeh et al. (2015a).

2.4.2. Population structure and individual relatedness and genome wide association

To estimate the structure of the collection, a Principal Component Analysis (PCA) and a Kinship relatedness matrix were carried out, as described in Desgroux et al. (2016), using genotyping data at 2,937 SNP markers mapped to distinct genetic positions according to Tayeh et al. (2015a) and the EMMA (efficient mixed-model association) method in the GAPIT (Genome Association and Prediction Integrated Tool) R package (Lipka et al. 2012). Regarding to eigenvalues, the three first principal components (PCs) of the PCA were included in the GWA analysis to take into account the structure of the collection.

GWA study used data obtained on the 260 lines of the collection: (i) LSMMeans of plant height (ShootL), shoot biomass (ShootB), nine RSA traits (NLatRoot, TRootL, LatRootL, TProjArea, RootDia, RootB, TB and RootB:TB) and three Aphanomyces resistance variables (RRI, DS and Br:TProjArea); (ii) genotyping data at 11,789 SNP markers over all the seven LGs. GWA analyses were performed using a modified version of the multi-locus mixed model (MLMM) R package (Segura et al. 2012), as described in Desgroux et al. (2016). The PCA matrix of population structure and the Kinship matrix were defined as cofactors in the MLMM (see the `mlmm_cof.r` R script at <https://sites.google.com/site/vincentosegura/mlmm>). Significant SNP markers were also used as cofactors in a forward/backward approach. A multiple-Bonferroni (`mBonf`) threshold of 4.58 (p -value of $2.6E-5$), taking into account 3,840 distinct genetic positions, was used to declare significant SNPs, as described in Desgroux et al. (2016). In each GWA analysis, the optimal MLMM step was determined as the largest stepwise mixed model regression in which all cofactors have $-\log(p\text{-value})$ above the `mBonf` threshold defined. Local LD analysis defined the confidence intervals (CIs) around significant associated markers detected by GWA study using Plink 1.9 software. Each CI was determined as the interval containing all markers in LD ($r^2 > 0.2$) with the significant associated marker, as described in Pascual et al. (2015a) and Desgroux et al. (2016).

2.4.3. Comparative mapping

The genetic map described in Bourion et al. (2010) was previously used to map QTL implied in pea root and aerial architecture (Bourion et al. 2010) and shares 141 common markers with the THMap map developed in this study. Several RSA and aerial architecture traits were

common to the present study and the one of Bourion et al. (2010): ShootL, ShootB, TB, TRootL, NLatRoot, RootB and BGB:TB, this last trait being the ratio of below ground biomass and total biomass. In the present study, seedlings were too young to carry nodules and in this case, BGB and RootB values were the same. Plant architecture QTL detected in Bourion et al. (2010) could thus be projected onto the present consensus marker map following the same method than described in Desgroux et al. (2016).

Visualization of loci detected by association and linkage mapping on the resulting comparative marker map was computed using MapChart 2.1 software (Voorrips 2002).

3. Results

3.1. Analysis of phenotypic data

Statistical analysis of plant architecture variables, obtained on the 266-pea-line collection at 8 and 14 days after sowing in Exp#1 and Exp#2, respectively, revealed highly significant G effects ($p\text{-value} < 0.001$) for all the variables (Table III.1). Heritabilities of root, shoot and overall plant architecture traits were high ($h^2 \geq 0.71$ and 0.8 in Exp#1 and Exp#2, respectively). Frequency distribution of LSMeans values for each variable tended to fit normal curves (Additional file III.2), except for ShootL which showed bimodal distribution in Exp#2. Statistical analysis of disease variables RRI, DS and Br:TProjArea, obtained on the 266 pea lines in Exp#3, revealed highly significant G effects ($p\text{-value} < 0.001$) and a high trait heritability ($h^2 > 0.82$). Frequency distribution of LSMeans values for the RRI and DS variables tended to show groups of resistant and susceptible lines and the one for Br:TProjArea tended to fit a normal curve (continuous distribution).

In Exp#1 and Exp#2, LSMeans of most of the RSA variables, except RootDia, were highly significantly and positively correlated between each other (corrected $p\text{-value} < 0.001$; $0.22 < r < 0.94$) (Table III.2). RootDia LSMeans were mostly significantly and negatively correlated with LSMeans of other RSA variables ($2.91E-13 < \text{corrected } p\text{-value} < 0.048$, $-0.44 < r < -0.13$), except with RootB LSMeans ($r \leq 0.39$). In both experiments, ShootL and ShootB LSMeans were significantly and positively correlated between each other (corrected $p\text{-value} < 0.001$, $r \leq 0.51$). They were significantly and positively correlated with LSMeans of most of the other RSA variables (corrected $p\text{-value} < 0.001$, $0.34 < r < 0.7$), except with RootDia LSMeans ($-0.49 \leq r \leq 0.15$). TB and RootB:TB LSMeans were significantly and positively correlated with most of the root and shoot architecture variables LSMeans ($0.15 < r < 0.98$),

Table III.1: Statistical analysis results of *Aphanomyces* resistance and plant architecture data in the 266-pea-line collection

Experiment (a)	Variable (b)	Number of lines (c)	Number of observations (c)	Mean (c)	Min (c)	Max (c)	SE (c)	Adjusted r ² model (d)	h ² (e)	G effect (f,g)	R effect (f,h)	P effect (f,i)
Exp#1	TProjArea	255	891	9.8	0	17.25	1.2	0.63	0.81	***	***	**
	RootDia	\$	884	0.8	0.58	1.1	0	0.71	0.87	***	***	ns
	TRootL		891	126.76	0	237.82	16.1	0.68	0.82	***	***	**
	NLatRoot	\$	889	39.49	4.78	72	3.9	0.64	0.84	***	***	ns
	LatRootL	\$	882	2.8	0.17	5.21	0.3	0.54	0.71	***	***	**
	RootB		882	42.33	2.69	69.32	4.1	0.66	0.85	***	***	*
	ShootB		870	50.14	7.31	101.52	4.3	0.75	0.89	***	***	*
	ShootL	\$	879	6.58	0.99	21.14	0.8	0.88	0.96	***	***	ns
	TB		868	92.58	9.97	161	7.5	0.70	0.85	***	***	*
	RootB:TB		867	0.46	0.28	0.61	0	0.73	0.9	***	**	ns
	Exp#2	TProjArea	\$	1527	26.18	5.75	44.7	1.9	0.54	0.87	***	***
RootDia			1522	0.66	0.51	0.82	0	0.66	0.9	***	***	***
TRootL		\$	1527	400.53	101.59	685.47	28.2	0.59	0.89	***	***	ns
NLatRoot			1536	46.8	22.32	81.47	2.7	0.68	0.92	***	***	ns
LatRootL			1525	8.34	3.93	14.33	0.6	0.46	0.8	***	***	ns
RootB		\$	1535	45.33	6.33	72.4	3.4	0.57	0.87	***	***	ns
ShootB		\$	1531	100.85	18.67	206.75	6.9	0.72	0.93	***	***	ns
ShootL		\$	1527	14.07	4.7	33.84	0.7	0.96	0.99	***	***	*
TB		\$	1530	146.14	25.03	271.86	9.5	0.68	0.92	***	***	ns
RootB:TB		\$	1531	0.31	0.19	0.41	0	0.62	0.9	***	***	ns
Exp#3		CC_RRI_RB84	\$	819	3.36	1.9	4.28	0.23	0.66	0.82	***	***
	CC_DS_RB84	\$	1602	3.37	1.62	4.26	0.21	0.56	0.86	***	***	ns
	CC_Br:TProjArea_RB84	\$	1586	0.48	0.35	0.6	0.02	0.60	0.84	***	***	***

(a) Exp#1: Plants were grown in a healthy greenhouse during 8 days before assessment, Exp#2: Plants were grown in a healthy climatic chamber during 14 days before assessment, Exp#3: Plants were grown in climatic chamber during 14 days and inoculated at the 7th day with a reference strain (RB84); (b) TProjArea: total projected area, RootDia: average root diameter, TRootL: total root length, NLatRoot: number of lateral roots, LatRootL: average length of lateral roots, RootB: root biomass, ShootB: shoot biomass, ShootL: shoot length, TB: total biomass, RootB:TB: root to total biomass ratio, CC: controlled conditions, RRI: Root rot index, DS: disease severity score, Br:TProjArea: percentage of brown projected area, \$' Residual normality statistics are not exact but graphs of residuals are correct; (c) Number of pea lines actually assessed, number of total observations, mean, minimum, maximum and standard error in the 266-pea-line collection; (d) adjusted r² coefficient of the model; (e) mean-based heritability (f) ANOVA significance p-value codes: 0 < '****' < 0.001 ≤ '***' < 0.01 ≤ '**' < 0.05 ≤ 'ns' < 1; (g): Genotype effect; (h): Replicate effect; (i): Plant effect

except. RootB:TB and shoot architecture LSMMeans which were significantly and negatively correlated between each other ($-0.42 < r < -0.74$). LSMMeans obtained for each variable in Exp#1 and Exp#2 were significantly and positively correlated between each other (corrected p -value < 0.001 , 0.38 for LatRootL $< r < 0.94$ for ShootL).

In Exp#3, LSMMeans of RRI and DS were significantly and positively correlated with LSMMeans of Br:TProjArea ($r = 0.7$; corrected p -value < 0.001). The regression curve of Br:TProjArea LSMMeans plotted against DS LSMMeans showed a coefficient (r^2) of 0.48 (Figure III.1). RRI, DS and Br:TProjArea LSMMeans (Exp#3) were significantly and negatively correlated with LSMMeans of most root growth and RSA variables (TProjArea, TRootL, NLatRoot, RootB:TB) in both Exp#1 and Exp#2 (corrected p -value < 0.05 , $-0.33 < r < -0.13$). They were significantly and positively correlated with LSMMeans of RootDia and ShootL (corrected p -value < 0.05 , $0.15 < r < 0.25$). In contrast, they were not correlated with LatRootL, ShootB and TB in both Exp#1 and Exp#2.

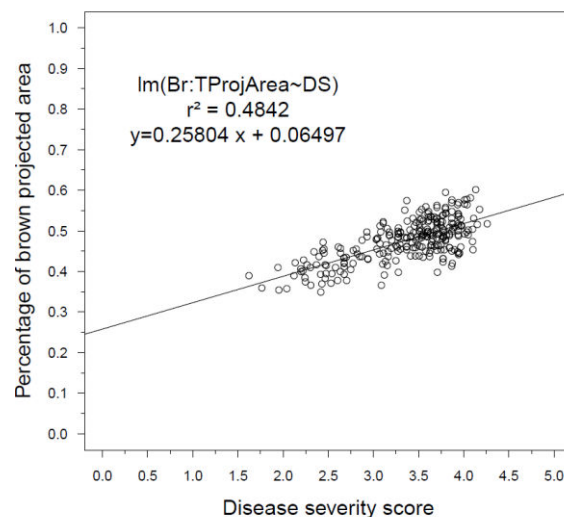


Figure III.1: Regression line between disease severity score and percentage of brown projected area.

LSMeans of disease severity (DS) score was plotted against those of percentage of brown projected area (Br:TProjArea). Function of the regression line ($y=0.25804x + 0.06497$) was calculated through the linear model Br:TProjArea~DS. The regression coefficient r^2 translate the relevance of the model.

Table III.2: Pearson correlation coefficients between experiments for morphology and resistance data variables (to be continued)

Exp#1

	Root morphology						Shoot morphology			Plant morphology	
	TProjArea	RootDia	TRootL	NLatRoot	LatRootL	RootB	ShootB	ShootL	TB	RootB:TB	
	1	-0.13 *	0.94 ***	0.75 ***	0.67 ***	0.83 ***	0.49 ***	ns	0.75 ***	0.39 ***	
Root morphology		1	-0.44 ***	-0.38 ***	0.25 ***	-0.21 ***	-0.49 ***	ns	0.43 ***		
			1	0.81 ***	0.65 ***	0.5 ***	ns	0.67 ***	0.21 **		
				1	0.58 ***	0.46 ***	ns	0.61 ***	0.21 **		
					1	0.34 ***	0.15 *	0.45 ***	ns		
Shoot morphology					1	0.45 ***	-0.34 ***	0.82 ***	0.59 ***		
						1	0.43 ***	0.87 ***	-0.42 ***		
Plant morphology							1	ns	-0.74 ***		
								1	ns		
									1		

Exp#2

	Root morphology						Shoot morphology			Plant morphology	
	TProjArea	RootDia	TRootL	NLatRoot	LatRootL	RootB	ShootB	ShootL	TB	RootB:TB	
	1	ns	0.94 ***	0.61 ***	0.56 ***	0.89 ***	0.68 ***	0.19 **	0.78 ***	0.16 *	
Root morphology		1	-0.27 ***	ns	0.39 ***	0.15 *	-0.35 ***	0.23 ***	0.29 ***		
			1	0.68 ***	0.72 ***	0.60 ***	0.29 ***	0.67 ***	ns		
				1	0.41 ***	0.40 ***	ns	0.43 ***	ns		
					1	0.34 ***	0.28 ***	0.41 ***	0.15 *		
Shoot morphology					1	0.70 ***	ns	0.83 ***	0.28 ***		
						1	0.51 ***	0.98 ***	-0.46 ***		
Plant morphology							1	0.43 ***	-0.56 ***		
								1	-0.28 ***		
									1		

Variables are coded as described in Table 1. Pearson correlation p -values: $0 < \text{***} < 0.001 \leq \text{**} < 0.01 \leq \text{*} < 0.05 \leq \text{ns} < 1$. Level of correlation is coded with a colour gradation scale from dark red ($r = 1$) to white ($r = 0$) and from white to dark green ($r = 1$): dark green. "ns": non-significant correlation (p -value > 0.05).

Table III.2: Pearson correlation coefficients between experiments for morphology and resistance data variables (to be continued)

		Exp#2											
		Root morphology					Shoot morphology					Plant morphology	
		TProjArea	RootDia	TRootL	NLatRoot	LatRootL	RootB	ShootB	ShootL	TB	RootB:TB		
Root morphology	TProjArea	0.56 ***	ns	0.52 ***	0.48 ***	0.15 *	0.52 ***	0.29 ***	-0.13 *	0.37 ***	0.25 ***		
	RootDia	-0.18 **	0.67 ***	-0.40 ***	-0.29 ***	-0.20 **	ns	ns	-0.42 ***	ns	ns		
	TRootL	0.56 ***	-0.14 *	0.61 ***	0.54 ***	0.20 **	0.43 ***	0.28 ***	ns	0.34 ***	0.16 *		
	NLatRoot	0.56 ***	ns	0.60 ***	0.74 ***	ns	0.42 ***	0.35 ***	ns	0.39 ***	ns		
	LatRootL	0.32 ***	ns	0.35 ***	ns	0.38 ***	0.28 ***	ns	ns	0.16 *	0.20 **		
Shoot morphology	RootB	0.50 ***	0.45 ***	0.34 ***	0.35 ***	ns	0.60 ***	0.33 ***	-0.34 ***	0.43 ***	0.30 ***		
	ShootB	0.56 ***	ns	0.52 ***	0.29 ***	0.33 ***	0.56 ***	0.77 ***	0.44 ***	0.76 ***	-0.29 ***		
	ShootL	0.14 *	-0.38 ***	0.26 ***	ns	0.29 ***	ns	0.40 ***	0.94 ***	0.33 ***	-0.47 ***		
Plant morphology	TB	0.62 ***	0.28 ***	0.51 ***	0.38 ***	0.25 ***	0.68 ***	0.66 ***	ns	0.71 ***	ns		
	RootB:TB	ns	0.40 ***	ns	0.14 *	-0.25 ***	ns	-0.31 ***	-0.75 ***	-0.21 **	0.56 ***		

Exp#1

Exp#3

		Disease	
		RRI	DS
Disease	RRI	1	0.97 ***
	DS		1
	Br:TProjArea		0.70 ***

Exp#3

Variables are coded as described in Table 1. Pearson correlation *p-values* : 0 < '***' < 0.001 ≤ '***' < 0.01 ≤ '**' < 0.05 ≤ '*' < 0.1 ≤ 'ns' < 1. Level of correlation is coded with a colour gradation scale from dark red (*r* = 1) to white (*r* = 0) and from white to dark green (*r* = 1): dark green (*r* = 1): dark green. "ns": non-significant correlation (*p-value* > 0.05).

Table III.2: Pearson correlation coefficients between experiments for morphology and resistance data variables (end)

Exp#3

	Disease			
	RRI	DS	Br:TProjArea	
Root morphology	TProjArea	-0.26 ***	-0.25 ***	ns
	RootDia	0.16 *	0.15 *	0.26 ***
	TRootL	-0.29 ***	-0.28 ***	-0.18 **
	NLatRoot	-0.33 ***	-0.32 ***	-0.14 *
	LatRootL	ns	ns	ns
Shoot morphology	RootB	-0.18 **	-0.19 **	ns
	ShootB	ns	ns	ns
	ShootL	0.23 ***	0.25 ***	ns
Plant morphology	TB	ns	ns	ns
	RootB:TB	-0.17 **	-0.17 **	ns

Exp#3

	Disease			
	RRI	DS	Br:TProjArea	
Root morphology	TProjArea	-0.13 *	-0.13	ns
	RootDia	0.17 **	0.15 *	0.35 ***
	TRootL	-0.18 **	-0.17 **	ns
	NLatRoot	-0.25 ***	-0.24 ***	ns
	LatRootL	ns	ns	ns
Shoot morphology	RootB	ns	ns	0.18 **
	ShootB	ns	ns	0.25 ***
	ShootL	0.23 ***	0.25 ***	ns
Plant morphology	TB	ns	ns	0.24 ***
	RootB:TB	-0.15 *	-0.14 *	-0.11

Variables are coded as described in Table 1. Pearson correlation *p-values*: 0 < '***' < 0.001 ≤ '**' < 0.01 ≤ '*' < 0.05 ≤ 'ns' < 1. Level of correlation is coded with a colour gradation scale from dark red (*r* = 1) to white (*r* = 0) and from white to dark green (*r* = -1): dark green. "ns": non-significant correlation (*p-value* > 0.05).

3.2. Genetic analysis

3.2.1. Linkage disequilibrium, structure, and kinship in the 266-pea-line collection

From the imputed genotyping data of the collection obtained with the filtered 11,789 SNP markers, the LD decay was estimated to range from 0.034 to 0.115 cM, depending on the LG, and averaged to 0.055 over all the LGs (Additional file III.3), which is lower than LD estimated by Desgroux et al. (2016) on 175 of the 266 pea lines of the collection.

From the subset of 2,937 SNP markers defined at non-redundant positions, PCA analysis of the collection identified three first PCs that explained a total of 18.64 % of the genetic variation in the 266-pea-line collection (Figure III.2). The first PC contributed to 7.70 % of the variation and the second and third PCs to 6.75 % and 4.20 %, respectively. Pea lines were clustered on PCs 1 and 2 in two major groups. The first group of 115 lines included 85 pea lines of the 104-pea-line-Root collection of Bourion (in prep), the sixteen varieties, ten new sources of resistance to *A. euteiches*, and five parental lines of the pea-Aphanomyces collection (Desgroux et al. 2016). This group of 115 pea lines included seven lines common to both Root- and pea-Aphanomyces collections. The second group included all other pea lines of the pea-Aphanomyces collection as well as seven pea lines of the Root collection. The first three PCs were added to the GWA model as cofactors to take this population structure into account. From the same subset of 2,937 SNP markers, the Kinship matrix of genetic similarities revealed a moderate relatedness among pea lines of the collection ($0.5 < r^2 < 0.75$ for the major part of the lines), as shown in Desgroux et al. (2016) (Figure III.3). The kinship relatedness matrix was also added to the GWA model as a cofactor.

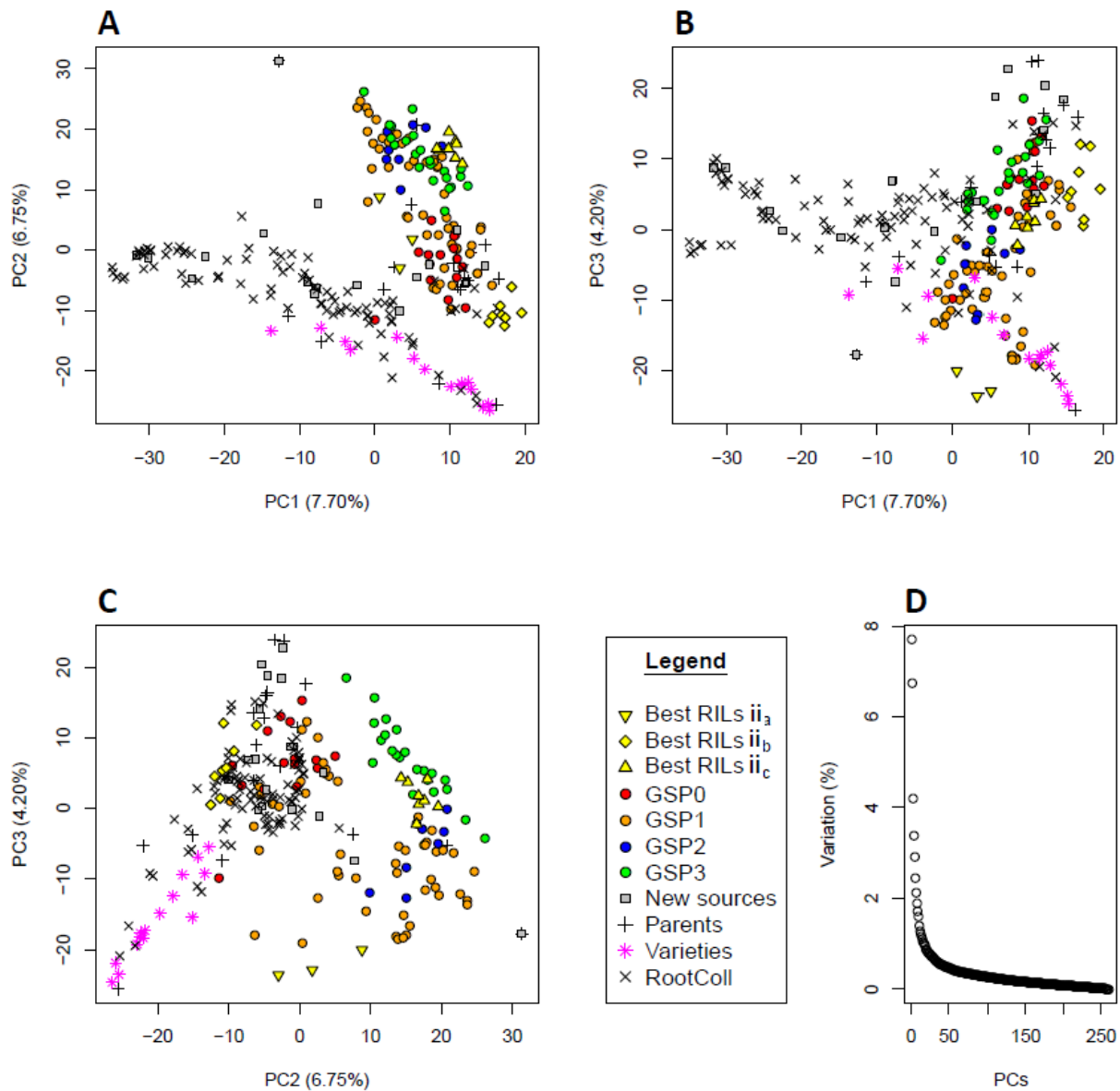


Figure III.2: Population structure of the 266-pea-line collection based on Principal Component Analysis (PCA)

PCA from GAPIT R package (Lipka et al. 2012) based on 2,937 markers. Pea lines of the 266-pea-line collection are represented on the first three principal components (A-C) that explain a total of 18.25 % of inertia (D). GSP0-3: lines from the GSP breeding program containing in their pedigree zero to three previously described major sources of resistance (Desgroux et al. 2016), respectively. Parents: Parental lines of GSP breeding programs and/or RIL populations. BRIL ii_a, ii_b, ii_c: Selected lines from the *Baccara* x *PI180693*, *Baccara* x *552* and *DSP* x *90-2131* RIL population, respectively; New sources: Selected accessions from a large INRA screening program; Varieties: Registered dry pea varieties; RootColl: pea lines of the root collection described in Bourion (in prep), without lines common to the pea-*Aphanomyces* collection (Desgroux et al. 2016). (D) Inertia contribution of each principal component (from PC1 to PC260).

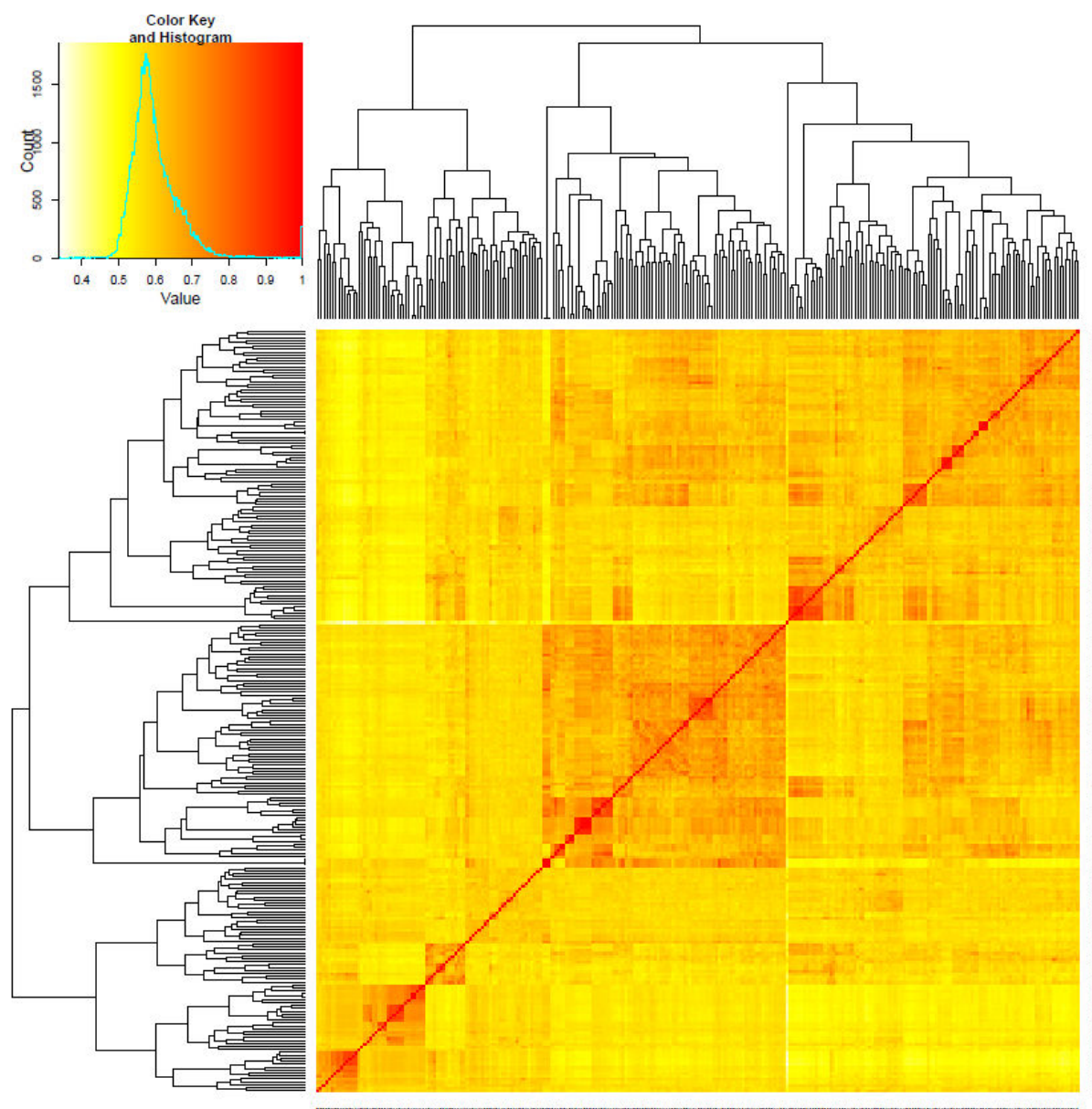


Figure III.3: Clustered heatmap of the Kinship matrix of the 266-pea-line collection

Kinship matrix from the GAPIT R package (Lipka et al. 2012) based on 2,937 SNP markers. Clustering is based on the UPGMA method. Colour gradient represent the degree of relationship between two given lines.

Table III.3: Number of markers detected by genome-wide association mapping for plant architecture and *Aphanomyces* resistance variables

Experiment (a)	Variable (b)	Number of markers (c)	Range of <i>p</i> -value (d)	Range of allelic effect (e)	% of phenotypic variance explained by			Unexplained variance (i)
					PCA (f)	Markers (g)	Kinship (h)	
Exp#1	TProjArea	8	2.74E-11 - 1.68E-05	0.74 - 1.94	29%	35%	0%	36%
	RootDia	0	-	-	35%	0%	30%	35%
	TRootL	1	3.29E-06	27.7	18%	11%	37%	34%
	NLatRoot	0	-	-	21%	0%	56%	23%
	LatRootL	2	5.63E-12 - 2.21E-06	0.37 - 0.70	1%	26%	22%	51%
	RootB	6	2.12E-08 - 1.97E-06	2.46 - 5.68	45%	23%	4%	28%
	ShootB	16	5.70E-19 - 4.05E-06	2.29 - 8.83	4%	73%	0%	23%
	ShootL	10	1.27E-33 - 2.01E-05	0.46 - 2.66	48%	40%	7%	5%
	TB	2	4.79E-06 - 1.08E-05	6.31 - 9.34	8%	20%	37%	35%
	RootB:TB	1	2.23E-07	0.03	68%	4%	14%	14%
Exp#2	TProjArea	0	-	-	3%	0%	76%	21%
	RootDia	14	5.71E-19 - 9.22E-06	0.009 - 0.029	47%	36%	6%	11%
	TRootL	3	1.19E-07 - 2.25E-05	28.89 - 33.82	4%	22%	61%	13%
	NLatRoot	1	4.96E-06	2.81	8%	10%	70%	12%
	LatRootL	1	2.66E-05	0.46	3%	5%	50%	42%
	RootB	4	6.25E-09 - 1.97E-05	2.54 - 6.74	10%	24%	34%	32%
	ShootB	12	2.37E-44 - 1.50E-05	5.42 - 29.59	3%	72%	0%	25%
	ShootL	5	2.14E-59 - 1.54E-05	0.94 - 7.96	53%	38%	7%	2%
	TB	7	9.04E-18 - 1.88E-05	7.33 - 28.19	2%	60%	10%	28%
	RootB:TB	2	1.60E-07 - 1.35E-05	0.01 - 0.02	23%	21%	32%	24%
Exp#3	RRI	1	2.78E-16	0.33	52%	14%	17%	18%
	DS	5	3.63E-22 - 2.21E-05	0.12 - 0.35	51%	23%	4%	22%
	Br:TProjArea	7	6.62E-11 - 1.57E-05	0.01 - 0.02	38%	28%	3%	31%

(a) Experiments are defined as in Table 1; (b) Variables are abbreviated as in Table 1; (c) number of markers used as cofactors at the optimal step of the multi-locus mixed model (MLMM) analysis; (d) Range of *p*-values of the significant markers, significance threshold is *p*-value < 3.4E-05 as described in Methods section; (e) Range of allelic effects of the significant markers in absolute values; Percentage of phenotypic variance explained by ; (f) the principal component analysis (PCA) matrix of the pea *Aphanomyces* collection structure, (g) all cofactor markers together, (h) the Kinship relatedness matrix among lines of the pea *Aphanomyces* collection, (i) the unexplained variance qualified as "missing heritability"

3.2.2. Genome wide association results and comparative mapping

A total of 89 SNP markers, included in 85 CIs on the seven LGs, were significantly associated with phenotypic variation observed in the collection for the 13 plant architecture and resistance variables scored in this study (Table III.3, Table III.4, Figure III.4 and Additional file III.4). The average size of the CIs defined around the significant SNPs was 0.89 cM, which is lower than the one previously observed in Desgroux et al. (2016) (average size = 1.03 cM). Twelve CIs were detected for at least two variables from the same or different experiments. Zero to 16 markers (ShootB variable in Exp#1) were significantly associated to each variable with a *p-value* that ranged from 2.14E-59 to 2.66E-05, depending on the marker. The most significant *p-values* were detected for ShootL and ShootB traits in both experiments. For each model, the set of markers retained as cofactors explained zero to 73% of the phenotypic variation, depending on the variable and experiment. Looking at disease variables, the most significant SNP marker was *Ps115429* detected with *p-values* that ranged from 6.62E-11 to 3.63E-22, depending on the variable.

A total of 76 CIs, were identified for all the plant architecture variables scored in the two experiments, including 11 CIs detected for at least two root, shoot or overall plant architecture variables. (i) Thirty three markers were detected for RSA traits (LatRootL, NLatRoot, RootB, RootDia, TRootL, TProjArea) located in 33 CIs, among which one was consistently detected for three root variables close to the *ELF3* gene region and the *PEAFLO* gene region on LGIII (involved in circadian clock function and flower morphogenesis, respectively). (ii) Thirty eight markers were detected for shoot architecture traits (ShootB, ShootL), located in 34 CIs, among which four were consistently detected for one or the two shoot variables in Exp#1 and/or Exp#2. These four consistent CIs were close to the *ELF3* and *PEAFLO* gene regions and to the *LE* gene region on the LGIII (involved in plant height), on the LGV and close to the *CRY* gene on the LGVII (involved in plant growth regulation); (iii) seven markers covering seven CIs for overall plant architecture traits (TB, RootB:TB), among which the *LE* gene region consistently detected for three variables scored in Exp#1 and/or Exp#2. The most consistent CI was located at the *LE* gene region detected with the highest significance (*p-value* > 10E-59 for ShootL), in which two markers were associated with eight variables from the two experiments. All of the 11 consistent CIs were detected for biomass variables on different parts of the plants, except two CIs on LGIII (31.2-31.6 cM) and LGVII (43.3-45.2 cM). Two and one of the 11 CIs were specific to shoot and root variables, respectively. Among the 11 CIs, four were associated with variables from the two experiments and only

Table III.4: Genetic markers associated with architecture and *Aphanomyces* resistance variables detected by genome-wide association mapping (to be continued)

LG (a)	Confidence interval (cM) (b,c)	Position (cM) (b)	SNP (d)	MAF (e)	Variable (f)	p-value (g)	Intercept (h)	Allelic effect (i)	SE (j)
I	1.2 - 2.5	2.2	PsCam017274_10719_335	0.43	ShootB-13_Exp#1	9.31E-07	25.07	2.51	0.50
	26.8 - 28.1	28.1	PsCam003814_2915_740	0.11	TRootL-13_Exp#2	2.52E-05	471.49	-33.55	7.81
	36.4 - 36.7	36.4	PsCam049439_32075_1874	0.38	ShootB-13_Exp#2	1.14E-08	68.99	5.64	0.95
	41.4 - 41.9	41.9	PsCam006242_4658_70	0.17	ShootB-13_Exp#2	1.50E-05	68.99	7.60	1.72
	46.1 - 47.3	47.3	PsCam003504_2708_311	0.18	CC_DS_RB84_Exp#3	2.04E-05	3.29	-0.12	0.03
	48.3 - 50.8	48.3	PsCam049210_31851_713	0.11	ShootB-13_Exp#2	1.16E-13	68.99	-16.77	2.13
	54.8 - 55.9	55.1	PsCam036755_21882_1944	0.39	RootB-13_Exp#1	1.55E-06	29.27	2.46	0.50
					TB-13_Exp#1	4.79E-06	69.55	6.31	1.35
	62.0 - 63.3	62	PsCam039096_24074_1470	0.15	TProjArea-13_Exp#1	1.64E-07	5.64	-0.93	0.17
	64.8 - 65.4	65	PsCam028096_16571_362	0.46	RootDia-13_Exp#2	6.24E-10	0.624	-0.012	0.002
		65.2	PsCam044095_28021_2731	0.18	ShootB-13_Exp#1	4.56E-07	25.07	3.02	0.58
	69.0 - 69.6	69.6	PsCam042750_26795_95	0.1	ShootB-13_Exp#1	9.37E-09	25.07	-4.89	0.82
	80.9 - 81.7	81.2	Ps105642	0.28	RootDia-13_Exp#2	2.42E-08	0.624	-0.013	0.002
	II	9.9 - 10.4	10.4	PsCam039960_24795_161	0.15	ShootL-13_Exp#2	2.56E-06	8.55	-1.84
12.9 - 14.8		13.8	PsCam001738_1431_519	0.4	CC_DS_RB84_Exp#3	3.09E-06	3.29	0.12	0.02
22.2-22.3		22.2	PsCam049309_31947_1306	0.17	ShootL-13_Exp#1	1.81E-06	6.60	-0.59	0.12
23.7 - 23.7		23.7	PsCam004807_3625_131	0.08	LatRootL-13_Exp#1	2.21E-06	1.39	0.37	0.08
38.4 - 39.0		38.8	PsCam027690_16273_464	0.26	RootDia-13_Exp#2	9.22E-06	0.624	0.010	0.002
40.5 - 41.0		40.7	PsCam046463_29895_916	0.16	TB-13_Exp#2	1.63E-09	61.12	16.77	2.67
					ShootB-13_Exp#2	1.63E-07	68.99	9.07	1.68
40.7 - 42.0		41.7	Ps163697	0.4	RootDia-13_Exp#2	9.71E-07	0.624	0.010	0.002
45.5 - 45.5		45.5	PsCam029325_17487_1726	0.18	RootDia-13_Exp#2	1.82E-06	0.624	-0.013	0.002

Description of genetic markers detected by genome-wide association (GWA) study on the 266-pea-line collection, using the Multi-Locus Mixed Model (MLMM) method (Segura et al 2012), for architecture and resistance to *A. euteiches* variables. (a) LG: Linkage group; (b) Genetic positions are estimated on the genetic marker map resulting from the projection of the 953 SNP (Boutet et al submitted) markers onto the THMap (Desgroux et al 2016); (c) A confidence interval (CI) is defined with all markers in linkage disequilibrium ($r^2 > 0.2$) with the associated marker; (d) PsCam- markers are SNP markers developed by Tayeh et al (2015), Ps1- markers are SNP markers developed by Boutet et al (submitted); (e) Minor allele frequency (MAF) of each detected marker in the 266-pea-line collection; (f) Variables are coded as described in Additional File 1; (g) significant p-value of the marker-trait association < 2.6E-05; (h) mean value of the trait corrected with the effect of all markers included in the MLMM analysis; (i) allelic effect of the marker on the variable, sign depends on the allele code; (j) standard error in the 266-pea-line collection.

Table III.4: Genetic markers associated with architecture and Aphanomyces resistance variables detected by genome-wide association mapping (to be continued)

LG (a)	Confidence interval (cM) (b,c)		SNP (d)	MAF (e)	Variable (f)	p-value (g)	Allelic effect			
	(b)	(c)					Intercept (h)	(i)	SE (j)	
II	46.3 - 46.3	46.3	<i>PsCam035741_20914_1106</i>	0.18	ShootL-13_Exp#1	4.51E-06	6.60	-0.63	0.13	
	48.4 - 49.6	49.5	<i>PsCam017051_10599_277</i>	0.18	RootDia-13_Exp#2	5.71E-19	0.624	-0.029	0.003	
	52.2 - 52.2	52.2	<i>PsCam044104_28029_2491</i>	0.23	CC_Br:TProjArea_RB84_Exp#3	1.11E-05	0.480	0.014	0.003	
	53.8 - 55.0	54.1	<i>PsCam049658_32282_3329</i>	0.12	RootDia-13_Exp#2	7.14E-06	0.624	0.015	0.003	
	61.3 - 63.1	62.5	<i>PsCam038222_23267_331</i>	0.39	ShootB-13_Exp#2	7.44E-06	68.99	-5.42	1.18	
	71.2 - 72.7	72.4	<i>PsCam012901_8704_1349</i>	0.46	CC_Br:TProjArea_RB84_Exp#3	1.57E-05	0.480	-0.011	0.002	
	112.1 - 112.6	112.6	<i>PsCam028844_17107_289</i>	0.18	ShootB-13_Exp#1	6.26E-10	25.07	8.83	1.37	
	112.8 - 114.1	112.9	<i>PsCam045947_29519_1131</i>	0.18	ShootB-13_Exp#1	4.05E-06	25.07	6.57	1.39	
	III	1.1 - 2.2	1.4	<i>PsCam023220_13093_2290</i>	0.46	ShootB-13_Exp#1	2.78E-08	25.07	-2.63	0.46
		22.0 - 22.0	22	<i>PsCam058589_38986_524</i>	0.05	ShootL-13_Exp#1	1.11E-06	6.60	-1.11	0.22
26.7 - 29.0		27.3	<i>PsCam038460_23494_691</i>	0.15	TB-13_Exp#2	9.79E-08	61.12	-17.56	3.20	
		27.8	<i>Ps160242</i>	0.15	ShootB-13_Exp#2	9.12E-06	68.99	-9.11	2.01	
31.2 - 31.6		31.2	<i>PsCam007407_5452_4389</i>	0.06	ShootB-13_Exp#1	7.38E-13	25.07	-6.20	0.81	
					LatRootL-13_Exp#1	5.63E-12	1.39	0.70	0.10	
					TProjArea-13_Exp#1	3.01E-10	5.64	1.94	0.30	
					TRootL-13_Exp#1	3.29E-06	71.81	27.74	5.82	
36.3 - 36.4		36.4	<i>PsCam039622_24530_225</i>	0.1	ShootB-13_Exp#1	7.27E-09	25.07	4.69	0.78	
40.4 - 40.6		40.4	<i>PsCam051159_33688_656</i>	0.31	ShootB-13_Exp#1	3.82E-06	25.07	2.29	0.48	
43.7 - 45.7	44.5	<i>PsCam045339_29025_1722</i>	0.48	TB-13_Exp#2	1.64E-05	61.12	7.58	1.72		
47.8 - 48.0	48	<i>PsCam000645_563_953</i>	0.29	ShootL-13_Exp#2	1.54E-05	8.56	0.99	0.22		
58.0 - 59.0	59	<i>PsCam003194_2488_239</i>	0.14	TProjArea-13_Exp#1	2.74E-11	5.64	1.78	0.26		
59.0 - 60.5	60.5	<i>PsCam045034_28765_3186</i>	0.11	RootB-13_Exp#1	2.12E-08	29.27	5.68	0.98		

Description of genetic markers detected by genome-wide association (GWA) study on the 266-pea-line collection, using the Multi-Locus Mixed Model (MLMM) method (Segura et al 2012), for architecture and resistance to *A. euteiches* variables. (a) LG: Linkage group; (b) Genetic positions are estimated on the genetic marker map resulting from the projection of the 953 SNP (Boutet et al submitted) markers onto the THMap (Desgroux et al 2016); (c) A confidence interval (CI) is defined with all markers in linkage disequilibrium ($r^2 > 0.2$) with the associated marker; (d) PsCam- markers are SNP markers developed by Tayeh et al (2015), Ps1- markers are SNP markers developed by Boutet et al (submitted); (e) Minor allele frequency (MAF) of each detected marker in the 266-pea-line collection; (f) Variables are coded as described in Additional File 1; (g) significant p-value of the marker-trait association $< 2.6E-05$; (h) mean value of the trait corrected with the effect of all markers included in the MLM analysis; (i) allelic effect of the marker on the variable, sign depends on the allele code; (j) standard error in the 266-pea-line collection.

Table III.4: Genetic markers associated with architecture and *Aphanomyces* resistance variables detected by genome-wide association mapping (to be continued)

LG (a)	Confidence interval (cM) (b,c)	Position (cM) (b)	SNP (d)	MAF (e)	Variable (f)	p-value (g)	Intercept (h)	Allelic effect (i)	SE (j)
III	75.7 - 76.5	76.4	<i>PsCam001691_1400_4718</i>	0.2	RootDia-13_Exp#2	1.33E-11	0.624	0.014	0.002
	89.4 - 90.0	89.4	<i>PsCam057723_38286_2075</i>	0.29	CC_Br:TProjArea_RB84_Exp#3	1.06E-06	0.480	0.013	0.002
	110.0 - 110.0	110	<i>PsCam049362_31999_267</i>	0.06	ShootL-13_Exp#1	1.56E-12	6.60	-1.41	0.19
					RootB-13_Exp#2	6.25E-09	40.38	-6.74	1.12
	131.0 - 133.0	132	<i>PsCam035617_20792_637</i>	0.29	ShootB-13_Exp#1	5.70E-19	25.07	7.60	0.78
		132.1	<i>PsCam037922_22979_691</i>	0.29	ShootL-13_Exp#2	2.14E-59	8.56	7.96	0.36
					ShootB-13_Exp#2	2.37E-44	68.99	29.59	1.70
					ShootL-13_Exp#1	1.27E-33	6.60	2.66	0.19
					TB-13_Exp#2	9.04E-18	61.12	28.19	3.04
					RootB-13_Exp#2	1.85E-08	40.38	6.52	1.12
IV	0.0 - 1.4	0.7	<i>PsCam037555_22634_340</i>	0.46	RootB-13_Exp#1	1.14E-06	29.27	-2.69	0.54
	2.0 - 2.0	2	<i>PsCam029391_17536_1186</i>	0.41	TRootL-13_Exp#2	1.19E-07	471.49	33.82	6.20
	14.5 - 16.2	16.2	<i>PsCam005487_4151_529</i>	0.48	RootDia-13_Exp#2	4.90E-07	0.624	0.010	0.002
	19.6 - 19.6	19.6	<i>PsCam006615_4905_1230</i>	0.06	ShootL-13_Exp#1	2.14E-08	6.60	1.17	0.20
	34.2 - 34.2	34.2	<i>PsCam039041_24025_293</i>	0.09	ShootB-13_Exp#1	1.51E-14	25.07	6.38	0.78
	50.0 - 50.0	50	<i>PsCam014592_9851_995</i>	0.46	ShootB-13_Exp#1	1.26E-06	25.07	2.37	0.48
	60.8 - 61.4	60.8	<i>PsCam045099_28822_422</i>	0.39	TB-13_Exp#2	1.88E-05	61.12	7.33	1.68
					LatRootL-13_Exp#2	2.36E-05	7.57	0.46	0.11
	79.4 - 79.7	79.4	<i>PsCam041012_25616_571</i>	0.34	ShootL-13_Exp#2	2.46E-08	8.56	-1.09	0.19
					CC_Br:TProjArea_RB84_Exp#3	3.90E-06	0.480	-0.013	0.003

Description of genetic markers detected by genome-wide association (GWA) study on the 266-pea-line collection, using the Multi-Locus Mixed Model (MLMM) method (Segura et al 2012), for architecture and resistance to *A. euteiches* variables. (a) LG: Linkage group; (b) Genetic positions are estimated on the genetic marker map resulting from the projection of the 953 SNP (Boutet et al submitted) markers onto the THMap (Desgroux et al 2016); (c) A confidence interval (CI) is defined with all markers in linkage disequilibrium ($r^2 > 0.2$) with the associated marker; (d) PsCam- markers are SNP markers developed by Tayeh et al (2015), Ps1- markers are SNP markers developed by Boutet et al (submitted); (e) Minor allele frequency (MAF) of each detected marker in the 266-pea-line collection; (f) Variables are coded as described in Additional File 1; (g) significant p-value of the marker-trait association $< 2.6E-05$; (h) mean value of the trait corrected with the effect of all markers included in the MLM analysis; (i) allelic effect of the marker on the variable, sign depends on the allele code; (j) standard error in the 266-pea-line collection.

Table III.4: Genetic markers associated with architecture and Aphanomyces resistance variables detected by genome-wide association mapping (to be continued)

LG (a)	Confidence interval (cM) (b,c)	Position (cM) (b)	SNP (d)	MAF (e)	Variable (f)	p-value (g)	Intercept (h)	Allelic effect (i)	SE (j)
VI	86.3 - 87.2	86.3	PsCam048995_31641_4404	0.25	RootDia-13_Exp#2	1.99E-06	0.624	0.010	0.002
	87.0 - 87.2	87	PsCam045027_28758_835	0.25	TRootL-13_Exp#2	1.22E-05	472.49	-28.89	6.47
	91.2 - 92.3	91.6	Ps160276	0.24	CC_DS_RB84_Exp#3	2.21E-05	3.29	0.12	0.03
	95.8 - 96.2	96.2	PsCam011015_7459_191	0.13	ShootB-13_Exp#2	1.48E-06	68.99	6.80	1.38
V	12.1 - 13.8	12.9	PsCam037366_22454_1257	0.35	RootDia-13_Exp#2	1.89E-06	0.624	-0.009	0.002
	18.6 - 18.6	18.6	PsCam001152_972_1401	0.06	ShootL-13_Exp#1	3.74E-07	6.60	1.00	0.19
	28.8 - 29.8	29.8	PsCam001634_1352_1418	0.22	CC_Br:TProjArea_RB84_Exp#3	2.80E-06	25.07	4.68	0.97
	31.6 - 31.6	31.6	PsCam038133_23182_1157	0.27	ShootB-13_Exp#2	9.77E-07	0.480	0.014	0.003
	31.6 - 33.5	32.1	PsCam050575_33149_2320	0.36	RootB-13_Exp#2	1.80E-07	68.99	6.79	1.26
	33.9 - 34.9	34.7	PsCam003629_2792_434	0.36	ShootB-13_Exp#1	1.97E-05	40.38	3.14	0.72
	36.9 - 38.6	37.8	PsCam035913_21070_488	0.35	ShootB-13_Exp#1	7.23E-16	25.07	-4.09	0.47
	39.7 - 41.5	41.1	PsCam037194_22296_2394	0.34	TB-13_Exp#2	1.99E-05	61.12	8.07	1.85
	52.9 - 53.2	52.9	PsCam035879_21036_718	0.36	CC_DS_RB84_Exp#3	6.93E-07	3.29	0.12	0.02
	72.6 - 74.7	73.3	PsCam042707_26754_2818	0.28	RootB-13_Exp#1	2.84E-07	29.27	-3.05	0.58
	75.5 - 75.8	75.7	PsCam036114_21263_3048	0.34	ShootL-13_Exp#1	2.01E-05	6.60	0.46	0.11
	81.6 - 82.7	82.5	Ps100801	0.45	RootDia-13_Exp#2	8.67E-07	0.624	0.011	0.002
	93.9 - 94.3	94.3	PsCam040651_25349_285	0.09	TProjArea-13_Exp#1	1.27E-05	5.64	-0.74	0.17
	94.5 - 94.9	94.7	PsCam035619_20794_2598	0.45	NLatRoot-13_Exp#2	4.96E-06	51.69	-2.81	0.60
	96.1 - 97.3	96.9	Ps105087	0.26	ShootB-13_Exp#2	6.97E-07	68.99	-9.17	1.80
						1.99E-06	5.64	-0.76	0.16
						9.20E-10	5.64	-1.25	0.20

Description of genetic markers detected by genome-wide association (GWA) study on the 266-pea-line collection, using the Multi-Locus Mixed Model (MLMM) method (Segura et al 2012), for architecture and resistance to *A. euteiches* variables. (a) LG; Linkage group; (b) Genetic positions are estimated on the genetic marker map resulting from the projection of the 953 SNP (Boutet et al submitted) markers onto the THMap (Desgroux et al 2016); (c) A confidence interval (CI) is defined with all markers in linkage disequilibrium ($r^2 > 0.2$) with the associated marker; (d) PsCam- markers are SNP markers developed by Tayeh et al (2015), Ps1- markers are SNP markers developed by Boutet et al (submitted); (e) Minor allele frequency (MAF) of each detected marker in the 266-pea-line collection; (f) Variables are coded as described in Additional File 1; (g) significant p-value of the marker-trait association $< 2.6E-05$; (h) mean value of the trait corrected with the effect of all markers included in the MLMM analysis; (i) allelic effect of the marker on the variable, sign depends on the allele code; (j) standard error in the 266-pea-line collection.

Table III.4: Genetic markers associated with architecture and *Aphanomyces* resistance variables detected by genome-wide association mapping (End)

LG (a)	Confidence interval (cM) (b,c)	Position (cM) (b)	SNP (d)	MAF (e)	Variable (f)	p-value (g)	Intercept (h)	Allelic effect (i)	SE (j)	
VI	16.5 - 17.8	17.3	<i>PsCam037834_22896_1142</i>	0.46	RootB:TB-13_Exp#2	1.35E-05	0.331	-0.011	0.002	
	19.3 - 21.0	20.9	<i>PsCam027216_15889_435</i>	0.4	CC_Br:TProjArea_RB84_Exp#3	1.85E-06	0.480	0.011	0.002	
	23.6 - 25.6	25.1	<i>PsCam059830_39901_263</i>	0.31	TProjArea-13_Exp#1	1.05E-06	5.64	1.78	0.26	
	28.0 - 30.4	29.9	<i>PsCam001371_1144_813</i>	0.27	RootB-13_Exp#1	1.92E-06	29.27	2.56	0.52	
	70.9 - 71.9	71.9	<i>PsCam004564_3420_65</i>	0.11	ShootB-13_Exp#2	3.74E-07	68.99	9.64	1.84	
					TB-13_Exp#2	6.33E-06	61.12	13.75	2.98	
	82.7-83.7	83.7	<i>PsCam017063_10609_1876</i>	0.37	ShootB-13_Exp#1	2.68E-08	25.07	2.83	0.49	
	VII	34.3 - 34.3	34.3	<i>PsCam034173_19641_100</i>	0.23	RootDia-13_Exp#2	2.60E-08	0.624	0.012	0.002
		43.3 - 45.2	43.6	<i>PsCam036149_21297_2851</i>	0.23	ShootL-13_Exp#1	5.14E-06	6.60	0.56	0.12
			44	<i>PsCam044214_28123_1477</i>	0.25	ShootL-13_Exp#2	1.13E-05	8.56	0.94	0.21
48.5 - 48.5		48.5	<i>PsCam038255_23296_791</i>	0.24	ShootB-13_Exp#2	9.85E-06				
65.9 - 66.8		66.2	<i>PsCam001741_1434_1146</i>	0.48	RootB-13_Exp#1	1.97E-06	29.27	2.65	0.54	
66.8 - 67.0		67	<i>PsCam055974_36869_2750</i>	0.07	ShootL-13_Exp#1	1.74E-05	6.60	-0.80	0.18	
70.8 - 71.5		71.3	<i>Ps115429</i>	0.25	CC_DS_RB84_Exp#3	3.63E-22	3.29	-0.35	0.03	
					CC_RRI_RB84_Exp#3	2.78E-16	3.5	-0.33	0.04	
					CC_Br:TProjArea_RB84_Exp#3	6.62E-11	0.480	-0.024	0.003	
					TProjArea-13_Exp#1	1.68E-05	5.64	0.85	0.19	
74.4 - 74.5	74.45	<i>PsCam034936_20271_288</i>	0.17	TB-13_Exp#1	1.08E-05	69.55	9.34	2.08		
87.9 - 87.9	87.9	<i>PsCam030019_17852_150</i>	0.26	RootDia-13_Exp#2	3.55E-07	0.624	0.012	0.002		
89.9 - 91.2	91.2	<i>PsCam021385_12015_245</i>	0.47	RootB-13_Exp#2	2.47E-05	40.38	2.54	0.59		
98.4 - 100.7	99.9	<i>PsCam038054_23106_222</i>	0.26	ShootB-13_Exp#1	1.19E-10	25.07	-4.04	0.60		

Description of genetic markers detected by genome-wide association (GWA) study on the 266-pea-line collection, using the Multi-Locus Mixed Model (MLMM) method (Segura et al 2012), for architecture and resistance to *A. euteiches* variables. (a) LG; (b) Linkage group; (c) Genetic positions are estimated on the genetic marker map resulting from the projection of the 953 SNP (Boutet et al submitted) markers onto the THMap (Desgroux et al 2016); (d) A confidence interval (CI) is defined with all markers in linkage disequilibrium ($r^2 > 0.2$) with the associated marker; (e) PsCam- markers are SNP markers developed by Tayeh et al (2015), Ps1- markers are SNP markers developed by Boutet et al (submitted); (f) Minor allele frequency (MAF) of each detected marker in the 266-pea-line collection; (g) Variables are coded as described in Additional File 1; (h) significant p-value of the marker-trait association < 2.6E-05; (i) mean value of the trait corrected with the effect of all markers included in the MLMM analysis; (j) allelic effect of the marker on the variable, sign depends on the allele code; (k) standard error in the 266-pea-line collection.

two (LGIII 131-133 cM and LGVII, 43.3-45.2 cM) were detected from the same variable(s) (Shoot B, RootB:TB and/or ShootL) in the two experiments. Among the 76 CIs associated with plant architecture variables, 50 overlapped with QTL detected for plant architecture by linkage analysis (LA) in Bourion et al. (2010), according to the projection conducted in this study. However, few overlaps were found for the same type of variables. One, three, five and four GWA-QTL detected for shoot variables overlap with LA-QTL detected for shoot variables on LGI, LGII, LGIII and LGVII, respectively. One, five, eight, three, two and four GWA-QTL detected for root variables overlapped with LA-QTL detected for root traits on LGI, LGII, LGIII, LGIV, LGV and LGVII, respectively. One GWA-QTL detected for TB overlapped with LA-QTL detected for BGB :TB on LGVII.

Eleven SNP markers, included in 11 CIs, were detected for resistance to *A. euteiches* on roots (RRI, DS, Br:TProjArea). One marker on LGVII (*Ps115429*, LGVII, 71 cM) was highly significantly detected for the three disease variables, and overlapped with the most consistent QTL for resistance to *Aphanomyces*, *Ae-Ps7.6*, detected by Hamon et al. (2013). The SNP *Ps115429* also mapped close to three CI identified in Desgroux et al. (2016) for resistance to *Aphanomyces* in controlled conditions (RB84 strain) and field RRI variables on LGVII. The other significant SNPs were localized in other consistent (*Ae-Ps2.2*, *Ae-Ps5.1*) or less consistent (*Ae-Ps1.1*, *Ae-Ps2.1*, *Ae-Ps3.2*, *Ae-Ps6.1*) genomic regions identified by Hamon et al. (2013) or in regions not previously identified by linkage analysis.

The projection of QTL previously detected for plant architecture (Bourion et al. 2010) onto the THMap of Desgroux et al. (2016) identified a total of 11 overlaps of large intervals with *Aphanomyces* resistance QTL. Even if correlations were highlighted between LSMeans of disease variables and some RSA traits, a single marker over the 89 SNPs detected for all traits was associated with both *A. euteiches* resistance (RRI, DS, Br:TProjArea) and RSA (TProjArea-13_8d) traits. At this marker, allelic effects for RSA and resistance were opposite, *i.e.* alleles contributing to resistance (smaller disease score) also contributed to a higher root projected area.

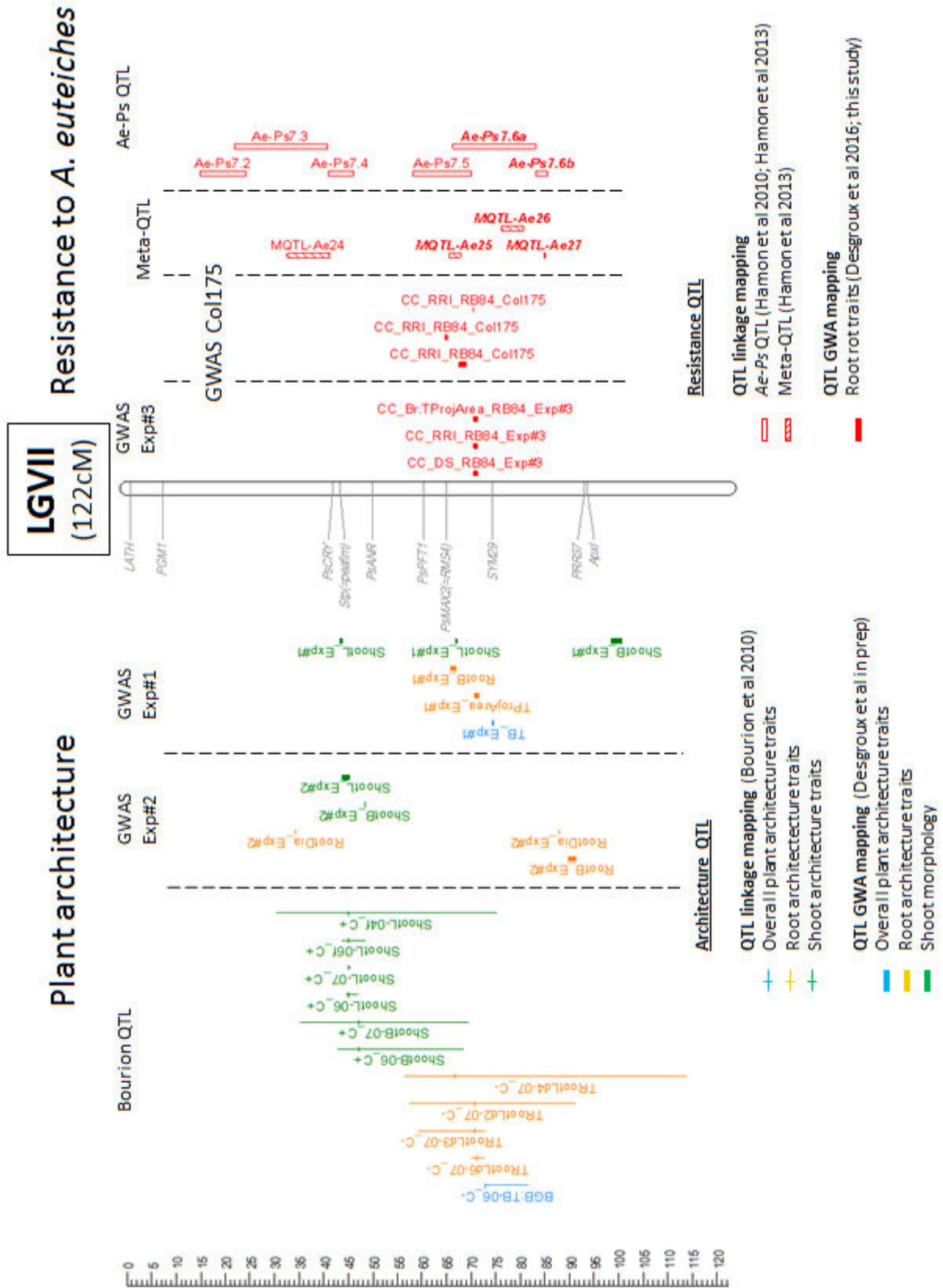


Figure III.4: Comparative genetic map of genome-wide association (GWA) and previously detected linkage quantitative trait loci (QTL) for resistance to *Aphanomyces euteiches* and plant architecture for linkage group VII

Figure on previous page

Legend on next page

The comparative genetic map was constructed from the projection of the 953 markers from Boutet et al. (2016) onto the consensus THMap from Desgroux et al. (2016). Only linkage group (LG) VII is shown. Its size is indicated in cM Haldane. Shoot-architecture-, Root-architecture-, Overall-plant-architecture and Resistance-associated markers and QTL are indicated in green, orange, blue and red, respectively. **To the right from each LG:** Confidence intervals (CIs) around significant Resistance-associated markers, based on linkage disequilibrium (LD) value $r^2 > 0.2$, identified in this study by GWA and name of the trait are indicated (GWAS Exp#3); CIs around significant controlled-conditions-Resistance-associated markers, identified by GWA in Desgroux et al. (2016), with the RB84 strain (GWAS Col175); Projected Meta-QTL (Meta-) and QTL (Ae-Ps QTL) described in Hamon et al. (2011) and Hamon et al. (2013), hatched bars represent Meta-QTL, while blank bars represent initial QTL before meta-analysis Hamon et al. (2011) and Hamon et al. (2013). The main *Aphanomyces* resistance QTL and Meta-QTL names are in bold italic. **To the left from each LG:** Genomic positions of cloned pea genes are indicated in grey; CIs around architecture-associated markers identified in this study (GWAS Exp#1 and GWAS Exp#2); Projected QTL for root, shoot and plant architecture traits described in Bourion et al. (2010) (Bourion QTL).

4. Discussion

This study describes the relationships between phenotypes and genetic determinants of plant (root and shoot) architecture and resistance to *A. euteiches* in pea using a 266-pea-line collection. It explored genetic diversity for both resistance to *A. euteiches* and plant architecture through GWA study.

4.1. A wide genetic diversity for plant architecture and resistance to *A. euteiches* in the 266-pea-line collection

In two experiments conducted in different controlled conditions free of pathogen, we confirmed the wide plant architecture diversity in *Pisum sativum*, especially on the root system, as identified since Mendel's works (1822-1884). The first study on root biomass and RSA diversity in the pea species reported a great diversity based on a collection of 330 accessions (McPhee 2005). Even if our collection was smaller (266 accessions), we confirmed this existing diversity in pea with an even wider range of values for some traits (ShootB, RootB, TRootL) at the same plant stage (14 days post sowing). Bourion et al. (2010) identified a genetic diversity of RSA traits among seven pea lines from the 9-leaf stage to the beginning of seed filling. We also demonstrated that genetic diversity of pea plant architecture was observable at earlier stages (8 days after sowing). For each plant architecture trait analysed, data between the two stages (8 and 14 days post sowing) were highly correlated ($r > 0.56$). Thus, we may hypothesize that plant architecture is set up very early in development process. As *A. euteiches* infection occurs at seedling stage, if relationship exists between RSA and resistance, it could establish from the first development stages.

In the 266-pea-line collection, we confirmed the large diversity of responses to *A. euteiches* and the high levels of resistance showed by several lines, among which *AeD99OSW-45-8-7*, as described previously (Desgroux et al. 2016). No additional pea resistant lines were identified with moderate or high levels of resistance in the Root collection.

4.2. Root growth and architecture traits correlated with resistance to *A. euteiches* in pea

In this study, the plant architecture traits measured were related to growth and architecture of root (TProjArea, TRootL, NLatRoot, LatRootL, RootB, RootDia), shoot (ShootB, ShootL) and overall (TB, RootB:TB) plant parts. As a result, most of these traits were highly significantly and positively correlated between each other, except the RootDia which was significantly negatively correlated with several root and shoot traits. Most of the RSA traits were negatively correlated with *Aphanomyces* root rot severity traits, *i.e.* resistance is correlated with higher root growth and architecture (root length, number, biomass). These correlations were most observed at the earliest plant stage studied (8 days post sowing). Among the most resistant pea lines, *A111* (*AeD99OSW-58-10-5*) and *A104* (*AeD99OSW-50-2-5*) had high root system bushiness (longer root system and higher number of lateral roots). However, some most resistant pea lines had moderate root system bushiness as *AeD99OSW-45-8-7*. In similar studies on the bean-*Fusarium* root rot pathosystem, a larger number of lateral roots was correlated to an increase of resistance (Kraft and Boge 2001; Román-Avilés et al. 2004) demonstrated on 12 pea lines that large-rooted pea lines (*i.e.* large total root length) were less susceptible to *Fusarium solani* in infested fields. Dias et al. (2004) demonstrated in a melon recombinant population that a wide, long and branched root system was related to an increased tolerance to vine collapse. In our study, the average root diameter was higher in pea lines susceptible to *A. euteiches*. Inconsistently, Snapp et al. (2003), demonstrated on 20 bean lines that roots with a larger diameter were more resistant to *Fusarium* root rot resistance in bean, but collection size may bias correlations. We could also conclude to this correlation in our study if we would have used a sub-set of pea lines from our collection. Interestingly, in aerial plant/pathogen pathosystems, the density of shoot plant parts was more often correlated with higher susceptibility (Calonnec et al. 2012). For example, Richard et al. (2013) showed that the leaf area index was positively correlated with leaf wetness duration after rainfall and dew. Longer leaf wetness duration was favourable to *Didymella pinodes* infection. Thus, we may hypothesize that RSA/root pathogens and aerial architecture/aerial epidemics have interactions.

4.3. Multiple genetic determinants of plant architecture poorly closely-linked to *Aphanomyces* resistance loci

GWA study confirmed the great number of genetic factors that control root and shoot architecture traits. Even if shoot length and shoot biomass were mainly explained by the *Le* gene on LGIII and the *Cry* gene on LGVII, four to fourteen other genomic regions contributed also to trait variations. A total of 74 CIs, not organ- nor experiment-specific, was actually associated with 13 root, shoot and overall plant variables scored in this study. Although large previous QTL CIs suggested common loci controlling both RSA and *Aphanomyces* resistance, small CIs identified in this study thanks to the resolution of GWA invalidated this hypothesis of common genetic factors. Only one CI was associated with both resistance and RSA traits. This result is not in accordance with the significant correlations found in this study between RSA traits and *Aphanomyces* resistance. As discussed in Desgroux et al. (2016), putative co-localization with underlying genes may have been missed in this study, due to the short CIs defined around significant SNPs. These short CIs were related to the low LD decay observed and/or to detection imprecisions due to GxE effects on the phenotype. Overlaps of CIs are highly dependent on the way CIs are defined. We used here one of the methodologies described to define a CI around a GWA marker, but other methodologies were described in the literature and we could have found overlaps if we would have used them. Further studies would be required to pinpoint genetic relationships between *Aphanomyces* root rot resistance and plant architecture loci in regions where co-localisations were observed. In other pathosystems, RSA traits correlated with root rot resistance have been shown to have different genetic control as in *F. solani* or *A. euteiches*/common bean pathosystems (Hagerty et al. 2015).

4.4. A major locus associated with resistance and intrinsic root system architecture

A co-localization was found between the three resistance variables and total root projected area, for which the same SNP marker (*Ps115429*) was detected. No previous study pinpointed such a precise co-localization for root-disease resistance and RSA traits. Furthermore, this SNP marker was detected with the most significant *p-values* for disease variables. It was located in the same genomic regions that major QTL detected by linkage mapping (Hamon et al. 2013) and by GWA mapping (Desgroux et al. 2016), confirming its major effect. Introduction of SNP markers from Boutet et al. (2016) enabled to pinpoint the genetic position of this locus, which made it interesting for marker assisted selection. The SNP

sequence mapped to an intron sequence of MTR_4g074875 gene coding for a MAP kinase on *Medicago truncatula* genome (Alves Carvalho S. Pers. Com.). In our study, different hypotheses can be put forward to explain the co-localization with root projected area, which could underly closely-linked genes or a single gene with pleiotropic effect on both traits. In wheat, a gene coding for a MAP-kinase was involved in tolerance to several abiotiques stresses, and root growth (Hao et al. 2015). MAP-kinases were also involved in plant immunity triggered by pathogen effectors (Sekhwal et al. 2015). Thus, hypothesis of pleiotropic gene involved in both root growth and resistance to *A. euteiches* is highly probable. Pea lines with a larger root projected area will conserve greater amount of healthy root segments to continue to catch water and nutrients, and to produce new root segments. This hypothesis was confirmed with the observation of greater number of second order ramifications on healthy root of pea lines with larger root systems than with smaller ones (personal observation data not shown). In our test conditions, it would also take more time to the pathogen to colonize and develop symptoms on larger root systems than on smaller ones. Thus, selecting for this locus could improve disease resistance and root system bushiness, conferring advantages even in the absence of *A. euteiches* to catch water and nutrients.

5. Conclusion

This study brought new tools for a better understanding of genetic links between root disease resistance and RSA. Pea lines with good levels of resistance to *A. euteiches* and bushy root system (larger number of roots and longer roots) were identified and could be interesting for breeders to improve resistance to *A. euteiches* in pea varieties. A SNP marker, detected for both improved resistance to *A. euteiches* and greater root projected area looked relevant to use for marker assisted selection. In further study, interest for phenotypic and genetic links between resistance and RSA at older stages should be addressed in order to explore stage effect. Even if intrinsic RSA was studied, a little is known about plant response to infection by *A. euteiches*. A pea line with a bushy root system could be highly impacted by infection and lost its fitness. A study is undergoing to explore root system and plant architectural response to infection by *A. euteiches*, its genetic control and its involvement in decreased yield losses.

Chapitre IV

Architecture induite en réponse
à l'infection et résistance à
Aphanomyces euteiches chez le
pois: analyse génétique
comparative et rôle dans
l'amélioration de la tolérance

Chapitre IV : Architecture induite en réponse à l'infection et résistance à *Aphanomyces euteiches* chez le pois: analyse génétique comparative et rôle dans l'amélioration de la tolérance

1. INTRODUCTION DU CHAPITRE

Dans les chapitres précédents, nous avons déchiffré le déterminisme génétique de la résistance partielle à *A. euteiches* dans une collection de pois (Chapitre II), puis comparé les régions génomiques contrôlant la résistance à *A. euteiches* à celles impliquées dans la morphologie racinaire, foliaire et globale de la plante, ce que nous avons appelé « architecture constitutive » (architecture des plantes saines) (Chapitre III). Dans ce chapitre, nous examinons la diversité des modifications de caractères d'architecture racinaire et aérienne des plantes infectées par *A. euteiches*, en comparaison à des plantes saines, ce que nous appelons dans ce chapitre « architecture induite », dans la collection de 266 lignées de pois. L'objectif est de déterminer les facteurs génétiques d'architecture induite qui seraient également associés à une réduction de symptômes racinaires et aériens provoqués par *A. euteiches*, ainsi qu'à une meilleure tolérance. La tolérance du pois à *A. euteiches* désigne dans cette thèse la capacité d'une lignée à subir de moindres pertes de rendement par rapport à des lignées ayant la même quantité de symptômes de pourriture racinaires suite à l'infection par *A. euteiches*.

L'étude de l'organisation génétique des régions impliquées dans la réponse à l'infection a été réalisée par génétique d'association génome entier utilisant un modèle mixte multi-locus dans la collection génotypée à l'aide de 13 204 marqueurs SNP. Les régions ainsi décelées ont été comparées avec celles détectées précédemment pour la résistance à *A. euteiches* et pour l'architecture constitutive en l'absence d'agent pathogène. Un sous-set de 15 lignées a ensuite été expérimenté, en parcelles de rendement fortement et faiblement infestées sur deux lieux et deux années (2014 et 2015), pour étudier relation entre l'architecture induite de la plante au champ, et le rendement en grain, en réponse à l'infection par *A. euteiches*.

Les résultats obtenus ont permis (i) de confirmer la diversité génétique de l'architecture induite de la plante en réponse à l'infection, ainsi que de mettre en évidence des corrélations

entre ces modifications architecturales et des variables de résistance intrinsèque; (ii) de déceler des facteurs génétiques communs contrôlant l'architecture induite et la quantité de symptômes de pourriture racinaire dus à *A. euteiches*; (iii) comparer le déterminisme génétique de l'architecture racinaire constitutive à celui de l'architecture induite par l'infection; (iii) d'identifier certains génotypes pour lesquels la combinaison d'un bon niveau de résistance intrinsèque et un système racinaire moins impacté par la maladie, permet de minimiser la perte de rendement à la récolte.

Ce chapitre est présenté sous la forme d'un article, en préparation, prévu pour soumission.

2. ARTICLE ASSOCIÉ

Les références bibliographiques mentionnées sont situées à la fin du manuscrit

Les fichiers additionnels sont présentés en Annexe 3

Disease-induced root architecture is genetically related to resistance to *Aphanomyces euteiches* in pea and contributes to reduce field disease impact

1. Introduction

Crops are subject to biotic stresses through their growing cycle which may affect yield and be more damaging in the context of climate change, which imply an increase of temperature and/or moisture. Plant tolerance to biotic stresses is the ability of a plant to maintain performance in the presence of disease symptoms (Ney et al. 2013). In many cases, the damage against which tolerance is measured is loss in grain yield. Tolerance differs from resistance, *i. e.* the ability of the plant to limit disease symptoms, and avoidance, *i. e.* its ability to escape pest or pathogen infection in time or in space (Ney et al. 2013). The ways of tolerance to diseases are multiple and expressed at different scales. The contribution of organs to the capture and use of resources depends on canopy and root system architecture (RSA), so the respective locations of disease and plant organs will have a strong effect on the crop's response to the disease. The determinants of plant tolerance variability within a species may thus be considered as a way to manage the effects of diseases and pests on crop growth and yield (Ney et al. 2013).

Mechanisms, biological processes and genetic determinism of plant response to pest or pathogen infection that may contribute to tolerance (yield maintenance) have poorly been reported, while those limiting symptoms have been more described. On aerial plant parts, dense canopies have long been shown to favor increase of symptoms in the case of aerial diseases (Calonnec et al. 2012). On root plant parts, Kraft and Boge (2001) showed that large-rooted pea genotypes (high total root length) produced more roots when infected by *Fusarium solani*, either in compacted soil or when cotyledon or a part of the root system were removed. The root system of resistant tomato plants infected by *Rhizoctonia solani* was characterized by the scarcity of short adventitious roots and the emergence of many short laterals, leading to a more branched root system (Berta et al. 2005; Sampo et al. 2007) and a reduction of mitotically active apices (Fusconi et al. 1999). A higher degree of branching was also reported on roots of resistant tomato plants infected with *Phytophthora nicotianae* var. *parasitica* (Trotta et al. 1996). In tomato and *Medicago sativa*, *R. solani* and *Pithium* sp. , respectively, have been shown to induce a more monopodial type of branching pattern, with

fewer orders of branching than uninfected controls (Larkin and Fravel 1998; Sampo et al. 2007). These results suggest the importance of root vigor and ability to preserve root vigor under adverse conditions in plant resistant (tolerant) response to root diseases.

The soil-borne oomycete pathogen *Aphanomyces euteiches* is responsible for high yield losses in highly infested pea (*Pisum sativum*) fields (Gaulin et al. 2007). This worldwide economically damaging pathogen causes honey brown root rot at seedling stage and resulted in shoot yellowing and dwarfism at later stage. For the past 20 years, scientists explored genetic control of pea resistance to *A. euteiches* using root rot scores and shoot yellowing scores as disease severity assessment (Desgroux et al. 2016; Hamon et al. 2011; Hamon et al. 2013; Pilet-Nayel et al. 2002, 2005). Breeding pea lines with increased level of resistance and interesting agronomic traits (Gritton 1990; Kraft 1992; Kraft and Coffman 2000a, b, c), was also mainly based on shoot yellowing scoring. Even if *Aphanomyces* resistance quantitative trait loci (QTL) were detected and resistant breeding lines developed, levels of resistance obtained were not sufficient to preserve yield in highly infested fields. Some *Aphanomyces* resistance QTL were detected when considering only shoot symptoms (Hamon et al. 2011; Hamon et al. 2013) and several pea lines displayed important root symptoms in controlled conditions but low shoot symptoms in field conditions (Pilet-Nayel Pers. Com.). Biological processes contributing to low impact of the disease on aerial parts of some pea lines are not known. They may include root vigor, through the induction of roots after infection, allowing plant nutrition to be preserved. In *Medicago truncatula*, Djéballi et al. (2009) and Bonhomme et al. (2014) identified some resistant accessions which increased lateral roots growth in infected versus healthy conditions. Plant architectural modifications under infection compared to pathogen free conditions were referred as disease-induced plant architecture in the present study. Disease-induced plant architecture lowly impacted, combined to intrinsic resistance, could be a way to preserve yield in infested fields. Recently, Desgroux et al. (in prep) showed little phenotypic correlations and few common genetic determinants between intrinsic root system architecture (RSA) traits (pathogen free controlled conditions) and resistance to *A. euteiches*. No information is available, neither on the relationships between genetic determinants of disease-induced RSA and resistance, nor on impact of intrinsic and disease-induced RSA on yield preservation. The ability of a pea line to preserve yield compared to other lines with the same level of symptoms is referred in the present study to plant tolerance.

The aim of this study was to determine relationships between the genetic control of disease-induced RSA and the one of root and shoot symptoms reduction, as well as its impact on plant

tolerance to *A. euteiches*. This study (i) explored the genetic diversity of disease-induced plant architecture at seedling stage under *A. euteiches* infection in controlled conditions, in a collection of 266 pea lines studied for *Aphanomyces* resistance and intrinsic plant architecture in a previous study (Desgroux et al. in prep); (ii) identified genomic regions associated with disease-induced architecture (DIA) traits by genome wide association (GWA) analysis and compared them to the ones reported for intrinsic architecture traits and resistance to *A. euteiches* in previous studies; and (iii) analyzed the relationships between disease-induced plant architecture, resistance and yield traits, in a field assay including 15 pea lines, contrasted for resistance and intrinsic or disease-induced RSA traits observed in controlled conditions.

2. Material and methods

2.1. Plant material

The 266-pea-line collection, previously described and genotyped using 14, 157 SNP markers (Desgroux et al. in prep), was used to analyse disease-induced root and shoot system architecture under *A. euteiches* infection in controlled conditions and for GWA analyses. Fifteen pea lines, composing the “Apharchi” subset, were chosen from the 266 lines for field experiments, based on their contrasted disease resistance and RSA phenotypes observed in this study and previous studies (Desgroux et al. in prep; Desgroux et al. 2016).

2.2. Controlled conditions experiment

2.2.1. Phenotyping in controlled conditions

Root, shoot and overall plant architecture of the 266 lines were measured on 14 day-old pathogen free or *Aphanomyces* infected plants from Exp#2 and Exp#3, respectively, described in Desgroux et al. (in prep) and conducted with three replicates.

The Exp#2 was carried out in a climatic chamber (25°C for 16h-day, 23°C for 8h-night) located at INRA-Le Rheu (Brittany, France). All the lines were assayed using a randomized complete block design with three replicates. Five seeds per replicate were sown in a pot (9 cm x 9 cm x 9.5 cm) for each pea line. Plants were grown in vermiculite substrate and watered as needed.

The Exp#3, was conducted in the same climatic chamber as Exp#2 (25°C for 16h-day, 23°C for 8h-night), as described in Desgroux et al. (2016). Seven-day old seedlings were inoculated with a pure-culture of the RB84 French reference strain of *A. euteiches* described in (Hamon

et al. 2011), referred as from pathotype I (Wicker et al. 2001). Inoculation was performed with a solution of 10^3 zoospores per plant, produced as described by Moussart et al. (2001). Fourteen-day old seedlings were uprooted and two seedlings per pot were carefully washed before plant architecture trait measurements.

In each experiment and replicate, two plants per line were scored, as described in Bourion et al. (2010), for tap root length (TapRootL), number of first lateral roots (NLatRoot), total projected area (TProjArea), total root length (TRootL), average root diameter (RootDia), average lateral root length (LatRootL), root biomass (RootB), shoot biomass (ShootB), plant height (ShootL), total biomass (TB), and root to total biomass ratio (RootB:TB), as described in Desgroux et al. (in prep).

The root, shoot and overall plant architecture measurements of Exp#2 and Exp#3 were used to assess disease-induced architecture (DIA). For each pea line, DIA of each trait was calculated, for each replicate, as the percentage of gain or loss between inoculated and non-inoculated conditions (X_{ratio}), using the formula: $X_{ratio} = (X_{I+} - X_{I-})/X_{I-}$, with X_{I+} , the average score obtained on the plants in one replicate in inoculated conditions (Exp#3) and X_{I-} , the average score obtained on the plants in one replicate in non-inoculated conditions (Exp#2, (Desgroux et al. in prep)). The DIA scorings were compared to disease severity (DS), Root Rot Index (RRI) and percentage of brown projected area (Br:TProjArea) scores for Aphanomyces resistance, previously obtained in Exp#3 and described in Desgroux et al. (in prep).

2.2.2. Statistical analysis of controlled condition data

Phenotypic datasets obtained were analysed using the R 3. 1. 1 program (R Core Team 2014). For each plant architecture variable scored in Exp#2 (Desgroux et al. in prep) and Exp#3 (this study), a linear model (LM) was used for a global analysis, including G (genotype), I (inoculation state), GxI (genotype x inoculation state interaction), R (replicate) and P (plant) as fixed factors. Plant architecture variables measured in inoculated conditions (Exp#3) were also analysed separately using a LM with G, R and P as fixed factors. DIA variables were analysed individually, using a LM with G and R as fixed factors. For each variable, normality of residuals and homogeneity of their variance, as well as mean-based heritability (h^2) in the global and individual analysis were estimated as described in Desgroux et al. (2016). Least Square Means (LSMeans) were calculated from each LM analysis and their frequency distribution were drawn, as described in Desgroux et al. (in prep). Pearson correlations

analysis and significance were also tested between LSM means of the different variables, including ratios, as described by Desgroux et al. (in prep).

2.3. Genetic analysis and comparative mapping

The genotyping dataset of 11, 789 SNP markers, Kinship matrix and PCA components described in Desgroux et al. (in prep) for the 266 pea-line collection, were used for GWA analysis.

GWA study used LSM means of DIA variables (inoculated vs. non-inoculated ratios) of all traits (NLatRoot, TProjArea, TRootL, RootDia, LatRootL, RootB, ShootB, ShootL, TB and RootB:TB). GWA analysis was performed as in Desgroux et al. (in prep), using a multi-locus mixed model (MLMM) with a multiple-Bonferroni (mBonf) threshold of 4.58 ($p\text{-value} \leq 2.6E-05$) to declare a SNP marker significant. Local linkage disequilibrium (LD) analysis defined the confidence intervals (CIs) around significant associated markers, using Plink 1.9 software (Chang et al. 2015; Purcell et al. 2007; Purcell and Chang 2015).

The same map as described in Desgroux et al. (in prep) was used for graphic representation of trait-associated loci using MapChart 2.1 software (Voorrips 2002). Significant loci identified in this study were compared to those previously detected for root, shoot and overall plant architecture at 8 and 14 days after sowing in non-inoculated conditions (intrinsic architecture), and for resistance to *A. euteiches* in (Desgroux et al. in prep; Desgroux et al. 2016), in order to visualize putative true co-localisations (the same marker detected, detected markers in LD, or CIs sharing common SNP markers) and overlaps (independent loci with overlapping CIs genetic positions).

2.4. Field experiment

2.4.1. Field assay

The Apharchi subset of 15 pea lines was evaluated in field conditions over two locations (Riec-sur-Belon (RI), Brittany, France and Dijon-Epoisses (DI), Burgundy, France) and two years (2014 and 2015). In each environment (location x year), the experiment was conducted both in a highly infested field by *A. euteiches* (Ae+) and in a lowly infested field (Ae-). Both fields were located a few hundred meters from each other, in order to ensure similar pedo-climatic conditions. In each field, the experiment was performed using a randomized complete block design with three replicates. In each replicate, each pea line was sown in two

adjacent micro-plots of 4.5 m² (3 m-long x 1.5 m-wide), with a density of 85 seeds.m², including one plot for scorings and the other one for yield evaluation.

In the scoring plots of the highly and lowly infested fields, plant resistance and architecture were measured two months after sowing (approximately 10 nodes stage) in two to four environments and yield components at seed maturity stage in two to three environments. Missing measurements in some environments were due to field fertilisation difficulties or bird damages.

- Two months post sowing (2Mps), 10 plants per plot were randomly up-rooted and scored for RRI as described in Desgroux et al. (2016). Five plants were chosen among the 10 plants, for root and shoot architecture scorings, including (i) taproot length (TapRootL), number of lateral roots (NLatRoot), shoot length (ShootL) and number of nodes (NNodes) of the longest stem, of the individual plants and (ii) root and shoot weights of the five pooled plants dried for 48h at 80°C (RootB and ShootB).
- At seed maturity (Sm), 10 plants per plot were randomly harvested and scored individually for the length of the longest branch (ShootL), and for five yield components: the total number of podded nodes (NPodNode), the total number of pods (NPod), the total number of seeds (NSeed) and the weight of pooled seeds (SeedW). Additional components were further calculated as the number of pod per podded node (NPod.PodNode = NPod / NPodNode), the number of seeds per pod (NSeed.Pod = NSeed / NPod) and the thousand seeds weight (TSW = SeedW * 100 / NSeed).

In three environments (DI_2014, RI_2014 and DI_2015), the yield plots of the highly and lowly infested fields were evaluated once or twice for disease resistance using the aerial decline index (ADI), as described in Desgroux et al. (2016), between the flowering and pre-maturity stages. The plots were harvested at seed maturity stage with a harvesting machine and pooled seeds were weighted (SeedYield).

2.4.2. Statistical analysis of field experiment

Phenotypic variables measured in field conditions on the Apharchi subset were analysed using a LM including G, Ae, E, GxE, GxAe, GxExAe, R and P when appropriate, as fixed factors, with G the genotype, Ae the infestation level of the soil by *A. euteiches* (Ae- and Ae+), E the environment, R the replicate, and P the plant. Normality of residuals and homogeneity of their variance was checked, as described for controlled conditions experiment. LSMeans were calculated from the LM analysis, for each pea line and infestation level. A Tuckey test was

applied on the model between Ae- and Ae+ on each pea line to determine if differences were significant ($p\text{-value} < 0.05$) between the levels of infestation (R function `cld` of package `multcompview`, (Graves et al. 2015)).

3. Results

3.1. Controlled conditions experiment

3.1.1. Analysis of phenotypic data

The global statistical analysis of plant architecture variables on the 266 pea lines in non-inoculated (Exp#2, (Desgroux et al. in prep)) and inoculated (this study) conditions revealed highly significant G, I and GxI effects ($p\text{-value} < 0.001$) for all variables, except for ShootL (I effect: $0.01 < p\text{-value} < 0.05$) (Table IV.1).

Statistical analysis of plant architecture scores, in inoculated conditions (Exp#3, data not shown), and of DIA scores (non-inoculated vs. inoculated conditions ratios), revealed a highly significant G effect ($p\text{-value} < 0.001$), except for RootDia ($p\text{-value} < 0.01$) and NLatRoot ($p\text{-value} > 0.05$) (Table IV.2). Heritability of the DIA variables was moderate for most of them ($0.31 < h^2 < 0.66$), and low for NLatRoot and RootDia ($0.11 < h^2 < 0.23$). Frequency distribution of LSM means of DIA scores tended to fit normal curves (Additional file IV.1). Most of the 266 pea lines showed negative DIA scores for most traits, except for RootDia. Negative DIA reflected losses of NLatRoot, TRootL, LatRootL, TProjArea, ShootB, RootB, TB and ShootL when pea lines were infested by *A. euteiches*. Interestingly, 2 to 21 % of the lines presented positive DIA scores, depending on the variables.

Table IV.1: Global statistical analysis results of plant architecture data in the 266-pea-line collection

Variable (a)	Number of pea lines (b)	Number of observations (b)	Adjusted r^2 model (c)	h^2 (d)	G effect (e,f)	I effect (e,g)	R effect (e,h)	G x I effect (e,i)	P effect (e,j)
TProjArea \$	261	3110	0.69	0.79	***	***	***	***	*
RootDia \$	261	3110	0.71	0.93	***	***	***	***	***
TRootL \$	261	3110	0.74	0.81	***	***	***	***	ns
NLatRoot \$	261	3110	0.70	0.93	***	**	***	***	ns
LatRootL \$	261	3110	0.54	0.66	***	***	***	***	ns
RootB \$	261	3110	0.67	0.79	***	***	*	***	ns
ShootB \$	261	3110	0.71	0.92	***	***	***	***	ns
ShootL \$	261	3110	0.95	0.99	***	*	***	***	**
TB \$	261	3110	0.70	0.9	***	***	***	***	ns
RootB:TB \$	261	3110	0.70	0.85	***	***	***	***	ns

(a) TProjArea: total projected area, RootDia: average root diameter, TRootL: total root length, NLatRoot: number of lateral roots, LatRootL: average length of lateral roots, RootB: root biomass, ShootB: shoot biomass, ShootL: shoot length, TB: total biomass, RootB:TB: root to total biomass ratio, '\$' Residual normality statistics are not exact but graphs of residuals are correct; (b) Number of pea lines actually assessed, number of total observations in the 266-pea-line collection; (c) adjusted r^2 coefficient of the model; (d) mean-based heritability; (e) ANOVA significance p-value codes: 0 < '***' < 0.001 ≤ '***' < 0.01 ≤ '**' < 0.05 ≤ 'ns' < 1; (f): Genotype effect; (g): Inoculation effect; (h): Replicate effect; (i) Genotype by Inoculation interaction effect; (j): Plant effect

Table IV.2: Statistical analysis results of disease-induced plant architecture data in the 266-pea-line collection

inoculated vs non-inoculated ratios (a)	Number of pea lines (b)	Number of observations (b)	Mean (b)	Min (b)	Max (b)	se (b)	Adjusted r ²		R effect (e,g)
							model (c)	h ² (d)	
TProjArea_ratio	261	777	-0.28	-0.76	0.45	0.1	0.39	0.64	***
RootDia_ratio	261	777	0.05	-0.09	0.21	0.04	0.09	0.23	ns
TRootL_ratio	261	777	-0.31	-0.76	0.38	0.09	0.40	0.66	***
NLatRoot_ratio	261	778	-0.18	-0.46	0.06	0.09	0.07	0.11	***
LatRootL_ratio	261	777	-0.16	-0.73	0.57	0.11	0.33	0.59	*
RootB_ratio	261	777	-0.25	-0.84	0.56	0.12	0.31	0.55	***
ShootB_ratio	261	777	-0.1	-0.46	0.4	0.11	0.18	0.37	***
ShootL_ratio	261	777	-0.07	-0.33	0.42	0.09	0.20	0.31	***
TB_ratio	261	776	-0.15	-0.56	0.45	0.1	0.21	0.41	***
RootB:TB_ratio	261	776	-0.13	-0.6	0.22	0.07	0.39	0.64	***

(a) Traits were coded as described in Table IV.1; (b) Number of pea lines actually assessed, number of total observations, means, minimum, aximum and standard error in the 266-pea-line collection; (c) adjusted r² coefficient of the model; (d) mean-based heritability; (e) ANOVA significance p-value codes: 0 < '***' < 0.001 ≤ '***' < 0.01 ≤ '**' < 0.05 ≤ 'ns' < 1; (f): Genotype effect; (g): Replicate effect

Table IV.3: Pearson correlation coefficients between plant architecture traits in inoculated and non-inoculated conditions and disease-induced architecture traits (to be continued)

	Inoculated											
	Root morphology						Shoot morphology			Plant morphology		
	TProjArea	RootDia	TRootL	NLatRoot	LatRootL	RootB	ShootB	ShootL	TB	RootB:TB		
TProjArea	1	-0.17 **	0.97 ***	0.75 ***	0.75 ***	0.92 ***	0.52 ***	ns	0.69 ***	0.44 ***		
RootDia		1	-0.40 ***	-0.49 ***	ns	ns	-0.17 **	-0.42 ***	-0.13 *	0.30 ***		
TRootL			1	0.82 ***	0.71 ***	0.83 ***	0.52 ***	ns	0.66 ***	0.35 ***		
NLatRoot				1	0.22 ***	0.57 ***	0.47 ***	0.14 *	0.55 ***	0.13 *		
LatRootL					1	0.74 ***	0.37 ***	ns	0.51 ***	0.43 ***		
RootB						1	0.54 ***	ns	0.73 ***	0.52 ***		
ShootB							1	0.60 ***	0.97 ***	-0.37 ***		
ShootL								1	0.47 ***	-0.67 ***		
TB									1	-0.15 *		
RootB:TB										1		

Inoculated

	Inoculated											
	Root morphology						Shoot morphology			Plant morphology		
	TProjArea	RootDia	TRootL	NLatRoot	LatRootL	RootB	ShootB	ShootL	TB	RootB:TB		
TProjArea	0.66 ***	ns	0.64 ***	0.58 ***	0.44 ***	0.57 ***	0.59 ***	0.20 **	0.65 ***	ns		
RootDia		1	-0.15 *	-0.26 ***	ns	0.27 ***	ns	-0.39 ***	ns	0.32 ***		
TRootL	0.61 ***	-0.34 ***	1	0.65 ***	0.39 ***	0.45 ***	0.56 ***	0.32 ***	0.60 ***	ns		
NLatRoot	0.54 ***	-0.34 ***	0.62 ***	1	ns	0.35 ***	0.38 ***	ns	0.41 ***	ns		
LatRootL	0.23 ***	ns	0.22 ***	ns	1	0.24 ***	0.33 ***	0.29 ***	0.34 ***	ns		
RootB	0.63 ***	0.27 ***	0.53 ***	0.39 ***	0.48 ***	0.67 ***	0.59 ***	ns	0.68 ***	0.25 ***		
ShootB	0.34 ***	ns	0.30 ***	0.32 ***	0.18 **	0.37 ***	0.86 ***	0.48 ***	0.81 ***	-0.33 ***		
ShootL	-0.19 **	-0.33 ***	ns	ns	-0.21 ***	-0.20 **	0.55 ***	0.98 ***	0.39 ***	-0.71 ***		
TB	0.44 ***	0.16 *	0.38 ***	0.36 ***	0.27 ***	0.47 ***	0.84 ***	0.40 ***	0.82 ***	-0.19 **		
RootB:TB	0.31 ***	0.20 **	0.24 ***	ns	0.35 ***	0.33 ***	-0.41 ***	-0.51 ***	-0.24 ***	0.77 ***		

Non-inoculated

Variables are coded as described in Table IV.1. Pearson correlation p -values : 0 < '***' < 0.001 ≤ '***' < 0.01 ≤ '**' < 0.05 ≤ 'ns' < 1. Level of correlation is coded with a colour gradation scale from dark red ($r = 1$) to white ($r = 0$) and from white to dark green ($r = 1$): dark green. "ns": non-significant correlation (p -value > 0.05).

Table IV.3: Pearson correlation coefficients between plant architecture traits in inoculated and non-inoculated conditions and disease-induced architecture traits (to be continued)

		Inoculated vs non-inoculated ratios												
		Root morphology					Shoot morphology					Plant morphology		
		TProjArea	RootDia	TRootL	NLatRoot	LatRootL	RootB	ShootB	ShootL	TB	RootB:TB			
Root morphology	TProjArea	1	ns	0.98 ***	0.51 ***	0.86 ***	0.86 ***	0.50 ***	0.53 ***	0.68 ***	0.65 ***			
	RootDia		1	-0.23 ***	-0.26 ***	ns	ns	ns	-0.24 ***	ns	ns			
	TRootL			1	0.54 ***	0.87 ***	0.84 ***	0.51 ***	0.56 ***	0.68 ***	0.62 ***			
	NLatRoot				1	ns	0.51 ***	0.41 ***	0.35 ***	0.49 ***	0.30 ***			
	LatRootL					1	0.71 ***	0.38 ***	0.48 ***	0.54 ***	0.57 ***			
	RootB						1	0.61 ***	0.53 ***	0.82 ***	0.73 ***			
Shoot morphology	ShootB							1	0.69 ***	0.95 ***	ns			
	ShootL								1	0.69 ***	ns			
Plant morphology	TB									1	0.22 ***			
	RootB:TB											1		

Inoculated vs non-inoculated ratios

Inoculated vs non-inoculated ratios

		Inoculated vs non-inoculated ratios											
		Root morphology					Shoot morphology					Plant morphology	
		TProjArea	RootDia	TRootL	NLatRoot	LatRootL	RootB	ShootB	ShootL	TB	RootB:TB		
Root morphology	TProjArea	ns	-0.15 *	ns	ns	ns	ns	ns	ns	ns	ns		
	RootDia	ns	ns	ns	ns	0.16 *	ns	-0.24 ***	-0.17 **	-0.21 ***	0.19 **		
	TRootL	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns		
	NLatRoot	0.14 *	-0.17 **	0.17 **	ns	0.25 ***	ns	ns	ns	ns	0.15 *		
	LatRootL	-0.22 ***	ns	-0.23 ***	0.18 **	-0.36 ***	ns	ns	ns	ns	ns		
	RootB	ns	ns	ns	ns	ns	ns	-0.13 *	ns	ns	0.13 *		
Shoot morphology	ShootB	-0.14 *	ns	-0.15 *	ns	-0.12 *	-0.13 *	-0.13 *	-0.19 ***	-0.13 *	ns		
	ShootL	-0.43 ***	ns	-0.42 ***	ns	-0.47 ***	-0.32 ***	ns	-0.14 *	ns	-0.52 ***		
Plant morphology	TB	ns	ns	ns	ns	ns	ns	-0.14 *	-0.17 **	-0.13 *	ns		
	RootB:TB	0.24 ***	-0.15 *	0.26 ***	ns	0.24 ***	0.14 *	ns	0.22 ***	ns	0.22 ***		

Non-inoculated

Variables are coded as described in Table IV.1. Pearson correlation p -values: $0 < *** < 0.001 \leq ** < 0.01 \leq * < 0.05 \leq ns < 1$. Level of correlation is coded with a colour gradation scale from dark red ($r = 1$) to white ($r = 0$) and from white to dark green ($r = 1$): dark green. "ns": non-significant correlation (p -value > 0.05).

Table IV.3: Pearson correlation coefficients between plant architecture traits in inoculated and non-inoculated conditions and disease-induced architecture traits (end)

	Inoculated vs non-inoculated ratios											
	Root morphology						Shoot morphology			Plant morphology		
	TProjArea	RootDia	TRootL	NLatRoot	LatRootL	RootB	ShootB	ShootL	TB	RootB:TB	RootB:TB	RootB:TB
Root morphology	TProjArea	0.67 ***	-0.19 **	0.68 ***	0.45 ***	0.56 ***	0.59 ***	0.40 ***	0.43 ***	0.53 ***	0.43 ***	0.53 ***
	RootDia	ns	0.38 ***	ns	-0.21 ***	ns	ns	-0.28 ***	-0.23 ***	0.18 **	-0.23 ***	0.18 **
	TRootL	0.63 ***	-0.28 ***	0.66 ***	0.48 ***	0.52 ***	0.57 ***	0.44 ***	0.46 ***	0.46 ***	0.46 ***	0.46 ***
	NLatRoot	0.38 ***	-0.28 ***	0.42 ***	0.44 ***	0.27 ***	0.35 ***	0.27 ***	0.29 ***	0.29 ***	0.29 ***	0.29 ***
	LatRootL	0.57 ***	ns	0.57 ***	0.24 ***	0.57 ***	0.53 ***	0.40 ***	0.41 ***	0.41 ***	0.45 ***	0.45 ***
Shoot morphology	RootB	0.64 ***	-0.13 *	0.64 ***	0.41 ***	0.54 ***	0.65 ***	0.31 ***	0.46 ***	0.61 ***	0.46 ***	0.61 ***
	ShootB	ns	ns	ns	0.16 **	ns	0.15 *	0.36 ***	0.34 ***	ns	0.34 ***	ns
Plant morphology	ShootL	-0.34 ***	ns	-0.32 ***	ns	-0.39 ***	-0.23 ***	0.24 ***	ns	-0.51 ***	ns	-0.51 ***
	TB	0.28 ***	ns	0.28 ***	0.25 ***	0.21 ***	0.32 ***	0.39 ***	0.41 ***	ns	0.41 ***	ns
	RootB:TB	0.57 ***	ns	0.56 ***	0.27 ***	0.52 ***	0.54 ***	ns	0.16 **	0.77 ***	0.16 **	0.77 ***

Inoculated

Inoculated

	Inoculated vs non-inoculated ratios					
	Disease			Disease		
	RRI	DS	Br:TProjArea	RRI	DS	Br:TProjArea
Root morphology	TProjArea	-0.59 ***	-0.57 ***	-0.33 ***	-0.57 ***	-0.39 ***
	RootDia	0.35 ***	0.31 ***	0.48 ***	0.31 ***	0.32 ***
	TRootL	-0.64 ***	-0.62 ***	-0.43 ***	-0.62 ***	-0.44 ***
	NLatRoot	-0.46 ***	-0.45 ***	-0.29 ***	-0.40 ***	-0.35 ***
	LatRootL	-0.47 ***	-0.46 ***	-0.31 ***	-0.49 ***	-0.30 ***
Shoot morphology	RootB	-0.54 ***	-0.53 ***	-0.22 ***	-0.61 ***	-0.41 ***
	ShootB	-0.13 *	ns	ns	-0.31 **	-0.33 ***
Plant morphology	ShootL	0.13 *	0.15 *	ns	-0.44 ***	-0.42 ***
	RootB:TB	-0.26 ***	-0.25 ***	ns	-0.46 ***	-0.40 ***
		-0.54 ***	-0.53 ***	-0.22 ***	-0.49 ***	-0.19 **

Variables are coded as described in Table IV.1. Pearson correlation p -values : 0 < '***' < 0.001 ≤ '**' < 0.01 ≤ '*' < 0.05 ≤ 'ns' < 1. Level of correlation is coded with a colour gradation scale from dark red ($r = 1$) to white ($r = 0$) and from white to dark green ($r = 1$): dark green. "ns": non-significant correlation (p -value > 0.05).

3.1.2. Correlations among variables

In inoculated condition, LSMeans of the root, shoot and overall plant architecture variables were positively and significantly correlated between each other (corrected p -value < 0.001 ; $0.37 < r < 0.97$), except for (i) RootDia and ShootL for which scores were negatively (corrected p -value < 0.05 ; $-0.49 < r < -0.13$) or non-significantly correlated with those of most of the other traits and for (ii) RootB:TB, which scores were negatively and significantly correlated with those of shoot traits and TB (corrected p -value < 0.05 ; $-0.67 < r < -0.15$) (Table IV.3). For a same variable, LSMeans in non-inoculated condition were moderately to highly correlated with those in inoculated condition (corrected p -value < 0.001 ; $0.51 < r < 0.98$).

LSMeans of DIA variables were significantly and positively correlated between each other for all variables (corrected p -value < 0.001 ; $0.22 < r < 0.98$), except for RootDia, which DIA scores were either not correlated or negatively correlated with those of the other traits (corrected p -value < 0.001 ; $-0.26 < r < -0.23$). Similar results were observed between LSMeans of DIA variables and scores obtained in inoculated condition, except for ShootL scores which were also negatively correlated with scores of most root variables (corrected p -value < 0.001 ; $-0.23 < r < -0.51$). Inversely, LSMeans of most root DIA variables were not correlated or poorly correlated (corrected p -value < 0.05 ; $-0.15 < r < 0.25$) with those of other traits in non-inoculated condition, except for RootDia and LatRootL which scores in this condition were significantly and negatively correlated with DIA values of shoot traits and most of root traits, respectively (corrected p -values < 0.01 ; $-0.36 < r < -0.17$). LSMeans of shoot variables in non-inoculated condition were negatively correlated with those of DIA for most variables (corrected p -values < 0.05 ; $-0.52 < r < -0.12$). LSMeans of disease scores were mostly highly significantly and negatively correlated with those of all DIA variables and scores in inoculated condition (corrected p -value < 0.001 ; $-0.64 < r < -0.13$), except for RootDia which scores were positively correlated with disease ones (corrected p -value < 0.001 , $0.30 < r < 0.48$). Susceptible pea lines had more losses of their morphologic traits in inoculated vs. non inoculated conditions than resistant ones, and increased their root diameter.

Table IV.4: Number of markers detected by genome-wide association mapping for disease-induced architecture traits

Variable (a)	Number of markers (b)	Range of <i>p-value</i> (c)	Range of allelic effect (d)		% of phenotypic variance		Unexplained variance (h)	
			PCA (e)	Markers (f)	Kinship (g)			
LatRootL_ratio	0	-	-	-	50%	0%	12%	38%
NLatRoot_ratio	1	1.18E-06	0.039	0.039	6%	9%	0%	85%
RootB:TB_ratio	0	-	-	-	48%	0%	19%	34%
RootB_ratio	1	7.38E-07	0.081	0.081	31%	6%	16%	47%
RootDia_ratio	0	-	-	-	9%	0%	17%	74%
ShootB_ratio	3	9.14E-07 - 1.78E-05	0.040 - 0.082	0.040 - 0.082	9%	19%	0%	72%
ShootL_ratio	2	3.34E-06 - 5.75E-06	0.028 - 0.036	0.028 - 0.036	21%	14%	5%	61%
TB_ratio	2	6.29E-06 - 1.62E-05	0.037 - 0.058	0.037 - 0.058	11%	13%	9%	66%
TProjArea_ratio	0	-	-	-	43%	0%	13%	44%
TRootL_ratio	10	3.70E-10 - 2.09E-05	0.032 - 0.062	0.032 - 0.062	45%	26%	0%	30%

(a) Variables are abbreviated as in Table IV.1, disease-induced architecture traits are inoculated vs. non-inoculated ratios; (b) number of markers used as cofactors at the optimal step of the multi-locus mixed model (MLMM) analysis; (c) Range of *p-values* of the significant markers, significance threshold is $p\text{-value} < 3.4E-05$ as described in Methods section; (d) Range of allelic effects of the significant markers in absolute values; Percentage of phenotypic variance explained by: (e) the principal component analysis (PCA) matrix of the pea *Aphanomyces* collection structure, (f) all cofactor markers together, (g) the Kinship relatedness matrix among lines of the pea *Aphanomyces* collection, (h) the unexplained variance qualified as "missing heritability"

3.1.3. Genome wide association study for plant response to *A. euteiches* infection

A total of 15 SNP markers in independent CIs on six LGs were significantly associated with phenotypic variations observed in the collection, *i. e.* for the 10 DIA variables of root, shoot and plant architecture (Table VI. 4 and Table IV.5). Eleven SNP markers were detected for DIA variables of three RSA traits (TRootL: 10; NLatRoot: 1, RootB: 1), including one marker detected for DIA of both TRootL and RootB (*Ps115429* on LGVII). These markers, included as co-factors in the MLM, explained 6 to 26% of the total variance of root DIA variables (Table IV.4). Four SNP markers were detected for DIA variables of shoot architecture traits, including one marker on LGIV detected for both ShootB and ShootL. These markers explained 14 to 19 % of the variance of shoot DIA variables. Two markers were detected for TB DIA variable, including one marker associated with ShootB DIA variable and the other one with RootB DIA variable. The two markers explained 13 % of the variance of TB DIA variable. No markers were associated to phenotypic variation for four DIA variables, which had a low heritability (RootDia) and/or a phenotypic variance mainly captured by structure and/or kinship in the collection (LatRootL, RootB:TB, TProjArea).

Table IV.5: Genetic markers associated with disease-induced architecture traits detected by genome-wide association mapping

LG (a)	Confidence interval		Position (cM) (b)	SNP (d)	MAF (e)	Variable (f)	p-value (g)	Intercept (h)	Allelic effect (i)	SE (j)
	(cM) (b,c)	(cM) (b)								
I	40.6 - 41.4	40.9	<i>PsCam005546_4188_964</i>	0.46	TRootL_ratio	3.25E-06	-0.189	0.036	0.008	
II	32.4 - 34.2	33.4	<i>Ps127697</i>	0.30	TRootL_ratio	2.09E-05	-0.189	-0.038	0.009	
	35.8 - 36.7	36.6	<i>PsCam037731_22801_1472</i>	0.32	TRootL_ratio	3.70E-10	-0.189	-0.053	0.008	
	38.7 - 39.0	38.9	<i>PsCam048060_30816_341</i>	0.10	TRootL_ratio	8.36E-06	-0.189	-0.050	0.011	
III	45.5 - 45.5	45.5	<i>PsCam050243_32836_1352</i>	0.19	ShootL_ratio	5.75E-06	-0.062	0.036	0.008	
	22.0 - 22.0	22.0	<i>PsCam058589_38986_524</i>	0.05	ShootB_ratio	1.78E-05	-0.015	-0.082	0.019	
IV	70.4 - 71.6	70.4	<i>PsCam043428_27437_755</i>	0.45	ShootB_ratio	2.73E-06	-0.015	-0.042	0.009	
	41.0 - 42.8	42.5	<i>PsCam021147_11860_302</i>	0.49	ShootB_ratio	9.14E-07	-0.015	-0.040	0.008	
	51.9 - 52.4	52.4	<i>PsCam036799_21921_219</i>	0.24	ShootL_ratio	3.34E-06	-0.062	-0.028	0.006	
V	110.5 - 111.6	110.8	<i>PsCam045667_29303_1927</i>	0.33	TB_ratio	6.29E-06	-0.142	-0.037	0.008	
	65.9 - 66.5	66.1	<i>PsCam050864_33422_321</i>	0.18	TRootL_ratio	1.92E-05	-0.189	0.036	0.008	
	93.5-93.5	93.5	<i>PsCam049170_31813_427</i>	0.43	TRootL_ratio	3.56E-06	-0.189	-0.032	0.007	
	110.6-110.6	110.6	<i>PsCam015054_10010_287</i>	0.33	TRootL_ratio	1.18E-06	-0.197	0.039	0.008	
VII	9.3 - 10.7	10.1	<i>PsCam036297_21444_870</i>	0.45	NLatRoot_ratio	4.09E-06	-0.189	-0.034	0.007	
	70.8 - 71.5	71.3	<i>Ps115429</i>	0.25	TRootL_ratio	8.97E-07	-0.189	-0.036	0.007	
					RootB_ratio	3.35E-06	-0.189	-0.034	0.007	
					TB_ratio	3.80E-09	-0.189	0.062	0.010	
					TB_ratio	7.38E-07	-0.296	0.081	0.016	
					TB_ratio	1.62E-05	-0.142	0.058	0.013	

Description of genetic markers detected by genome-wide association (GWA) study on the 266-pea-line collection, using the Multi-Locus Mixed Model (MLMM) method (Segura et al 2012), for inoculated vs non-inoculated ratios of plant architecture traits. (a) LG: Linkage group; (b) Genetic positions are estimated on the genetic marker map resulting from the projection of the 953 SNP (Boutet et al submitted) markers onto the THMap (Desgroux et al 2016); (c) A confidence interval (CI) is defined with all markers in linkage disequilibrium ($r^2 > 0.2$) with the associated marker; (d) *PsCam*- markers are SNP markers developed by Tayeh et al (2015), *Ps1*- markers are SNP markers developed by Boutet et al (submitted); (e) Minor allele frequency (MAF) of each detected marker in the 266-pea-line collection; (f) Variables are coded as described in Table IV.1, disease-induced architecture traits are inoculated vs. non-inoculated ratios; (g) significant p-value of the marker-trait association $< 2.6E-05$; (h) mean value of the trait corrected with the effect of all markers included in the MLMM analysis; (i) allelic effect of the marker on the variable, sign depends on the allele code; (j) standard error in the 266-pea-line collection.

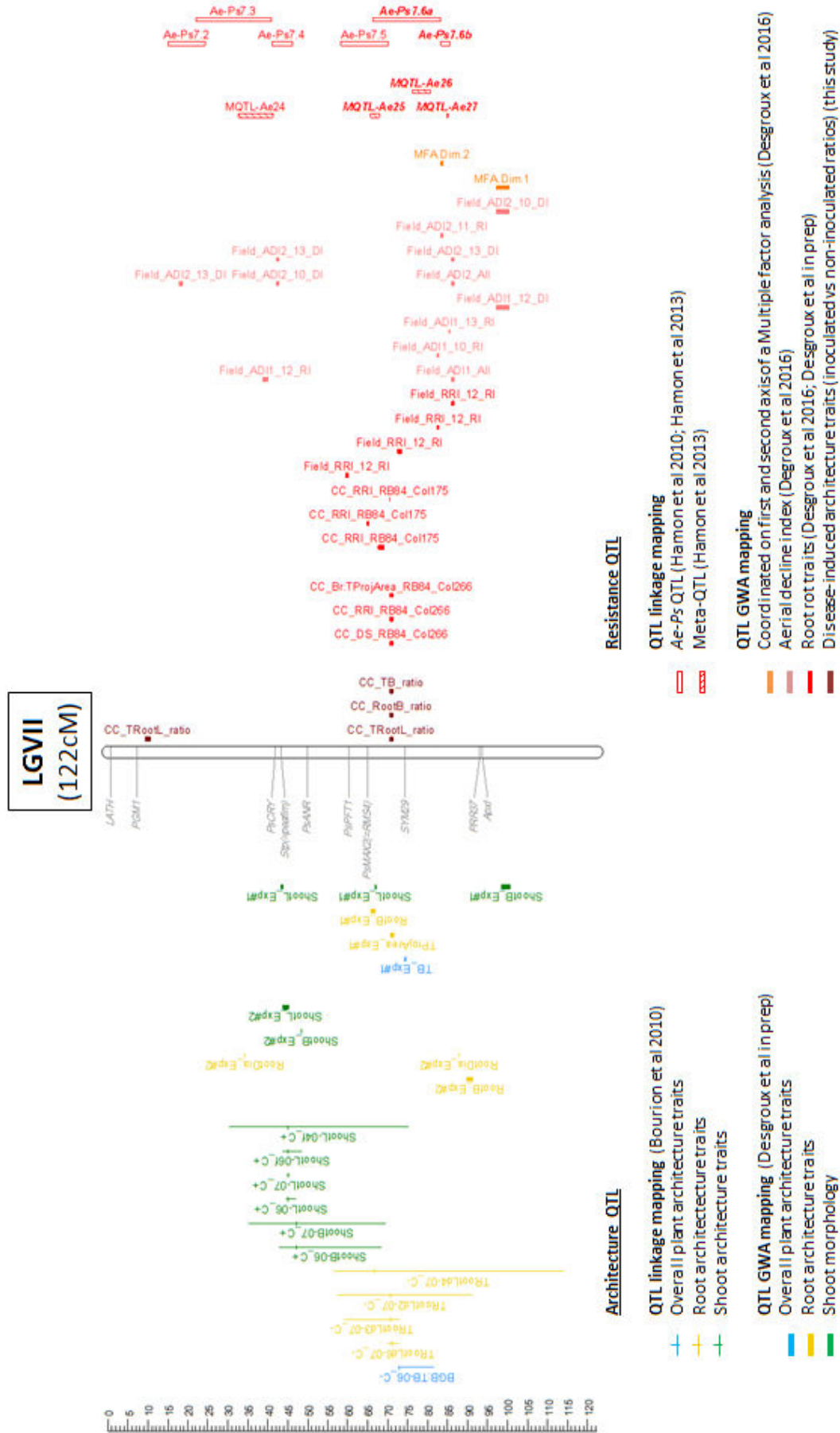
3.1.4. Comparative mapping with previous studies for plant architecture and Aphanomyces resistance

A single locus, at the *Ps115429* SNP marker on LGVII, was detected for three DIA variables in this study and three other disease variables (RRI, DS and Br:TProjArea) measured in Desgroux et al. (in prep). This locus overlapped with the major QTL *Ae-Ps7. 6* previously identified by Hamon et al. (2013). It was located close to markers associated with resistance to *A. euteiches* in controlled (RB84 strain) and field (Riec, 2012) conditions, recently identified by GWA in the pea *Aphanomyces* collection (Desgroux et al. 2016). The *Ps115429* SNP marker was not included in the study conducted by Desgroux et al. (2016), but in this study, we validated its association with RRI score variation previously observed in controlled condition experiments with the RB84 *A. euteiches* strain (data not shown). *Ps115429* was also detected by Desgroux et al. (in prep) for the total projected area variable in non-inoculated controlled conditions 8 days after sowing.

All the other loci detected for plant DIA variables overlapped with genomic regions that have already been described as controlling resistance to *A. euteiches* in pea by QTL Linkage analyses and/or association mapping (*Ae-Ps1. 3*, *Ae-Ps2. 2*, *Ae-Ps3. 1*, *Ae-Ps4. 2*, *Ae-Ps4. 3*, bottom of LGIV close to *RGA2. 75* gene, *Ae-Ps5. 2*, *Ae-Ps5. 3* and *Ae-Ps7. 2*; Figure IV.1 and Additional file IV.2). These overlapping regions were all previously detected for Aerial Decline Index (ADI) variables and five of them were also identified for root rot variables (Desgroux et al. 2016; Hamon et al. 2011; Hamon et al. 2013).

In addition to the *Ps115429*-associated locus, nine loci associated with plant architecture DIA variables from this study (one on LGI, four in LGII, two in LGIII and two in LGIV) overlapped with plant architecture traits previously detected in Bourion et al. (2010) and/or Desgroux et al. (2016), whatever the organ on which the trait was evaluated. Especially, the SNP detected for ShootL DIA variable at 45. 5 cM on LGII (*PsCam050243_32836_1352*) was in LD ($r^2 = 0.34$) with a locus detected for RootDia in healthy conditions 14 days after sowing (*PsCam029325_17487_1726*) by Desgroux et al. (in prep). The SNP detected for ShootB DIA variable at 22. 0 cM on LGIII (*PsCam058589_38986_524*) was the same than the one detected for ShootL in non-inoculated controlled conditions 8 days after sowing by Desgroux et al. (in prep).

The loci detected for DIA variables in this study were not associated with plant morphological genes reported in the literature. The SNPs detected for root or shoot DIA



Legend on next page

Figure on previous page

Figure IV.1: Comparative genetic map of genome-wide association (GWA) and previously detected linkage quantitative trait loci (QTL) for resistance to *Aphanomyces euteiches*, constitutive plant architecture and inoculated vs non-inoculated ratios of plant architecture for linkage group VII

The comparative genetic map was constructed from the projection of the 953 markers from Boutet et al. (2016) onto the consensus THMap from Desgroux et al. (2016). Only linkage group (LG) VII is shown (the other LGs are in additional file 1). Its size is indicated in cM Haldane. **To the right from each LG:** Confidence intervals (CIs) around significant markers associated with inoculated vs non-inoculated ratios of architectural traits, based on linkage disequilibrium (LD) value $r^2 > 0.2$, CIs around significant Resistance-associated markers, identified by GWA in Desgroux et al. (in prep); CIs around significant Resistance-associated markers, identified by GWA in Desgroux et al. (2016), with the RB84 strain (GWAS Col175); Projected Meta-QTL (Meta-) and QTL (Ae-Ps QTL) described in Hamon et al. (2011) and Hamon et al. (2013), hatched bars represent Meta-QTL, while blank bars represent initial QTL before meta-analysis Hamon et al. (2011) and Hamon et al. (2013). The main *Aphanomyces* resistance QTL and Meta-QTL names are in bold italic. **To the left from each LG:** Genomic positions of cloned pea genes are indicated in grey; CIs around architecture-associated markers identified in Desgroux et al. (in prep); Projected QTL for root, shoot and plant architecture traits described in Bourion et al. (2010). ADI: Aerial Decline Index; BGB:TB: below ground to total biomass ratio; Br:TProjArea: % of brown root projected area; CC: Controlled conditions; Exp#1: Greenhouse non-inoculated controlled conditions at 8 days post sowing; Exp#2: Climate chamber non-inoculated controlled conditions at 14 days post sowing; DS: Disease severity score; MFA: Multiple factor analysis; RB84: French reference strain of *A. euteiches*; RootB: Root biomass; RootDia: Average root diameter; RRI: Root Rot Index; ShootB: Shoot biomass; ShootL: Shoot length; TB: Total biomass; TProjArea: Total root projected area; TRootL: Total root length

variables on LGII were not in LD with the SNP designed in the sequence of the *A* gene controlling anthocyanins content (*PsbHLH*). However, the SNP detected for the ShootB DIA variable at 70.4 cM on LGIII (*PsCam043428_27437_755*) was in LD with the *SYM7* gene ($r^2 = 0.26$) controlling symbiosis with *Rhizobium*. The SNP detected for the NLatRoot DIA variable at 66.1 cM on LGV (*PsCam050864_33422_321*) was close to the *RGAI.5* gene.

3.2. Field experiment of the Apharchi pea line-subset

3.2.1. Subset selection for field experiment

Among the 15 pea lines of the Apharchi subset, four classes of lines were defined (Table IV.6). The class 1 included four lines with high levels of partial resistance to *A. euteiches* both in controlled conditions and in the field. All of them showed favourable DIA scores for most variables from this study and two of them also showed favourable intrinsic RSA scores (phenotypes positively correlated with disease resistance) from Desgroux et al. (in prep). The class 2 included two lines showing partial resistance in controlled conditions but susceptibility in the field. Those lines presented favourable scores for most variables of DIA and/or intrinsic plant architecture traits. The class 3 included five lines which did not show partial resistance

Table IV.6: Description of the Apharchi subset and its behaviour for resistance, intrinsic and disease-induced architecture

Line name	Class (b)	Disease Field (c)							Disease CC (d)					Intrinsic architecture (e)					disease-induced architecture traits (f)														
		Field_RRI_All			Field_AD1_All		Field_AD2_All		MFA.Dim.1	CC_RRI_13	DS_13	Br:ProjArea_13	Exp#1		Exp#2			RootDia_ratio		RootArea_ratio		RootDia_ratio		RootArea_ratio		RootTB_ratio		ShootTB_ratio		ShootL_ratio		RootTB_ratio	
		Field_RRI_All	Field_AD1_All	Field_AD2_All								RootDia	RootTB	NatRoot	RootDia	RootTB	NatRoot	RootDia_ratio	RootArea_ratio	RootTB_ratio	NatRoot_ratio	LatRoot_ratio	RootTB_ratio	ShootTB_ratio	ShootL_ratio	RootTB_ratio	ShootTB_ratio	ShootL_ratio	RootTB_ratio	ShootTB_ratio			
AeD99OSW-50-2-5 A104	1	2.1	2.1	3.3	-3.48				2.0	2.1	0.423	0.667	188.7	49	0.612	425.0	50	2%	6%	-4%	-5%	7%	10%	2%	-2%	5%	2%	10%	7%	10%	2%	-2%	5%
AeD99OSW-45-8-7 A100	1	1.8	2.0	3.3	-3.91				2.0	2.2	0.401	0.782	107.0	35	0.625	362.3	43	3%	6%	-4%	-12%	15%	19%	12%	6%	14%	6%	19%	15%	19%	12%	6%	14%
AeD9904-7 A009	1	1.8	2.2	4.1	-2.39				2.6	2.4	0.370	0.906	117.9	39	0.687	415.7	50	-14%	1%	-14%	-15%	-6%	-12%	-19%	-16%	-17%	-14%	-12%	-6%	-12%	-19%	-16%	-17%
AeD99OSW-53-3-5 A106	1	2.0	2.1	3.8	-2.66				2.8	2.7	0.421	0.726	186.7	48	0.625	367.6	46	17%	11%	4%	3%	6%	29%	-3%	7%	7%	11%	6%	6%	29%	-3%	7%	7%
PI180693 A116	2	2.4	3.2	4.4	-1.17				2.6	2.4	0.473	0.674	194.4	60	0.598	537.1	75	-18%	5%	-21%	-11%	-10%	-20%	-9%	-4%	-12%	-18%	-11%	-10%	-20%	-9%	-4%	-12%
E11 A130	2	2.9	3.0	4.7	0.13				2.5	2.7	0.393	0.717	112.1	34	0.582	368.3	40	4%	5%	-4%	-10%	9%	30%	37%	2%	36%	4%	5%	9%	30%	37%	2%	36%
AeD99OSW-37-3-4 A092	3	2.4	1.8	3.2	-1.99				3.4	3.4	0.463	0.780	76.5	45	0.621	469.3	51	-26%	12%	-33%	-13%	-22%	-16%	4%	4%	-2%	-26%	-13%	-22%	-16%	4%	4%	-2%
90-2131 A003	3	2.1	2.5	4.6	-0.94				3.2	3.4	0.459	0.742	174.0	57	0.654	583.7	64	-27%	3%	-32%	-17%	-20%	-30%	-22%	-8%	-25%	-27%	3%	-32%	-17%	-20%	-30%	-22%
AeA95-335.M.51.3 A025	3	2.4	2.1	4.1	-1.18				3.5	3.5	0.462	0.813	104.3	43	0.620	401.3	54	-40%	2%	-41%	-37%	-6%	-38%	-22%	-3%	-27%	-40%	2%	-41%	-37%	-6%	-38%	-22%
552 A127	3	2.4	2.1	3.9	-1.66				3.8	3.9	0.571	0.940	90.7	31	0.706	420.2	46	-40%	6%	-44%	-27%	-26%	-41%	-7%	-14%	-40%	6%	-44%	-27%	-26%	-41%	-7%	
BA55.88 MR08	3	2.5	2.6	4.0	-0.90				4.0	4.0	0.445	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
AeD99OSW-4-6-4 A064	4	3.3	3.6	5.8	2.19				3.3	3.3	0.493	0.926	76.3	24	0.696	324.9	31	-26%	-5%	-22%	5%	-25%	-13%	-9%	-5%	-11%	-26%	-5%	-22%	5%	-25%	-13%	
Isard A123	4	2.9	4.0	6.1	2.75				3.6	3.5	0.499	0.810	68.6	28	0.660	375.3	35	-41%	7%	-44%	-15%	-34%	-31%	-20%	-12%	-24%	-41%	7%	-44%	-15%	-34%	-31%	
Baccara A125	4	3.4	3.9	5.5	3.15				4.0	3.5	0.525	0.879	103.7	33	0.719	359.5	37	-36%	12%	-43%	-24%	-27%	-39%	-33%	-20%	-35%	-36%	12%	-43%	-24%	-27%	-39%	
VAVILOV-D261=K4088	-	NA	NA	NA	NA				3.7	3.8	0.506	0.697	218.9	57	0.605	609.6	72	-41%	4%	-43%	-18%	-31%	-43%	-9%	-4%	-18%	-41%	4%	-43%	-18%	-31%		

Color scale (g)

unfavorable score

favorable score

(a) Line codes as described in Desgroux et al (2016); (b) class of pea lines based on resistance levels in controlled conditions and field resistance. 1: partially resistant in controlled and field conditions; 2: partially resistant in controlled conditions and susceptible in field conditions; 3: susceptible in controlled conditions and partially resistant in field conditions; 4: susceptible in controlled and field conditions; (c) disease scores in field conditions described in Desgroux et al (2016); (d) disease score in controlled conditions described in Desgroux et al (in prep); (e) intrinsic root architecture in controlled conditions described in Desgroux et al (in prep), Exp#1: Greenhouse non-inoculated controlled conditions at 8 days post sowing, Exp#2: Climate chamber non-inoculated controlled conditions at 14 days post-sowing; (f): inoculated vs non-inoculated ratios of architectural traits described in this study; (g) red to green color scale from unfavorable to favorable values regarding to resistance level; names of variables are coded as described in Table IV.1, disease-induced architecture traits are inoculated vs. non-inoculated ratios

in controlled conditions but moderate to high levels of resistance in the field. In this class, plant architecture scorings were variable according to the lines. Two of them had intermediate to favourable scores for most of the intrinsic or disease-induced RSA traits. The class 4 included three susceptible lines in both controlled conditions and in the field, for whom RSA phenotypes were negatively correlated with resistance. One pea line was not classified since its level of resistance was not evaluated in the field. The line, susceptible in controlled conditions, showed yet favourable and unfavourable values, for intrinsic and DIA, respectively.

3.2.2. Level of soil infection by *A. euteiches*

Statistical analysis of the RRI variable on the Apharphi subset over all the environments revealed that Ae- field conditions were slightly infected by *A. euteiches* and they were significantly less infested than Ae+ field conditions (Tuckey test p -value < 0.05; Table IV.7; Figure IV.2). This result confirmed that the two field conditions well represented low and high level of soil infection by *A. euteiches*.

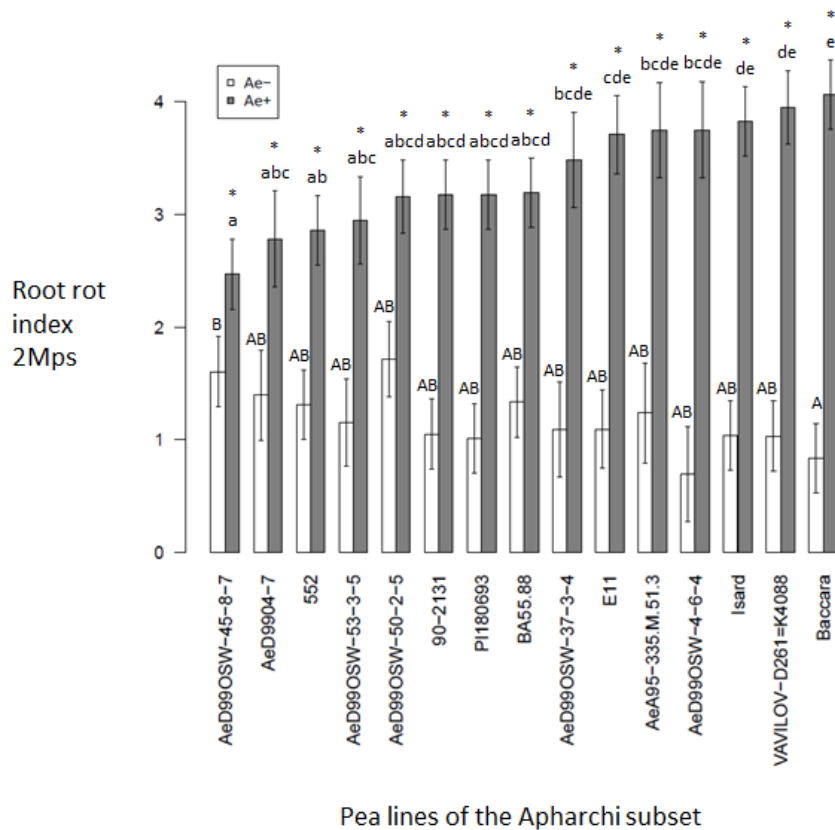


Figure IV.2: Root rot index of the Apharphi subset in lowly and highly infested fields

*: significant difference between lowly (Ae-) and highly (Ae+) infested fields (p -value < 0.05); For each level of infestation, lines sharing the same letter are not significantly different (p -value < 0.05)

Table IV.7: Statistical analysis results of plant architecture traits data in the Apharchi subset

Disease resistance	Trait (a)	Stage (b)	Environments (c)		Number of pea lines (d)	G effect (e,f)	Ae effect (e,g)	E effect (e,h)	R effect (e,i)	P effect (e,j)	G x Ae effect (e,k)	G x E effect (e,l)	Ae x E effect (e,m)	G x E x Ae effect (e,n)
			RI_14, DI_14, RI_15, DI_15	RI_14, DI_14, RI_15										
resistance	RRI	2Mps	RI_14, DI_14, RI_15, DI_15	DI_14, DI_15	15	*	***	***	ns	-	***	ns	ns	ns
	AD1	Flo	RI_14, DI_14, DI_15	RI_14, DI_14, DI_15	15	***	-	***	***	-	-	***	-	-
	AD2	Flo	RI_14, DI_14, DI_15	RI_14, DI_14, DI_15	15	***	-	***	**	-	-	**	-	-
Root morphology	TapRootL	2Mps	RI_14, DI_14, RI_15	RI_14, DI_14, RI_15	15	***	***	***	**	ns	*	ns	ns	ns
	NLatRoot	2Mps	RI_14, DI_14, RI_15	RI_14, DI_14, RI_15	15	***	***	***	***	ns	ns	***	***	***
	RootB	2Mps	DI_14, DI_15	DI_14, DI_15	15	***	***	***	ns	-	**	**	ns	ns
Shoot morphology	ShootL	2Mps	DI_14, RI_14	DI_14, RI_14	15	***	***	***	ns	ns	***	**	***	*
	NNode	2Mps	DI_14, RI_14	DI_14, RI_14	15	***	ns	***	ns	ns	***	**	***	**
	ShootB	2Mps	DI_14, RI_14	DI_14, RI_14	15	***	***	***	ns	-	ns	ns	ns	ns
Yield and its components	SeedYield	Sm	DI_14, DI_15	DI_14, DI_15	15	***	***	ns	ns	-	***	ns	***	ns
	NPodNode	Sm	RI_14, DI_14, DI_15	RI_14, DI_14, DI_15	15	***	***	***	**	*	***	***	***	***
	NPod.PodNode	Sm	RI_14, DI_14, DI_15	RI_14, DI_14, DI_15	15	***	***	***	***	ns	***	***	***	***
	NSeed.Pod	Sm	RI_14, DI_14, DI_15	RI_14, DI_14, DI_15	15	***	***	***	***	ns	***	***	***	***
	TSW	Sm	RI_14, DI_14, DI_15	RI_14, DI_14, DI_15	15	***	***	***	ns	ns	***	***	***	***
						1847	***	***	***	ns	ns	***	***	***

(a) Traits were coded as described in Table IV.1; (b) Stage of plant while measuring traits, 2Mps: two months post sowing; Flo: flowering; Sm: Seed maturity; (c) environments where traits were actually assessed, RI: Riec-sur-Belon, France; DI: Dijon-Epoisses, France; 14: 2014; 15: 2015; (d) Number of pea lines actually assessed and number of total observations in the 266-pea-line collection; (e) ANOVA significance p-value codes: 0 < * < ** < *** < 0.001 < **** < 0.01 < ** < 0.05 < ns < 1; (f) Genotype effect; (g) Aphanomyces infestation of the field level effect; (h) Environment effect; (i) Replicate effect; (j) Plant effect when accurate; (k) Genotype by infestation level of the field interaction effect; (l) Genotype by environment interaction effect; (m) Infestation level of the field by environment interaction effect; (n) Genotype by environment by infestation level of the field triple interaction effect

3.2.3. Disease impact on pea root and shoot architecture

Two months post sowing, the level of soil infection by *A. euteiches* had a significant effect (p -value < 0.001) on the six root and shoot architecture variables scored (Table IV.7). Genotype and environment effects were highly significant (p -value < 0.001) for the six variables. Interactions between G and/or Ae and E (p -value < 0.01) were also observed for NLatRoot, RootB, ShootL and NNode. For most of the pea lines of the Apharchi subset, smaller tap root, less lateral roots and less root biomass production was observed in highly infested fields than in lowly infested fields (Figure IV.3). However, root scores of some lines, especially BA55. 88, were not or little reduced in highly infested fields compared to lowly infested ones.

Few lines had significantly different shoot scores (p -value < 0.05) between highly and lowly infested fields for ShootL and NNode. A few of the most resistant pea lines showed a longer shoot part and a larger number of nodes when infested by *A. euteiches*, while the most susceptible lines showed a smaller shoot part and a smaller number of nodes (Figure IV.4). Inversely, infection by *A. euteiches* led to a significant (p -value < 0.05) loss of shoot biomass for most of pea lines of the Apharchi subset (Figure IV.4). The resistant lines *PII80693* and *AeD99OSW-50-2-5* did not show any significant loss of shoot biomass in highly compared to lowly infested fields.

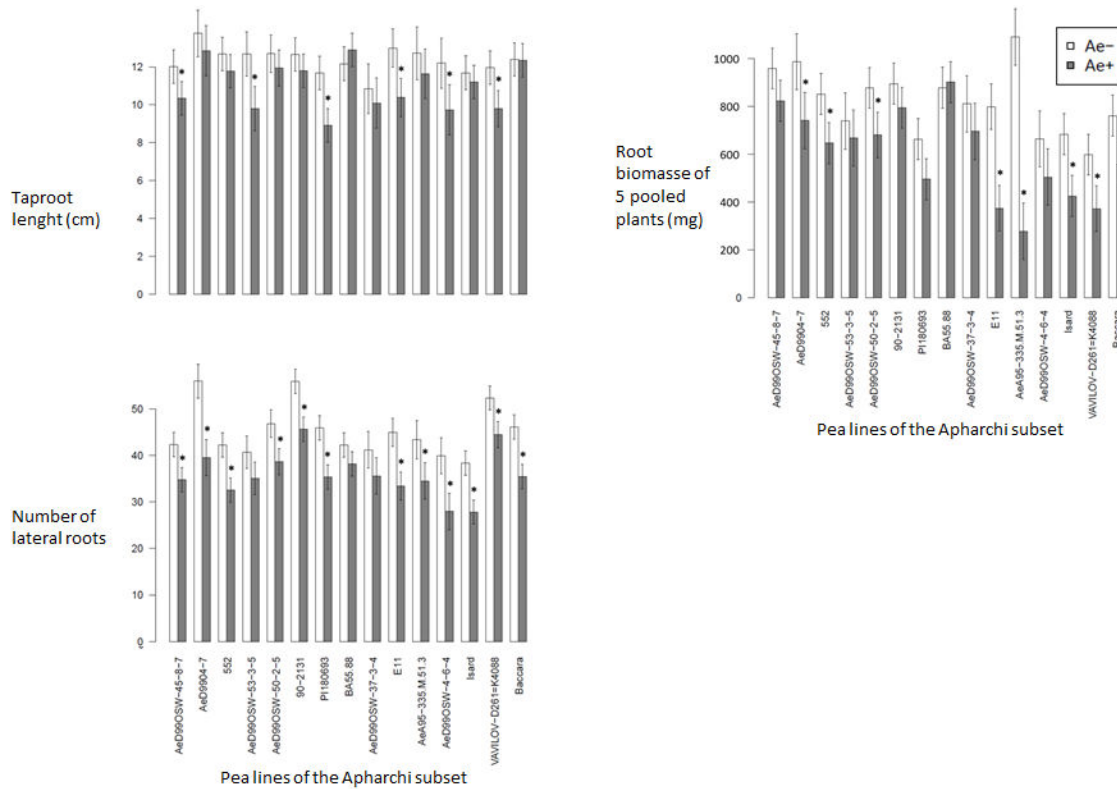


Figure IV.3: Root architecture traits of the Apharchi panel in lowly and highly *Aphanomyces* infested fields

*: significant differences between lowly (Ae-) and highly (Ae+) infested fields (p -value < 0.05)

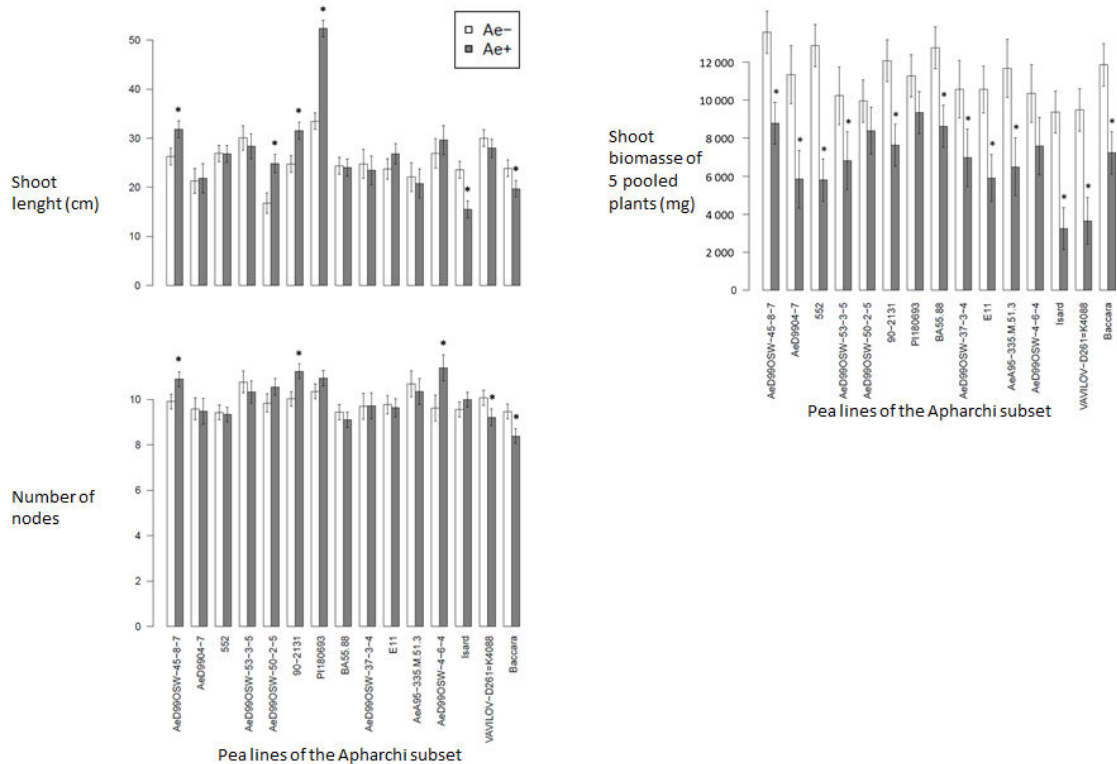


Figure IV.4: Shoot architecture traits of the Apharchi panel in lowly and highly *Aphanomyces* infested fields

*: significant differences between lowly (Ae-) and highly (Ae+) infested fields (p -value < 0.05)

3.2.4. Disease impact on pea yield and its components

Infection by *A. euteiches* had a highly significant (p -value < 0.001) effect on yield loss of the 15 pea lines of the Apharchi subset over the two environments (Table IV.7). *A. euteiches* impact depended on pea lines and was high on the most resistant pea lines (-55% for *AeDOSW-45-8-7*) and very high for the most susceptible pea lines (-99% for *VAVILOV-D261=K4088*) (Figure IV.5).

The yield components studied were impacted by *A. euteiches* on all the lines (Ae effect p -value < 0.001). Genotype effects and GxAe and/or E interactions were observed on the yield components (p -value < 0.001; Table IV.7). NPodNode was the yield component the most impacted by *A. euteiches* on all pea lines of the Apharchi subset (Figure IV.5). Impact of the disease on NPod, PodNode and NSeed. Pod depended on the pea lines and tended to be higher on susceptible lines (p -value < 0.05). TSW was significantly lower for almost all pea lines of the subset in highly than in lowly infested fields, except for *VAVILOV-D261=K4088*.

4. Discussion

In this study, we report impacts of *Aphanomyces* disease on plant architecture and comparison of genetic determinants controlling plant resistance and root morphological response to disease, in a large pea-line collection.

4.1. Plant architecture induced by infection is highly correlated with resistance and less related to intrinsic architecture

In this study, we showed that **root and shoot DIA traits under infected conditions** by *A. euteiches* vs. uninfected conditions **was highly correlated to intrinsic resistance** level in the pea collection studied. Root and shoot architecture of resistant pea lines were actually less affected than in susceptible ones, as indicated by highly significant negative correlations observed between susceptibility and most of the DIA variables (inoculated vs. non-inoculated ratios) studied and positive correlations between susceptibility and root diameter ratios. Positive correlations between susceptibility and root diameter DIA scores mean that the average root diameter of susceptible pea lines increased with infection. Root diameter, at the young stage of observation, was mainly composed of taproot diameter and was balanced by smaller diameter of lateral roots. Thus, a higher average root diameter in this study could reflect a higher global root diameter, but more probably reflected a smaller number of shorter lateral roots. The lower impact of the disease on plant architecture in resistant lines may be

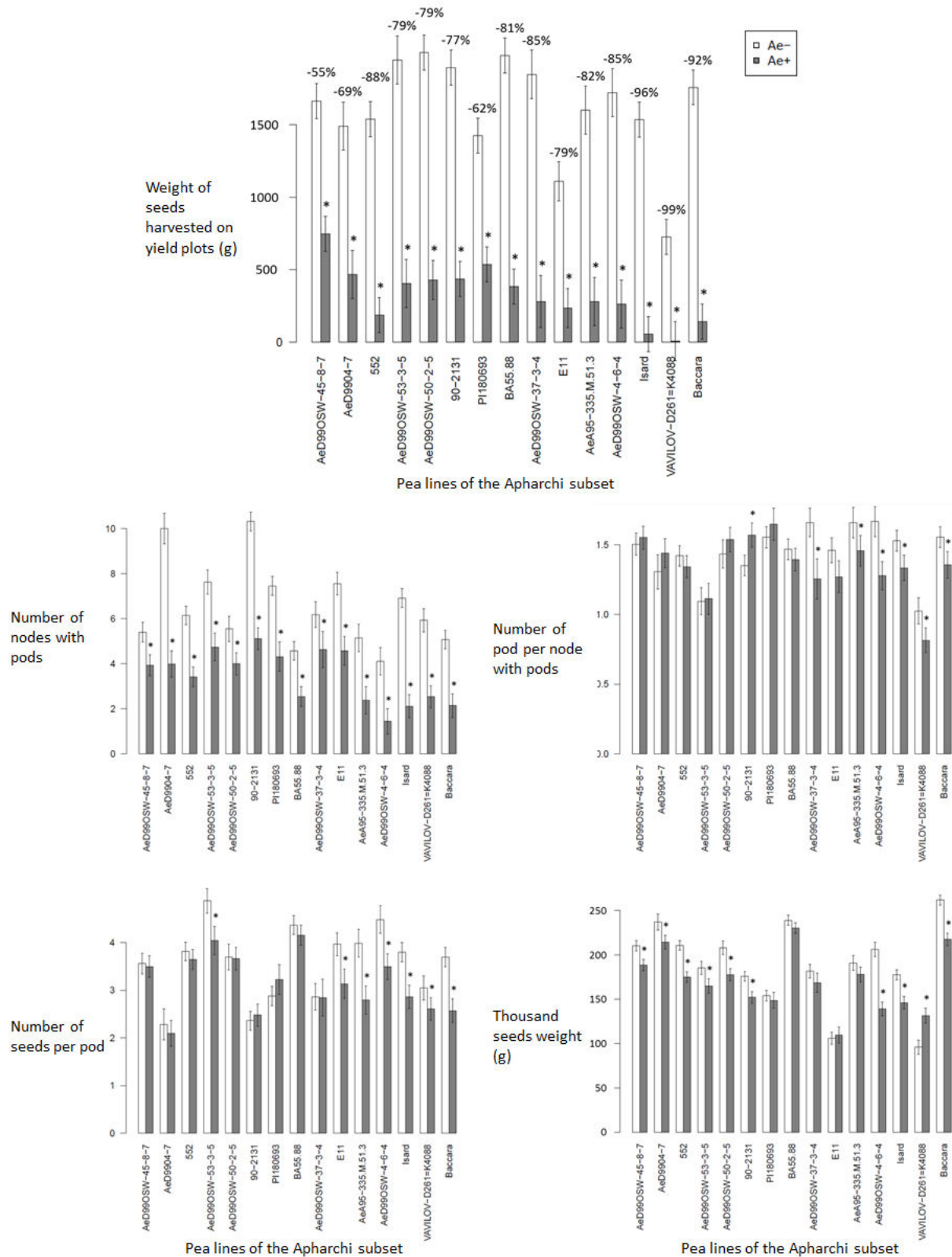


Figure IV.5: Yield traits of the Apharchi panel in weakly and highly *Aphanomyces* infested fields

*: significant differences between lowly (Ae-) and highly (Ae+) infested fields (p -value < 0.05)

most likely a consequence of resistance-mediated mechanisms. However, a wide genetic variability was observed in the collection for several DIA variables (total root length, root biomass). For these variables, several pea lines (*AeA95-376-M. 33. 1*, *AeD99OSW-27-7-6* and *AeD99OSW-53-3-5* for example) had positive DIA values, meaning that they had higher architecture score values in inoculated than in non-inoculated conditions. Those lines had the ability to preserve their root system or induce more roots or biomass as a response process to infection. These results are in accordance with previous reports on *M. truncatula* accessions responses to infection by *A. euteiches* (Djébali et al. 2009), or on *Pisum sativum* accessions responses to infection by *Fusarium solani* (Kraft and Boge 2001). We may hypothesize that root vigor contribute to higher tolerance of the plant to the disease.

We also showed that **plant architecture induced by *A. euteiches* infection was poorly to moderately correlated with intrinsic plant architecture** in the collection. Intrinsic and DIA traits are two distinct types of traits, the first one representing the phenotypic plasticity of the plant in response to the infection and the other one its intrinsic characteristics in pathogen-free conditions. More root or shoot score losses in inoculated vs. non inoculated conditions were observed for pea lines with smaller intrinsic shoot and/or lateral root length variables. This observation is in accordance with the higher impact of the disease often observed on sickly plants, and with the previous results from Desgroux et al. (in prep) reporting that a denser intrinsic RSA is associated with a higher resistance. More losses were observed for most plant DIA variables when biomass allocation to the root system was higher (positive correlation between DIA variables and RootB:TB scores). Genetic variability was observed in the collection for the two types of traits, with pea lines having favorable scores for intrinsic and disease-induced architecture (*AeD99OSW-50-2-5* for example) or for only one of the two types of trait (*AeD99OSW-53-3-5* for example).

4.2. Most of genomic regions controlling disease-induced plant architecture were common to those controlling resistance but different from those implied in intrinsic plant architecture

This study showed that **genetic determinisms of resistance and DIA are mainly common**. All the 15 markers associated with six DIA variables, located over six LGs, overlapped with genomic regions previously identified for field and/or controlled condition resistance, by GWAS (Desgroux et al. 2016) or QTL mapping analysis (Hamon et al. 2013). A causal relationship between resistance and DIA would most likely explain the common genetic determinism observed, suggesting single causal genes in some overlapping genomic regions.

However, it would be interesting to identify marker haplotypes at loci controlling DIA that would be specifically found in pea lines able to preserve their architecture under infection. Linkages between several genes may also underlie the common genomic regions and explain the shared genetic determinism. Two of the 15 loci, located on LGVII (position 71.3 cM) and LGIV (position 42.5 cM) were consistently detected for DIA variables of three traits. The first locus, at marker *Ps115429* on LGVII was detected with major allelic effects in Desgroux et al. (in prep) for three disease variables, and overlapped with the major QTL *Ae-Ps7.6* described by Hamon et al. (2011) and Hamon et al. (2013). This finding greatly confirms the importance of the major QTL *Ae-Ps7.6* and provides a closely-linked marker to the QTL to reduce its interval by fine mapping. The second locus, at marker *PsCam021147_11860_302* on LGIV, overlapped with *Ae-Ps4.2*, a minor QTL detected exclusively for ADI (Hamon et al. 2011; Hamon et al. 2013). It was detected at an early stage only for shoot DIA traits. ADI is a disease resistance score based on shoot yellowing and shoot dwarfism. Thus, we may hypothesize that this locus, expressed at an early stage, may act on limiting dwarfism symptoms. In this study, 14 of the 15 loci associated with DIA variables overlapped with QTL mapped for ADI variables. These overlaps suggest a role of DIA in the reduction of disease impact on the aerial part of the plant.

Genetic determinisms of disease-induced and intrinsic plant architecture were quite different. Nine of the 15 loci identified for DIA traits overlapped with QTL controlling intrinsic plant architecture traits previously detected by GWAS (Desgroux et al. in prep) or QTL mapping (Bourion et al. 2010). Different genetic control is not surprising, as correlations between LSM means of those traits were low or not significant. On LGII, a cluster of root and shoot DIA loci was detected in a genomic region including the *A* gene controlling anthocyanins content. However, the detected loci were not in LD with the SNP designed in the *A* gene, as suggested previously for resistance by Desgroux et al. (2016). On LGIII, a SNP was detected for ShootB DIA trait in LD with the SNP designed in the *Sym7* gene, coding for a transcription factor involved in nodulation and overlapped with QTL detected for shoot traits in Bourion et al. (2010). We hypothesized that processes involved in *Rhizobium* recognition during symbiosis could also be involved in root pathogens recognition. For further research, it would be interesting to study infection in parallel to symbiosis and to compare processes involved in both mechanisms.

4.3. Resistance and disease-induced plant architecture were related to a lower impact of Aphanomyces root rot on yield in pea

The relationship between root architecture and tolerance to *A. euteiches*, suggested in this study by overlaps between genomic regions controlling DIA and ADI variables, was explored in a subset of 15 pea lines contrasted for resistance and plant architecture traits according to Desgroux et al. (2016) and Desgroux et al. (in prep).

This study showed the impact of *A. euteiches* on pea architecture traits, such as taproot length, number of lateral roots and root and shoot biomass, under highly infested field conditions. Variable levels of impacts were observed according to the 15 lines and the four classes of lines defined in this study. Particularly, some lines with moderate levels of resistance (*AeD9904-7* and *90-2131*) had better ability to preserve their root system from high losses due to the disease, resulting in lower losses of yield components than those of lines with similar levels of susceptibility. Description and selection of marker haplotypes associated with disease-induced RSA traits in these lines would be of high interest to increase tolerance to the disease by breeding. Vigor of root system is of great importance for plants as it allows them to anchor in the soil and to absorb water and nutrient for growth. The impact of Aphanomyces root rot on RSA decreases the plant ability to absorb water and nutrients, resulting in an unfavorable impact on shoot vigor of different lines, as shown in this study on length, number of nodes and biomass at early stage. Intrinsic resistance was not sufficient on its own to decrease Aphanomyces impact on yield. Thus, further pea varieties improvement should take into account both intrinsic resistance and disease-induced RSA to improve tolerance to *A. euteiches*, thanks to SNP markers detected in the present study. Further research should improve knowledge on the best marker haplotypes to combine on this purpose.

5. Conclusion

This study brought new tools for a better understanding of genetic links between root disease resistance and disease-induced RSA. A SNP marker, previously detected for both improved resistance to *A. euteiches* and greater root projected area, was also detected in this study associated to less impacted plant architecture in controlled conditions. Thus, this SNP marker looked relevant to use for marker assisted selection. Pea lines, with good levels of resistance to *A. euteiches* and/or with less impacted plant architecture traits and yield after infection, were identified and could be interesting for breeders to improve resistance to *A. euteiches* in pea varieties.

Chapitre V

Synthèse, discussion générale
et perspectives

Chapitre V : Synthèse, discussion générale et perspectives

La compréhension des **bases génétiques** et des **mécanismes de résistance partielle** aux maladies chez les plantes, ainsi que de l'implication de l'**architecture** des organes ou des compartiments affectés par la maladie, dans les processus de réduction des épidémies représentent des enjeux scientifiques majeurs. Cette connaissance permettra le développement de variétés agronomiques présentant des résistances et/ou des tolérances efficaces pour lutter durablement contre les bioagresseurs.

Plus particulièrement, l'étude du pathosystème pois/*A. euteiches* présente un intérêt économique et agronomique majeur pour le développement de la culture du pois, et plus généralement des légumineuses, en France et en Europe. En effet, la stratégie la plus appropriée pour lutter contre *A. euteiches* repose sur l'exploitation des caractéristiques génétiques de la plante, notamment la résistance partielle, associée à des pratiques culturales limitant le développement de la pression pathogène.

Ma thèse s'insère dans le cadre de la compréhension des relations entre résistance, tolérance et architecture racinaire afin de proposer à terme des variétés combinant des caractères de résistance partielle et d'architecture pour mieux lutter contre *A. euteiches*.

Dans le cadre de ce travail de thèse, j'ai contribué à cet enjeu de recherche en tentant de répondre aux questions de recherche suivantes : (i) Existe-t-il une diversité génétique de facteurs impliqués dans la résistance à *A. euteiches* ? (ii) L'architecture racinaire constitutive de la plante est-elle corrélée à la résistance du pois à *A. euteiches* via un contrôle génétique commun ? (iii) L'architecture racinaire « induite » est-elle corrélée à la résistance et partage-t-elle des déterminants génétiques communs ? peut-elle être impliquée dans la tolérance de la plante à la maladie ?

Ce chapitre résume tout d'abord les principaux résultats acquis au cours de ma thèse permettant d'apporter des éléments de réponse aux questions de recherche posées. Dans un deuxième temps, il discute de l'apport de ces résultats pour la communauté scientifique et la sélection variétale et soulève des hypothèses sur (i) les mécanismes sous-jacents à la résistance ; (ii) l'implication de l'architecture racinaire dans le contrôle génétique de la résistance et la tolérance aux agents pathogènes racinaires ; (iii) l'amélioration génétique de la

résistance et de la tolérance du pois à *A. euteiches*. Enfin, ce chapitre propose des perspectives et nouvelles pistes de recherche pour améliorer les connaissances sur le rôle de l'architecture racinaire dans la résistance et la tolérance aux maladies racinaires, ainsi que sur les mécanismes sous-jacents, en mettant en relation les programmes de recherche déjà en cours.

1. Synthèse des résultats de la thèse

Mon travail de thèse repose sur plusieurs jeux de données acquis en amont ou au cours de la thèse :

- des données phénotypiques sur la collection « Aphanomyces » de 175 lignées de pois obtenues en conditions contrôlées pour la résistance à deux souches d'*A. euteiches* de référence (Ae109 et RB84) en 2009 ;
- des données phénotypiques obtenues sur cette même collection au champ sur quatre années (2010-2013) et deux à trois sites par an pour la résistance à *A. euteiches* ;
- des données phénotypiques sur la collection conjointe « Aphanomyces-Racines » de 266 lignées de pois (incluant les 175 précédentes) obtenues en conditions contrôlées pour la résistance à une souche d'*A. euteiches* (RB84) en 2013 ;
- des données phénotypiques obtenues sur cette même collection de 266 lignées en conditions contrôlées pour l'architecture de la plante en présence et en absence d'infection par *A. euteiches* en 2013, respectivement à un et deux stades de développement de la plante,
- des données phénotypiques sur un sous-set de 15 lignées de la collection conjointe, obtenues au champ sur deux années (2014-2015) et un à deux sites, pour la résistance, l'architecture de la plante et le rendement sur des parcelles faiblement ou fortement infestées par *A. euteiches* ;
- des données génotypiques sur l'ensemble du matériel végétal précédemment cité, obtenues à partir de la puce GENOPEA Illumina Infinium développée récemment (Tayeh et al. 2015a), incluant 13,2k marqueurs SNP couvrant l'ensemble du génome ;
- des données génotypiques sur l'ensemble du matériel issues d'un génotypage KASP™ de 953 marqueurs SNP développés récemment plus particulièrement dans les zones de QTL de résistance aux maladies précédemment identifiées chez le pois (Boutet et al. 2016).

1.1. Variabilité génétique et corrélations entre variables de résistance à *A. euteiches* et d'architecture racinaire de la plante chez le pois

Les jeux de données phénotypiques cités précédemment ont été exploités afin de comparer la variabilité génétique observée chez *P. sativum* pour la résistance à *A. euteiches*, l'architecture constitutive de la plante et son architecture induite en réponse à l'infection par *A. euteiches*. Une grande variabilité génétique a été mise en évidence pour la résistance à *A. euteiches* au champ et en conditions contrôlées, mais également pour plusieurs caractères d'architecture constitutive et induite de la plante, au sein des collections étudiées. Bien qu'aucune nouvelle source de résistance à *A. euteiches* n'ait été détectée, certains génotypes, issus du croisement entre plusieurs sources de résistance identifiées dans des études précédentes, ont montré un niveau de résistance partielle supérieur à celui de leurs parents respectifs.

La comparaison des données phénotypiques de résistance et d'architecture de la plante montre des corrélations entre notes de résistance et de caractères d'architecture constitutive de la plante. Ainsi, les résultats ont permis de montrer qu'un système racinaire plus développé (en longueur et nombre de racines) était corrélé à un niveau plus élevé de résistance partielle du pois à *A. euteiches* en conditions contrôlées.

La comparaison des données phénotypiques de résistance à celles d'architecture de la plante obtenues en conditions inoculées et non inoculées a permis de mettre en évidence des corrélations élevées entre le niveau de résistance et la moindre perte de système racinaire, concernant notamment la longueur des racines, la biomasse racinaire, le nombre de racines latérales, la biomasse aérienne ou encore la part de la biomasse allouée au système racinaire. Certains génotypes, comportant des niveaux élevés ou non de résistance partielle, ont même présenté un système racinaire plus développé en conditions inoculées qu'en conditions non inoculées. Ces résultats ont suggéré que l'architecture induite pourrait constituer un caractère d'intérêt pour limiter les symptômes au niveau aérien et les pertes de rendement, en permettant de conserver une certaine capacité de la plante à puiser ses ressources.

La comparaison des données phénotypiques d'architecture constitutive et induite de la plante a montré généralement de faibles niveaux de corrélations, suggérant davantage une indépendance de ces deux types de caractères. Des génotypes présentant des systèmes racinaires développés n'ont pas forcément préservé leur système racinaire en réponse à l'infection. Ceux ayant des systèmes racinaires constitutifs peu développés les ont le plus souvent réduits en réponse à l'infection.

Table V.1 : Synthèse et comparaison du nombre de locus détectés pour chaque variable de résistance et d'architecture chez le pois

Variables		Résistance/Architecture induite		Architecture constitutive		Variables		Toutes variables		INR conditions contrôlées coll_175		INR champ coll_175		IDA champ coll_175		Note maladie individuelle coll_266		% de nécrose racinaire par analyse d'image coll_266		Architecture induite coll_266 (toutes var.)	
		Nombre de locus détectés par GWAS		Nombre de locus détectés par GWAS partageant au moins un SNP commun entre caractères de résistance / architecture induite et d'architecture constitutive (Nombre de locus chevauchants)		Nombre de locus détectés par GWAS partageant au moins un SNP commun entre caractères de résistance / architecture induite et d'architecture constitutive (Nombre de locus chevauchants)		Nombre de locus détectés par GWAS		Nombre de locus détectés par GWAS		Nombre de locus détectés par GWAS		Nombre de locus détectés par GWAS		Nombre de locus détectés par GWAS		Nombre de locus détectés par GWAS		Nombre de locus détectés par GWAS	
Variables	Nombre total de locus GWAS	Chevauchant avec les QTL de résistance précédents (a)	Chevauchant avec les QTL d'architecture précédents (b)																		
Nombre total de locus GWAS		74	6	15	33	5	7	15													
Chevauchant avec QTL de résistance précédents (a)		56	5	10	25	4	4	12													
Chevauchant avec QTL d'architecture précédents (b)		44	3	10	20	2	6	8													
Toutes variables		76	52	48					9 (16)	0 (1)	2 (3)	3 (6)	1 (1)	1 (2)	3 (5)						
TProjArea		8	7	4					2 (3)	0 (1)	0 (1)	1 (1)	1 (1)	1 (1)	1 (1)						
RootDia		14	11	11					1 (3)			0 (1)			1 (2)						
TRootL		4	2	3					0 (1)		0 (1)										
NLatRoot		1	1	0																	
LatRootL		3	2	1					0 (1)		0 (1)										
RootB		10	5	8					0 (1)					0 (1)							
ShootB		26	15	16					2 (5)		1 (1)	1 (2)		0 (1)	0 (1)						
ShootL		13	9	9					1 (2)					0 (1)	1 (1)						
TB		9	8	5					2 (4)			2 (3)		0 (1)							
RootB:TB		2	1	1					1 (2)		1 (1)			0 (1)							

Nombre de locus détectés par GWAS partageant au moins un SNP commun entre caractères de résistance / architecture induite et d'architecture constitutive (Nombre de locus chevauchants)

TProjArea : aire projetée totale ; RootDia : diamètre moyen des racines ; TRootL : Longueur total de racine ; NLatRoot : nombre de racines latérales ; LatRootL : longueur moyenne des racines latérales ; RootB : biomasse racinaire ; ShootB : biomasse aérienne ; ShootL : hauteur de plante ; TB : biomasse totale ; RootB:TB : % de biomasse allouée au système racinaire ; INR : indice de nécrose racinaire ; IDA : indice de dépérissement aérien ; coll_175 : collection de 175 lignées ; coll_266 : collection de 266 lignées. (a) QTL de résistance décrit dans Hamon et al. (2011) et Hamon et al. (2013) ; (b) QTL d'architecture décrit dans Bourion et al. (2010)

1.2. Cartographie comparée des locus de résistance à *A. euteiches* et d'architecture racinaire de la plante

La diversité génétique observée au sein des collections étudiées, pour les caractères de résistance et d'architecture mesurés ainsi que le génotypage dense de l'ensemble de ces lignées ont permis d'effectuer des études de génétique d'association génome entier en utilisant un modèle multi-locus, pour l'ensemble de ces caractères. Ces études ont montré que la résistance à *A. euteiches* et l'architecture de la plante chez le pois étaient des caractères fortement polygéniques, contrôlés par un grand nombre de locus. Au total, 74 et 76 locus ont été détectés par GWAS pour l'ensemble des variables étudiées, respectivement, de résistance ou d'architecture induite, et d'architecture constitutive de la plante (Table V.1).

L'étude comparative de ces locus identifiés par GWAS, avec les QTL détectés dans les études précédentes pour la résistance et l'architecture (respectivement, Hamon et al 2013 et Bourion et al 2010), a montré de nombreuses régions chevauchantes entre résistance et architecture (Table V.1). Dans cette étude, parmi les 76 locus détectés pour l'architecture, 52 sont chevauchants avec des intervalles assez larges de QTL de résistance identifiés précédemment (Hamon et al. 2011; Hamon et al. 2013). Parmi les 74 locus détectés dans cette étude pour la résistance et l'architecture induite, 44 sont chevauchants avec les intervalles assez larges de QTL d'architecture détectés précédemment (Bourion et al. 2010). L'analyse de génétique d'association, utilisant un génotypage dense et exploitant des recombinaisons historiques, a permis de diminuer les intervalles de confiance des QTL. Ainsi, la comparaison de l'ensemble des locus identifiés par GWA dans cette étude pour toutes les variables de résistance et d'architecture étudiées, a montré que seuls 16 d'entre eux ont été détectés à la fois pour la résistance et l'architecture, dont 9 partagent au moins un marqueur en commun (Table V.1). Seul l'un des 16 locus a été détecté avec le même marqueur à la fois pour la résistance ou l'architecture induite et un caractère d'architecture racinaire constitutive. Ces résultats ont montré davantage d'interdépendance entre le déterminisme génétique de la résistance et celui de l'architecture induite qu'avec celui de l'architecture constitutive.

Les travaux réalisés pendant la thèse ont donc permis (i) de préciser les intervalles de confiance des QTL de résistance à *A. euteiches* et d'architecture aérienne et racinaire chez le pois ; (ii) de mettre en évidence des relations génétiques entre allèles de résistance et d'architecture, en relation avec les niveaux de corrélations modérés et élevés entre données de résistance et, respectivement, d'architecture constitutive et induite ; et (iii) d'identifier des marqueurs moléculaires et haplotypes de marqueurs, utiles en sélection assistée par marqueurs

pour la création de nouvelles variétés de pois combinant des allèles de résistance, d'architecture induite et d'architecture constitutive défavorables à la maladie.

1.3. Exploitation de la variabilité génétique pour diminuer les pertes de rendement

L'impact de l'architecture constitutive et induite de la plante, en combinaison ou non avec la résistance, pour limiter les symptômes aériens et pertes de rendement a été étudié dans un sous-set de 15 génotypes contrastés pour les différents caractères étudiés. Les résultats d'évaluation de ces génotypes au champ ont montré un impact significatif de la maladie sur l'architecture de la plante et le rendement en conditions fortement infestées par rapport aux conditions faiblement infestées. Néanmoins, certains génotypes, sélectionnés pour leur niveau de résistance et/ou leur architecture constitutive et/ou induite défavorable(s) à la maladie en conditions contrôlées n'ont pas montré de valeurs significativement différentes entre conditions faiblement et fortement infestées, pour plusieurs variables d'architecture racinaire et de composantes de rendement. Ce constat peut s'expliquer par la résistance mais également potentiellement par l'architecture chez les génotypes ne présentant pas des niveaux élevés de résistance en conditions contrôlées (classe 3). Le niveau de perte de rendement dans les parcelles fortement infestées est apparu donc dépendant : (i) du niveau de résistance intrinsèque de la lignée de pois ; (ii) de l'architecture constitutive du système racinaire produit (biomasse et longueur de racines en conditions saines) ; et (iii) du niveau de réponse induite de l'architecture du système racinaire face à l'infection, pour un même niveau de résistance ou sensibilité initial.

2. Apports des résultats de la thèse en recherche et en sélection

Ce travail de thèse a permis de préciser l'organisation structurale des locus contrôlant la résistance partielle du pois à *A. euteiches* et l'architecture de la plante, par rapport aux études précédentes (Bourion et al. 2010; Hamon et al. 2011; Hamon et al. 2013).

D'une façon générale, les résultats obtenus permettent d'apporter des connaissances sur :

- ✓ l'interdépendance ou l'indépendance des QTL de résistance aux maladies et d'architecture, au niveau racinaire chez les végétaux ;
- ✓ les hypothèses sur les fonctions sous-jacentes aux QTL de résistance aux maladies et d'architecture de la plante chez les végétaux ;

- ✓ la combinaison de caractères de résistance et d'architecture, et les marqueurs associés potentiellement utilisables en sélection, pour la création de variétés de pois résistantes et/ou tolérantes à *A. euteiches*.

Ces connaissances sont primordiales pour l'utilisation de combinaisons d'allèles d'architecture et de résistance aux maladies en sélection, plus particulièrement dans le cas des pathosystèmes racinaires, peu renseignés dans la littérature.

2.1. Relation entre architecture racinaire et résistance chez les pathosystèmes racinaires

L'analyse des relations entre caractères de résistance et d'architecture des plantes ont fait l'objet de peu d'études. **Au niveau aérien**, l'effet de l'architecture constitutive de la plante et du couvert végétal sur le développement d'une épidémie a été le plus souvent rapporté (Tivoli et al. 2012b). Des caractères architecturaux influençant directement les symptômes de la maladie, en créant un environnement moins favorable au développement des épidémies ou en permettant d'éviter les infections (échappement), ont été observés. Ainsi, chez les espèces annuelles comme le pois et la pomme de terre, la porosité et la hauteur du couvert, la surface foliaire et la sénescence tardive des plantes ont été identifiées comme des caractères d'intérêt pour limiter les épidémies d'ascochytose (Richard et al. 2013) ou de mildiou (Zarzynska and Szutkowska 2013). Chez le riz, le nombre de talles et leur hauteur ont été corrélés à la résistance au virus de la panachure jaune (Albar et al. 1998). Chez les espèces pérennes, les portes-greffes nains ou l'élagage contribuent à la résistance aux maladies (Gainza et al. 2015; Simon et al. 2006). Des colocalisations génomiques entre QTL de résistance aux maladies aériennes et QTL d'architecture ou de développement ont également été identifiées. Ainsi, chez le pois, des QTL de résistance à *D. pinodes* ont été localisés dans les mêmes régions génomiques que des QTL de date de floraison, de hauteur de plante ou de taille de stipules (Giorgetti 2013; Prioul et al. 2004). Chez le haricot, des QTL de résistance à *S. sclerotinium* ont été identifiés dans des régions comportant des QTL de tenue de tige, de hauteur de la plante ou de longueur d'entre-nœuds (Miklas et al. 2013).

Au niveau racinaire, très peu de références sont disponibles sur la relation entre les maladies et l'architecture de la plante. Les quelques études réalisées sur plusieurs pathosystèmes rapportent que la résistance aux agents pathogènes telluriques serait améliorée par des racines plus nombreuses et/ou plus longues (Bonhomme et al. 2014; Cichy et al. 2007; Dias et al. 2004; Graham et al. 2011; Hagerly et al. 2015; Laffont et al. 2015; Román-Avilés et al. 2004;

Snapp et al. 2003). Notre étude a permis de montrer qu'au niveau racinaire, ***l'architecture constitutive*** pouvait ne pas avoir un impact aussi important pour limiter les maladies telluriques qu'au niveau aérien. Ceci peut s'expliquer par les caractéristiques différentes des épidémies aériennes et telluriques (notamment la dissémination) et des facteurs environnementaux les influençant. Dans notre étude, une corrélation entre différents caractères de « densité » du système racinaire et de résistance a été observée. Celle-ci peut s'expliquer par le fait qu'un système racinaire plus développé peut être moins vite colonisé, dès lors que des mécanismes de défense limitent la progression de l'agent pathogène, et peut conserver une plus grande capacité à puiser ses ressources. Plusieurs colocalisations entre QTL de résistance et QTL d'architecture constitutive ont été identifiées dans notre étude. Ces résultats pourraient être confirmés à des stades plus avancés de développement de la plante dans l'hypothèse où l'architecture constitutive d'un génotype varierait avec le stade de la plante. On peut également imaginer que d'autres caractéristiques constitutives des racines peuvent impacter davantage le développement de la maladie. Chez *M. truncatula*, la modification de la composition des parois cellulaires par l'accumulation de composés phénoliques solubles a été observée chez un génotype résistant en réponse à l'infection par *A. euteiches* (Djébali et al. 2009). Chez *A. thaliana*, la production par les racines de toxines, comme les glucosinolates ou la camalexine, participeraient à la résistance à *Phytophthora brassicae* (Schlaeppli et al. 2010) et *Plasmodiophora brassicae* (Lemarié et al. 2015).

Notre étude a également montré que la plasticité architecturale ou ***architecture induite*** de la plante pouvait constituer une caractéristique importante de la plante pour lutter contre les infections sur le compartiment racinaire. Par exemple, la capacité des racines à constituer des organes de réserve pour les assimiler après l'attaque d'un agent pathogène a été démontrée chez plusieurs espèces telles que le maïs, le ray-grass ou le peuplier (Costes et al. 2013). Par ailleurs, plusieurs exemples ont démontré que les racines ne sont pas seulement des récepteurs passifs de produits de stockage, mais peuvent changer leurs relations puit-source après infection par un agent pathogène (Erb et al. 2009). La vigueur racinaire en réponse à une infection, qui se traduit par un développement racinaire plus élevé en conditions infestées qu'en conditions saines, est également une caractéristique de plasticité de la plante qui a été observée chez des génotypes résistants de *M. truncatula* inoculés par *A. euteiches* (Bonhomme et al. 2014) et de haricot infestés par *Fusarium solani* (Kraft and Boge 2001). Cette plasticité architecturale pourrait contribuer à une meilleure tolérance de la plante en conditions infectées, par le développement d'organes ou de caractéristiques (croissance,

morphologie) lui permettant de préserver sa capacité à puiser ses ressources. Dans notre étude, des corrélations phénotypiques élevées et des colocalisations de QTL entre caractères d'architecture induite et de résistance ont été mises en évidence. Cette interdépendance des déterminismes génétiques peut s'expliquer par des liaisons génétiques entre locus contrôlant les deux types de caractère et/ou par des effets pléiotropes de gènes causaux identiques, codant pour des fonctions impliquées dans des voies métaboliques hormonales et de défense interconnectées. En particulier, les cytokinines ont émergé récemment comme des voies régulatrices affectant les interactions plantes-micro-organismes (Choi et al. 2011). Chez *M. truncatula*, des mutants *cre1* de récepteurs aux cytokinines ont montré un niveau de symptômes réduits, corrélés à une formation accrue de racines latérales, en réponse à l'infection par *A. euteiches* (Laffont et al. 2015). La voie métabolique médiée par la mutation *CRE1* serait nécessaire à l'activation des divisions cellulaires corticales (Plet et al. 2011), mécanisme de défense observé en réponse à l'inoculation par *A. euteiches* chez un génotype résistant (Djébali et al. 2009).

2.2. Hypothèses sur les mécanismes biologiques sous-jacents aux QTL de résistance aux maladies chez les plantes

Dans notre étude, la précision de l'analyse de génétique d'association réalisée, ainsi que l'usage de marqueurs SNP ancrés au répertoire de gènes identifié récemment chez le pois (Tayeh et al. 2015a), a permis d'identifier des gènes candidats sous-jacents aux locus de résistance et d'architecture détectés. Ces gènes candidats permettent d'émettre quelques hypothèses sur les mécanismes sous-jacents aux QTL de résistance.

L'interdépendance génétique des locus de résistance et d'architecture précédemment discutée suggère que des gènes impliqués dans la **signalisation** des voies de défense et de développement de la plante pourraient sous-tendre des QTL de résistance. De plus, dans notre étude, certains marqueurs SNP associés à des locus de résistance correspondent à des gènes des voies de signalisation de l'acide jasmonique et de l'éthylène. Plusieurs études ont en effet mis en évidence l'implication de ces deux phytohormones dans la réponse des plantes aux interactions compatibles avec des agents pathogènes. Lemarie et al. (2015) ont ainsi montré que les voies de signalisation de l'acide jasmonique et de l'acide salicylique sont impliquées dans la résistance d'*A. thaliana* à la hernie des crucifères. Hosseini et al. (2015) ont également montré l'implication des voies de signalisation hormonales de l'acide jasmonique et de l'éthylène dans la réponse compatible du pois à *A. euteiches*. Les gènes impliqués dans ces voies de signalisation sont le plus souvent mis en évidence pour des résistances à des agents

pathogènes nécrotrophes (Buscaill and Rivas 2014; Pieterse et al. 2012). Étant donné que le cycle biologique d'*A. euteiches* comporte une phase nécrotrophe, les gènes candidats impliqués dans ces voies de signalisation sont d'autant plus intéressants.

Dans notre étude, certains QTL détectés pour la résistance à *A. euteiches* ont également colocalisé avec des gènes d'architecture et de développement du pois au niveau aérien. Certains de ces gènes pourraient avoir des effets pléiotropes sur la résistance, comme le gène *af* déterminant la forme des feuilles et associé à la résistance à *A. euteiches* au champ. Le gène *af* serait impliqué dans la signalisation hormonale de l'auxine, les génotypes feuillus présentant un niveau d'auxine plus important que les génotypes *afila* (DeMason et al. 2013). Chez le pois, Hosseini et al. (2015) ont montré que la voie de signalisation de l'auxine était activée dans le cas d'une interaction compatible avec *A. euteiches*. Chez *A. thaliana*, la voie de signalisation de l'auxine serait également impliquée dans la résistance à des oomycètes comme *Phythium irregulare* et *Hyaloperonospora arabidopsidis* (Tiryaki and Staswick 2002; Wang et al. 2007).

Des gènes sous-jacents aux QTL de résistance aux maladies et d'architecture de la plante pourraient être impliqués dans la **reconnaissance** de l'agent pathogène. En effet, le marqueur *Ps115429*, identifié pour les deux types de caractères, serait situé dans un intron d'un gène codant pour une MAP-Kinase, par homologie avec le génome de *M. truncatula*. Chez le blé, un gène codant pour une MAP-kinase a été identifié comme impliqué dans la tolérance à plusieurs stress abiotiques et dans la croissance racinaire (Hao et al. 2015). Les MAP-Kinases sont également impliquées dans l'immunité de type PTI en réponse à des effecteurs pathogènes (Sekhwal et al. 2015). Ces résultats suggèrent qu'un gène de reconnaissance pourrait engendrer des cascades de signaux, activant ainsi des voies métaboliques communes à la défense de la plante et à la croissance racinaire.

2.3. Combinaison d'allèles de résistance et d'architecture pour la sélection

Ce travail de thèse a permis d'apporter des outils et informations pour combiner des allèles de résistance et d'architecture afin de créer des variétés résistantes et/ou tolérantes à *A. euteiches*. Ces outils et informations incluent :

- des **marqueurs SNP associés à des locus d'intérêt** pour la résistance et l'architecture. Jusqu'à présent, seuls des marqueurs SSR étaient proposés pour la sélection assistée par marqueurs. Les marqueurs SNP faciliteront le génotypage d'un grand nombre de lignées par rapport aux marqueurs SSR ;

- des **intervalles de confiance de QTL** de résistance et d'architecture **plus précis** grâce à la génétique d'association. Cette précision pourra permettre de limiter la sélection d'allèles défavorables étroitement liés aux QTL d'intérêt, comme le gène *A* lié à la teneur en anthocyanes. En effet, jusqu'à présent, la résistance apportée par l'un des allèles favorables au QTL *Ae-Ps2.2* était très liée à une forte teneur en anthocyanes dans les tissus. Les anthocyanes sont des composés antinutritionnels qui ne sont pas intéressants pour la culture de pois protéagineux à destination de l'alimentation animale ou humaine. La rupture de cette liaison va permettre de faciliter la sélection pour la résistance à ce locus indépendamment de la teneur en anthocyanes ;
- des **haplotypes de marqueurs SNP** pour tracer des allèles favorables à la résistance. Une analyse des haplotypes de marqueurs dans 14 régions génomiques détectées de manière consistante, a permis d'identifier des haplotypes favorables pour la résistance à *A. euteiches*. Ces haplotypes représentent des outils, inexistant jusqu'à ce jour, qui permettront de faciliter grandement l'intégration des QTL de résistance dans les schémas de sélection.

Les résultats obtenus dans cette thèse permettent également d'apporter des connaissances sur :

- la **stabilité des locus** de résistance et d'architecture en fonction des environnements ou des conditions de test. Comme décrit par Hamon (2010), les QTL détectés pour la résistance à *A. euteiches* dépendent souvent de l'environnement ou de la souche testée. Cependant, plusieurs locus (*I.5*, *II.1*, *III.3*, *III.4* et *VII.16*) ont été détectés de façon stable selon les environnements et présentent donc un intérêt pour la sélection de variétés résistantes. De même, des QTL d'architecture différents ont pu être détectés selon les conditions de test, comme décrit par Bourion et al. (2010). Seuls les locus à effets forts, comme le gène de longueur d'entre-nœud *LE* ont été détectés de manière stable dans toutes les conditions ;
- la **diversité des allèles** (ou haplotypes de marqueurs) de résistance et d'architecture et le contenu de chaque lignée de la collection en haplotypes favorables. Ces résultats apportent des éléments pour le choix de géniteurs d'intérêt à utiliser comme géniteurs dans les schémas de sélection. Les QTL détectés à l'aide de populations biparentales sont limités aux allèles présents chez les deux parents des populations. L'analyse d'association, réalisée à partir d'un panel de diversités, permet d'augmenter potentiellement le nombre d'allèles différents. Dans cette étude, l'exploitation de cette

diversité a permis de détecter de nouveaux QTL (II.7, IV.9 à IV.13, VII.18), encore jamais détectés auparavant, notamment pour la résistance du pois à *A. euteiches* ;

- l'effet du **cumul de locus de résistance** pour le choix de combinaisons de locus/haplotypes permettant d'augmenter le niveau de résistance partielle. Notre étude a montré que le cumul de QTL permettait d'augmenter les niveaux de résistance partielle à *A. euteiches* chez le pois. Ce constat est cohérent avec les résultats de (Lavaud et al. 2015) sur le cumul de QTL dans des lignées quasi-isogéniques pour le même pathosystème. Chez le blé, le cumul d'allèles favorables à la résistance à *Puccinia graminis* f. sp. *tritici* a également été corrélé à un meilleur niveau de résistance (Letta et al. 2013). Dans notre étude, le QTL majeur *Ae-Ps7.6*, dont l'intervalle a été reprecisé, est un locus de choix du fait de son effet élevé mais son usage en combinaison avec d'autres QTL à effets plus faibles est à recommander en perspective de sa gestion durable dans les futures variétés résistantes. Chez le piment, le cumul de QTL de résistance a été corrélé à un moindre contournement de la résistance aux nématodes à galles de la famille des *Meloidogynes* (Djian-Caporalino et al. 2014) ;
- l'effet du **cumul de caractères de résistance et d'architecture** sur le rendement au sein des 15 génotypes expérimentés au champ pour le choix de combinaisons de locus/haplotypes de résistance et d'architecture permettant de limiter les pertes de rendement. Certains génotypes de niveau de résistance moyen (*AeD9904-7* and *90-2131*) ont subi moins de pertes de système racinaire et de rendement dans les parcelles fortement infestées, par rapport à des lignées de même niveau de résistance. Dans la version finale de l'article en préparation (chapitre IV), une étude de la composition en allèles de résistance et d'architecture des 15 génotypes sera réalisée et mise en lien avec les phénotypes décrits pour la résistance, l'architecture et les pertes de rendement. Il serait alors possible d'identifier des allèles d'architecture racinaire constitutive et induite permettant d'augmenter la tolérance, et ainsi limiter les pertes de rendement dans des parcelles fortement infestées.

3. Perspectives de recherche

Ce travail de thèse a apporté des connaissances qui pourront servir de base pour la poursuite de différentes pistes de travaux méthodologiques ou de recherche sur le pathosystème pois/*A. euteiches*. Ces perspectives de travaux incluent : (i) l'amélioration de la précision du

phénotypage des symptômes de maladie racinaire ; (ii) la cartographie fine de QTL de résistance et d'architecture dans les zones génomiques communes ; et (iii) l'étude des mécanismes biologiques sous-jacents aux QTL de résistance et d'architecture.

3.1. Améliorer la qualité la précision et le débit du phénotypage

Mon travail de thèse a participé à développer une nouvelle technique de phénotypage des symptômes racinaires dû à *A. euteiches* chez le pois par analyse d'image. Cependant, il reste des imprécisions de détection des symptômes par confusion avec des ombres sur les images. Des travaux sont en cours à l'INRA de Dijon afin de développer un algorithme permettant de distinguer les symptômes réels des ombres.

De plus, la prise d'image du système racinaire nécessite l'étalement des racines sur un support transparent, processus long et fastidieux. Avec le développement récent de plates-formes de phénotypage racinaire à haut débit, telle que celle de l'INRA de Dijon (4PMI) ou celle de l'institut de bio et géoscience de Jülich en Allemagne (*Jülich Plant Phenotyping Center*), il est possible d'envisager l'acquisition automatisée d'images pour observer l'évolution de l'architecture racinaire conjointement à celle des symptômes de façon dynamique et non destructive. Un test non destructeur d'analyse de la cinétique de développement des symptômes de la maladie sur papiers imbibés a notamment été développé à l'INRA de Rennes (Moussart A, Pers. Com.). Dans cette thèse, la relation entre l'architecture et la résistance n'a été étudiée qu'à un stade précoce (14 jours après semis) sur l'ensemble de la population. Avec l'usage d'une méthode de test non destructive, il serait intéressant d'étudier cette relation à des stades plus avancés. La comparaison de l'architecture des racines au moment de l'inoculation, puis son évolution comparée en conditions inoculée et non-inoculée grâce à un tel dispositif, permettrait d'évaluer le rôle de l'architecture des racines au moment de l'inoculation puis la différence de dynamique de croissance des racines saines et infectée. Par ailleurs, d'autres types de dispositifs que celui utilisé dans la thèse seraient également intéressants à tester, tels que l'usage de pots de taille plus élevée afin de favoriser davantage le développement racinaire en conditions inoculées. Ce type de dispositif (contenants plus profonds et plus larges) a été testé en début de thèse mais aurait nécessité davantage de temps pour adapter le protocole d'inoculation sur un grand nombre de génotypes.

Avec l'augmentation de la précision des méthodes de phénotypage, il sera également intéressant d'observer d'autres caractères d'architecture plus précis (nombre de ramifications secondaires, angle d'insertion des racines, densité de racines latérales dans les premiers

centimètres), mais également des caractères de développement racinaire (vitesse de production et de croissance des racines latérales) ou de symbiose (nombre et taille de nodosités). En effet, chez le haricot un nombre élevé de racines latérales dans le premier centimètre a été corrélé à un niveau supérieur de résistance à *F. solani* (Snapp et al. 2003). Les voies métaboliques impliquées dans les interactions compatibles avec les agents pathogènes sont parfois identiques à celles impliquées dans les processus de symbioses. Chez *M. truncatula*, le gène *MtCRE1* est à la fois requis pour la fixation symbiotique de *Rhizobium* et impliqué dans les processus de résistance à *A. euteiches* (Laffont et al. 2015), suggérant des mécanismes communs. À un niveau plus fin, des caractères de composition structurale des racines pourraient également être explorés pour leur relation avec la résistance et l'architecture constitutive. Ainsi, dans le projet PeaMUST, une collection de 240 lignées de pois sera phénotypée pour la composition pariétale des cellules racinaires, ainsi que pour la résistance à *A. euteiches* et l'architecture, en vue d'analyse GWA.

3.2. Préciser les liens génétiques entre locus de résistance et d'architecture

La précision de localisation génomique des locus impliqués dans la résistance du pois à *A. euteiches* et dans l'architecture a été grandement améliorée dans le cadre des travaux de cette thèse. Toutefois, l'analyse GWAS n'ayant pas été réalisée sur une collection à base génétique très large et avec une densité très élevée de marqueurs étant donné la décroissance de DL, il est possible que les intervalles des locus de résistance et d'architecture identifiés soient encore imprécis. La co-ségrégation des allèles de résistance et d'architecture dans les zones de colocalisation de QTL pourrait être confirmée, soit en augmentant la taille de la collection d'association et/ou de sa densité de génotypage soit en identifiant et caractérisant des individus recombinants dans ces zones. Dans le cadre du projet PeaMUST, des individus recombinants ont été identifiés et ont permis de réduire l'intervalle de confiance du QTL majeur *Ae-Ps7.6*, qui a été re-déTECTÉ dans ce travail de thèse pour différents caractères de résistance et d'architecture constitutive et induite. Dans cet intervalle, le marqueur SNP *Ps115429*, identifié par GWAS dans notre étude de manière consistante pour différentes variable de résistance et d'architecture, sera utilisé pour génotyper les individus recombinants dans l'intervalle réduit. Les individus recombinants permettant de confirmer l'intervalle réduit du QTL *Ae-Ps7.6* pourront également être phénotypés pour les caractères d'architecture associés à cet intervalle, afin de confirmer ou non la co-ségrégation des allèles de résistance et d'architecture dans l'intervalle très réduit.

3.3. Explorer les mécanismes moléculaires associés à la résistance et à l'architecture

Ma thèse a permis d'émettre des hypothèses quant aux mécanismes moléculaires sous-jacents aux déterminants communs de la résistance à *A. euteiches* et de l'architecture. Si la co-ségrégation d'allèles de résistance et d'architecture est confirmée dans les zones des QTL communs, notamment à *Ae-Ps7.6*, une analyse de gènes différentiellement exprimés en réponse à l'infection pourrait être effectuée sur couples de NIL ne différant que par la zone cible. La validation de gènes candidats positionnels (identifiés par cartographie fine), expressionnels (identifiés par analyse de transcriptomes) et/ou fonctionnels (connus pour leur fonction associée à la résistance et à l'architecture) pourra être envisagée chez le pois par extinction de gènes (VIGS, *virus induced gene silencing*) ou par recherche et phénotypage de mutants de TILLING (*target induced local lesions in genomes*). Notamment, suite aux résultats récents de Laffont et al. (2015), il serait intéressant d'identifier et caractériser des mutants de pois, porteurs d'orthologues de gènes de récepteurs aux cytokinines de *M. truncatula*. Chez d'autres pathosystèmes, des gènes fonctionnels, impliqués dans les voies phytohormonales de l'acide jasmonique ou de l'éthylène, ou la modification des parois cellulaires, ont également été mis en évidence et constitueraient des candidats d'intérêt.

4. Conclusion générale

Au cours de cette thèse, nous avons développé une approche de génétique d'association pour étudier l'interdépendance des déterminants génétiques de la résistance à *A. euteiches* et de l'architecture racinaire chez le pois. Cette interdépendance génétique a été mise en évidence entre la résistance et certains caractères de développement racinaire intrinsèque ou de vigueur racinaire en réponse à l'infection. La combinaison de la résistance aux différents caractères d'architecture a montré une efficacité chez certains géotypes pour augmenter la tolérance du pois à *A. euteiches* en limitant les pertes de rendement.

Les résultats de la thèse apportent des connaissances originales sur la génétique comparative de la résistance aux maladies et de l'architecture de la plante au niveau racinaire, ainsi que des outils (géniteurs, marqueurs, haplotypes) et des éléments de choix de QTL, en vue de créer des variétés de pois combinant des allèles de résistance et d'architecture permettant de lutter durablement contre *A. euteiches*.

Références bibliographiques

Références bibliographiques

- Adu MO, Chatot A, Wiesel L, Bennett MJ, Broadley MR, White PJ, Dupuy LX (2014) A scanner system for high-resolution quantification of variation in root growth dynamics of *Brassica rapa* genotypes. *J Exp Bot* 65:2039-2048
- Albar L, Lorieux M, Ahmadi N, Rimbault I, Pinel A, Sy AA, Fargette D, Ghesquiere A (1998) Genetic basis and mapping of the resistance to rice yellow mottle virus. I. QTLs identification and relationship between resistance and plant morphology. *Theor Appl Genet* 97:1145-1154
- Allmaras RR, Fritz VA, Pflieger FL, Copeland SM (2003) Impaired internal drainage and *Aphanomyces euteiches* root rot of pea caused by soil compaction in a fine-textured soil. *Soil and Tillage Research* 70:41-52
- Alves-Carvalho S, Aubert G, Carrère S, Cruaud C, Brochot A-L, Jacquin F, Klein A, Martin C, Boucherot K, Kreplak J, de Silva C, Moreau S, Gamas P, Wincker P, Gouzy J, Burstin J (2015) Full-length de novo assembly of RNA-seq data in pea (*Pisum sativum* L.) provides a gene expression atlas and gives insights into root nodulation in this species. *Plant J* 84:1-19
- Andrivon D, Giorgetti C, Baranger A, Calonnec A, Cartolaro P, Faivre R, Guyader S, Lauri PE, Lescourret F, Parisi L, Ney B, Tivoli B, Sache I (2013) Defining and designing plant architectural ideotypes to control epidemics? *Eur J Plant Pathol* 135:611-617
- Arsenault J, Poulcur S, Messier C, Guay R (1995) WinRHIZO™, a root-measuring system with a unique overlap correction method. 92nd Annual Meeting of the American Society for Horticultural Science / 40th Annual Congress of the Canadian Society for Horticultural Science. HortScience, Montréal, Quebec, Canada, p 906
- Badis Y, Bonhomme M, Lafitte C, Huguet S, Balzergue S, Dumas B, Jacquet C (2015) Transcriptome analysis highlights preformed defences and signalling pathways controlled by the prAe1 quantitative trait locus (QTL), conferring partial resistance to *Aphanomyces euteiches* in *Medicago truncatula*. *Mol Plant Pathol* 16:973-986
- Bardol N, Ventelon M, Mangin B, Jasson S, Loywick V, Couton F, Derue C, Blanchard P, Charcosset A, Moreau L (2013) Combined linkage and linkage disequilibrium QTL mapping in multiple families of maize (*Zea mays* L.) line crosses highlights complementarities between models based on parental haplotype and single locus polymorphism. *Theor Appl Genet* 126:2717-2736
- Bengough AG, Gordon DC, Al-Menaie H, Ellis RP, Allan D, Keith R, Thomas WTB, Forster BP (2004) Gel observation chamber for rapid screening of root traits in cereal seedlings. *Plant and Soil* 262:63-70
- Benjamin JG, Nielsen DC (2004) A method to separate plant roots from soil and analyze root surface area. *Plant and Soil* 267:225-234
- Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate - A practical and powerful approach to multiple testing. *J R Stat Soc series B Stat Methodol* 57:289-300
- Berta G, Sampo S, Gamalero E, Massa N, Lemanceau P (2005) Suppression of *Rhizoctonia* root-rot of tomato by *Glomus mossae* BEG12 and *Pseudomonas fluorescens* A6RI is associated with their effect on the pathogen growth and on the root morphogenesis. *Eur J Plant Pathol* 111:279-288
- Bonhomme M, Andre O, Badis Y, Ronfort J, Burgarella C, Chantret N, Prosperi J-M, Briskine R, Mudge J, Debelle F, Navier H, Miteul H, Hajri A, Baranger A, Tiffin P, Dumas B, Pilet-Nayel M-L, Young ND, Jacquet C (2014) High-density genome-wide association mapping implicates an F-box encoding gene in *Medicago truncatula* resistance to *Aphanomyces euteiches*. *New Phytol* 201:1328-1342
- Bourion V (in prep) Genetic diversity for partner choice in a core collection of pea accessions inoculated by a mix of five *Rhizobium leguminosarum* ssp. *viciae* genotypes

- Bourion V, Laguerre G, Depret G, Voisin A, Salon C, Duc G (2007) Genetic variability in nodulation and root growth affects nitrogen fixation and accumulation in pea. *Ann Bot* 100:589-598
- Bourion V, Rizvi SM, Fournier S, de Larambergue H, Galmiche F, Marget P, Duc G, Burstin J (2010) Genetic dissection of nitrogen nutrition in pea through a QTL approach of root, nodule, and shoot variability. *Theor Appl Genet* 121:71-86
- Boutet G, Alves-Carvalho S, Falque M, Peterlongo P, Lhuillier E, Bouchez O, Lavaud C, Pilet-Nayel M, Riviere N, Baranger A (2016) SNP discovery and genetic mapping using Genotyping by Sequencing of whole genome genomic DNA from a pea RIL population. *BMC Genomics*
- Brogie KE, Butler KH, Butruille MG, Da SCA, Frey TJ, Hawk JA, Jaqueth JS, Jones ES, Multani DS, Wolters PJ (2006) Polynucleotides and methods for making plants resistant to fungal pathogens. Google Patents
- Burstin J, Marget P, Huart M, Moessner A, Mangin B, Duchene C, Desprez B, Munier-Jolain N, Duc G (2007) Developmental genes have pleiotropic effects on plant morphology and source capacity, eventually impacting on seed protein content and productivity in pea. *Plant Physiol* 144:768-781
- Burstin J, Salloignon P, Chabert-Martinello M, Magnin-Robert JB, Siol M, Jacquin F, Chauveau A, Pont C, Aubert G, Delaitre C, Truntzer C, Duc G (2015) Genetic diversity and trait genomic prediction in a pea diversity panel. *BMC Genomics* 16:105
- Burton AL, Johnson JM, Foerster JM, Hirsch CN, Buell CR, Hanlon MT, Kaeppler SM, Brown KM, Lynch JP (2014) QTL mapping and phenotypic variation for root architectural traits in maize (*Zea mays* L.). *Theor Appl Genet* 127:2293-2311
- Buscaill P, Rivas S (2014) Transcriptional control of plant defence responses. *Curr Opin Plant Biol* 20:35-46
- Calonnec A, Burie JB, Langlais M, Guyader S, Saint-Jean S, Sache I, Tivoli B (2012) Impacts of plant growth and architecture on pathogen processes and their consequences for epidemic behaviour. *Eur J Plant Pathol* 135:479-497
- Cannesan MA, Durand C, Burel C, Gangneux C, Lerouge P, Ishii T, Laval K, Follet-Gueye ML, Driouich A, Vire-Gibouin M (2012) Effect of arabinogalactan proteins from the root caps of pea and *Brassica napus* on *Aphanomyces euteiches* zoospore chemotaxis and germination. *Plant Physiol* 159:1658-1670
- Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ (2015) Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* 4:7
- Charmet G (2011) De la détection de QTL à la génétique d'association. *Le Sélectionneur Français* 62:17-29
- Choi J, Choi D, Lee S, Ryu C-M, Hwang I (2011) Cytokinins and plant immunity: old foes or new friends? *Trends Plant Sci* 16:388-394
- Cichy KA, Snapp SS, Kirk WW (2007) Fusarium root rot incidence and root system architecture in grafted common bean lines. *Plant and Soil* 300:233-244
- Clark RT, MacCurdy RB, Jung JK, Shaff JE, McCouch SR, Aneshansley DJ, Kochian LV (2011) Three-dimensional root phenotyping with a novel imaging and software platform. *Plant Physiol* 156:455-465
- Cook DE, Lee TG, Guo X, Melito S, Wang K, Bayless AM, Wang J, Hughes TJ, Willis DK, Clemente TE, Diers BW, Jiang J, Hudson ME, Bent AF (2012) Copy number variation of multiple genes at Rhg1 mediates nematode resistance in soybean. *Science* 338:1206-1209
- Costes E, Lauri PE, Simon S, Andrieu B (2013) Plant architecture, its diversity and manipulation in agronomic conditions, in relation with pest and pathogen attacks. *Eur J Plant Pathol* 135:455-470
- Crosson P, Moussart A (2011) *Aphanomyces* - Un test pour cultiver du pois sans craintes. *Perspectives agricoles* 378:68-71

- Dalmais M, Schmidt J, Le Signor C, Moussy F, Burstin J, Savoie V, Aubert G, Brunaud V, de Oliveira Y, Guichard C, Thompson R, Bendahmane A (2008) UTILLdb, a *Pisum sativum* in silico forward and reverse genetics tool. *Genome Biology* 9:R43
- Davis DW (1995) MN 144, MN 313, and MN 314: Garden pea lines resistant to root rot caused by *Aphanomyces euteiches* Drechs. *HortScience* 30:639-640
- de Dorlodot S, Forster B, Pagès L, Price A, Tuberosa R, Draye X (2007) Root system architecture: opportunities and constraints for genetic improvement of crops. *Trends Plant Sci* 12:474-481
- Deacon JW, Saxena G (1998) Germination triggers of zoospore cysts of *Aphanomyces euteiches* and *Phytophthora parasitica*. *Mycological Research* 102:33-41
- Delory BM, Baudson C, Brostaux Y, Lobet G, Jardin P, Pagès L, Delaplace P (2015) archiDART: an R package for the automated computation of plant root architectural traits. *Plant and Soil* 398:351-365
- DeMason DA, Chetty V, Barkawi LS, Liu X, Cohen JD (2013) *Unifoliata-Afila* interactions in pea leaf morphogenesis. *American Journal of Botany* 100:478-495
- Desgroux A, Baudais V, Riviere JP, Aubert G, Tayeh N, Moussart A, Mangin P, Aubert V, de Larambergue H, Vetel P, Burstin J, Baranger A, Manzanares-Dauleux M, Pilet-Nayel ML, Bourion V (in prep) Genetic determinism of root growth and architecture traits shares a common locus with *Aphanomyces* root rot resistance QTL in pea
- Desgroux A, L'Anthoene V, Roux-Duparque M, Riviere JP, Aubert G, Tayeh N, Moussart A, Mangin P, Vetel P, Piriou C, McGee RJ, Coyne CJ, Burstin J, Baranger A, Manzanares-Dauleux M, Bourion V, Pilet-Nayel ML (2016) Genome-Wide association mapping of partial resistance to *Aphanomyces euteiches* in pea. *BMC Genomics*
- Dias RdCS, Picó B, Espinos A, Nuez F (2004) Resistance to melon vine decline derived from *Cucumis melo* ssp. *agrestis*: genetic analysis of root structure and root response. *Plant Breeding* 123:66-72
- Ding J, Zhang L, Chen J, Li X, Li Y, Cheng H, Huang R, Zhou B, Li Z, Wang J, Wu J (2015) Genomic dissection of leaf angle in maize (*Zea mays* L.) using a four-way cross mapping population. *PLoS One* 10:e0141619
- Djébali N, Aribi S, Taamalli W, Arraouadi S, Aouani ME, Badri M (2013) Natural variation of *Medicago truncatula* resistance to *Aphanomyces euteiches*. *Eur J Plant Pathol* 135:831-843
- Djébali N, Jauneau A, Ameline-Torregrosa C, Chardon F, Jaulneau V, Mathé C, Bottin A, Cazaux M, Pilet-Nayel M-L, Baranger A, Aouani ME, Esquerré-Tugayé M-T, Dumas B, Huguet T, Jacquet C (2009) Partial resistance of *Medicago truncatula* to *Aphanomyces euteiches* is associated with protection of the root stele and is controlled by a major QTL rich in proteasome-related genes. *Mol Plant Microbe Interact* 22:1043-1055
- Djian-Caporalino C, Palloix A, Fazari A, Marteu N, Barbary A, Abad P, Sage-Palloix A-M, Mateille T, Risso S, Lanza R, Taussig C, Castagnone-Sereno P (2014) Pyramiding, alternating or mixing: comparative performances of deployment strategies of nematode resistance genes to promote plant resistance efficiency and durability. *BMC Plant Biol* 14:1-13
- Domoney C, Knox M, Moreau C, Ambrose M, Palmer S, Smith P, Christodoulou V, Isaac PG, Hegarty M, Blackmore T, Swain M, Ellis N (2013) Exploiting a fast neutron mutant genetic resource in *Pisum sativum* (pea) for functional genomics. *Functional Plant Biology* 40:1261-1270
- Downie HF, Adu MO, Schmidt S, Otten W, Dupuy LX, White PJ, Valentine TA (2015) Challenges and opportunities for quantifying roots and rhizosphere interactions through imaging and image analysis. *Plant Cell Environ* 38:1213-1232
- Erb M, Lenk C, Degenhardt J, Turlings TCJ (2009) The underestimated role of roots in defense against leaf attackers. *Trends Plant Sci* 14:653-659

- FAOSTAT (2015) FAOSTAT data base. Food and agriculture organisation of the United Nations - statistic division
- Fondevilla S, Satovic Z, Rubiales D, Moreno MT, Torres AM (2008) Mapping of quantitative trait loci for resistance to *Mycosphaerella pinodes* in *Pisum sativum* subsp. *syriacum*. *Mol Breed* 21:439-454
- Fu D, Uauy C, Distelfeld A, Blechl A, Epstein L, Chen X, Sela H, Fahima T, Dubcovsky J (2009) A kinase-START gene confers temperature-dependent resistance to wheat stripe rust. *Science* 323:1357-1360
- Fukuoka S, Saka N, Koga H, Ono K, Shimizu T, Ebana K, Hayashi N, Takahashi A, Hirochika H, Okuno K, Yano M (2009) Loss of function of a proline-containing protein confers durable disease resistance in rice. *Science* 325:998-1001
- Fukuoka S, Yamamoto S-I, 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
- Fusconi A, Gnani E, Trotta A, Berta G (1999) Apical meristems of tomato roots and their modifications induced by arbuscular mycorrhizal and soilborne pathogenic fungi. *New Phytol* 142:505-516
- Gainza F, Opazo I, Guajardo V, Meaz P, Ortiz M, Pinochet J, Muñoz C (2015) Rootstock breeding in *Prunus* species: Ongoing efforts and new challenges. *Chilean journal of agricultural research* 75:6-16
- Gaulin E, Jacquet C, Bottin A, Dumas B (2007) Root rot disease of legumes caused by *Aphanomyces euteiches*. *Mol Plant Pathol* 8:539-548
- Giorgetti C (2013) Part relative de l'architecture et de la résistance partielle dans le contrôle génétique du ralentissement des épidémies d'ascochytose chez le pois (*Pisum sativum* L.). *AgroCampus Ouest, Rennes*
- Goffinet B, Gerber S (2000) Quantitative Trait Loci: A Meta-analysis. *Genetics* 155:463
- Graham J, Hackett CA, Smith K, Woodhead M, MacKenzie K, Tierney I, Cooke D, Bayer M, Jennings N (2011) Towards an understanding of the nature of resistance to *Phytophthora* root rot in red raspberry. *Theor Appl Genet* 123:585-601
- Graves S, Piepho H-P, Selzer L, Dorai-Raj S (2015) multcompView: Visualisations of Paired Comparisons. R package version 0.1-7.
- Gritton ET (1990) Registration of five pea root rot resistant germplasm lines of processing pea. *Crop Sci* 30:1166-1167
- Gupta PK, Kulwal PL, Jaiswal V (2014) Association mapping in crop plants: opportunities and challenges. *Adv Genet* 85:109-147
- Hagerty CH, Cuesta-Marcos A, Cregan PB, Song Q, McClean P, Noffsinger S, Myers JR (2015) Mapping and root rot resistance and root architecture quantitative trait loci in common bean. *Crop Sci* 55:1969
- Hamon C (2010) Variabilité structurale des déterminants génétiques de la résistance à *Aphanomyces euteiches* chez le pois (*Pisum sativum* L.) et l'espèce légumineuse modèle *Medicago truncatula*. *AgroCampus Ouest, Rennes*
- Hamon C, Baranger A, Coyne CJ, McGee RJ, Le Goff I, L'Anthoene V, Esnault R, Riviere JP, Klein A, Mangin P, McPhee KE, Roux-Duparque M, Porter L, Miteul H, Lesne A, Morin G, Onfroy C, Moussart A, Tivoli B, Delourme R, Pilet-Nayel ML (2011) New consistent QTL in pea associated with partial resistance to *Aphanomyces euteiches* in multiple French and American environments. *Theor Appl Genet* 123:261-281
- Hamon C, Baranger A, Miteul H, Lecointe R, Le Goff I, Deniot G, Onfroy C, Moussart A, Prosperi JM, Tivoli B, Delourme R, Pilet-Nayel ML (2010) A complex genetic network involving a broad-spectrum locus

- and strain-specific loci controls resistance to different pathotypes of *Aphanomyces euteiches* in *Medicago truncatula*. *Theor Appl Genet* 120:955-970
- Hamon C, Coyne CJ, McGee RJ, Lesné A, Esnault R, Mangin P, Hervé M, Le Goff I, Deniot G, Roux-Duparque M, Morin G, McPhee K, Delourme R, Baranger A, Pilet-Nayel M (2013) QTL meta-analysis provides a comprehensive view of the moderately low diversity of loci controlling partial resistance to *Aphanomyces euteiches* in four pea sources of resistance. *BMC Plant Biol* 13:45
- Hao L, Wen Y, Zhao Y, Lu W, Xiao K (2015) Wheat mitogen-activated protein kinase gene TaMPK4 improves plant tolerance to multiple stresses through modifying root growth, ROS metabolism, and nutrient acquisitions. *Plant Cell Rep* 34:2081-2097
- Hascoët E, Jaminon O, deaux C, Blassiau C, Bahrman N, Bochard AM (2014) Toward fine mapping of frost tolerance QTL in pea. 2nd PeaMUST Annual Meeting, Dijon
- Hodge A, Berta G, Doussan C, Merchan F, Crespi M (2009) Plant root growth, architecture and function. *Plant and Soil* 321:153-187
- Hossain S, Bergkvist G, Berglund K, Mårtensson A, Persson P (2012) *Aphanomyces* pea root rot disease and control with special reference to impact of Brassicaceae cover crops. *Acta Agriculturae Scandinavica Section B-Soil and Plant Science* 62:477-487
- Hosseini S, Elfstrand M, Heyman F, Jensen DF, Karlsson M (2015) Deciphering common and specific transcriptional immune responses in pea towards the oomycete pathogens *Aphanomyces euteiches* and *Phytophthora pisi*. *BMC Genomics* 16
- Hughes TJ, Grau CR (2007) *Aphanomyces* root rot (common root rot) of legumes. *The plant Health Instructor*
- Hurni S, Scheuermann D, Krattinger SG, Kessel B, Wicker T, Herren G, Fitze MN, Breen J, Presterl T, Ouzunova M, Keller B (2015) The maize disease resistance gene Htn1 against northern corn leaf blight encodes a wall-associated receptor-like kinase. *Proc Natl Acad Sci U S A* 112:8780-8785
- Jones FR, Drechsler C (1925) Root rot of peas caused by *Aphanomyces euteiches*. *J Agric Res* 30:293-325
- Kirchhof G (1992) Measurement of root length and thickness using a hand-held computer scanner. *Field Crops Research* 29:79-88
- Kover PX, Caicedo AL (2001) The genetic architecture of disease resistance in plants and the maintenance of recombination by parasites. *Molecular Ecology* 10:1-16
- Kraft JM (1992) Registration of 90-2079, 90-2131, and 90-2322 pea germplasms. *Crop Sci* 32:1076
- Kraft JM, Boge WL (2001) Root characteristics in pea in relation to compaction and fusarium root rot. *Plant disease* 85:936-940
- Kraft JM, Coffman VA (2000a) Registration of 96-2052, 96-2058, 96-2068, 96-2198, and 96-2222 pea germplasms. *Crop Sci* 40:301-302
- Kraft JM, Coffman VA (2000b) Registration of 97-261 and 97-2154 pea germplasms. *Crop Sci* 40:302-303
- Kraft JM, Coffman VA (2000c) Registration of 97-363, 97-2170, and 97-2162 pea germplasms. *Crop Sci* 40:303
- Krattinger SG, Lagudah ES, Spielmeier W, Singh RP, Huerta-Espino J, McFadden H, Bossolini E, Selter LL, Keller B (2009) A putative ABC transporter confers durable resistance to multiple fungal pathogens in wheat. *Science* 323:1360-1363
- Laffont C, Rey T, Andre O, Novero M, Kazmierczak T, Debelle F, Bonfante P, Jacquet C, Frugier F (2015) The CRE1 cytokinin pathway is differentially recruited depending on *Medicago truncatula* root environments and negatively regulates resistance to a pathogen. *PLoS One* 10

- Lander ES, Botstein D (1989) Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* 121:185-199
- Larkin RP, Fravel DR (1998) Efficacy of various fungal and bacterial biocontrol organisms for control of Fusarium wilt of tomato. *Plant disease* 82:1022-1028
- Lavaud C, Lesné A, Piriou C, Le Roy G, Boutet G, Moussart A, Poncet C, Delourme R, Baranger A, Pilet-Nayel ML (2015) Validation of QTL for resistance to *Aphanomyces euteiches* in different pea genetic backgrounds using Near Isogenic Lines. *Theor Appl Genet*
- Lebowitz RJ (1988) Digital image-analysis measurement of root length and diameter. *Environmental and Experimental Botany* 28:267-273
- Lecoeur J (2005) Développement végétatif. In: Murier-Jolain N (ed) *Agrophysiologie du pois protéagineux*. QUAE, Paris, pp 27-36
- Lejeune-Hénaut I, Hanocq E, Béthencourt L, Fontaine V, Delbreil B, Morin J, Petit A, Devaux R, Boilleau M, Stempniak JJ, Thomas M, Lainé AL, Foucher F, Baranger A, Burstin J, Rameau C, Giauffret C (2008) The flowering locus Hr colocalizes with a major QTL affecting winter frost tolerance in *Pisum sativum* L. *Theor Appl Genet* 116:1105-1116
- Lemarie S, Robert-Seilaniantz A, Lariagon C, Lemoine J, Marnet N, Jubault M, Manzanares-Dauleux MJ, Gravot A (2015) Both the jasmonic acid and the salicylic acid Pathways contribute to resistance to the biotrophic clubroot agent *Plasmodiophora brassicae* in *Arabidopsis*. *Plant Cell Physiol* 56:2158-2168
- Lemarié S, Robert-Seilaniantz A, Lariagon C, Lemoine J, Marnet N, Levrel A, Jubault M, Manzanares-Dauleux MJ, Gravot A (2015) Camalexin contributes to the partial resistance of *Arabidopsis thaliana* to the biotrophic soilborne protist *Plasmodiophora brassicae*. *Front Plant Sci* 6
- Lenth RV, Hervé M (2015) lsmeans: Least-Squares Means. R package version 2.15.
- Letta T, Maccaferri M, Badebo A, Ammar K, Ricci A, Crossa J, Tuberosa R (2013) Searching for novel sources of field resistance to Ug99 and Ethiopian stem rust races in durum wheat via association mapping. *Theor Appl Genet* 126:1237-1256
- 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*. *Proc Natl Acad Sci U S A* 103:7994-7999
- Lipka AE, Tian F, Wang Q, Peiffer J, Li M, Bradbury PJ, Gore MA, Buckler ES, Zhang Z (2012) GAPIT: genome association and prediction integrated tool. *Bioinformatics* 28:2397-2399
- Liu S, Kandath PK, Warren SD, Yeckel G, Heinz R, Alden J, Yang C, Jamai A, El-Mellouki T, Juvale PS, Hill J, Baum TJ, Cianzio S, Whitham SA, Korkin D, Mitchum MG, Meksem K (2012) A soybean cyst nematode resistance gene points to a new mechanism of plant resistance to pathogens. *Nature* 492:256-260
- Lobet G, Pages L, Draye X (2011) A novel image-analysis toolbox enabling quantitative analysis of root system architecture. *Plant Physiol* 157:29-39
- Lockwood J, Ballard J (1960) Evaluation of pea introductions for resistance to *Aphanomyces* and *Fusarium* root rots. *Michigan Quaterly Bulletin* 42:704-713
- Madoui MA, Labadie K, Aury JM, Coyne CJ, Warkentin TD, Kreplak J, Lichtenzweig J, Doležel J, Aubert G, Bergès H, Wincker P, Burstin J (2015) The international pea genome sequencing project: sequencing and assemble progresses. *Plant and Animal Genome XXIII*, San Diego, CA
- Malamy JE (2005) Intrinsic and environmental response pathways that regulate root system architecture. *Plant Cell and Environment* 28:67-77

- Malvick DK, Percich JA (1999) Identification of *Pisum sativum* germ plasm with resistance to root rot caused by multiple strains of *Aphanomyces euteiches*. *Plant disease* 83:51-54
- Manosalva PM, Davidson RM, Liu B, Zhu X, Hulbert SH, Leung H, Leach JE (2009) A germin-Like protein gene family functions as a complex quantitative trait locus conferring broad-spectrum disease resistance in rice. *Plant Physiol* 149:286-296
- McDonald MR, Gossen BD, Kora C, Parker M, Boland G (2013) Using crop canopy modification to manage plant diseases. *Eur J Plant Pathol* 135:581-593
- McPhee K (2005) Variation for seedling root architecture in the core collection of pea germplasm. *Crop Sci* 45:1758-1763
- Miklas PN, Porter LD, Kelly JD, Myers JR (2013) Characterization of white mold disease avoidance in common bean. *Eur J Plant Pathol* 135:525-543
- Moore JW, Herrera-Foessel S, Lan C, Schnippenkoetter W, Ayliffe M, Huerta-Espino J, Lillemo M, Viccars L, Milne R, Periyannan S, Kong X, Spielmeier W, Talbot M, Bariana H, Patrick JW, Dodds P, Singh R, Lagudah E (2015) A recently evolved hexose transporter variant confers resistance to multiple pathogens in wheat. *Nat Genet* 47:1494-1498
- Morin J, Mautord J, Sarda X, Canoy AS, Anglares S, Deniot G, Muel F, Lejeune-Hénaut I, Pilet-Nayel ML, Baranger A (2008) Transcript profiling during the *Aphanomyces euteiches* infection in pea (*Pisum sativum*). ANR-Genoplante meeting, Poster, Arles, France
- Moussart A, Devaux C, Muel F, Pilet-Nayel M, Baranger A, Tivoli B, Roux-Duparque M (2007) Improving partial resistance to *Aphanomyces* root rot in GSP breeding program. 3rd International *Aphanomyces* Workshop on Legumes, Rennes, Brittany, France, pp 46-47
- Moussart A, Even MN, Tivoli B (2008) Reaction of genotypes from several species of grain and forage legumes to infection with a French pea isolate of the oomycete *Aphanomyces euteiches*. *Eur J Plant Pathol* 122:321-333
- Moussart A, Wicker E, Duparque M, Rouxel F (2001) Development of an efficient screening test for pea resistance to *Aphanomyces euteiches*. 4th European conference on grain legumes, Cracow, Poland
- Nagel KA, Putz A, Gilmer F, Heinz K, Fischbach A, Pfeifer J, Faget M, Blossfeld S, Ernst M, Dimaki C, Kastenholz B, Kleinert A-K, Galinski A, Scharr H, Fiorani F, Schurr U (2012) GROWSCREEN-Rhizo is a novel phenotyping robot enabling simultaneous measurements of root and shoot growth for plants grown in soil-filled rhizotrons. *Functional Plant Biology* 39:891-904
- Ney B, Bancal MO, Bancal P, Bingham II, Foulkes J, Gouache D, Paveley N, Smith J (2013) Crop architecture and crop tolerance to fungal diseases and insect herbivory. Mechanisms to limit crop losses. *Eur J Plant Pathol* 135:561-580
- Orman-Ligeza B, Civava R, de Dorlodot S, Draye X (2014) Root system architecture. In: Morte A, Varma A (eds) *Root engineering : Basic and applied concepts*. Springer, Berlin, Germany, pp 39-56
- Papavizas GC, Ayers WA (1974) *Aphanomyces* species and their root diseases in pea and sugarbeet: a review. United States Dept of Agriculture Technical bulletin n°1485:1-158
- Pascual L, Albert E, Sauvage C, Duangjit J, Bouchet J-P, Bitton F, Desplat N, Brunel D, Le Paslier M-C, Ranc N, Bruguier L, Chauchard B, Verschave P, Causse M (2015a) Dissecting quantitative trait variation in the resequencing era: complementarity of bi-parental, multi-parental and association panels. *Plant Sci*
- Pascual L, Desplat N, Huang BE, Desgroux A, Bruguier L, Bouchet JP, Le QH, Chauchard B, Verschave P, Causse M (2015b) Potential of a tomato MAGIC population to decipher the genetic control of quantitative traits and detect causal variants in the resequencing era. *Plant Biotechnol J* 13:565-577
- Perez-Perez JM, Esteve-Bruna D, Micol JL (2010) QTL analysis of leaf architecture. *J Plant Res* 123:15-23

- Pieterse CMJ, Van der Does D, Zamioudis C, Leon-Reyes A, Van Wees SCM (2012) Hormonal Modulation of Plant Immunity. *Annu Rev Cell Dev Biol* 28:489-521
- Pilet-Nayel L, Coyne CJ, Hamon C, Lesne A, Le Goff I, Esnault R, Lecointe R, Roux-Duparque M, McGee RJ, Mangin P, McPhee K, Moussart A, Baranger A (2007) Understanding genetics of partial resistance to *Aphanomyces* root rot in pea for new breeding prospects. 3rd International *Aphanomyces* Workshop on Legumes, Rennes, Brittany, France, pp 34-39
- Pilet-Nayel ML, Muehlbauer FJ, McGee RJ, Kraft JM, Baranger A, Coyne CJ (2002) Quantitative trait loci for partial resistance to *Aphanomyces* root rot in pea. *Theor Appl Genet* 106:28-39
- Pilet-Nayel ML, Muehlbauer FJ, McGee RJ, Kraft JM, Baranger A, Coyne CJ (2005) Consistent Quantitative Trait Loci in Pea for Partial Resistance to *Aphanomyces euteiches* Isolates from the United States and France. *Phytopathology* 95:1287-1293
- Pilet ML, Duplan G, Archipiano M, Barret P, Baron C, Horvais R, Tanguy X, Lucas MO, Renard M, Delourme R (2001) Stability of QTL for field resistance to blackleg across two genetic backgrounds in oilseed rape. *Crop Sci* 41
- Plet J, Wasson A, Ariel F, Le Signor C, Baker D, Mathesius U, Crespi M, Frugier F (2011) MtCRE1-dependent cytokinin signaling integrates bacterial and plant cues to coordinate symbiotic nodule organogenesis in *Medicago truncatula*. *Plant Journal* 65:622-633
- Poland JA, Balint-Kurti PJ, Wisser RJ, Pratt RC, Nelson RJ (2009) Shades of gray: the world of quantitative disease resistance. *Trends Plant Sci* 14:21-29
- Prioul S, Frankewitz A, Deniot G, Morin G, Baranger A (2004) Mapping of quantitative trait loci for partial resistance to *Mycosphaerella pinodes* in pea (*Pisum sativum* L.), at the seedling and adult plant stages. *Theor Appl Genet* 108:1322-1334
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155:945-959
- Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, de Bakker PIW, Daly MJ, Sham PC (2007) PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 81:559-575
- Purcell SM, Chang CC (2015) PLINK 1.9.
- Qiu D, Xiao J, Ding X, Xiong M, Cai M, Cao Y, Li X, Xu C, Wang S (2007) OsWRKY13 mediates rice disease resistance by regulating defense-related genes in salicylate- and jasmonate-dependent signaling. *Mol Plant Microbe Interact* 20:492-499
- R Core Team (2014) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria
- Revelle W (2015) psych: procedures for psychological, psychometric, and personality research. R package version 1.5.1.
- Richard B, Bussière F, Langrume C, Rouault F, Jumel S, Faivre R, Tivoli B (2013) Effect of pea canopy architecture on microclimate and consequences on ascochyta blight infection under field conditions. *Eur J Plant Pathol* 135:509-524
- Rogers ED, Benfey PN (2015) Regulation of plant root system architecture: implications for crop advancement. *Current Opinion in Biotechnology* 32:93-98
- Román-Avilés B, Snapp SS, Kelly JD (2004) Assessing root traits associated with root rot resistance in common bean. *Field Crops Research* 86:147-156

- Roux-Duparque M, Boitel C (2004) Breeding peas for resistance to *Aphanomyces* root rot: current main outputs of three breeding programmes. 5th Eur Conf Grain Legumes, Dijon, France, p 133
- Salon C, Jeudy C, Bernard C, Mougel C, Coffin A, Bourion V, Moreau D, Palavioux K, Baussart C (2014) Rhizotron et utilisations de celui-ci. Google Patents
- Sampo S, Avidano L, Berta G (2007) Morphogenetic effects induced by pathogenic and non pathogenic *Rhizoctonia solani* Kuhn strains on tomato roots. *Caryologia* 60:141-145
- Sax K (1923) The Association of Size Differences with Seed-Coat Pattern and Pigmentation in PHASEOLUS VULGARIS. *Genetics* 8:552-560
- Schlaeppli K, Abou-Mansour E, Buchala A, Mauch F (2010) Disease resistance of *Arabidopsis* to *Phytophthora brassicae* is established by the sequential action of indole glucosinolates and camalexin. *Plant J* 62:840-851
- Schwender H, Fritsch A (2013) scribe: analysis of high-dimensional categorical data such as SNP data. R package version 1.3.3.
- Segura V, Vilhjalmsson BJ, Platt A, Korte A, Seren U, Long Q, Nordborg M (2012) An efficient multi-locus mixed-model approach for genome-wide association studies in structured populations. *Nat Genet* 44:825-830
- Sekhwal MK, Li P, Lam I, Wang X, Cloutier S, You FM (2015) Disease Resistance Gene Analogs (RGAs) in Plants. *Int J Mol Sci* 16:19248-19290
- Simon S, Lauri PE, Brun L, Defrance H, Sauphanor B (2006) Does manipulation of fruit-tree architecture affect the development of pests and pathogens? A case study in an organic apple orchard. *The Journal of Horticultural Science and Biotechnology* 81:765-773
- Snapp S, Kirk W, Román-Avilés B, Kelly JD (2003) Root traits play a role in integrated management of *Fusarium* root rot in snap beans. *HortScience* 32:187-191
- Sorin C, Bussell JD, Camus I, Ljung K, Kowalczyk M, Geiss G, McKhann H, Garcion C, Vaucheret H, Sandberg G, Bellini C (2005) Auxin and Light Control of Adventitious Rooting in Arabidopsis Require ARGONAUTE1. *Plant Cell* 17:1343-1359
- Sosnowski O, Charcosset A, Joets J (2012) BioMercator V3: an upgrade of genetic map compilation and quantitative trait loci meta-analysis algorithms. *Bioinformatics* 28:2082-2083
- Tar'an B, Warkentin T, Somers DJ, Miranda D, Vandenberg A, Blade S, Bing D (2004) Identification of quantitative trait loci for grain yield, seed protein concentration and maturity in field pea (*Pisum sativum* L.). *Euphytica* 136:297-306
- Tayeh N, Aluome C, Falque M, Jacquin F, Klein A, Chauveau A, Bérard A, Houtin H, Rond C, Kreplak J, Boucherot K, Martin C, Baranger A, Pilet-Nayel M-L, Warkentin TD, Brunel D, Marget P, Le Paslier M-C, Aubert G, Burstin J (2015a) Development of two major resources for pea genomics: the GenoPea 13.2K SNP Array and a high density, high resolution consensus genetic map. *Plant J*
- Tayeh N, Aubert G, Pilet-Nayel M-L, Lejeune-Hénaut I, Warkentin TD, Burstin J (2015b) Genomic tools in pea breeding programs: status and perspectives. *Front Plant Sci* 6
- Terres Univia (2014) Oléagineux et plantes riches en protéines. Chiffres Clés:1-20
- Tiryaki I, Staswick PE (2002) An *Arabidopsis* mutant defective in jasmonate response is allelic to the auxin-signaling mutant *axr1*. *Plant Physiol* 130:887-894
- Tivoli B, Andrivon D, Baranger A, Calonnec A, Jeger M (2012a) Foreword: plant and canopy architecture impact on disease epidemiology and pest development. *Eur J Plant Pathol* 135:453-454

- Tivoli B, Calon nec A, Richard B, Ney B, Andrivon D (2012b) Current knowledge on plant/canopy architectural traits that reduce the expression and development of epidemics. *Eur J Plant Pathol* 135:471-478
- Trotta A, Varese GC, Gnavi E, Fusconi A, Sampo S, Berta G (1996) Interactions between the soilborne root pathogen *Phytophthora nicotianae* var *parasitica* and the arbuscular mycorrhizal fungus *Glomus mosseae* in tomato plants. *Plant and Soil* 185:199-209
- UNIP, ARVALIS (2014) Pois protéagineux de printemps et d'hiver. Guide de culture 2014-2015:1-40
- Voorrips RE (2002) MapChart: Software for the Graphical Presentation of Linkage Maps and QTLs. *J Hered* 93:77-78
- Wachsman G, Sparks EE, Benfey PN (2015) Genes and networks regulating root anatomy and architecture. *New Phytol* 208:26-38
- Wallace JG, Larsson SJ, Buckler ES (2014) Entering the second century of maize quantitative genetics. *Heredity* 112:30-38
- Wang D, Pajerowska-Mukhtar K, Culler AH, Dong X (2007) Salicylic acid inhibits pathogen growth in plants through repression of the auxin signaling pathway. *Curr Biol* 17:1784-1790
- Wang G-L, Ruan D-L, Song W-Y, Sideris S, Chen L, Pi L-Y, Zhang S, Zhang Z, Fauquet C, Gaut BS, Whalen MC, Ronald PC (1998) Xa21D encodes a receptor-like molecule with a leucine-rich repeat domain that determines race-specific recognition and is subject to adaptive evolution. *Plant Cell* 10:765-779
- Wang Q, Liu Y, He J, Zheng X, Hu J, Liu Y, Dai H, Zhang Y, Wang B, Wu W, Gao H, Zhang Y, Tao X, Deng H, Yuan D, Jiang L, Zhang X, Guo X, Cheng X, Wu C, Wang H, Yuan L, Wan J (2014) STV11 encodes a sulphotransferase and confers durable resistance to rice stripe virus. *Nat Commun* 5
- Warkentin TD (2015) Pea. In: De Ron AM (ed) Grain Legumes. Springer-Verlag New York
- Weaver JE (1919) The ecological relations of roots. Papers of John E Weaver (1884 - 1956) Paper 13
- Weeden NF, McGee RJ, Grau CR, Muehlbauer FJ (2000) A gene influencing tolerance to common root rot is located on linkage group IV. *Pisum Genetics* 32:53-55
- Wellensiek SJ (1925) Genetic monograph on pisum. [s.n.], Wageningen
- White FF, Frommer W (2015) Deciphering durable resistance one R gene at a time. *Nat Genet* 47:1376-1377
- Wicker E (2001) Diversité des populations françaises d'*Aphanomyces euteiches* Drechs, agent de la pourriture racinaire du pois : variabilité pathogène et moléculaire. Ecole Nationale Supérieure Agronomique de Rennes, Rennes
- Wicker E, Hullé M, Rouxel F (2001) Pathogenic characteristics of isolates of *Aphanomyces euteiches* from pea in France. *Plant Pathology* 50:433-442
- Wicker E, Moussart A, Duparque M, Rouxel F (2003) Further contributions to the development of a differential set of pea cultivars (*Pisum sativum*) to investigate the virulence of isolates of *Aphanomyces euteiches*. *Eur J Plant Pathol* 109:47-60
- Wicker E, Rouxel F (2001) Specific behaviour of French *Aphanomyces euteiches* Drechs. populations for virulence and aggressiveness on pea, related to isolates from Europe, America and New Zealand. *Eur J Plant Pathol* 107:919-929
- Willoquet L, Jumel S, Lemarchand E (2007) Spatio-temporal development of pea root rot disease through secondary infections during a crop cycle. *Journal of Phytopathology* 155:623-632

- Yu JM, Pressoir G, Briggs WH, Bi IV, Yamasaki M, Doebley JF, McMullen MD, Gaut BS, Nielsen DM, Holland JB, Kresovich S, Buckler ES (2006) A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nat Genet* 38:203-208
- Zappala S, Mairhofer S, Tracy S, Sturrock CJ, Bennett M, Pridmore T, Mooney SJ (2013) Quantifying the effect of soil moisture content on segmenting root system architecture in X-ray computed tomography images. *Plant and Soil* 370:35-45
- Zarzynska K, Szutkowska M (2013) Developmental differences, yield and late blight (*Phytophthora infestans*) infection of potato plants grown under organic and conventional systems. *Journal of Agricultural Science and Technology A* 3:281-289
- Zeng ZB (1994) Precision mapping of quantitative trait loci. *Genetics* 136:1457-1468
- Zheng Z, Qamar SA, Chen Z, Mengiste T (2006) Arabidopsis WRKY33 transcription factor is required for resistance to necrotrophic fungal pathogens. *Plant J* 48:592-605

Annexes

Annexe 1: Additional files du Chapitre II

Additional File II.1

Description of the pea-Aphanomyces collection

(a) Line categories, as referred in Figure 1B and Additional File 9 in legends; (b) Line codes in the Biological resource center (CRB) database of INRA Dijon, France; (c) Country codes ISO3166-1; (d) Cv: Cultivar, Bl: breeding line, Gmp: germplasm, Lv: local variety, Wa: wild accession; (e) Fd: fodder, Gd: garden; (f) Ws: winter sown, Ss: spring sown; (g) Nl: normal leaves, af: afila leaves, (h) W: white, Pu: purple, Pi: pink; (i) Wr: wrinkled, Sm: smooth; #: parents of GSP breeding programs; \$: parents of RIL populations

Line code	Line category (a)	Common Name	CRB Code (b)	Country of origin (c)	Breeding company or donor	Registration year	Cultivation status (d)	End use (e)	Sowing type (f)	Foliage type (g)	Flower colour (h)	Seed type (i)	Pedigree	Reference
A009	GSP1	AeD9904-7		FRA	GSP	-	Gmp	Dry Ss	Ss	af	W	Sm	(Baccara x P1180693) x (Capella x Mn314)	[37,38]
A010	GSP0	AeB9703730		FRA	GSP	-	Gmp	Dry Ss	Ss	af	W	Sm		[37,38]
A015	GSP1	AeA95-299.M.32.1		FRA	GSP	-	Gmp	Dry Ss	Ss	Ni	W	Sm	Baccara x 90-2131	[37,38]
A016	GSP1	AeA95-304.M.4.4.5		FRA	GSP	-	Gmp	Dry Ss	Ss	af	W	Sm	Baccara x 90-2131	[37,38]
A017	GSP1	AeA95-306.NZ.13.2		FRA	GSP	-	Gmp	Dry Ss	Ss	af	W	Sm	Baccara x 90-2131	[37,38]
A018	GSP1	AeA95-315.M.43.5		FRA	GSP	-	Gmp	Dry Ss	Ss	af	Pu	Wr	P1180693 x Capella	[37,38]
A019	GSP1	AeA95-317.M.71.2		FRA	GSP	-	Gmp	Dry Ss	Ss	Ni	Pu	Wr	P1180693 x Capella	[37,38]
A020	GSP1	AeA95-320.M.71.2		FRA	GSP	-	Gmp	Dry Ss	Ss	Ni	Pi	Sm	P1180693 x Capella	[37,38]
A021	GSP1	AeA95-321.M.33.3		FRA	GSP	-	Gmp	Dry Ss	Ss	af	Pu	Sm	P1180693 x Capella	[37,38]
A022	GSP1	AeA95-325.M.34.1		FRA	GSP	-	Gmp	Dry Ss	Ss	af	Pu	Sm	P1180693 x Capella	[37,38]
A023	GSP1	AeA95-326.M.7.1.1		FRA	GSP	-	Gmp	Dry Ss	Ss	af	Pi	Sm	P1180693 x Capella	[37,38]
A024	GSP1	AeA95-327.M.24.7		FRA	GSP	-	Gmp	Dry Ss	Ss	af	Pi	Sm	P1180693 x Capella	[37,38]
A025	GSP2	AeA95-335.M.51.3		FRA	GSP	-	Gmp	Dry Ss	Ss	Ni	Pi	Sm	(Baccara x 90-2131) x (P1180693 x Capella)	[37,38]
A026	GSP2	AeA95-343.M.61.5		FRA	GSP	-	Gmp	Dry Ss	Ss	af	W	Sm	(Baccara x 90-2131) x (P1180693 x Capella)	[37,38]
A027	GSP0	AeA95-359.2002.1.4		FRA	GSP	-	Gmp	Dry Ss	Ss	af	Pi	Sm	(Capella x Mistral) x (Rustic x Victor)	[37,38]
A028	GSP2	AeA95-376.M.33.1		FRA	GSP	-	Gmp	Dry Ss	Ss	af	Pi	Sm	(P1269802 x Baccara) x (90-2131 x P1180693)	[37,38]
A029	GSP0	AeB97-709.11.1		FRA	GSP	-	Gmp	Dry Ss	Ss	af	W	Sm	200 x Capella	[37,38]
A030	GSP0	AeB97-710.10.5		FRA	GSP	-	Gmp	Dry Ss	Ss	af	W	Sm	200 x Capella	[37,38]
A031	GSP0	AeB97-712.11.1		FRA	GSP	-	Gmp	Dry Ss	Ss	af	W	Sm	200 x Capella	[37,38]
A032	GSP1	AeB97-722.14.1		FRA	GSP	-	Gmp	Dry Ss	Ss	af	W	Sm	(200 x Capella) x (90-2131 x Rustic)	[37,38]
A033	GSP1	AeB97-723.3.2		FRA	GSP	-	Gmp	Dry Ss	Ss	af	W	Sm	(90-2131 x Rustic) x (200 x Capella)	[37,38]
A034	GSP1	AeB97-724.2.3		FRA	GSP	-	Gmp	Dry Ss	Ss	af	W	Sm	(90-2131 x Rustic) x (200 x Capella)	[37,38]
A035	GSP1	AeB97-725.11.2		FRA	GSP	-	Gmp	Dry Ss	Ss	af	W	Sm	(90-2131 x Rustic) x (200 x Capella)	[37,38]
A036	GSP1	AeB97-729.2.2		FRA	GSP	-	Gmp	Dry Ss	Ss	Ni	W	Sm	(Rustic x 200) x (Capella x 90-2131)	[37,38]
A037	GSP1	AeB97-730.10.2		FRA	GSP	-	Gmp	Dry Ss	Ss	af	W	Sm	(45 x 90-2131) x (200 x Baccara)	[37,38]
A038	GSP1	AeB97-736.3.10		FRA	GSP	-	Gmp	Dry Ss	Ss	af	W	Sm	(45 x 90-2131) x (200 x Baccara)	[37,38]
A039	GSP1	AeB97-738.2.3		FRA	GSP	-	Gmp	Dry Ss	Ss	Ni	W	Sm	(45 x 90-2131) x (200 x Baccara)	[37,38]
A040	GSP1	AeB97-739.15.2		FRA	GSP	-	Gmp	Dry Ss	Ss	Ni	W	Sm	(45 x 90-2131) x (200 x Baccara)	[37,38]
A041	GSP1	AeB97-740.11.1		FRA	GSP	-	Gmp	Dry Ss	Ss	Ni	W	Sm	(90-2131 x Rustic) x (200 x Capella)	[37,38]
A042	GSP1	AeB97-741.9.6		FRA	GSP	-	Gmp	Dry Ss	Ss	Ni	W	Sm	(45 x 90-2131) x (200 x Baccara)	[37,38]
A043	GSP1	AeB97-745.2.1		FRA	GSP	-	Gmp	Dry Ss	Ss	af	W	Sm	(45 x 90-2131) x (200 x Baccara)	[37,38]
A044	GSP1	AeB97-747.3.7		FRA	GSP	-	Gmp	Dry Ss	Ss	af	W	Wr	(45 x 90-2131) x (200 x Baccara)	[37,38]
A045	GSP1	AeD99QU-04-3p-7-1		FRA	GSP	-	Gmp	Dry Ss	Ss	Ni	Pu	Sm	(Baccara x P1180693) x (Capella x Mn314)	[37,38]
A046	GSP1	AeD99QU-04-3sp-9-2		FRA	GSP	-	Gmp	Dry Ss	Ss	Ni	Pu	Sm	(Baccara x P1180693) x (Capella x Mn314)	[37,38]
A047	GSP1	AeD99QU-04-4-6-1		FRA	GSP	-	Gmp	Dry Ss	Ss	Ni	Pu	Sm	(Baccara x P1180693) x (Capella x Mn314)	[37,38]
A048	GSP1	AeD99QU-04-6-6-3		FRA	GSP	-	Gmp	Dry Ss	Ss	af	W	Sm	(Baccara x P1180693) x (Capella x Mn314)	[37,38]
A049	GSP1	AeD99QU-04-7-3-1		FRA	GSP	-	Gmp	Dry Ss	Ss	af	W	Sm	(Baccara x P1180693) x (Capella x Mn314)	[37,38]
A050	GSP1	AeD99QU-04-8b-5-2		FRA	GSP	-	Gmp	Dry Ss	Ss	af	W	Sm	(Baccara x P1180693) x (Capella x Mn314)	[37,38]
A051	GSP1	AeD99QU-04-8v-2-5		FRA	GSP	-	Gmp	Dry Ss	Ss	af	Pu	Sm	(Baccara x P1180693) x (Capella x Mn314)	[37,38]
A052	GSP1	AeD99QU-04-9-10-1		FRA	GSP	-	Gmp	Dry Ss	Ss	af	W	Sm	(Baccara x P1180693) x (Capella x Mn314)	[37,38]
A053	GSP1	AeD99QU-04-14-10-4		FRA	GSP	-	Gmp	Dry Ss	Ss	af	W	Sm	(Rustic x 200) x (Mn314 x 90-2131)	[37,38]
A054	GSP1	AeD99QU-04-15-8-1		FRA	GSP	-	Gmp	Dry Ss	Ss	af	W	Sm	(Rustic x 200) x (Mn314 x 90-2131)	[37,38]
A055	GSP1	AeD99QU-04-17-8-1		FRA	GSP	-	Gmp	Dry Ss	Ss	af	W	Sm	(Rustic x 200) x (Mn314 x 90-2131)	[37,38]
A056	GSP1	AeD99QU-04-18af-10-2		FRA	GSP	-	Gmp	Dry Ss	Ss	af	W	Sm	(Rustic x 200) x (Mn314 x 90-2131)	[37,38]
A057	GSP1	AeD99QU-04-21-9-1		FRA	GSP	-	Gmp	Dry Ss	Ss	af	W	Sm	(Baccara x P1180693) x (Capella x Mn314)	[37,38]
A058	GSP1	AeD99QU-04-22-10-3		FRA	GSP	-	Gmp	Dry Ss	Ss	af	W	Sm	(Baccara x P1180693) x (Capella x Mn314)	[37,38]
A059	GSP1	AeD99QU-04-23-8-1		FRA	GSP	-	Gmp	Dry Ss	Ss	af	W	Sm	(Rustic x 200) x (Mn314 x 90-2131)	[37,38]
A060	GSP1	AeD99QU-04-24-3-2		FRA	GSP	-	Gmp	Dry Ss	Ss	af	W	Sm	(Rustic x 200) x (Mn314 x 90-2131)	[37,38]
A061	GSP0	AeD99OSW- 1-10-8		FRA	GSP	-	Gmp	Dry Ss	Ss	af	W	Wr	(Rustic x 200) x (Capella x Mn314)	[37,38]

(i) Selected lines from GSP breeding program

Line code	Line category (a)	Common Name	CRB Code		Country of origin	Breeding company or donor	Registration year	Cultivation status	End use (e)	Sowing type (f)	Foliage type (g)	Flower colour (h)	Seed type (i)	Pedigree	Reference
			(b)	(c)											
A062	GSP0	AeD99OSW- 2-7-8	FRA	FRA	GSP	-	Gmp	Dry	Ss	af	W	Wr	(Rustic x 200) x (Capella x Mn314)	[37,38]	
A063	GSP0	AeD99OSW- 3-8-7	FRA	FRA	GSP	-	Gmp	Dry	Ss	af	W	Wr	(Rustic x 200) x (Capella x Mn314)	[37,38]	
A064	GSP0	AeD99OSW- 4-6-4	FRA	FRA	GSP	-	Gmp	Dry	Ss	af	W	Sm	(Rustic x 200) x (Capella x Mn314)	[37,38]	
A065	GSP0	AeD99OSW- 5-9-2	FRA	FRA	GSP	-	Gmp	Dry	Ss	af	W	Wr	(Rustic x 200) x (Capella x Mn314)	[37,38]	
A066	GSP0	AeD99OSW- 6-2-3	FRA	FRA	GSP	-	Gmp	Dry	Ss	af	W	Sm	(Rustic x 200) x (Capella x Mn314)	[37,38]	
A067	GSP0	AeD99OSW- 7-6-7	FRA	FRA	GSP	-	Gmp	Dry	Ss	Nl	W	Sm	(Rustic x 200) x (Capella x Mn314)	[37,38]	
A068	GSP0	AeD99OSW- 8-8-1	FRA	FRA	GSP	-	Gmp	Dry	Ss	af	W	Sm	(Rustic x 200) x (Capella x Mn314)	[37,38]	
A069	GSP0	AeD99OSW- 9-9-2	FRA	FRA	GSP	-	Gmp	Dry	Ss	af	W	Sm	(Rustic x 200) x (Capella x Mn314)	[37,38]	
A070	GSP0	AeD99OSW- 10-10-6	FRA	FRA	GSP	-	Gmp	Dry	Ss	af	W	Sm	(Rustic x 200) x (Capella x Mn314)	[37,38]	
A071	GSP0	AeD99OSW- 11-7-4	FRA	FRA	GSP	-	Gmp	Dry	Ss	af	W	Sm	(Rustic x 200) x (Capella x Mn314)	[37,38]	
A072	GSP1	AeD99OSW- 12-18-4	FRA	FRA	GSP	-	Gmp	Dry	Ss	Nl	Pl	Wr	(Baccara x P1180693) x (Capella x Mn314)	[37,38]	
A073	GSP1	AeD99OSW- 13-9-1	FRA	FRA	GSP	-	Gmp	Dry	Ss	af	Pl	Sm	(Baccara x P1180693) x (Capella x Mn314)	[37,38]	
A074	GSP1	AeD99OSW- 14-8-8	FRA	FRA	GSP	-	Gmp	Dry	Ss	af	W	Wr	(Baccara x P1180693) x (Capella x Mn314)	[37,38]	
A075	GSP1	AeD99OSW- 15-7-4	FRA	FRA	GSP	-	Gmp	Dry	Ss	Nl	W	Sm	(Baccara x P1180693) x (Capella x Mn314)	[37,38]	
A076	GSP1	AeD99OSW- 16-1-10	FRA	FRA	GSP	-	Gmp	Dry	Ss	af	W	Wr	(Baccara x P1180693) x (Capella x Mn314)	[37,38]	
A077	GSP1	AeD99OSW- 18-6-5	FRA	FRA	GSP	-	Gmp	Dry	Ss	Nl	W	Sm	(Baccara x P1180693) x (Capella x Mn314)	[37,38]	
A078	GSP1	AeD99OSW- 20-3-7	FRA	FRA	GSP	-	Gmp	Dry	Ss	af	W	Sm	(Baccara x P1180693) x (Capella x Mn314)	[37,38]	
A079	GSP1	AeD99OSW- 22-3-7	FRA	FRA	GSP	-	Gmp	Dry	Ss	Nl	W	Sm	(Baccara x P1180693) x (Capella x Mn314)	[37,38]	
A080	GSP1	AeD99OSW- 23-7-4	FRA	FRA	GSP	-	Gmp	Dry	Ss	Nl	W	Sm	(Baccara x P1180693) x (Capella x Mn314)	[37,38]	
A081	GSP1	AeD99OSW- 24-10-5	FRA	FRA	GSP	-	Gmp	Dry	Ss	af	W	Sm	(Baccara x P1180693) x (Capella x Mn314)	[37,38]	
A082	GSP1	AeD99OSW- 25-9-4	FRA	FRA	GSP	-	Gmp	Dry	Ss	af	W	Sm	(Baccara x P1180693) x (Capella x Mn314)	[37,38]	
A083	GSP1	AeD99OSW- 26-3-8	FRA	FRA	GSP	-	Gmp	Dry	Ss	Nl	W	Wr	(Baccara x P1180693) x (Capella x Mn314)	[37,38]	
A084	GSP1	AeD99OSW- 27-7-6	FRA	FRA	GSP	-	Gmp	Dry	Ss	af	W	Sm	(Baccara x P1180693) x (Capella x Mn314)	[37,38]	
A085	GSP2	AeD99OSW- 28-8-4	FRA	FRA	GSP	-	Gmp	Dry	Ss	af	Pu	Sm	(90-2131 x P1180693) x (Capella x Mn314)	[37,38]	
A086	GSP2	AeD99OSW- 29-3-3	FRA	FRA	GSP	-	Gmp	Dry	Ss	af	W	Sm	(90-2131 x P1180693) x (Capella x Mn314)	[37,38]	
A087	GSP2	AeD99OSW- 30-4-4	FRA	FRA	GSP	-	Gmp	Dry	Ss	Nl	W	Sm	(90-2131 x P1180693) x (Capella x Mn314)	[37,38]	
A088	GSP2	AeD99OSW- 33-7-5	FRA	FRA	GSP	-	Gmp	Dry	Ss	af	W	Sm	(90-2131 x P1180693) x (Capella x Mn314)	[37,38]	
A089	GSP2	AeD99OSW- 34-7-2	FRA	FRA	GSP	-	Gmp	Dry	Ss	Nl	-	Sm	(90-2131 x P1180693) x (Capella x Mn314)	[37,38]	
A090	GSP3	AeD99OSW- 35-9-2	FRA	FRA	GSP	-	Gmp	Dry	Ss	Nl	W	Wr	(90-2131 x P1180693) x 552	[37,38]	
A091	GSP3	AeD99OSW- 36-9-10	FRA	FRA	GSP	-	Gmp	Dry	Ss	Nl	Pu	Wr	(90-2131 x P1180693) x 552	[37,38]	
A092	GSP3	AeD99OSW- 37-3-4	FRA	FRA	GSP	-	Gmp	Dry	Ss	Nl	W	Sm	(90-2131 x P1180693) x 552	[37,38]	
A093	GSP3	AeD99OSW- 38-5-4	FRA	FRA	GSP	-	Gmp	Dry	Ss	Nl	W	Wr	(90-2131 x P1180693) x 552	[37,38]	
A094	GSP3	AeD99OSW- 39-5-5	FRA	FRA	GSP	-	Gmp	Dry	Ss	Nl	W	Wr	(90-2131 x P1180693) x 552	[37,38]	
A095	GSP3	AeD99OSW- 40-3-3	FRA	FRA	GSP	-	Gmp	Dry	Ss	Nl	Pl	Sm	(90-2131 x P1180693) x 552	[37,38]	
A096	GSP3	AeD99OSW- 41-7-6	FRA	FRA	GSP	-	Gmp	Dry	Ss	Nl	Pl	Wr	(90-2131 x P1180693) x 552	[37,38]	
A097	GSP3	AeD99OSW- 42-3-1	FRA	FRA	GSP	-	Gmp	Dry	Ss	Nl	W	Sm	(90-2131 x P1180693) x 552	[37,38]	
A098	GSP3	AeD99OSW- 43-9-8	FRA	FRA	GSP	-	Gmp	Dry	Ss	Nl	W	Sm	(90-2131 x P1180693) x 552	[37,38]	
A099	GSP3	AeD99OSW- 44-19-1	FRA	FRA	GSP	-	Gmp	Dry	Ss	Nl	W	Sm	(90-2131 x P1180693) x 552	[37,38]	
A100	GSP3	AeD99OSW- 45-8-7	FRA	FRA	GSP	-	Gmp	Dry	Ss	Nl	Pu	Sm	(90-2131 x P1180693) x 552	[37,38]	
A101	GSP3	AeD99OSW- 47-6-1	FRA	FRA	GSP	-	Gmp	Dry	Ss	Nl	W	Sm	(90-2131 x P1180693) x 552	[37,38]	
A102	GSP3	AeD99OSW- 48-5-4	FRA	FRA	GSP	-	Gmp	Dry	Ss	Nl	W	Sm	(90-2131 x P1180693) x 552	[37,38]	
A103	GSP3	AeD99OSW- 49-5-7	FRA	FRA	GSP	-	Gmp	Dry	Ss	Nl	Pu	Sm	(90-2131 x P1180693) x 552	[37,38]	
A104	GSP3	AeD99OSW- 50-2-5	FRA	FRA	GSP	-	Gmp	Dry	Ss	Nl	W	Sm	(90-2131 x P1180693) x 552	[37,38]	
A105	GSP3	AeD99OSW- 51-2-10	FRA	FRA	GSP	-	Gmp	Dry	Ss	Nl	W	Sm	(90-2131 x P1180693) x 552	[37,38]	
A106	GSP3	AeD99OSW- 53-3-5	FRA	FRA	GSP	-	Gmp	Dry	Ss	Nl	Pu	Sm	(90-2131 x P1180693) x 552	[37,38]	
A107	GSP3	AeD99OSW- 54-10-5	FRA	FRA	GSP	-	Gmp	Dry	Ss	Nl	W	Sm	(90-2131 x P1180693) x 552	[37,38]	
A108	GSP3	AeD99OSW- 55-6-4	FRA	FRA	GSP	-	Gmp	Dry	Ss	Nl	W	Sm	(90-2131 x P1180693) x 552	[37,38]	
A109	GSP3	AeD99OSW- 56-19-4	FRA	FRA	GSP	-	Gmp	Dry	Ss	Nl	Pl	Sm	(90-2131 x P1180693) x 552	[37,38]	
A110	GSP3	AeD99OSW- 57-9-8	FRA	FRA	GSP	-	Gmp	Dry	Ss	Nl	W	Sm	(90-2131 x P1180693) x 552	[37,38]	

(i) Selected lines from GSP breeding program

Line code	Line category (a)	Common Name	CRB Code (b)	Country of origin (c)	Breeding company or donor	Registration year	Cultivation status (d)	End use (e)	Sowing type (f)	Foliage type (g)	Flower colour (h)	Seed type (i)	Pedigree	Reference
A111	GSP3	AeD99OSW- 58-10-5		FRA	GSP	-	Gmp	Dry Ss	-	Ni W	W	Sm	(90-2131 x P1180693) x 552	[37,38]
A112	GSP3	AeD99OSW- 59-7-4		FRA	GSP	-	Gmp	Dry Ss	-	Ni W	W	Sm	(90-2131 x P1180693) x 552	[37,38]
A113	GSP3	AeD99OSW- 60-1-3		FRA	GSP	-	Gmp	Dry Ss	-	Ni W	W	Sm	(90-2131 x P1180693) x 552	[37,38]
MR01	BRILs ia	BAP8.10		FRA	INRA	-	Gmp	-	-	af W	W	Sm	Baccara x P1180693	[41]
MR02	BRILs ia	BAP8.43		FRA	INRA	-	Gmp	-	-	Ni W	W	Sm	Baccara x P1180693	[41]
MR03	BRILs ia	BAP8.128		FRA	INRA	-	Gmp	-	-	Ni W	W	Sm	Baccara x P1180693	[41]
MR04	BRILs ia	BAP8.184		FRA	INRA	-	Gmp	-	-	Ni W	W	Sm	Baccara x P1180693	[41]
MR05	BRILs ib	BA55.8		FRA	INRA	-	Gmp	-	-	af W	W	Sm	Baccara x 552	[41]
MR06	BRILs ib	BA55.26		FRA	INRA	-	Gmp	-	-	Ni W	W	Sm	Baccara x 552	[41]
MR07	BRILs ib	BA55.52		FRA	INRA	-	Gmp	-	-	Ni W	W	Sm	Baccara x 552	[41]
MR08	BRILs ib	BA55.88		FRA	INRA	-	Gmp	-	-	Ni W	W	Sm	Baccara x 552	[41]
MR09	BRILs ib	BA55.161		FRA	INRA	-	Gmp	-	-	Ni W	W	Sm	Baccara x 552	[41]
MR10	BRILs ib	BA55.171		FRA	INRA	-	Gmp	-	-	Ni W	W	Sm	Baccara x 552	[41]
MR11	BRILs ib	BA55.233		FRA	INRA	-	Gmp	-	-	Ni W	W	Sm	Baccara x 552	[41]
MR12	BRILs ib	BA55.231		FRA	INRA	-	Gmp	-	-	Ni W	W	Sm	Baccara x 552	[41]
MR13	BRILs ic	RIL 846-07		USA, WA	USDA	2012	Gmp	Gd	-	Ni W	W	Sm	DSP x 90-2131	[42,43]
MR14	BRILs ic	RIL 847-08		USA, WA	USDA	2012	Gmp	Gd	-	Ni W	W	Sm	DSP x 90-2131	[42,43]
MR15	BRILs ic	RIL 847-22		USA, WA	USDA	2012	Gmp	Gd	-	Ni W	W	Sm	DSP x 90-2131	[42,43]
MR16	BRILs ic	RIL 847-28		USA, WA	USDA	2012	Gmp	Gd	-	Ni W	W	Sm	DSP x 90-2131	[42,43]
MR17	BRILs ic	RIL 847-45		USA, WA	USDA	2012	Gmp	Gd	-	Ni W	W	Sm	DSP x 90-2131	[42,43]
MR18	BRILs ic	RIL 847-50		USA, WA	USDA	2012	Gmp	Gd	-	Ni W	W	Sm	DSP x 90-2131	[42,43]
MR19	BRILs ic	RIL 847-68		USA, WA	USDA	2012	Gmp	Gd	-	Ni W	W	Sm	DSP x 90-2131	[42,43]
MR20	BRILs ic	RIL 847-53		USA, WA	USDA	2012	Gmp	Gd	-	Ni W	W	Sm	DSP x 90-2131	[42,43]
#\$ A001	Parents	90-2079	CRB119	USA, WA	USDA, Kraft	1992	Gmp	Gd Ss	af W	W	W	Sm	[(69-1004 x Recette) x 790-213122]	[32]
# A002	Parents	0-2100		USA, WA	USDA, Kraft	-	Gmp	Ss	Ni W	W	W	Sm		[37,38]
#\$ A003	Parents	90-2131	CRB135	USA, WA	USDA, Kraft	1992	Gmp	Gd Ss	Ni W	W	W	Sm	[(Small Sieve Freezer' x C-165) x (Early Perfection 3040' x C-165) x PH-91-3] x 74SN4} x P1180693	[29,32]
# A006	Parents	MISTRAL	CRB103	FRA	Sersem	1991	Cv	Dry Ws	-	Ni W	W	Sm		[37,38]
# A007	Parents	RUSTIC		FRA	Sersem	1992	Cv	-	-	af W	W	Sm		[37,38]
# A008	Parents	0-2093	CRB147	USA, WA	USDA, Kraft	-	Gmp	-	Ss	Ni W	W	Sm		[37,38]
# A011	Parents	Mn314	CRB088	USA, MN	Univ. Min. Davis	-	Gmp	Gd	-	af W	W	Sm		[33,37,38]
# A012	Parents	P1269802	CRB079	GBR	USDA, USA, WA	-	Gmp	-	-	Ni Pu	W	Sm		[37,38]
# A013	Parents	200	CRB085	USA, WI	Univ. Wis., Gritton	-	Gmp	Gd	-	Ni W	W	Sm		[37,38]
# A014	Parents	45	CRB083	USA, WI	Univ. Wis., Gritton	-	Gmp	Gd	-	Ni W	W	Sm		[37,38]
# A114	Parents	CAPELLA	CRB127	SWE	Lantmännen ek. för.	-	Cv	Dry Ss	af W	W	W	Sm		[29]
#\$ A115	Parents	DSP (Dark skin perfection)	CRB134	USA, WA	USDA, US	1995	Cv	Gd Ss	Ni W	W	W	Sm		-
#\$ A116	Parents	PH180693	CRB128	DEU	USDA, US	-	Gmp	-	-	Ni Pi	W	Sm		[29]
# A120	Parents	VICTOR		USA	Pioneer	1992	Cv	Dry Ws	Ni W	W	W	Sm	Progreita x (FrimaxPrint)	[37,38]
#\$ A125	Parents	BACCARA	CRB423	FRA	F. Desprez	1991	Cv	Dry Ss	af W	W	W	Sm	Progreita x (FrimaxPrint)	[37,38]
#\$ A126	Parents	PUGET	CRB412	GBR	Brotherton	-	Cv	Gd Ss	Ni W	W	W	Sm	(USA&11197 x Finale) x Finale	[29]
#\$ A127	Parents	552	CRB130	USA, WI	Univ. Wis., Gritton	-	Gmp	Gd Ss	Ni W	W	W	Sm		[29]
# A142	Parents	MN313	CRB129	USA, MN	Univ. Min. Davis	1995	Gmp	Gd Ss	Ni W	W	W	Sm		[29,33]
A004	New sources	0-2042	CRB144	USA, WA	USDA, Kraft	-	Gmp	-	Ss	Ni W	W	Sm		[36]
A117	New sources	0-2016	CRB143	USA, WA	USDA, Kraft	-	Gmp	-	Ss	Ni W	W	Sm		[36]
A128	New sources	ROSABLENDE				-	Wa	Fd	-	Ni Pi	W	Sm		[36]

(i)

(ii) RILs from INRA and USDA research programs

program

(iii) Parents of RIL populations and/or GSP breeding

RA large

Line code	Line category (a)	Common Name	CRB Code (b)	Country of origin (c)	Breeding company or donor	Registration year	Cultivation status (d)	End use (e)	Sowing type (f)	Foliage type (g)	Flower colour (h)	Seed type (i)	Pedigree	Reference
A129	New sources	VK4888 SPARTANETS		RUS	VIR, RU	-	Wa	-	-	NI	Pi	Sm	-	[36]
A130	New sources	E11	CRB365	EGY	-	-	Wa	-	-	NI	Pu	Sm	-	[36]
A131	New sources	L2782.1		USA, NY	Blixt	-	Wa	-	-	NI	-	Sm	-	[36]
A132	New sources	LISA		DEU	Dr Hans Roif Spath	2002	Cv	Fd	Ss	NI	Pu	Sm	-	[36]
A133	New sources	ZP141 (=V281)	CRB326	ESP	Univ. Valladolid	1979	Lv	Gd	Ss	NI	W	Sm	-	[36]
A134	New sources	0-2073	CRB146	USA, WA	USDA, Kraft	-	Gmp	-	Ss	NI	W	Sm	-	[36]
A135	New sources	100	CRB162	-	-	-	-	Ws	-	NI	W	Sm	-	[36]
A136	New sources	E186	CRB374	-	-	-	Wa	-	-	NI	Pu	Sm	-	[36]
A137	New sources	ETHIOPIA-32 CGN 3255 (=PI196026)		NLD	CGN	-	Wa	-	-	NI	Pu	Sm	-	[36]
A138	New sources	FRIGLOO		FRA	-	-	Cv	Gd	Ws	NI	Pu	Sm	-	[36]
A139	New sources	GAT 1259		-	-	-	-	Fd	-	NI	Pu	Sm	-	[36]
A140	New sources	GSP1023	CRB086	USA, WI	Univ. Wis, Gritton	-	Bl	Gd	-	NI	W	Wr	-	[36]
A141	New sources	LOCAL (=bluish-green, small,B-563)	CRB202	RUS	VIR	-	Wa	-	-	NI	Pu	Sm	-	[36]
A143	New sources	NEPAL A		NPL	-	-	Cv	-	-	NI	Pu	Sm	-	[36]
A144	New sources	NO. 9845 (=PI 196032)		ETH	-	-	Cv	-	-	NI	Pu	Sm	-	[36]
A145	New sources	W 18903		USA, WI	Univ. Wis, Gritton	1990	Gmp	Gd	Ss	NI	W	Wr	-	[31,36]
A146	New sources	W 18905		USA, WI	Univ. Wis, Gritton	1990	Gmp	Gd	Ss	NI	W	Wr	-	[31,36]
A005	Varieties	Odeon		DNK	Danisco Seed	2007	Cv	Dry	Ss	af	W	Sm	-	-
A118	Varieties	Enduro		FRA	F.Desprez	2007	Cv	Dry	Ws	af	W	Sm	-	-
A119	Varieties	Lumina	CRB010	NLD	Cebeco Zaden BV	2001	Cv	Dry	Ss	af	W	Sm	-	-
A121	Varieties	Astirid		DNK	Danisco Seed	2007	Cv	Dry	Ws	af	W	Sm	-	-
A122	Varieties	Pawnee		FRA	Sersem	2006	Cv	Dry	Ws	af	W	Sm	-	-
A123	Varieties	ISARD		FRA	A. Obention	2005	Cv	Dry	Ws	af	W	Sm	-	-
A124	Varieties	EDEEN		DNK	Danisco Seed	2004	Cv	Dry	Ss	af	W	Sm	-	-
A147	Varieties	886/01		FRA	INRA	-	Bl	Dry	Ws	af	W	Sm	-	-
A148	Varieties	DOVE HR		FRA	INRA	-	Bl	Dry	Ws	-	-	Sm	-	-
A149	Varieties	Alxor		FRA	Momont	2007	Cv	Dry	Ss	af	W	Sm	-	-
A150	Varieties	Amical		FRA	Laboulet	2002	Cv	Dry	Ss	af	W	Sm	-	-
A151	Varieties	Antares		FRA	Sersem	2007	Cv	Dry	Ss	af	W	Sm	-	-
A152	Varieties	Beilde		FRA	F.Desprez	2006	Cv	Dry	Ss	af	W	Sm	-	-
A153	Varieties	Margnan		FRA	Unisigma	2007	Cv	Dry	Ss	af	W	Sm	-	-
A154	Varieties	Safranos		NLD	Cebeco Zaden BV	2007	Cv	Dry	Ss	af	W	Sm	-	-
A155	Varieties	Hardy		FRA	Sersem	2001	Cv	Dry	Ss	af	W	Sm	-	-

(iv) New sources of partial resistance from INI screening program

(v) European varieties

Additional File II.2

**Simple sequence repeat (SSR) markers used to genotype the
pea-Aphanomyces collection**

(a) main Aphanomyces resistance quantitative trait loci (QTL) defined in [43]

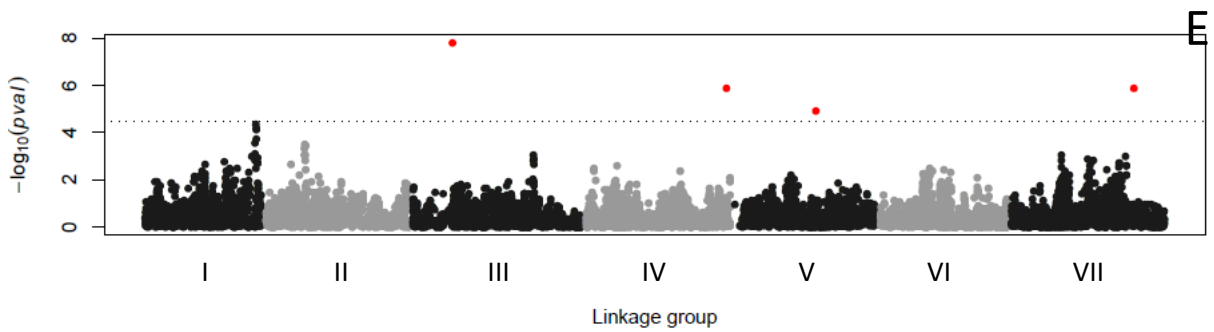
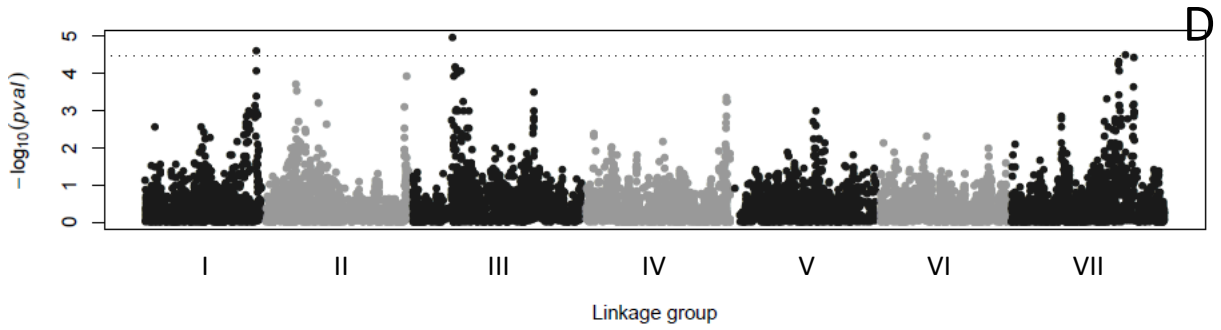
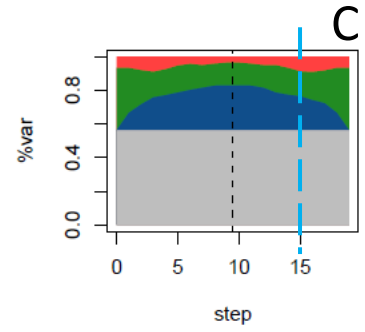
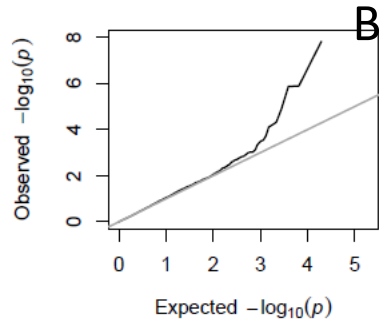
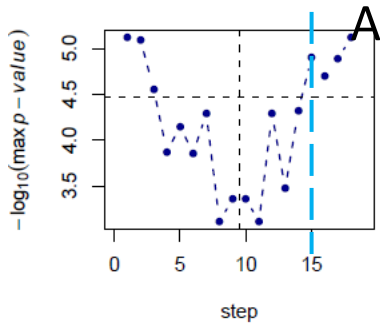
LG	Marker	Position on the consensus map of Hamon <i>et al.</i> , [43] (cM)	Marker associated with Aphanomyces resistance, earliness or height QTL detected by linkage mapping (a)	Number of alleles in the Pea-Aphanomyces collection
I	AA121	-	<i>Ae-Ps1.3</i>	10
	AD147	53.09		14
	AB56	103.71	<i>Ae-Ps1.2</i>	16
	D21	107.3	<i>Ae-Ps1.2</i>	15
	AF016458	124.82	<i>MQTL-Ae3</i> <i>Flo-Ps1.2</i> <i>Ae-Ps1.2</i> <i>MQTL-Ae3</i>	7
II	AD148	45.76		15
	AA372.1	70.36	<i>Ae-Ps2.2</i> <i>MQTL-Ae5/Ae6</i>	14
	AB112	78.17	<i>Ae-Ps2.2</i> <i>MQTL-Ae5/Ae6</i> <i>Flo-Ps2.2</i>	12
	AD83	82.04	<i>MQTL-Ae5/Ae6</i>	4
	AB33	82.24	<i>MQTL-Ae5/Ae6</i>	13
	AB149	169.78		10
	AB40	182.7	<i>Flo-Ps2.3</i>	7
	III	AA175	35.2	<i>MQTL-Ae8/Ae9</i>
AB28		44.69	<i>Flo-Ps3.1</i>	7
AB92		-	<i>MQTL-Ae8/Ae9</i>	10
AD57		66.6	<i>Ae-Ps3.1</i> <i>MQTL-Ae8/Ae9</i>	6
A6		133.39		4
AA278		133.81		5
AA5		162.02	<i>Ae-Ps3.2</i>	8
AD270		-		14
IV	AA174	28.35	<i>Ae-Ps4.1</i> <i>MQTL-Ae12</i>	13
	AD186	42.14	<i>Ae-Ps4.1</i> <i>MQTL-Ae12</i>	15
	AB31	65.78		15
	AA122	153.02	<i>Ae-Ps4.4</i> <i>Flo-Ps4.1</i>	14
	AC22	171.05	<i>MQTL-Ae15</i>	7
V	AA81	0	<i>Ae-Ps5.1</i> <i>MQTL-Ae16/Ae17</i>	7
	AD68	12.81	<i>MQTL-Ae16/Ae17</i>	4
	AB47	24.75	<i>MQTL-Ae16/Ae17</i>	7
	SugTrans	45.82		2
	AA475	59.21		11
	AA460	154.57	<i>Flo-Ps5.1</i>	6
VI	AC74	31		7
	AD51	32.42	<i>Ae-Ps6.1</i>	7
	AD59	92.23		6
	AA335	-		7
VII	AA90	50.9	<i>Ae-Ps7.2</i>	10
	AD56	96.44		11
	AA317	150.44	<i>Ae-Ps7.5</i> <i>Ae-Ps7.6</i>	12
	AA505	149.45		9
	AB27	172.72	<i>Ae-Ps7.6</i> <i>MQTL-Ae25/Ae26</i>	9
	AB136	182.51	<i>MQTL-Ae25/Ae26</i>	9
	AB122	184.08	<i>Ae-Ps7.6</i> <i>MQTL-Ae25/Ae26</i>	9
	AD53	187.52	<i>Ae-Ps7.6</i>	13
	AA387	188.88	<i>Ae-Ps7.6</i> <i>MQTL-Ae25/Ae26</i>	10
	AB101	202.06	<i>Ae-Ps7.6</i> <i>MQTL-Ae25/Ae26</i> <i>MQTL-Morpho1</i>	11
	Clpser	-		2

Additional File II.3

Example of MLM R package plot outputs for analysis of the MFA.Dim.1

variable

(A): Bonferroni correction of the highest marker p -value for each step of the forward and backward analysis. Optimal step (dashed blue vertical line) is determined as the largest stepwise mixed model regression in which all cofactors have $-\log_{10}(p\text{-value})$ above the mBonf threshold (dashed black horizontal line). (B): QQ-plot at the optimal step. (C): Partition of variance for each forward and backward step. Variance explained by population structure (PCA; grey); all markers used as cofactors (blue); Kinship matrix (green) and unexplained variance (= missing heritability; red). The dashed blue vertical line represents the optimal step according to mBonf threshold. (D) Genome-wide Manhattan plot without marker cofactors. The dotted line represents the multiple-Bonferroni threshold above which markers are considered as significant. (E): Genome-wide Manhattan plot at the optimal step with marker cofactors (4 red dots). The dotted line represents the multiple-Bonferroni threshold above which markers are considered as significant.



Additional File II.4

Statistical analysis results of Aphanomyces resistance, earliness and height scoring data in the pea-Aphanomyces collection

(a) CC: controlled condition experiment, Field: Infested field experiment, RRI: Root rot index, ADI1: First aerial decline index, ADI2: Second aerial decline index, FLO1: date to 50 % bloom, FLO2: date to 100 % bloom, RIPE: date to 100 % dried plants, HT: height of plants, 09 to 13: year of experiment, All: global variables over all field environments, DI: Dijon-Epoisses, Burgundy, France, RI: Riec-sur-Belon, Brittany, France, REN: Rennes-Le Rheu, Brittany, France, KEN: Kendrick, ID, USA, '\$' Residual normality statistics are not exact but graphs of residuals are correct, (b) Mean, minimum, maximum and standard error in the pea-Aphanomyces collection; (c) ANOVA significance *p-value* codes: $0 < \text{'***'} < 0.001 \leq \text{'**'} < 0.01 \leq \text{'*'} < 0.05 \leq \text{'ns'} < 1$; (d): Genotype effect; (e): Replicate effect; (f): Environment effect; (g): Genotype x Environment interaction effect; (h) mean-based heritability

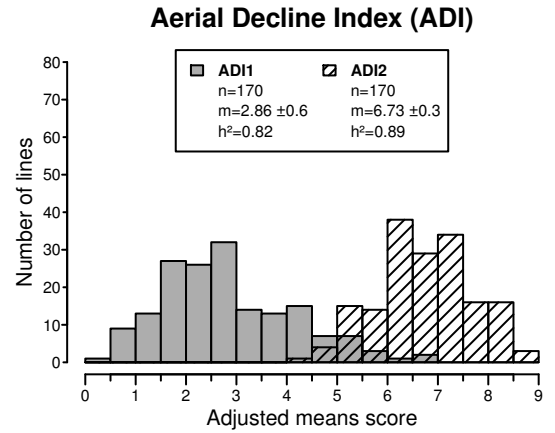
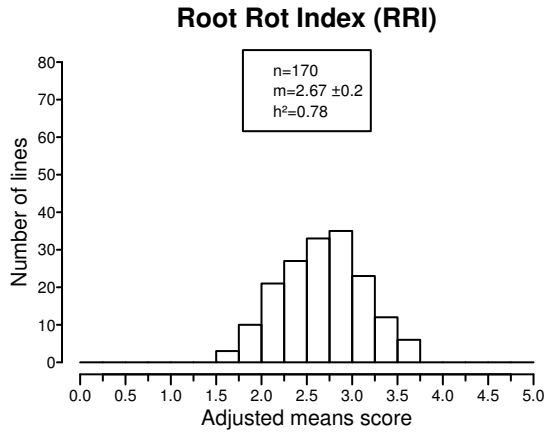
Variable (a)	Number of lines	Number of observations	Mean (b)	Min (b)	Max (b)	SE (b)	Adjusted r^2 model (c,d)	G effect (c,e)	E effect (c,f)	GxE effect (e,g)	h^2 (h)
CC_RRI_09_Ae109 \$	158	631	2.8	0.33	3.83	0.2	0.73 ***	***			0.91
CC_RRI_09_RB84	158	623	3.38	2	4.59	0.2	0.64 ***	***			0.86
CC_RRI_13_RB84	175	520	3.19	1.84	4.35	0.3	0.6 ***	***			0.78
Field_RRI_All \$	175	4115	2.7	1.79	3.66	0.3	0.65 ***	***	***	***	0.92
Field_RRI_DI_10	170	492	2.67	1.54	3.7	0.2	0.6 ***	***			0.78
Field_RRI_DI_11	175	523	2.87	1.83	4.04	0.2	0.71 ***	***			0.88
Field_RRI_DI_12	175	525	2.56	1.63	3.93	0.2	0.56 ***	***			0.78
Field_RRI_DI_13	175	525	1.95	1.08	3.59	0.3	0.36 ***	***			0.61
Field_RRI_RI_10	172	478	2.94	1.55	4.11	0.3	0.45 ***	***			0.65
Field_RRI_RI_11	175	522	3.18	0.95	4.8	0.3	0.52 ***	***			0.73
Field_RRI_RI_12	175	511	3.51	2.04	4.55	0.2	0.57 ***	***			0.78
Field_RRI_RI_13	175	525	1.93	0.77	3.15	0.4	0.12 **	ns			0.28
Field_ADI1_All \$	155	4640	3	1.78	4.63	0.4	0.77 ***	***	***	***	0.81
Field_ADI1_DI_10	170	491	2.86	0.33	6.67	0.6	0.61 ***	ns			0.82
Field_ADI1_DI_11	175	519	4.7	1.72	8.09	0.4	0.79 ***	ns			0.92
Field_ADI1_DI_12 \$	175	522	1.94	0.4	4.1	0.3	0.68 ***	*			0.86
Field_ADI1_DI_13	175	525	2.76	0.74	5.07	0.4	0.61 ***	***			0.82
Field_ADI1_RI_10	173	461	3.47	1.81	5.67	0.6	0.23 ***	***			0.35
Field_ADI1_RI_11	174	480	3.09	1.67	4.58	0.4	0.34 ***	**			0.57
Field_ADI1_RI_12	174	514	3.99	0	6.31	0.4	0.71 ***	*			0.88
Field_ADI1_RI_13	175	525	1.66	0.48	3.34	0.4	0.3 ***	***			0.44
Field_ADI1_KEN_12	174	519	1.81	0.81	2.98	0.3	0.25 ***	***			0.47
Field_ADI2_All \$	175	4117	4.68	3.19	6.27	0.5	0.8 ***	***	***	***	0.84
Field_ADI2_DI_10	170	489	6.73	4.47	8.69	0.3	0.73 ***	*			0.89
Field_ADI2_DI_12	175	517	4.03	0.56	7.19	0.3	0.89 ***	ns			0.96
Field_ADI2_DI_13	175	513	5.05	1.9	7.23	0.5	0.65 ***	***			0.83
Field_ADI2_RI_10	162	351	4.68	3.57	5.52	0.4	0.18 *	ns			0.29
Field_ADI2_RI_11	172	452	4.51	2.63	6.1	0.5	0.34 ***	***			0.55
Field_ADI2_RI_12	175	523	6.07	4.01	8.34	0.4	0.59 ***	***			0.81
Field_ADI2_RI_13	175	523	3.8	1.9	5.9	0.4	0.45 ***	***			0.7
Field_ADI2_KEN_12	174	522	2.5	1.11	3.67	0.3	0.3 ***	***			0.55
FLO1_REN_All \$	175	1020	144.87	129.6	165.5	0.6	0.93 ***	***	ns	***	0.9
FLO1_REN_10 \$	175	489	144.85	131.19	167.33	1	0.9 ***	***			0.96
FLO1_REN_12 \$	175	525	144.87	127	163.67	0.4	0.99 ***	*			1
FLO2_REN_All \$	175	1023	172.34	162.17	186.83	0.6	0.94 ***	***	***	***	0.89
FLO2_REN_10 \$	175	492	169.68	160	182.67	1	0.86 ***	**			0.95
FLO2_REN_12 \$	175	525	175.1	160	191	0.6	0.97 ***	ns			0.99
RIPE_REN_All \$	175	1008	201.22	190.81	208.43	0.7	0.94 ***	**	***	***	0.87
RIPE_REN_10 \$	175	498	194.92	185.4	205.33	1.2	0.71 ***	ns			0.87
RIPE_REN_12 \$	173	512	207.52	191.94	212.19	0.7	0.83 ***	***			0.93
HT_REN_All \$	175	1015	98.49	62.08	178.33	2.2	0.94 ***	***	***	***	0.95
HT_REN_10 \$	175	496	94.18	60	165	3.3	0.92 ***	***			0.97
HT_REN_12	175	517	102.81	60	191.67	2.8	0.96 ***	***			0.98

Additional File II.5

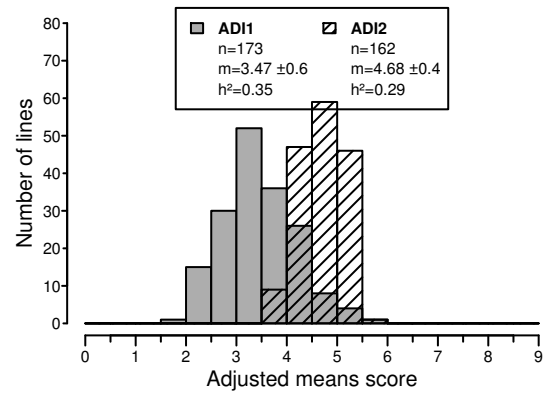
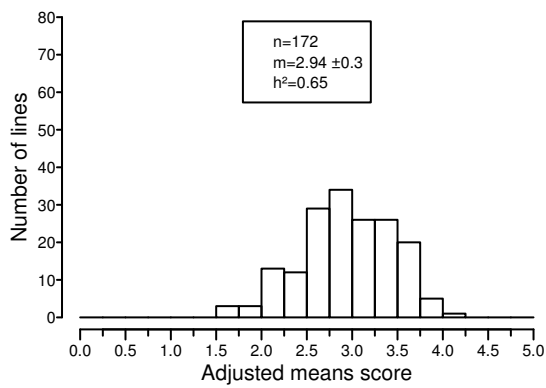
Frequency distribution of least square means obtained for *Aphanomyces* resistance, earliness and plant height in the pea-*Aphanomyces* collection

Least square means were obtained from analysis of variance for three *Aphanomyces* resistance traits (Root Rot Index, first and second Aerial Decline Indexes, coded RRI, ADI1 and ADI2, respectively), three earliness traits (dates to 50 % bloom, 100 % bloom and 100 % dried plants, coded FLO1, FLO2 and RIPE, respectively) and plant height (HT). *Aphanomyces* resistance traits were assessed over nine infested field environments and against two reference strains of *A. euteiches* (RB84 and Ae109) in controlled conditions. Earliness and height were assessed in two healthy environments. n: total number of pea lines assessed; m: mean \pm standard deviation of the pea-*Aphanomyces* collection; h²: mean-based heritability.

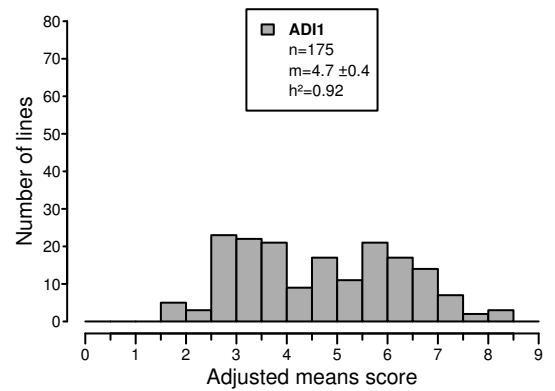
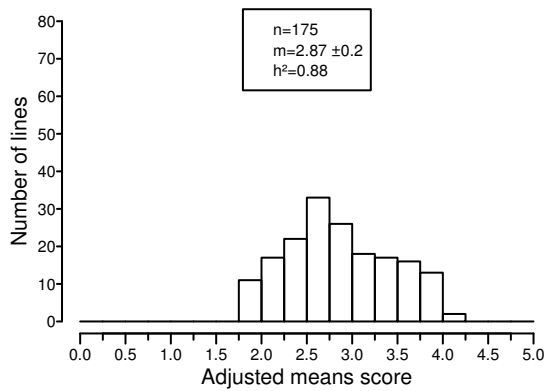
2010
Dijon
(FR)



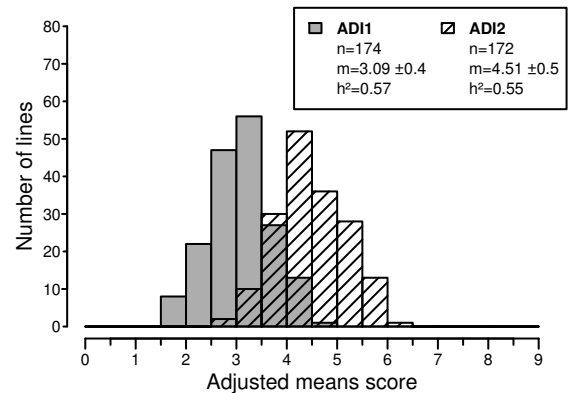
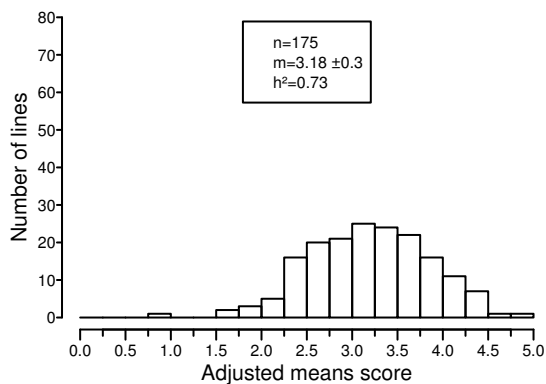
2010
Riec
(FR)



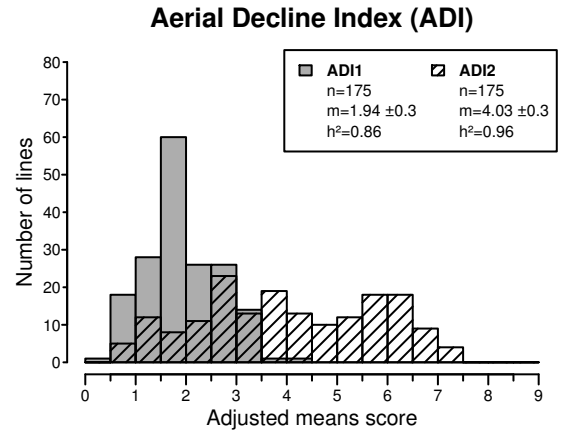
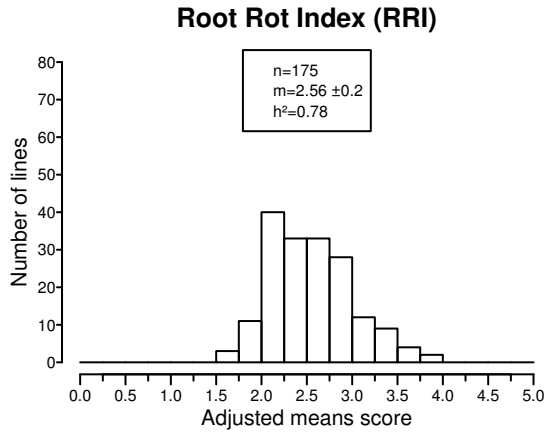
2011
Dijon
(FR)



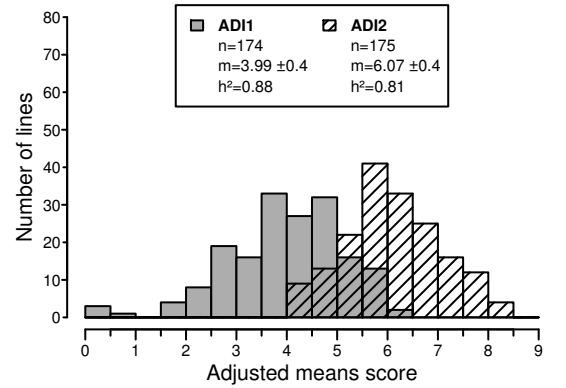
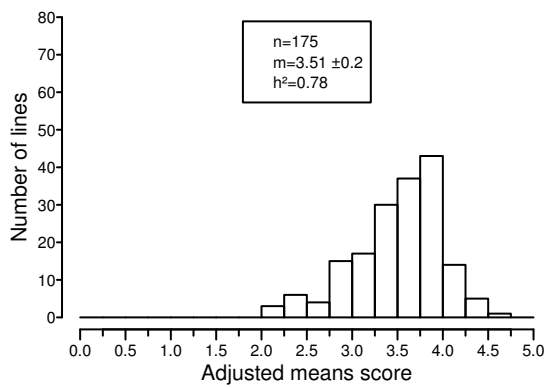
2011
Riec
(FR)



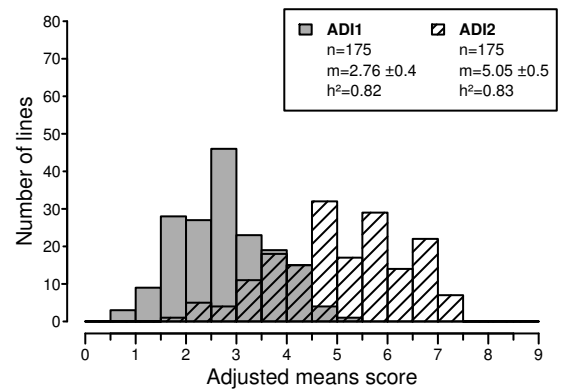
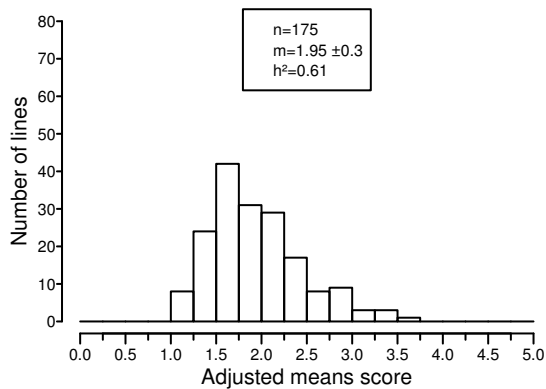
**2012
Dijon
(FR)**



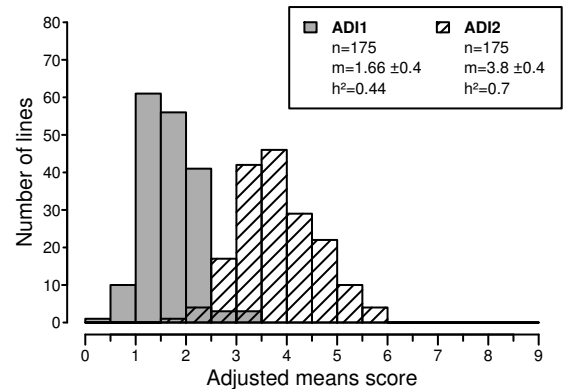
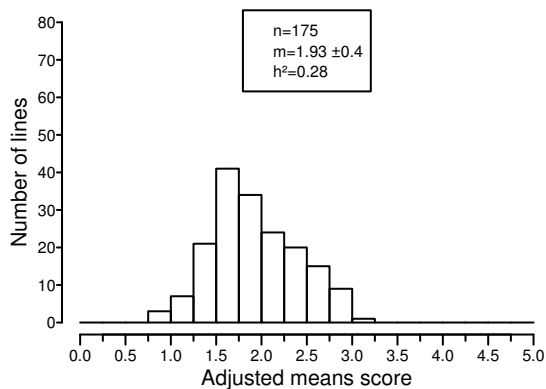
**2012
Riec
(FR)**



**2013
Dijon
(FR)**



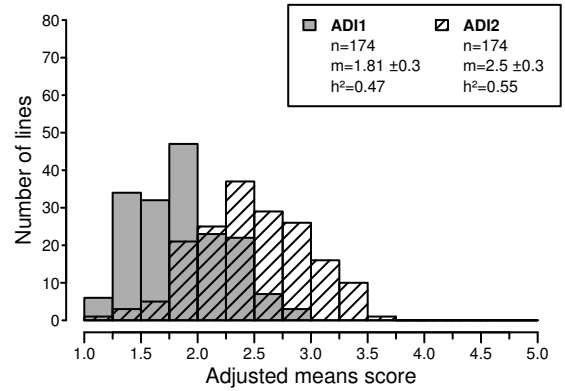
**2013
Riec
(FR)**



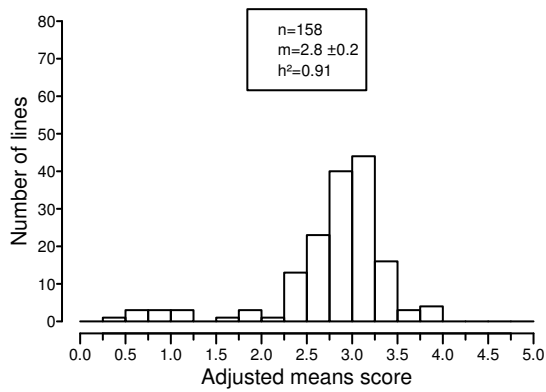
Root Rot Index (RRI)

2012
Kendrick
(USA)

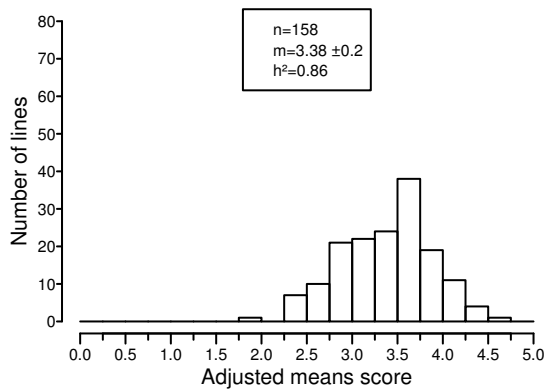
Aerial Decline Index (ADI)



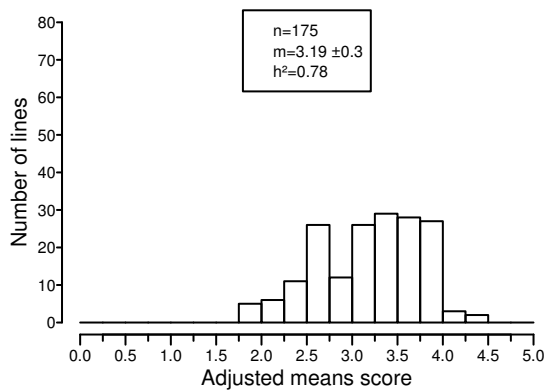
2009
Ae109
(CC)

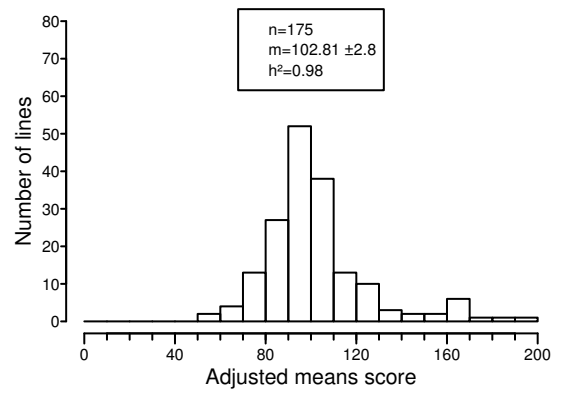
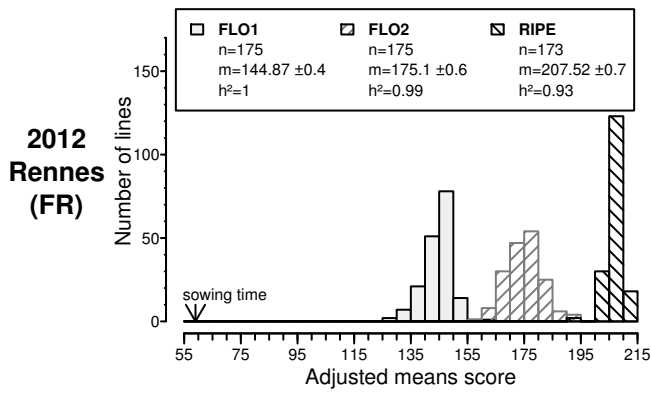
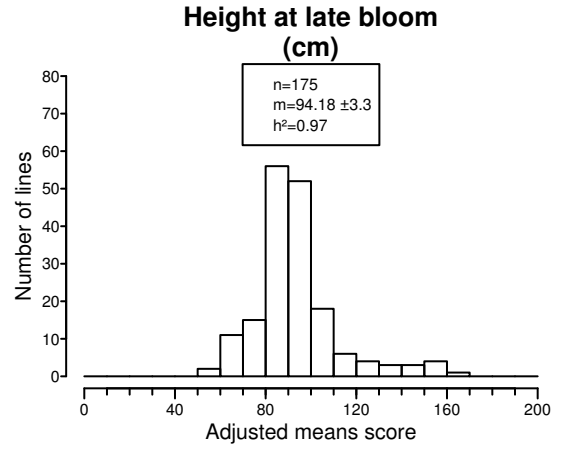
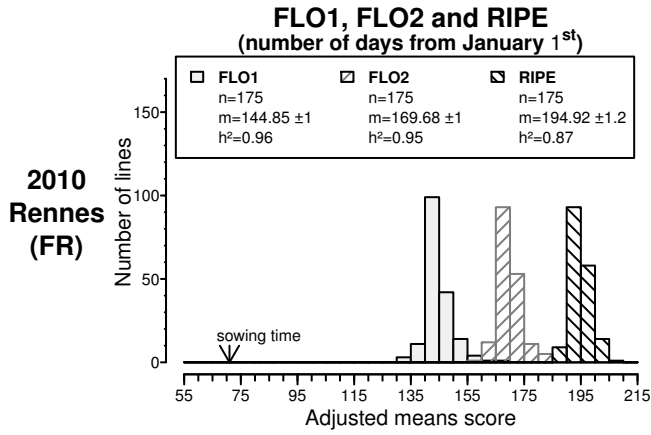


2009
RB84
(CC)



2013
RB84
(CC)

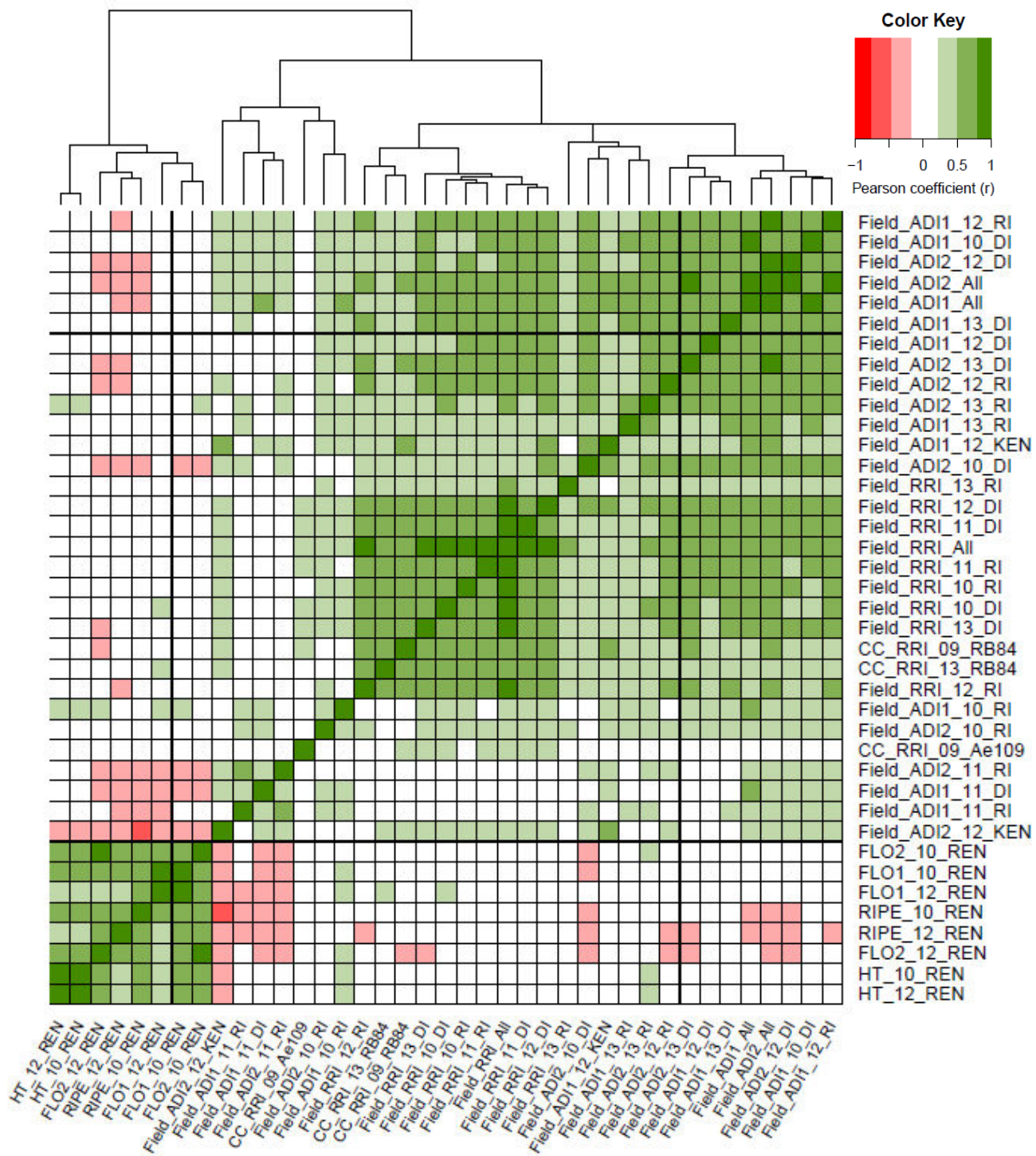




Additional File II.6

Clustered heatmap of correlation coefficients between disease resistance, earliness and height variables data

Variables are coded as described in Additional File 4. Only correlations with a *p-value* < 0.05 are shown. Colours represent level of correlation, coded as follow: dark red: $r < -0.8$; medium red: $-0.8 < r < -0.5$; light red: $-0.5 < r < -0.2$; white: $-0.2 < r < 0.2$ or *p-value* > 0.05; light green: $0.2 < r < 0.5$; medium green: $0.5 < r < 0.8$; dark green: $r > 0.8$. Clustering is based on the UPGMA method.



Additional File II.7

Variable contribution in the two first Multiple Factor Analysis (MFA)

principal components

(a) Variables are coded as described in Additional File 4, (b) percentage of variable contribution in the construction of the MFA axis, ‘*’: the six highest contributions; (c) \cos^2 indicates the representation quality of the variable on the MFA axis, ‘*’: $0.2 < \cos^2 < 0.5$, ‘**’: $\cos^2 > 0.5$; (d) correlation between variable and MFA axis, ‘*’: $0.5 < r^2 < 0.8$, ‘**’: $r^2 > 0.8$

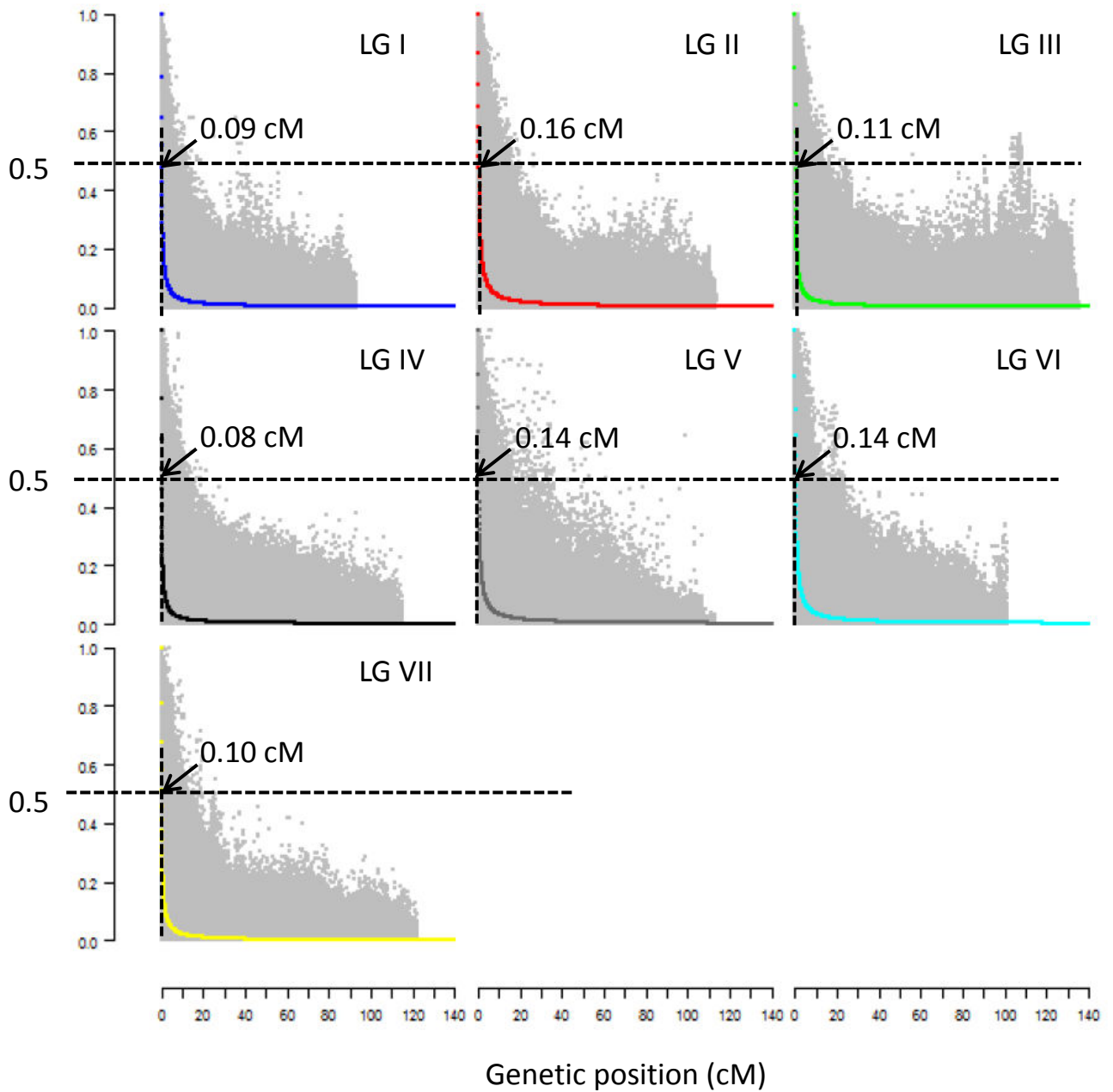
Dim1 (44.13%)				Dim2 (11.89%)			
Variables (a)	Contribution (b)	Cos ² (c)	Correlation (d)	Variables	Contribution (b)	Cos ² (c)	Correlation (d)
Field_AD11_11_DI	0.37	0.058	0.24	Field_AD12_12_RI	0.01	0.000	0.02
Field_AD11_11_RI	0.48	0.075	0.27	Field_RRI_13_RI	0.03	0.002	-0.04
Field_AD12_11_RI	0.76	0.108	0.33	Field_AD11_12_KEN	0.17	0.007	0.08
Field_AD12_10_RI	1.03	0.161	0.40	Field_AD12_13_DI	0.18	0.007	0.08
Field_AD11_10_RI	1.19	0.186	0.43	Field_RRI_13_DI	0.42	0.021	-0.14
Field_AD12_12_KEN	1.38	0.196	0.44	Field_AD12_12_KEN	0.45	0.017	0.13
Field_RRI_13_RI	1.46	0.268 *	0.52 *	Field_RRI_12_DI	0.58	0.029	-0.17
CC_RRI_09_Ae109	1.68	0.103	0.32	Field_RRI_11_DI	0.64	0.032	-0.18
Field_AD11_13_RI	1.98	0.307 *	0.55 *	Field_RRI_11_RI	0.74	0.037	-0.19
Field_RRI_12_RI	2.40	0.443 *	0.67 *	Field_AD12_10_RI	0.76	0.032	0.18
Field_RRI_10_RI	2.55	0.472 *	0.69 *	Field_RRI_10_RI	0.81	0.040	-0.20
Field_AD11_12_KEN	2.63	0.408 *	0.64 *	Field_AD11_13_DI	0.90	0.038	0.19
Field_RRI_13_DI	2.78	0.511 **	0.72 *	Field_RRI_All	0.95	0.047	-0.22
Field_AD12_10_DI	2.95	0.418 *	0.65 *	Field_RRI_10_DI	1.12	0.056	-0.24
Field_RRI_10_DI	3.15	0.585 **	0.76 *	Field_AD11_12_DI	1.39	0.058	0.24
Field_RRI_11_RI	3.45	0.634 **	0.80 **	Field_RRI_12_RI	1.62	0.080	-0.28
Field_RRI_12_DI	3.46	0.638 **	0.80 **	Field_AD12_13_RI	1.82	0.070	0.26
Field_RRI_11_DI	3.57	0.656 **	0.81 **	Field_AD11_12_RI	1.90	0.080	0.28
Field_AD11_12_RI	3.59	0.560 **	0.75 *	Field_AD11_10_RI	1.94	0.082	0.29
Field_AD11_13_DI	3.63	0.563 **	0.75 *	Field_AD12_12_DI	2.00	0.077	0.28
Field_AD11_12_DI	3.67	0.570 **	0.75 *	Field_AD11_13_RI	2.15	0.090	0.30
Field_AD12_13_RI	3.87	0.551 **	0.74 *	Field_AD11_10_DI	3.24	0.133	0.36
Field_AD12_12_RI	3.90	0.555 **	0.74 *	Field_AD12_All	3.24	0.124	0.35
Field_AD11_10_DI	3.92	0.598 **	0.77 *	Field_AD11_11_DI	3.91	0.163	0.40
Field_AD12_12_DI	4.00	0.569 **	0.75 *	Field_AD12_10_DI	4.34	0.165	0.41
Field_RRI_All	4.40 *	0.809 **	0.90 **	Field_AD11_All	5.42 *	0.226 *	0.48
Field_AD11_All	4.40 *	0.683 **	0.83 **	Field_AD11_11_RI	7.71 *	0.324 *	0.57 *
Field_AD12_13_DI	4.64 *	0.661 **	0.81 **	Field_AD12_11_RI	7.76 *	0.298 *	0.55 *
Field_AD12_All	5.67 *	0.807 **	0.90 **	CC_RRI_09_Ae109	10.94 *	0.180	-0.42
CC_RRI_13_RB84	7.60 *	0.426 *	0.65 *	CC_RRI_09_RB84	15.62 *	0.255 *	-0.50 *
CC_RRI_09_RB84	9.44 *	0.572 **	0.76 *	CC_RRI_13_RB84	17.25 *	0.261 *	-0.51 *

Additional File II.8

Linkage disequilibrium (LD) decay in the pea-Aphanomyces collection

Coloured curves represent the estimated LD decay for each linkage group (LG). Dashed horizontal lines represent half of the maximum LD value (LD threshold=0.5). Arrows represent the LD decay rate, as the estimated genetic distance (cM) to drop to the LD threshold on each LG.

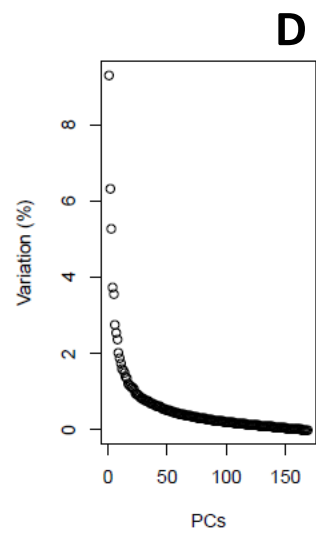
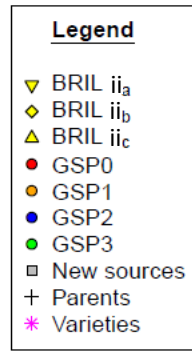
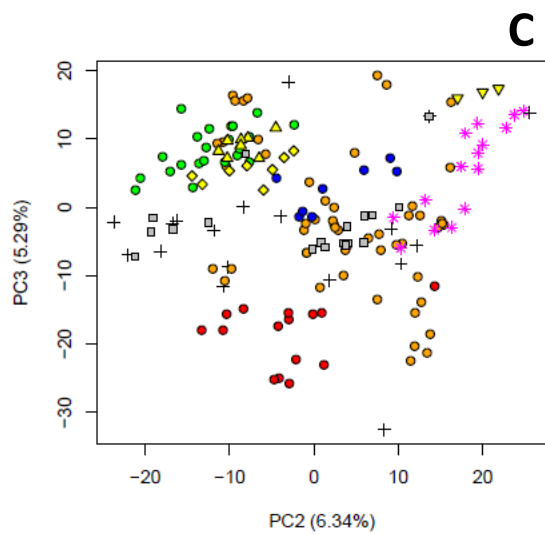
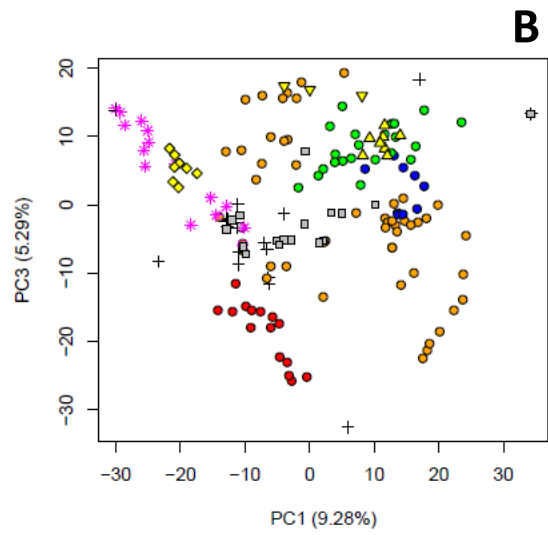
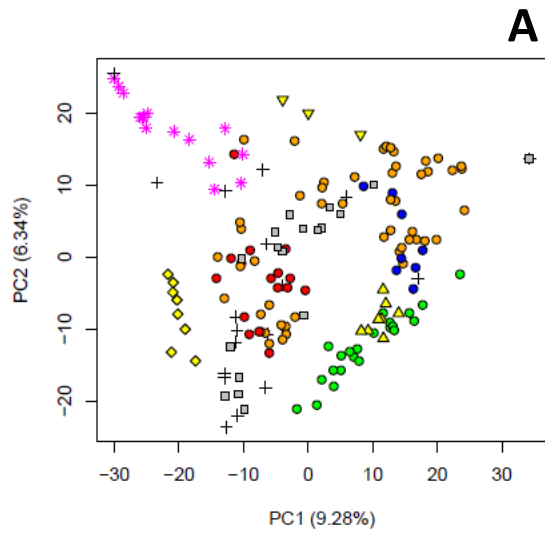
Linkage disequilibrium
 r^2



Additional File II.9

Population structure of the pea-Aphanomyces collection based on Principal Component Analysis (PCA)

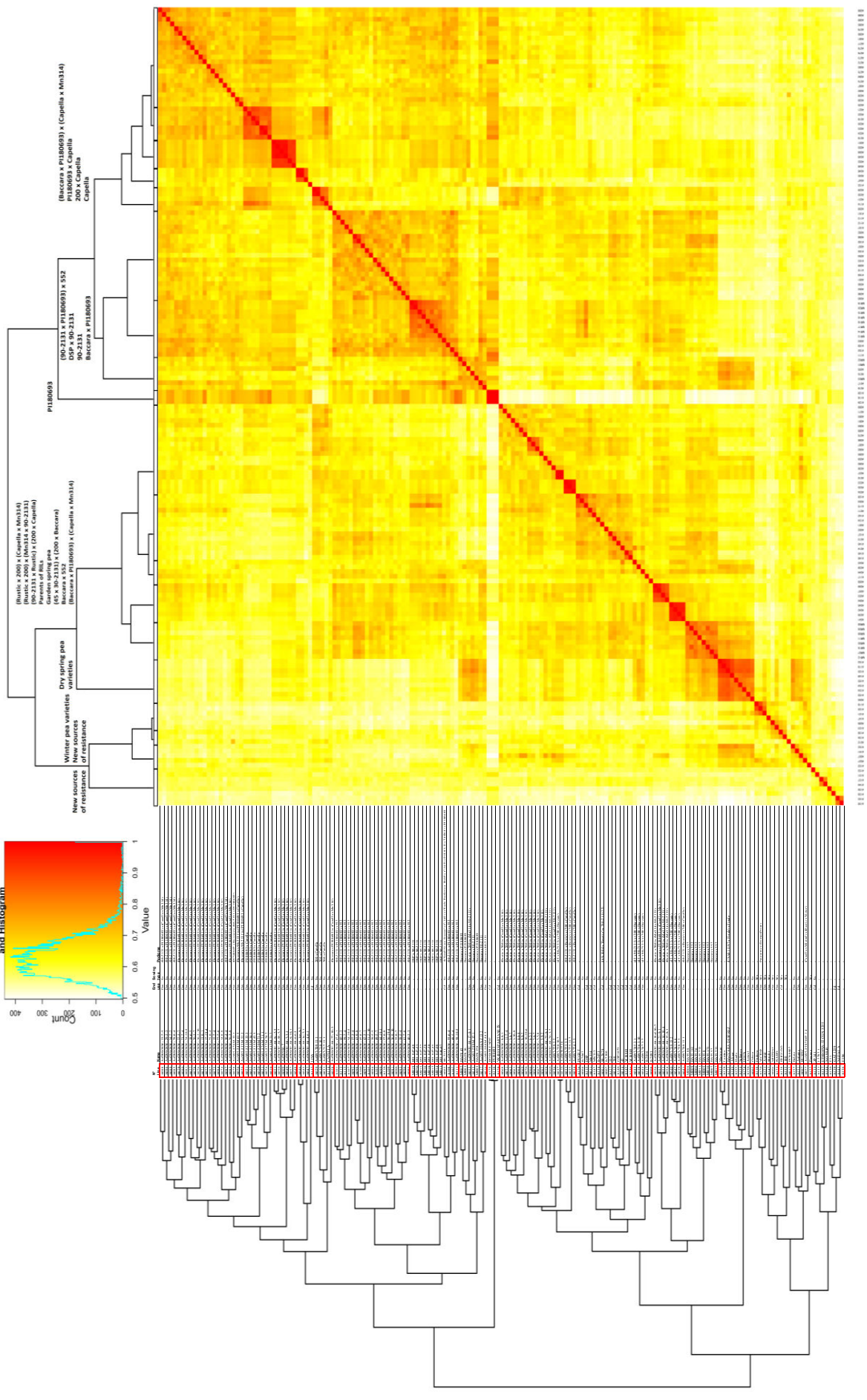
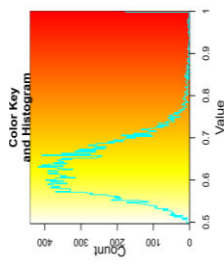
PCA from GAPIT R package [75] based on 2,937 markers. Distributions of pea lines of the collection are represented on the first three principal components (A-C), which explain a total of 20.91 % of inertia (D). Categories of pea lines are described in Figure 1 and Additional File 1. (D) Inertia contribution of each principal component (from PC1 to PC169).



Additional File II.10

Clustered heatmap of the Kinship matrix

Kinship matrix from the GAPIT R package [75] based on 2,937 SNP markers. Clustering is based on the UPGMA method. Colours represent the degree of relationship between two given lines. Pea line information is described as in Additional File 1. Lines sharing same pedigree or end use or sowing type are well clustered thus the Kinship matrix efficiently represents the relationships between individuals.



Additional File II.11

Genetic loci associated with *Aphanomyces* resistance, earliness and plant height detected by genome-wide association mapping using the MLM method

(a) Genetic positions are estimated on the genetic map resulting from the projection of the marker map from Hamon et al. [43] onto the consensus map from Tayeh et al. [49]; (b) A linkage disequilibrium (LD) block is designed with all markers in LD ($r^2 > 0.2$) with the associated marker. LD blocks are named in consecutive numerical order following their linkage group (LG) name and their genetic position; (c) *Aphanomyces* resistance and earliness Meta-QTL (quantitative trait loci) described in [43]; (d) Initial QTL for *Aphanomyces* resistance and earliness before meta-analysis described in [41, 43]; (e) Trait-associated marker(s) and their genetic position(s) in the LD blocks. *PsCam*- markers are SNPs developed by Tayeh et al [49], *Af* is a morphological gene involved in leaves shape, other markers are SSRs (detected allele in brackets); (f) Minor allele frequency (MAF) of each detected marker in the collection; (g) Variables are coded as described in Additional File 4; (h) significant *p-value* of the marker-trait association $< 2.5E-5$; (i) mean value of the trait corrected with the effect of all markers included in the MLM analysis; (j) allelic effect of the marker on the variable, the sign depends on the allele code; (k) standard error at the marker around the mean value of the trait in the collection.

LD block name	LD block interval (cM)	Aphanomyces resistance or earliness Meta-QTL co-localizing with LD block (c)	Aphanomyces resistance, earliness or height QTL detected by linkage mapping co-localizing with LD block (d)	Trait-associated marker (e)	Genetic position (cM) (e)	
LG	(a,b)	(a)	(c)	(d)	(e)	
I	I.1	38.5 - 39	<i>MQTL-Morpho1</i>	<i>Ae-Ps1.3, Flo-Ps1.1</i>	<i>PsCam047171_30305_572</i>	38.7
	I.2	73.85 - 74.3	<i>MQTL-Ae2</i>	<i>Ae-Ps1.2, Flo-Ps1.2</i>	<i>AB56 (379)</i>	73.85
	I.3	82.4 - 84.3	<i>MQTL-Morpho2</i>	<i>Ae-Ps1.2, Flo-Ps1.2</i>	<i>PsCam045028_28759_2059</i>	83.1
	I.4	85.6 - 87.2	<i>MQTL-Ae3, MQTL-Morpho2</i>	<i>Ae-Ps1.2, Flo-Ps1.2</i>	<i>Af</i>	85.6
	I.5	88.2 - 90.4	<i>MQTL-Ae3, MQTL-Morpho2</i>	<i>Ae-Ps1.2, Flo-Ps1.2</i>	<i>PsCam027153_15841_364</i> <i>PsCam044835_28587_1878</i>	88.2 89.1
II	II.1	41 - 42.7		<i>Ae-Ps2.2</i>	<i>PsCam033694_19257_1741</i>	42.1
					<i>PsCam011257_7650_1194</i>	42.7
	II.2	45.2 - 45.5	<i>MQTL-Ae6</i>	<i>Ae-Ps2.2, Flo-Ps2.2</i>	<i>PsCam059543_39701_683</i>	45.2
	II.3	45.5 - 46.3	<i>MQTL-Ae6</i>	<i>Ae-Ps2.2, Flo-Ps2.2</i>	<i>AB33 (388)</i>	45.9
					<i>PsCam035741_20914_1106</i>	46.3
	II.4	46.9 - 47.7	<i>MQTL-Ae6</i>	<i>Ae-Ps2.2</i>	<i>PsCam042382_26450_3953</i>	47.4
	II.5	67.6 - 68.1			<i>PsCam043114_27143_892</i>	68.1
II.6	101.8 - 103.7		<i>Flo-Ps2.3</i>	<i>PsCam009577_6385_812</i>	103.3	
II.7	112.75 - 112.9			<i>PsCam010896_7359_1404</i>	112.75	
III	III.1	8.5 - 10.2			<i>PsCam037306_22398_891</i>	9.6
	III.2	22.4 - 27.8	<i>MQTL-Ae8, MQTL-Morpho4</i>	<i>Ae-Ps3.1a, Flo-Ps3.1</i>	<i>PsCam017207_10682_389</i>	22.4
					<i>PsCam049584_32211_1847</i>	22.4
					<i>PsCam043138_27166_413</i>	27.3
	III.3	31 - 32.6	<i>MQTL-Ae9</i>	<i>Ae-Ps3.1a</i>	<i>PsCam035590_20765_380</i>	31.6
	III.4	32.6 - 32.7	<i>MQTL-Ae9</i>	<i>Ae-Ps3.1a</i>	<i>PsCam023109_13003_2421</i>	32.7
	III.5	32.8 - 33.9	<i>MQTL-Ae9</i>	<i>Ae-Ps3.1a</i>	<i>PsCam011358_7733_760</i>	33.8
	III.6	37.3 - 38.4		<i>Ae-Ps3.1a</i>	<i>PsCam034645_20018_1507</i>	38.3
	III.7	38.7 - 39.2		<i>Ae-Ps3.1a</i>	<i>PsCam056354_37186_567</i>	39.2
	III.8	43.5 - 43.7		<i>Ae-Ps3.1a</i>	<i>AB92 (370)</i>	43.5
	III.9	43.8 - 46.15		<i>Ae-Ps3.1a</i>	<i>PsCam030789_18160_162</i>	44.6
	III.10	62.5 - 64.1	<i>MQTL-Ae10</i>	<i>Ae-Ps3.1b</i>	<i>PsCam045079_28804_1964</i>	63.1
	III.11	64.4 - 67.1	<i>MQTL-Ae10</i>	<i>Ae-Ps3.1b</i>	<i>PsCam012833_8640_2918</i>	65.9
	III.12	70.4 - 72.6	<i>MQTL-Ae10</i>		<i>PsCam012453_8457_2872</i>	72.1
	III.13	76.4 - 78.1		<i>Ae-Ps3.2, HT-Ps3.1</i>	<i>PsCam051345_33851_551</i>	76.8
	III.14	91.1 - 91.7		<i>Ae-Ps3.2</i>	<i>PsCam011571_7892_294</i>	91.5
	III.15	96.5 - 97.4		<i>Ae-Ps3.2</i>	<i>PsCam043293_27312_1261</i> <i>PsCam011467_7813_1433</i>	96.8 97.1
	III.16	100.4 - 103.7		<i>Ae-Ps3.2</i>	<i>PsCam043018_27052_1394</i>	101.2
III.17	130.6 - 133			<i>PsCam038309_23348_747</i> <i>PsCam035617_20792_637</i>	132 132	
III.18	134 - 134.6			<i>PsCam034256_19681_301</i>	134.6	
IV	IV.1	3.7 - 3.8			<i>PsCam000404_355_3214</i>	3.8
	IV.2	5.9 - 7.5		<i>Ae-Ps4.1</i>	<i>PsCam050850_33409_1793</i>	5.9
	IV.3	10.6 - 12.3		<i>Ae-Ps4.1</i>	<i>PsCam036087_21236_2392</i>	11.4
	IV.4	22.1 - 22.7		<i>Ae-Ps4.1</i>	<i>PsCam003906_2976_359</i>	22.1
	IV.5	61.6 - 61.6	<i>MQTL-Ae14</i>	<i>Ae-Ps4.3, Ae-Ps4.4</i>	<i>PsCam039023_24009_852</i>	61.6
	IV.6	67.2 - 67.9		<i>Ae-Ps4.4</i>	<i>PsCam018415_11127_149</i>	67.6
	IV.7	68.7 - 69.3		<i>Ae-Ps4.4</i>	<i>PsCam001188_1003_335</i>	69.3
	IV.8	76.1 - 76.2		<i>Ae-Ps4.5, Flo-Ps4.1</i>	<i>PsCam036906_22019_1424</i>	76.2
	IV.9	105.7 - 106.4		<i>HT-Ps4.1</i>	<i>PsCam038042_23094_246</i>	106.2
	IV.10	109.3 - 110.7		<i>HT-Ps4.1</i>	<i>PsCam000016_12_4286</i>	109.8
	IV.11	112.05 - 112.5		<i>HT-Ps4.1</i>	<i>PsCam030702_18121_172</i> <i>PsCam030723_18124_1544</i>	112.3 112.3
	IV.12	112.5 - 112.8		<i>HT-Ps4.1</i>	<i>PsCam042808_26850_709</i> <i>PsCam043503_27507_1032</i>	112.6 112.6
	IV.13	115.4 - 115.4		<i>HT-Ps4.1</i>	<i>PsCam053795_35577_1738</i>	115.4
V	V.1	64.8 - 65.3		<i>Ae-Ps5.2</i>	<i>PsCam038363_23400_360</i>	65
	V.2	75.8 - 76.1	<i>MQTL-Ae18</i>	<i>Ae-Ps5.2, HT-Ps5.1</i>	<i>PsCam000007_4_886</i>	75.9
	V.3	88 - 88	<i>MQTL-Ae19</i>	<i>Ae-Ps5.3</i>	<i>PsCam042225_26317_1763</i>	88
VI	VI.1	17.4 - 17.5	<i>MQTL-Ae21</i>	<i>Ae-Ps6.1</i>	<i>AC74 (137)</i>	17.5
	VI.2	39.1 - 39.8	<i>MQTL-Ae22</i>	<i>Ae-Ps6.2</i>	<i>PsCam013085_8857_1728</i>	39.3

MAF (%) (f)	Nb associated traits (Aphanomyces resistance + Earliness/Height)		Disease associated traits (g)		Morphological or developmental associated traits (g)		Intercept (i)	Allelic effect (j)	SE (k)
				p-value (h)		p-value (h)			
24	1 + 0	Field_ADI1_10_RI	2.25E-05	-	-	4.11	-0.26	0.06	
8	1 + 0	Field_ADI1_13_RI	3.32E-05	-	-	2.19	0.25	0.06	
37	1 + 0	Field_ADI2_10_DI	9.09E-08	-	-	8.52	-0.35	0.06	
46	1 + 0	Field_RRI_12_DI	4.76E-06	-	-	2.52	0.15	0.03	
43	1 + 0	Field_ADI1_12_DI	2.53E-05	-	-	1.76	-0.21	0.05	
44	5 + 0	Field_ADI1_All	7.86E-11	-	-	3.26	-0.29	0.04	
		Field_ADI1_13_DI	2.53E-10	-	-	3.64	-0.43	0.06	
		Field_ADI1_12_RI	2.61E-06	-	-	4.50	-0.45	0.09	
		Field_ADI2_All	8.27E-06	-	-	4.76	-0.22	0.05	
		MFA.Dim.2	1.01E-05	-	-	0.58	-0.36	0.08	
41	4 + 0	Field_ADI1_12_DI	1.98E-07	-	-	1.76	0.26	0.05	
		Field_ADI2_11_RI	7.77E-07	-	-	4.01	0.33	0.06	
		Field_ADI1_All	2.43E-06	-	-	3.26	0.20	0.04	
		Field_ADI2_All	9.46E-06	-	-	4.76	0.23	0.05	
22	0 + 1	-	-	FLO1_12_REN	3.31E-06	133.38	-1.72	0.36	
8	0 + 1	-	-	FLO1_12_REN	6.29E-16	133.38	12.80	1.42	
8	0 + 1	-	-	FLO1_10_REN	1.95E-17	152.56	-5.86	0.61	
9	0 + 1	-	-	FLO1_12_REN	2.15E-09	133.38	-8.40	1.32	
26	0 + 1	-	-	FLO1_10_REN	6.02E-08	152.56	2.18	0.38	
6	0 + 1	-	-	FLO2_10_REN	3.90E-10	177.17	2.53	0.38	
14	0 + 1	-	-	RIPE_10_REN	7.74E-06	207.37	-1.09	0.24	
17	1 + 0	Field_RRI_13_DI	7.51E-06	-	-	2.01	-0.21	0.04	
13	1 + 0	Field_ADI1_12_DI	2.24E-05	-	-	1.76	-0.32	0.07	
6	0 + 2	-	-	RIPE_10_REN	9.94E-13	207.37	-2.88	0.37	
		-	-	FLO1_12_REN	7.64E-09	133.38	-3.42	0.56	
7	0 + 1	-	-	FLO2_10_REN	8.50E-12	177.17	-2.88	0.39	
7	1 + 3	Field_ADI1_12_DI	1.13E-10	-	-	1.76	0.61	0.09	
		-	-	FLO2_12_REN	1.18E-10	188.78	-5.35	0.78	
		-	-	FLO1_10_REN	1.12E-06	152.56	-2.58	0.51	
		-	-	RIPE_12_REN	4.63E-06	212.13	-2.15	0.45	
48	4 + 0	Field_RRI_All	7.86E-11	-	-	2.38	0.17	0.02	
		MFA.Dim.1	1.51E-08	-	-	-1.78	0.56	0.09	
		Field_RRI_11_RI	9.92E-08	-	-	2.52	0.23	0.04	
		Field_RRI_10_DI	5.78E-06	-	-	2.38	0.17	0.04	
42	3 + 0	Field_RRI_12_RI	1.23E-06	-	-	3.35	0.17	0.03	
		Field_RRI_12_DI	6.19E-06	-	-	2.52	0.14	0.03	
		Field_RRI_10_RI	1.81E-05	-	-	2.48	0.19	0.04	
37	0 + 1	-	-	RIPE_10_REN	1.63E-06	207.37	-1.21	0.24	
42	2 + 0	Field_ADI2_13_DI	3.25E-12	-	-	5.48	-0.45	0.06	
		Field_ADI1_13_DI	1.16E-08	-	-	3.64	-0.36	0.06	
23	1 + 0	Field_ADI1_10_RI	6.79E-06	-	-	4.11	0.27	0.06	
22	1 + 0	Field_ADI2_10_DI	1.94E-07	-	-	8.52	0.36	0.07	
35	2 + 0	Field_ADI1_All	4.91E-07	-	-	3.26	-0.22	0.04	
		Field_ADI1_13_RI	2.62E-06	-	-	2.19	-0.14	0.03	
20	1 + 0	Field_ADI1_12_RI	8.98E-06	-	-	4.50	-0.54	0.12	
21	0 + 1	-	-	HT_10_REN	8.33E-06	95.79	-6.48	1.41	
16	1 + 0	Field_RRI_11_RI	6.58E-06	-	-	2.52	0.26	0.06	
14	0 + 1	-	-	RIPE_10_REN	1.52E-06	207.37	-1.27	0.25	
23	1 + 0	Field_ADI1_13_RI	6.83E-06	-	-	2.19	0.16	0.04	
33	1 + 0	Field_ADI1_10_RI	1.24E-06	-	-	4.11	-0.28	0.06	
40	1 + 0	Field_ADI2_13_DI	4.47E-06	-	-	5.48	0.27	0.06	
43	0 + 1	-	-	RIPE_10_REN	4.23E-07	207.37	-0.94	0.18	
6	0 + 1	-	-	HT_12_REN	1.62E-23	92.13	34.61	2.94	
8	0 + 1	-	-	HT_10_REN	3.76E-26	95.79	30.65	2.41	
5	0 + 1	-	-	FLO1_12_REN	4.33E-07	133.38	2.92	0.55	
10	1 + 0	Field_RRI_11_DI	5.90E-07	-	-	3.24	-0.29	0.06	
5	1 + 0	Field_ADI2_12_KEN	2.55E-05	-	-	1.76	0.34	0.08	
46	1 + 0	Field_RRI_13_DI	2.17E-05	-	-	2.01	0.16	0.04	
42	1 + 0	Field_ADI2_13_DI	1.69E-05	-	-	5.48	-0.28	0.06	
8	1 + 0	CC_RRI_09_Ae109	1.60E-06	-	-	1.37	-0.27	0.05	
14	1 + 0	Field_ADI1_13_RI	3.85E-08	-	-	2.19	-0.22	0.04	
16	1 + 0	Field_RRI_13_DI	6.00E-08	-	-	2.01	0.26	0.05	
8	1 + 0	CC_RRI_09_Ae109	8.82E-28	-	-	1.37	0.81	0.06	
46	1 + 0	Field_ADI2_10_DI	1.27E-05	-	-	8.52	-0.26	0.06	
49	1 + 0	Field_ADI1_12_DI	5.02E-06	-	-	1.76	-0.22	0.05	
24	1 + 0	Field_RRI_12_DI	7.46E-06	-	-	2.52	-0.17	0.04	
25	1 + 0	Field_ADI1_12_KEN	2.64E-05	-	-	1.83	0.18	0.04	
23	1 + 0	Field_RRI_All	8.10E-06	-	-	2.38	0.13	0.03	
25	1 + 0	MFA.Dim.1	1.35E-06	-	-	-1.78	0.51	0.10	
32	1 + 0	CC_RRI_13_RB84	3.89E-06	-	-	2.73	0.17	0.03	
43	1 + 0	MFA.Dim.1	1.26E-05	-	-	-1.78	0.42	0.09	
34	1 + 0	Field_ADI2_13_DI	2.34E-05	-	-	5.48	0.26	0.06	
11	1 + 0	Field_ADI2_10_DI	8.57E-07	-	-	8.52	-0.46	0.09	
27	1 + 0	Field_RRI_10_RI	2.37E-05	-	-	2.48	0.16	0.04	
25	1 + 0	Field_ADI2_13_DI	3.36E-06	-	-	5.48	0.30	0.06	

LD block name (a,b)	LD block interval (cM) (a)	Aphanomyces resistance or earliness Meta-QTL co-localizing with LD block (c)	Aphanomyces resistance, earliness or height QTL detected by linkage mapping co-localizing with LD block (d)	Trait-associated marker (e)	Genetic position (cM) (e)
VI.3	84.2 - 85		<i>Ae-Ps6.4</i>	AA335 (229)	85
VI.4	87.4 - 88.4			<i>PsCam000034_29_2253</i>	88.3
VI.5	99.8 - 100.7			<i>PsCam009689_6479_476</i>	100
VII					
VII.1	17.8 - 18.8		<i>Ae-Ps7.2</i>	<i>PsCam033911_19410_3600</i>	18.8
VII.2	38.8 - 40.1	MQTL-Ae24	<i>Ae-Ps7.3</i>	<i>PsCam000133_116_1343</i>	39.9
VII.3	42.3 - 42.8		<i>Ae-Ps7.4</i>	<i>PsCam037992_23044_806</i>	42.6
				<i>PsCam039102_24080_996</i>	42.6
				<i>PsCam034444_19834_358</i>	42.6
				<i>PsCam023077_12977_819</i>	42.8
VII.4	49.8 - 51.3			<i>PsCam046153_29672_848</i>	50.5
VII.5	57.5 - 58.1			<i>PsCam049829_32444_1381</i>	57.7
VII.6	59.7 - 60.3		<i>Ae-Ps7.5</i>	<i>PsCam011881_8108_513</i>	59.7
VII.7	65 - 65.5	MQTL-Ae25	<i>Ae-Ps7.5</i>	<i>PsCam022983_12922_354</i>	65.5
VII.8	67.75 - 69.2	MQTL-Ae25	<i>Ae-Ps7.5, Ae-Ps7.6a</i>	<i>PsCam011266_7658_455</i>	68.8
VII.9	70.1 - 70.55		<i>Ae-Ps7.5, Ae-Ps7.6a</i>	<i>PsCam043617_27613_310</i>	70.3
VII.10	70.7 - 70.8		<i>Ae-Ps7.6a</i>	<i>PsCam000043_37_605</i>	70.8
VII.11	72.7 - 73.7		<i>Ae-Ps7.6a</i>	<i>PsCam044935_28677_577</i>	73.1
VII.12	80.5 - 80.9	MQTL-Ae26	<i>Ae-Ps7.6a</i>	<i>PsCam051731_34191_118</i>	80.9
VII.13	82.6 - 83.1		<i>Ae-Ps7.6a</i>	AA387 (469)	83
				AA387 (480)	83
				AA387 (501)	83
VII.14	83.5 - 84.1		<i>Ae-Ps7.6a, Ae-Ps7.6b</i>	AB101 (390)	83.5
VII.15	85.5 - 85.8		<i>Ae-Ps7.6b</i>	<i>PsCam050334_32923_389</i>	85.5
VII.16	86.2 - 86.9			<i>PsCam002799_2218_177</i>	86.4
VII.17	93 - 94.1			<i>PsCam042979_27014_298</i>	93
VII.18	97.4 - 100.7			<i>PsCam034448_19837_782</i>	97.9
				<i>PsCam005071_3847_951</i>	98
				<i>PsCam030828_18189_3163</i>	98.5
				<i>PsCam049302_31940_190</i>	98.5

MAF (%) (f)	Nb associated traits (Aphanomyces resistance + Earliness/Height)	Disease associated traits (g)	p-value (h)	Morphological or developmental associated traits (g)	p-value (h)	Intercept (i)	Allelic effect (j)	SE (k)
30	0 + 1	-	-	FLO2_10_REN	1.03E-09	177.17	-1.59	0.25
31	1 + 0	Field_ADI1_12_DI	6.89E-07	-	-	1.76	-0.30	0.06
25	1 + 0	Field_ADI2_13_DI	4.47E-07	-	-	5.48	-0.37	0.07
26	1 + 0	Field_ADI2_13_DI	5.74E-07	-	-	5.48	0.32	0.06
49	1 + 0	Field_ADI1_12_RI	3.99E-07	-	-	4.50	0.53	0.10
45	1 + 0	Field_ADI2_10_DI	3.14E-07	-	-	8.52	0.35	0.07
49	0 + 1	-	-	HT_12_REN	2.55E-05	92.13	8.10	1.87
47	0 + 1	-	-	HT_10_REN	5.00E-06	95.79	-6.42	1.36
44	1 + 0	Field_ADI2_13_DI	2.10E-09	-	-	5.48	-0.42	0.07
43	0 + 1	-	-	FLO2_10_REN	2.06E-06	177.17	-0.91	0.19
42	0 + 2	-	-	FLO2_12_REN	3.69E-08	188.78	-2.47	0.43
				FLO2_10_REN	2.21E-06	177.17	-1.00	0.20
8	1 + 0	Field_RRI_12_RI	1.50E-05	-	-	3.35	-0.24	0.05
39	1 + 0	CC_RRI_09_RB84	5.40E-12	-	-	3.02	0.28	0.04
45	1 + 0	CC_RRI_13_RB84	7.34E-11	-	-	2.73	0.26	0.04
21	0 + 1	-	-	FLO2_10_REN	1.05E-06	177.17	-1.09	0.22
26	1 + 0	CC_RRI_13_RB84	1.02E-07	-	-	2.73	0.19	0.03
47	1 + 0	Field_RRI_12_RI	7.09E-07	-	-	3.35	0.19	0.04
18	0 + 1	-	-	FLO1_12_REN	4.99E-06	133.38	1.55	0.33
15	1 + 0	Field_RRI_12_RI	3.14E-07	-	-	3.35	0.29	0.06
15	1 + 0	Field_ADI1_10_RI	3.12E-08	-	-	4.11	-0.40	0.07
12	0 + 2	-	-	RIPE_10_REN	1.44E-12	207.37	-1.99	0.26
				FLO2_10_REN	9.52E-11	177.17	-2.08	0.30
11	2 + 0	Field_ADI2_11_RI	3.61E-06	-	-	4.01	0.43	0.09
		MFA.Dim.2	1.48E-05	-	-	0.58	0.48	0.11
41	1 + 0	Field_ADI1_13_RI	4.81E-06	-	-	2.19	-0.13	0.03
42	4 + 4	-	-	RIPE_10_REN	5.02E-13	207.37	-1.53	0.19
				FLO1_10_REN	1.24E-17	152.56	-4.68	0.49
				RIPE_12_REN	2.69E-17	212.13	-3.39	0.36
				FLO2_12_REN	2.68E-15	188.78	-5.49	0.63
		Field_ADI1_All	9.49E-08	-	-	3.26	0.32	0.06
		Field_ADI2_All	1.15E-07	-	-	4.76	0.38	0.07
		Field_RRI_12_RI	2.79E-07	-	-	3.35	0.26	0.05
		Field_ADI2_13_DI	1.79E-05	-	-	5.48	0.40	0.09
12	0 + 1	-	-	FLO1_12_REN	4.30E-10	133.38	2.93	0.44
34	1 + 0	Field_ADI1_12_DI	1.71E-06	-	-	1.76	0.25	0.05
32	1 + 0	MFA.Dim.1	1.28E-06	-	-	-1.78	0.45	0.09
15	1 + 1	Field_ADI2_10_DI	7.56E-08	-	-	8.52	-0.44	0.08
				FLO2_10_REN	1.63E-06	177.17	1.36	0.27
23	0 + 1	-	-	FLO1_10_REN	1.97E-05	152.56	-1.82	0.41

Additional File II.12

Information on the connectivity between projected maps

(a) HMap: consensus map from Hamon et al. [43]; (b) TMap: consensus map from Tayeh et al. [49]; (c) number of common markers between HMap and TMap for each linkage group (LG); (d): Number of inversions of marker positions on each LG; (e) number of inversions of marker positions in confidence intervals of the quantitative trait loci (QTL) described by Hamon et al. [43].

LG	Number of markers			Inversions (d)	Inversion in QTL (e)
	HMap (a)	TMap (b)	Common (c)		
I	56	1715	13	4	4
II	65	2145	20	3	3
III	91	2676	18	3	2
IV	137	2254	22	9	9
V	64	1983	24	8	6
VI	94	1993	18	7	3
VII	111	2526	29	8	7

Additional File II.13

Description of linkage disequilibrium (LD) block haplotypes in the pea-Aphanomyces collection

(a) LD blocks are named as described in Additional File 11; (b) Each LD block is subdivided into marker haplotypes according to its allelic composition. Marker haplotypes are named with the LD block name followed by an Arabic numeral. Only haplotypes without missing values or heterozygous markers are shown;

For each LD block: (c) list of markers significantly detected by genome-wide association in the LD block; (d) the first line is a list of the markers included in the LD block (detected markers in bold font and their markers in (LD) in plain font); (e) the second line is the genetic positions of the listed markers on the genetic map described in Additional File 11; (f) following lines are the pairwise LD r^2 values between each marker defined in the LD block and the marker detected by GWA; (g) For SNP markers, 'AA' and 'BB' are allele codes of Genome Studio analysis while for SSR markers, amplicon lengths are indicated. For each Aphanomyces resistance associated marker, enhancing allele(s) is in bold dark green font and unfavourable allele(s) in bold dark red font; (h) the first line is the list of the variables significantly associated with a marker in the LD block, variables are described as in Additional File 4; (i) next lines are *p-values* of the marker-trait associations; (j) grey or coloured lines are mean phenotypic values \pm standard error for marker haplotypes carried by more than 5 % of the lines from the pea-Aphanomyces collection, letters indicate significantly different means (Tukey-HSD, $\alpha < 0.05$); dark green and red haplotypes: favourable and unfavourable haplotypes of the LD block, respectively, as defined in the Material and Methods section of the manuscript.

LD block Name (a)	Haplotype (b)	Associated marker(s) in the LD block (c)	Associated marker(s) (bold) and marker(s) in LD (plain) (d)	Genetic position (cM) (e)	LD r ² value of each marker with associated marker(s) (f)	Allelic composition of LD-block pattern (g)	Number of lines carrying the LD block haplotype			
I	I.5	PScam027153_15841_364 PScam044835_28587_1878	PScam005033_3817_61	88.2	0.83	AA	1			
			PScam027153_15841_364	88.2	1.00	AA	1			
			PScam042815_26857_431	88.2	0.91	AA	1			
			PScam024866_14152_261	88.9	0.40	BB	1			
			PScam037259_22356_1396	88.9	0.73	BB	1			
			PScam01366_7739_322	89.1	0.27	BB	1			
			PScam038113_23162_1036	89.1	0.39	BB	1			
			PScam044835_28587_1878	89.1	1.00	BB	1			
			PScam044855_28602_1561	89.1	0.55	BB	1			
			PScam059295_39494_4455	89.1	0.28	AA	1			
			PScam035316_20506_2839	89.4	0.29	BB	1			
			PScam035496_20678_3576	89.9	0.51	AA	1			
			PScam000400_351_2330	90.4	0.39	AA	1			
			PScam000507_444_666	90.4	0.71	AA	1			
			I.5.01	AA	BB	BB	BB	BB	AA	5
			I.5.02	AA	BB	BB	BB	BB	BB	15
			I.5.03	AA	BB	BB	BB	BB	BB	50
			I.5.04	BB	AA	AA	AA	AA	AA	2
			I.5.05	AA	AA	AA	AA	AA	AA	1
			I.5.06	AA	BB	BB	BB	BB	BB	1
			I.5.07	AA	BB	BB	BB	BB	BB	2
			I.5.08	AA	BB	BB	BB	BB	BB	1
			I.5.09	AA	BB	BB	BB	BB	BB	1
			I.5.10	AA	BB	BB	BB	BB	BB	1
I.5.11	BB	AA	AA	AA	AA	AA	15			
I.5.12	BB	AA	AA	AA	AA	AA	1			
I.5.13	BB	AA	AA	AA	AA	AA	1			
I.5.14	BB	AA	AA	AA	AA	AA	37			
I.5.15	BB	AA	AA	AA	AA	AA	2			
I.5.16	BB	AA	AA	AA	AA	AA	2			
I.5.17	BB	AA	AA	AA	AA	AA	23			
I.5.18	BB	AA	AA	AA	AA	AA	1			
I.5.19	BB	AA	AA	AA	AA	AA	2			
I.5.20	BB	AA	AA	AA	AA	AA	1			
I.5.21	BB	AA	AA	AA	AA	AA	1			
I.5.22	BB	AA	AA	AA	AA	AA	1			
I.5.23	AA	BB	BB	BB	BB	BB	1			
I.5.24	AA	BB	BB	BB	BB	BB	1			

LG	LD block Name (a)	Haplotype (b)	Associated marker(s) in the LD block (c)	Associated marker(s) (bold) and marker(s) in LD (plain) (d) Genetic position (cM) (e) LD r^2 value of each marker with associated marker(s) (f) Allelic composition of LD-block pattern (g)	Number of lines carrying the LD block haplotype
III	III.3		PsCam0355590_20765_380	PsCam055250_36400_562 PsCam035590_20765_380 PsCam039309_24261_68 PsCam02271_1824_370	7 80 37 1 7 6 27
		III.3.01		BB AA	
		III.3.02		BB AA	
		III.3.03		AA AA	
		III.3.04		AA BB	
		III.3.05		BB AA	
		III.3.06		BB BB	
		III.3.07		BB BB	
III	III.4		PsCam023109_13003_2421	PsCam037297_22389_303 PsCam023109_13003_2421	73 45 54
		III.4.01		AA AA	
		III.4.02		AA BB	
		III.4.03		BB BB	

LG	LD block Name (a)	Haplotype (b)	Associated marker(s) in the LD block (c)	Mean (±standard error) of associated traits for significantly represented LD-block patterns and group of mean comparison (j)	p-value of association between associated marker and traits (i)	Associated traits with one of the markers in the LD-block (h)					
III	III.3										
				Field_RRI_All	7.86E-11	1.51E-08	MFA Dim.1	Field_RRI_11_RI	9.92E-08	5.78E-06	Field_RRI_10_DI
		III.3.01	PsCam035590_20765_380	2.51 ± 0.04 a	-0.74 ± 0.16 a	2.92 ± 0.06 a	2.51 ± 0.05 a				
		III.3.02		2.72 ± 0.08 b	-0.15 ± 0.28 a	3.14 ± 0.11 a	2.69 ± 0.08 a				
		III.3.03									
		III.3.04									
		III.3.05									
III.3.06											
III.3.07				3.17 ± 0.08 c	2.04 ± 0.32 b	3.89 ± 0.12 b	3.11 ± 0.09 b				
III	III.4										
				Field_RRI_12_RI	1.23E-06	6.19E-06	Field_RRI_12_DI	1.81E-05	Field_RRI_10_RI		
		III.4.01	PsCam023109_13003_2421	3.29 ± 0.05 a	2.41 ± 0.05 a	2.71 ± 0.06 a					
		III.4.02		3.72 ± 0.09 b	2.78 ± 0.08 b	3.23 ± 0.09 c					
		III.4.03		3.60 ± 0.08 b	2.55 ± 0.08 a	3.00 ± 0.09 b					

LG	LD block Name (a)	Haplotype (b)	Associated marker(s) in the LD block (c)	Associated marker(s) (bold) and marker(s) in LD (plain) (d) Genetic position (cM) (e) LD r ² value of each marker with associated marker(s) (f) Allelic composition of LD-block pattern (g)	Number of lines carrying the LD block haplotype		
III	III.6	III.6.01 III.6.02 III.6.03 III.6.04 III.6.05 III.6.06 III.6.07 III.6.08 III.6.09 III.6.10	PsCam034645_20018_1507	PsCam037941_22996_137	AA	63	
				PsCam000088_75_1456	AA		AA
				PsCam045339_29025_1722	AA		BB
				PsCam046219_29720_864	AA		BB
				PsCam001039_886_135	AA		AA
				PsCam030789_18160_162	AA		BB
				PsCam056250_37103_770	AA		AA
				PsCam016014_10253_143	AA		AA
				PsCam013984_9520_1022	AA		AA
				PsCam043953_27903_3736	AA		AA
				PsCam037415_22499_621	AA		AA
				PsCam001871_1531_550	AA		AA
				PsCam034645_20018_1507	BB		BB
III	III.9	III.9.01 III.9.02 III.9.03 III.9.04 III.9.05 III.9.06 III.9.07 III.9.08 III.9.09 III.9.10 III.9.11 III.9.12 III.9.13 III.9.14 III.9.15	PsCam030789_18160_162	PsCam058449_38886_549	AA	60	
				PsCam000125_108_427	AA		BB
				PsCam045339_29025_1722	AA		BB
				PsCam046219_29720_864	AA		BB
				PsCam001039_886_135	AA		BB
				PsCam030789_18160_162	AA		BB
				PsCam056250_37103_770	AA		AA
				PsCam016014_10253_143	AA		AA
				PsCam013984_9520_1022	AA		AA
				PsCam043953_27903_3736	AA		AA
				PsCam037415_22499_621	AA		AA
				PsCam001871_1531_550	AA		AA
				PsCam034645_20018_1507	BB		BB
III	III.9	III.9.01 III.9.02 III.9.03 III.9.04 III.9.05 III.9.06 III.9.07 III.9.08 III.9.09 III.9.10 III.9.11 III.9.12 III.9.13 III.9.14 III.9.15	PsCam030789_18160_162	PsCam058449_38886_549	AA	60	
				PsCam000125_108_427	AA		BB
				PsCam045339_29025_1722	AA		BB
				PsCam046219_29720_864	AA		BB
				PsCam001039_886_135	AA		BB
				PsCam030789_18160_162	AA		BB
				PsCam056250_37103_770	AA		AA
				PsCam016014_10253_143	AA		AA
				PsCam013984_9520_1022	AA		AA
				PsCam043953_27903_3736	AA		AA
				PsCam037415_22499_621	AA		AA
				PsCam001871_1531_550	AA		AA
				PsCam034645_20018_1507	BB		BB

LG	LD block Name (a)	Haplotype (b)	Associated marker(s) in the LD block (c)	Mean (\pm standard error) of associated traits for significantly represented LD-block patterns and group of mean comparison (j)	Associated traits with one of the markers in the LD-block (h) p-value of association between associated marker and traits (i)	
III	III.6					
				Field_AD12_13_DI		
				Field_AD11_13_DI		
				3.25E-12	1.16E-08	
				4.32 \pm 0.13 a	2.18 \pm 0.09 a	
				4.70 \pm 0.26 ab	2.60 \pm 0.18 ab	
				6.52 \pm 0.32 c	3.79 \pm 0.22 c	
				6.16 \pm 0.27 c	3.64 \pm 0.19 c	
				5.23 \pm 0.24 b	2.92 \pm 0.16 b	
III	III.9					
				Field_AD1_All		
				Field_AD11_13_RI		
				4.91E-07	2.62E-06	
				2.49 \pm 0.14 a	1.36 \pm 0.1 a	
				2.67 \pm 0.17 ab	1.61 \pm 0.13 ab	
				3.27 \pm 0.22 c	1.72 \pm 0.17 ab	
				3.11 \pm 0.19 bc	1.79 \pm 0.14 b	
				3.22 \pm 0.16 c	1.70 \pm 0.11 b	
				2.89 \pm 0.23 ac	1.82 \pm 0.17 ab	

LG	LD block Name (a)	Haplotype LD block (b)	Associated marker(s) in the LD block (c)	Mean (±standard error) of associated traits for significantly represented LD-block patterns and group of mean comparison (j)	Associated traits with one of the markers in the LD-block (h) p-value of association between associated marker and traits (i)	
IV	IV.11			Field_RRI_12_DI	7.46E-06	
				Field_AD11_12_KEN	2.64E-05	
			PsCam030702_18121_172 PsCam030723_18124_1544			
		IV.11.01				
		IV.11.02				
		IV.11.03				
		IV.11.04		2.43 ± 0.06 a	1.73 ± 0.05 a	
		IV.11.05		2.70 ± 0.09 bc	1.93 ± 0.08 ab	
		IV.11.06				
		IV.11.07		2.93 ± 0.15 c	2.11 ± 0.13 b	
		IV.11.08				
		IV.11.09				
		IV.11.10				
	IV.11.11					
	IV.11.12					
	IV.11.13			2.51 ± 0.09 ab	1.83 ± 0.08 ab	
IV	IV.12			Field_RRI_All	8.10E-06	
				MFA_Dim_1	1.35E-06	
			PsCam042808_26850_709 PsCam043503_27507_1032			
		IV.12.01				
		IV.12.02		2.69 ± 0.1 a	0.09 ± 0.41 a	
		IV.12.03		2.85 ± 0.13 a	0.51 ± 0.5 a	
		IV.12.04		2.74 ± 0.12 a	0.03 ± 0.46 a	
		IV.12.05				
		IV.12.06				
		IV.12.07			2.63 ± 0.13 a	-0.28 ± 0.5 a
	IV.12.08					

LG	LD block Name (a)	Haplotype (b)	Associated marker(s) in the LD block (c)	Mean (±standard error) of associated traits for significantly represented LD-block patterns and group of mean comparison (j)	p-value of association between associated marker and traits (i)	Associated traits with one of the markers in the LD-block (h)
VII	VII.3			HT_10_REN 5.00E-06	2.55E-05	HT_12_REN
			PsCam034444_19834_358 PsCam037992_23044_806 PsCam039102_24080_996 PsCam023077_12977_819	3.14E-07	2.10E-09	Field_AD12_13_DI
				Field_AD12_10_DI		
				93.13 ± 3.7 ab 92.31 ± 4.59 ab	6.63 ± 0.24 a 6.38 ± 0.29 a	5.09 ± 0.29 ab 4.26 ± 0.36 a
				82.5 ± 5.03 a	6.44 ± 0.32 a	4.82 ± 0.4 ab
					89.57 ± 6.45 a	
					103.62 ± 5.88 ab	
					96.15 ± 4.75 ab	
					103.24 ± 5.26 b	
				95.78 ± 4.1 b	6.77 ± 0.26 a	5.08 ± 0.32 b

LG	LD block Name (a) (b)	Haplotype LD block (c)	Associated marker(s) in the LD block (c)	Associated marker(s) (bold) and marker(s) in LD (plain) (d) Genetic position (cM) (e) LD r^2 value of each marker with associated marker(s) (f) Allelic composition of LD-block pattern (g)	Number of lines carrying the LD block haplotype
VII	VII.14		AB101 (390)	<p>83.5 84.1 PsCam007709_5623_138 1.00 0.28</p> <p>364 AA 366 AA 368 AA 376 AA 378 AA 380 AA 362 BB 364 BB 368 BB 376 BB 380 BB 384 BB 392 BB 390 BB</p>	18 4 61 27 1 1 6 3 2 1 21 1 1 15
VII	VII.16		PsCam002799_2218_177 PsCam056236_37092_1071	<p>86.2 86.4 86.4 86.9 0.59 1.00 0.43 0.23 0.27</p> <p>PsCam034407_19800_790 PsCam002799_2218_177 PsCam038378_23415_721 PsCam056236_37092_1071 PsCam039084_24063_2038</p> <p>AA BB AA AA AA BB AA AA AA BB AA AA BB BB AA AA BB AA AA AA BB AA BB BB BB AA AA AA BB AA AA BB BB AA AA BB</p>	25 20 30 3 1 1 68 16 1

LG	LD block Name (a)	Haplotype (b)	Associated marker(s) in the LD block (c)	Associated traits with one of the markers in the LD-block (h)					
				Mean (±standard error) of associated traits for significantly represented LD-block patterns and group of mean comparison (j)	p-value of association between associated marker and traits (i)				
VII	VII.18			Field_AD12_12_DI	MFA_Dim.1	Field_AD12_10_DI	FLO2_10_REN	FLO1_10_REN	
			PsCam034448_19837_782 PsCam005071_3847_951 PsCam030828_18189_3163 PsCam049302_31940_190	1.71E-06	1.28E-06	7.56E-08	1.63E-06	1.97E-05	
				VII.18.01					
				VII.18.02					
				VII.18.03					
				VII.18.04					
				VII.18.05	1.79 ± 0.11 a	-0.24 ± 0.31 a	6.32 ± 0.17 a	169.91 ± 0.83 a	145.56 ± 0.92 ab
				VII.18.06					
				VII.18.07					
				VII.18.08					
				VII.18.09	1.81 ± 0.13 a	-0.47 ± 0.36 a	6.71 ± 0.2 a	168.95 ± 0.98 a	144.14 ± 1.09 ab
				VII.18.10					
				VII.18.11					
				VII.18.12					
				VII.18.13					
				VII.18.14					
				VII.18.15					
				VII.18.16					
				VII.18.17					
				VII.18.18					
				VII.18.19					
				VII.18.20					
				VII.18.21	2.64 ± 0.17 b	1.58 ± 0.47 b	7.36 ± 0.26 b	171.32 ± 1.26 a	146.99 ± 1.41 b
				VII.18.22					
				VII.18.23					
				VII.18.24	1.47 ± 0.18 a	-0.38 ± 0.48 a	6.17 ± 0.26 a	170.69 ± 1.31 a	142.46 ± 1.46 a
				VII.18.25					
				VII.18.26					

Additional File II.14

Marker haplotype composition of the pea-Aphanomyces collection

(a) Name of marker haplotype carried by each line as described in Additional File 13, at each of the 14 consistent linkage disequilibrium (LD) blocks detected; dark green and red: favourable and unfavourable marker haplotypes, respectively, as defined in the Material and Methods section of the manuscript; (b) coordinates of each line of the collection on the first Multiple Factor Analysis principal component; the pea lines of the collection are ranked according to these coordinates ; (c) Composition of each line in favourable and unfavourable haplotypes, is described in Additional file 13; (d) marker haplotypes with either a missing value or heterozygous score for at least one marker of a LD block.

Line code	Name	Haplotype composition of the pea-Aphanomyces collection lines at consistent LD-blocks (a)												MFA.Dim.1 (b)	Number of favorable haplotypes (c)	Number of unfavorable haplotypes (c)	Number of missing haplotypes (d)	
		I.5	II.1	III.3	III.4	III.6	III.9	III.15	IV.11	IV.12	III.9	III.13	III.14					III.16
A100	AeD990SW-45-8-7	I.5.03	II.1.02	III.3.02	III.4.01	III.6.01	III.9.04	III.15.13	IV.11.04	IV.12.02	III.3.19	III.13.11	III.14.11	III.16.07	III.18.09	12	1	0
A104	AeD990SW-50-2-5	I.5.03	II.1.06	III.3.02	III.4.01	III.6.01	III.9.04	III.15.13	IV.11.04	IV.12.02	III.3.19	III.13.11	III.14.11	III.16.07	III.18.09	8	0	6
A101	AeD990SW-47-6-1	I.5.03	II.1.06	III.3.02	III.4.01	III.6.01	III.9.11	III.15.13	IV.11.13	IV.12.07	III.3.19	III.13.11	III.14.14	III.16.07	III.18.09	9	4	0
A110	AeD990SW-57-9-8	I.5.20	II.1.06	III.3.02	III.4.01	III.6.01	III.9.04	III.15.22	IV.11.04	IV.12.04	III.3.04	III.13.11	III.14.14	III.16.07	III.18.12	10	1	0
A103	AeD990SW-49-5-7	I.5.03	II.1.02	III.3.02	III.4.01	III.6.01	III.9.04	III.15.03	IV.11.04	IV.12.02	III.3.19	III.13.11	III.14.03	III.16.07	III.18.09	13	1	0
A098	AeD990SW-43-9-8	I.5.02	II.1.01	III.3.02	III.4.01	III.6.10	III.9.12	III.15.13	IV.11.04	IV.12.02	III.3.19	III.13.10	III.14.11	III.16.07	III.18.09	8	1	0
A039	AeB97-738.2.3	I.5.01	II.1.17	III.3.02	III.4.01	III.6.01	III.9.02	III.15.02	IV.11.05	IV.12.01	III.3.07	III.13.05	III.14.03	III.16.07	III.18.06	6	0	0
A095	AeD990SW-40-3-3	I.5.03	II.1.02	III.3.03	III.4.03	III.6.10	III.9.12	III.15.03	IV.11.04	IV.12.04	III.3.19	III.13.08	III.14.03	III.16.07	III.18.09	7	1	0
A106	AeD990SW-53-3-5	NA	NA	NA	III.4.01	III.6.01	III.9.11	III.15.13	IV.11.04	IV.12.04	III.3.19	III.13.11	III.14.03	NA	III.18.09	8	0	7
A102	AeD990SW-48-5-4	I.5.02	II.1.06	III.3.02	III.4.01	III.6.01	III.9.04	III.15.13	IV.11.13	IV.12.07	III.3.19	III.13.11	III.14.13	III.16.07	III.18.09	9	2	1
A108	AeD990SW-55-6-4	I.5.02	II.1.17	III.3.03	III.4.03	III.6.10	III.9.11	III.15.13	IV.11.13	IV.12.07	III.3.19	III.13.11	III.14.13	III.16.07	III.18.11	9	1	1
A096	AeD990SW-41-7-6	I.5.17	II.1.02	III.3.02	III.4.01	III.6.01	III.9.04	III.15.22	IV.11.04	IV.12.04	III.3.04	III.13.08	III.14.11	III.16.07	III.18.09	8	1	2
A109	AeD990SW-56-19-4	I.5.03	II.1.18	III.3.02	III.4.01	III.6.01	III.9.04	III.15.22	IV.11.13	IV.12.07	III.3.04	III.13.11	III.14.03	III.16.07	III.18.12	10	1	0
A050	AeD99QU-04-88-5-2	I.5.14	II.1.19	III.3.02	III.4.01	III.6.02	III.9.05	III.15.01	IV.11.05	IV.12.03	III.3.04	III.13.05	III.14.04	III.16.01	III.18.09	5	0	0
A051	AeD99QU-04-88-2-5	I.5.14	II.1.19	III.3.02	III.4.01	III.6.02	III.9.05	III.15.01	IV.11.05	IV.12.03	III.3.04	III.13.05	III.14.04	III.16.01	III.18.09	5	0	0
A042	AeB97-741.9.6	I.5.01	II.1.17	III.3.02	III.4.01	III.6.01	III.9.02	III.15.02	IV.11.05	IV.12.01	III.3.07	III.13.05	III.14.03	III.16.07	III.18.07	6	0	0
A079	AeD990SW-22-3-7	I.5.17	II.1.13	III.3.01	III.4.02	III.6.01	III.9.11	III.15.13	IV.11.04	IV.12.04	III.3.04	III.13.08	III.14.03	III.16.07	III.18.09	7	2	0
A009	AeD9904-7	I.5.14	II.1.19	III.3.02	III.4.01	III.6.02	III.9.05	III.15.01	IV.11.05	IV.12.03	III.3.04	III.13.05	III.14.03	III.16.07	III.18.09	7	2	0
A105	AeD990SW-51-2-10	I.5.02	II.1.02	III.3.03	III.4.03	III.6.10	III.9.11	III.15.18	IV.11.04	IV.12.04	III.3.04	III.13.11	III.14.11	III.16.07	III.18.02	9	1	0
A040	AeB97-739.15.2	I.5.01	II.1.15	III.3.02	III.4.01	III.6.01	III.9.02	III.15.02	IV.11.05	IV.12.01	III.3.07	III.13.05	III.14.03	III.16.07	III.18.06	6	1	0
A111	AeD990SW-58-10-5	I.5.17	II.1.06	III.3.02	III.4.01	III.6.01	III.9.12	III.15.13	IV.11.13	IV.12.07	III.3.04	III.13.08	III.14.03	III.16.07	III.18.09	7	2	1
A092	AeD990SW-37-3-4	I.5.02	II.1.06	III.3.02	III.4.01	III.6.01	III.9.04	III.15.13	IV.11.13	IV.12.07	III.3.04	III.13.08	III.14.03	III.16.07	III.18.09	11	1	0
A041	AeB97-740.11.1	I.5.01	II.1.17	III.3.02	III.4.01	III.6.01	III.9.02	III.15.02	IV.11.05	IV.12.03	III.3.07	III.13.05	III.14.03	III.16.07	III.18.06	6	0	0
A113	AeD990SW-60-1-3	I.5.03	II.1.06	III.3.02	III.4.01	III.6.01	III.9.04	III.15.17	IV.11.04	IV.12.04	III.3.19	III.13.12	III.14.14	III.16.08	NA	8	3	1
A020	AeA95-320.M71.2	I.5.17	II.1.02	III.3.02	III.4.03	III.6.02	III.9.05	III.15.17	IV.11.04	IV.12.04	III.3.03	III.13.05	III.14.04	III.16.01	III.18.05	5	1	0
A019	AeA95-317.M.71.2	I.5.17	II.1.02	III.3.02	III.4.03	III.6.02	III.9.05	III.15.17	IV.11.04	IV.12.04	III.3.04	III.13.05	III.14.04	III.16.01	III.18.05	5	1	1
A077	AeD990SW-18-6-5	I.5.17	II.1.13	III.3.02	III.4.01	III.6.10	III.9.11	III.15.11	IV.11.13	IV.12.03	III.3.04	III.13.08	III.14.03	III.16.07	III.18.09	6	2	1
A049	AeD99QU-04-7-3-1	I.5.14	II.1.19	III.3.02	III.4.01	III.6.02	III.9.05	III.15.01	IV.11.05	IV.12.03	III.3.04	III.13.05	III.14.04	III.16.01	III.18.09	5	0	0
A127	552	I.5.03	II.1.06	III.3.02	III.4.01	III.6.01	III.9.01	III.15.03	IV.11.05	IV.12.03	III.3.19	III.13.11	III.14.11	III.16.07	III.18.18	9	1	1
A045	AeD99QU-04-3p-7-1	I.5.17	II.1.02	III.3.02	III.4.01	III.6.01	III.9.10	III.15.01	IV.11.05	IV.12.03	III.3.07	III.13.05	III.14.04	III.16.01	III.18.09	5	0	1
A047	AeD99QU-04-6-1	I.5.17	II.1.02	III.3.02	III.4.01	III.6.01	III.9.10	III.15.01	IV.11.13	IV.12.07	III.3.19	III.13.05	III.14.04	III.16.01	III.18.09	6	2	0
A078	AeD990SW-20-3-7	I.5.14	II.1.13	III.3.02	III.4.01	III.6.01	III.9.10	III.15.22	IV.11.05	IV.12.03	III.3.04	III.13.08	III.14.03	III.16.03	III.18.09	7	0	0
A046	AeD99QU-04-3sp-9-2	I.5.17	II.1.02	III.3.02	III.4.01	III.6.01	III.9.10	III.15.01	IV.11.05	IV.12.03	III.3.04	III.13.05	III.14.04	III.16.01	III.18.09	6	0	1
A048	AeD99QU-04-6-6-3	I.5.14	II.1.19	III.3.02	III.4.01	III.6.02	III.9.05	III.15.11	IV.11.04	IV.12.04	III.3.04	III.13.05	III.14.04	III.16.01	III.18.09	6	0	0
MR19	RIL 847-68	I.5.03	II.1.02	III.3.02	III.4.03	III.6.02	III.9.11	III.15.13	IV.11.13	IV.12.07	III.3.19	III.13.11	III.14.03	III.16.07	III.18.25	4	4	2
A052	AeD99QU-04-9-10-1	I.5.14	II.1.19	III.3.02	III.4.01	III.6.02	III.9.05	III.15.01	IV.11.04	IV.12.04	III.3.04	III.13.05	III.14.04	III.16.01	III.18.09	6	0	0
A031	AeB97-712.11.1	I.5.14	II.1.04	III.3.02	III.4.01	III.6.01	III.9.04	III.15.17	IV.11.05	IV.12.03	III.3.03	III.13.05	III.14.11	III.16.07	III.18.05	8	1	0
A035	AeB97-725.11.2	I.5.02	II.1.04	III.3.02	III.4.01	III.6.01	III.9.04	III.15.17	IV.11.05	IV.12.03	III.3.19	III.13.05	III.14.03	III.16.07	III.18.05	10	1	0
A022	AeA95-325.M.34.1	I.5.14	II.1.02	III.3.02	III.4.03	III.6.02	III.9.05	III.15.17	IV.11.04	IV.12.04	III.3.03	III.13.05	III.14.04	III.16.01	III.18.05	5	1	0
A117	O-2016	I.5.06	NA	III.3.03	III.4.03	III.6.10	III.9.11	III.15.03	IV.11.07	IV.12.02	III.3.19	III.13.15	III.14.03	III.16.07	III.18.05	2	2	2
MR13	RIL 846-07	I.5.02	NA	III.3.03	III.4.03	III.6.10	III.9.11	III.15.13	IV.11.13	IV.12.07	III.3.19	III.13.17	III.14.03	III.16.07	III.18.24	4	4	1
MR18	RIL 847-50	I.5.03	NA	III.3.03	III.4.03	III.6.08	III.9.09	III.15.02	IV.11.06	IV.12.04	III.3.19	III.13.17	III.14.11	III.16.07	III.18.25	5	1	2
MR06	BA55.26	NA	NA	III.3.02	III.4.01	III.6.01	III.9.04	III.15.01	IV.11.13	IV.12.07	III.3.04	III.13.11	III.14.11	III.16.07	III.18.05	2	3	3
A072	AeD990SW-12-18-4	I.5.17	II.1.02	III.3.02	III.4.01	III.6.01	III.9.10	III.15.22	IV.11.13	IV.12.01	III.3.18	III.13.06	III.14.03	III.16.03	III.18.09	6	0	0
A028	AeA95-376.M.33.1	I.5.11	II.1.12	III.3.07	III.4.02	III.6.02	III.9.05	III.15.10	IV.11.05	IV.12.04	III.3.04	III.13.05	III.14.03	III.16.02	III.18.05	5	3	3
A055	AeD99QU-04-17-8-1	I.5.14	II.1.15	III.3.07	III.4.02	III.6.02	III.9.05	III.15.10	IV.11.05	IV.12.03	III.3.03	III.13.05	III.14.01	III.16.02	III.18.09	3	0	0
A024	AeA95-327.M.24.7	I.5.14	II.1.02	III.3.02	III.4.03	III.6.02	III.9.05	III.15.22	IV.11.04	IV.12.04	III.3.04	III.13.05	III.14.14	III.16.08	III.18.05	6	2	0
A129	VK4888 SPARTANETS	I.5.17	II.1.02	III.3.05	III.4.02	III.6.04	III.9.12	III.15.22	IV.11.13	IV.12.07	III.3.04	III.13.05	III.14.14	III.16.08	III.18.05	6	2	0
A025	AeA95-335.M.51.3	I.5.02	II.1.02	III.3.02	III.4.03	III.6.02	III.9.05	III.15.22	IV.11.04	IV.12.04	III.3.19	III.13.06	III.14.03	III.16.03	III.18.09	6	0	0
A116	PH180693	I.5.17	II.1.02	III.3.05	III.4.02	III.6.04	III.9.12	III.15.22	IV.11.13	IV.12.07	III.3.04	III.13.05	III.14.14	III.16.08	III.18.05	3	3	3
MR15	RIL 847-22	I.5.03	NA	III.3.03	III.4.03	III.6.10	III.9.11	III.15.13	IV.11.04	IV.12.04	III.3.04	III.13.15	III.14.14	III.16.08	III.18.05	8	1	0
A008	O-2093	I.5.03	II.1.20	III.3.06	III.4.02	III.6.01	III.9.05	III.15.13	IV.11.04	IV.12.04	III.3.19	III.13.17	III.14.03	III.16.07	III.18.24	5	3	1
A054	AeD99QU-04-15-8-1	I.5.02	II.1.04	III.3.02	III.4.01	III.6.01	III.9.04	III.15.08	IV.11.06	IV.12.02	III.3.19	III.13.05	III.14.03	III.16.07	III.18.09	4	1	0
A080	AeD990SW-23-7-4	I.5.17	II.1.13	III.3.03	III.4.03	III.6.10	III.9.11	III.15.17	IV.11.13	IV.12.07	III.3.03	III.13.08	III.14.03	III.16.07	III.18.09	4	3	0

Line code	Name	Haplotype composition of the pea-Aphanomyces collection lines at consistent LD-blocks (a)																MFA.Dim.1 (b)	Number of favorable haplotypes (c)	Number of unfavorable haplotypes (c)	Number of missing haplotypes (d)
		I.5	II.1	III.3	III.4	III.6	III.9	III.15	IV.11	IV.12	III.3	III.13	III.14	III.16	III.18						
A081	AeD990SW- 24-10-5	I.5.14	II.1.19	III.3.03	III.4.03	III.6.10	III.9.11	III.15.11	IV.11.05	IV.12.03	III.3.04	III.13.08	III.14.03	III.16.07	III.18.09	-0.9604	5	1	0		
A003	90-2131	I.5.02	II.1.17	III.3.03	III.4.03	III.6.10	III.9.11	III.15.13	IV.11.13	IV.12.07	III.3.19	III.13.03	III.14.03	III.16.07	III.18.09	-0.9394	4	3	0		
A089	AeD990SW- 34-7-2	NA	NA	III.3.02	NA	III.6.01	III.9.04	III.15.22	IV.11.04	NA	III.3.19	III.13.08	III.14.14	III.16.08	III.18.09	-0.9384	5	3	4		
MR08	BASS5.88	I.5.03	III.3.03	III.4.01	III.4.01	III.6.01	III.9.04	III.15.18	IV.11.04	NA	III.3.19	III.13.17	NA	III.18.24	NA	-0.9004	1	8	8		
A107	AeD990SW- 54-10-5	I.5.17	II.1.06	III.3.03	III.4.03	III.6.10	III.9.11	III.15.18	IV.11.04	IV.12.04	III.3.19	III.13.17	III.14.03	III.16.07	III.18.10	-0.8994	6	2	0		
A128	ROSALENDE	I.5.17	II.1.02	III.3.05	III.4.02	III.6.04	III.9.12	III.15.22	IV.11.13	IV.12.07	III.3.04	NA	III.14.14	III.16.08	III.18.18	-0.8976	3	4	1		
A083	AeD990SW- 26-3-8	I.5.17	II.1.19	III.3.03	III.4.03	III.6.02	III.9.11	III.15.22	IV.11.04	IV.12.08	III.3.19	III.13.08	III.14.03	III.16.07	III.18.09	-0.8458	2	2	0		
A145	W/8903	I.5.03	II.1.04	III.3.02	III.4.01	III.6.01	III.9.09	III.15.02	IV.11.07	IV.12.02	III.3.19	III.13.09	III.14.07	III.16.07	III.18.21	-0.8416	6	3	2		
A097	AeD990SW- 42-3-1	I.5.03	II.1.06	III.3.03	III.4.03	III.6.01	III.9.11	III.15.22	IV.11.13	IV.12.07	III.3.19	III.13.09	III.14.03	III.16.07	III.18.21	-0.8230	5	3	1		
A056	AeD990SW- 04-18of-10-2	I.5.03	II.1.04	III.3.02	III.4.01	III.6.01	III.9.11	III.15.13	IV.11.13	IV.12.07	III.3.19	III.13.05	III.14.03	III.16.03	III.18.09	-0.8131	7	3	0		
A073	200	I.5.03	II.1.04	III.3.02	III.4.01	III.6.01	III.9.11	III.15.03	IV.11.13	IV.12.07	III.3.19	III.13.05	III.14.03	III.16.03	III.18.09	-0.7897	7	3	0		
A053	AeD990SW- 14-10-4	I.5.03	II.1.04	III.3.02	III.4.01	III.6.01	III.9.11	III.15.03	IV.11.13	IV.12.07	III.3.19	III.13.05	III.14.03	III.16.03	III.18.09	-0.7482	7	3	0		
A134	0-2073	I.5.03	II.1.15	III.3.02	III.4.01	III.6.01	III.9.11	III.15.23	IV.11.04	IV.12.04	III.3.19	III.13.09	III.14.04	III.16.03	III.18.20	-0.7410	6	3	1		
A091	AeD990SW- 36-9-10	I.5.02	II.1.02	III.3.03	III.4.03	III.6.10	III.9.11	III.15.12	IV.11.13	IV.12.07	III.3.04	III.13.08	III.14.11	III.16.07	III.18.09	-0.7396	7	2	0		
A017	AeA95-306.NZ.13.2	I.5.03	II.1.15	III.3.03	III.4.03	III.6.10	III.9.11	III.15.01	IV.11.13	IV.12.07	III.3.07	III.13.05	III.14.03	III.16.07	III.18.05	-0.7171	4	4	4		
MR20	RIL 847-53	I.5.03	NA	III.3.03	III.4.03	III.6.09	III.9.11	III.15.13	IV.11.04	IV.12.04	III.3.19	III.13.09	III.14.03	III.16.07	III.18.26	-0.5955	5	3	2		
A131	L2782.1	NA	NA	NA	III.4.03	NA	III.9.08	NA	IV.11.09	NA	III.3.19	III.13.07	III.14.04	III.16.07	III.18.19	-0.5906	0	0	9		
MR14	RIL 847-08	I.5.03	NA	III.3.03	III.4.03	III.6.10	III.9.11	III.15.02	IV.11.12	IV.12.07	III.3.19	III.13.17	III.14.03	III.16.07	III.18.24	-0.5879	5	3	1		
A034	AeB97-724.2.3	I.5.14	II.1.04	III.3.02	III.4.01	III.6.01	III.9.04	III.15.03	IV.11.05	IV.12.03	III.3.19	III.13.05	III.14.03	III.16.07	III.18.05	-0.5750	9	1	0		
A099	AeD990SW- 44-19-1	I.5.08	II.1.17	III.3.02	III.4.01	III.6.01	III.9.04	III.15.22	NA	NA	III.3.19	III.13.11	III.14.14	III.16.08	III.18.09	-0.5653	6	3	2		
MR17	RIL 847-45	I.5.02	NA	III.3.03	III.4.03	III.6.08	III.9.11	III.15.02	IV.11.04	IV.12.04	III.3.19	III.13.18	III.14.03	III.16.07	III.18.24	-0.5302	2	2	1		
MR07	BASS5.52	I.5.03	NA	III.3.02	III.4.01	III.6.01	III.9.04	III.15.01	IV.11.04	IV.12.04	III.3.07	III.13.17	III.14.11	III.16.07	III.18.24	-0.5172	6	1	1		
A140	GSP1023	I.5.03	II.1.04	III.3.02	III.4.01	III.6.01	III.9.03	III.15.02	IV.11.04	IV.12.04	III.3.19	III.13.17	III.14.11	III.16.07	III.18.21	-0.5130	9	2	1		
MR11	BASS5.233	I.5.01	NA	III.3.07	III.4.02	III.6.09	III.9.11	III.15.12	IV.11.04	IV.12.02	III.3.07	III.13.17	III.14.06	III.16.02	III.18.24	-0.4866	5	5	1		
A070	AeD990SW- 10-10-6	I.5.03	II.1.15	III.3.02	III.4.01	III.6.01	III.9.03	III.15.17	IV.11.04	IV.12.04	III.3.11	III.13.05	III.14.03	III.16.03	III.18.09	-0.4399	2	2	0		
A014	45	I.5.03	II.1.15	III.3.02	III.4.01	III.6.01	III.9.06	III.15.02	IV.11.04	IV.12.04	III.3.19	III.13.05	III.14.07	III.16.07	III.18.05	-0.4341	9	2	0		
MR16	RIL 847-28	I.5.03	NA	III.3.03	III.4.03	III.6.10	III.9.11	III.15.02	IV.11.04	IV.12.04	III.3.19	III.13.17	III.14.03	III.16.03	III.18.24	-0.4244	5	3	1		
A021	AeA95-321.M.33.3	I.5.14	II.1.02	III.3.02	III.4.03	III.6.02	III.9.05	III.15.17	IV.11.04	IV.12.04	III.3.04	III.13.05	III.14.04	III.16.01	III.18.05	-0.2708	1	1	0		
A023	AeA95-326.M.7.1.1	I.5.14	II.1.02	III.3.02	III.4.03	III.6.02	III.9.05	III.15.22	IV.11.04	IV.12.04	III.3.04	III.13.05	III.14.14	III.16.08	III.18.05	-0.2619	6	1	0		
A132	USA	I.5.17	NA	III.3.05	III.4.02	III.6.04	III.9.12	III.15.23	IV.11.10	IV.12.04	III.3.04	III.13.05	III.14.04	III.16.01	III.18.19	-0.2473	2	1	2		
A032	AeB97-722.14.1	I.5.14	NA	III.3.02	III.4.01	III.6.01	III.9.04	III.15.01	IV.11.05	IV.12.04	III.3.04	III.13.05	III.14.03	III.16.01	III.18.05	-0.2401	0	0	6		
MR12	BASS5.231	I.5.03	NA	III.3.02	III.4.01	III.6.01	III.9.04	III.15.07	IV.11.05	IV.12.04	III.3.07	III.13.17	III.14.08	III.16.07	III.18.24	-0.2394	7	1	1		
A094	AeD990SW- 39-5-5	I.5.17	II.1.06	III.3.03	III.4.03	III.6.10	III.9.11	III.15.18	IV.11.04	IV.12.04	III.3.19	III.13.08	III.14.11	III.16.07	III.18.09	-0.1676	6	2	0		
MR10	BASS5.171	I.5.03	NA	III.3.02	III.4.01	III.6.01	III.9.04	III.15.01	IV.11.04	IV.12.04	III.3.19	III.13.17	III.14.11	III.16.07	III.18.24	-0.1452	9	2	0		
A073	AeD990SW- 13-9-1	I.5.14	II.1.02	III.3.03	III.4.03	III.6.01	III.9.10	III.15.22	IV.11.05	IV.12.03	III.3.18	III.13.07	III.14.03	III.16.03	III.18.09	-0.1157	5	0	0		
MR02	BAP8.43	I.5.17	NA	III.3.05	III.4.02	III.6.04	III.9.12	III.15.01	IV.11.13	IV.12.07	III.3.04	III.13.17	III.14.14	III.16.08	III.18.24	-0.0660	2	5	1		
A133	ZP141	I.5.22	II.1.11	III.3.01	III.4.02	III.6.01	III.9.04	III.15.04	IV.11.07	IV.12.02	III.3.01	III.13.01	III.14.12	III.16.04	III.18.20	-0.0472	0	1	5		
A146	W/8905	I.5.03	II.1.04	III.3.02	III.4.01	III.6.01	III.9.09	III.15.02	IV.11.07	IV.12.02	III.3.19	III.13.09	III.14.11	III.16.07	III.18.21	-0.0433	8	2	1		
A015	AeA95-299.M.32.1	I.5.02	II.1.15	III.3.07	III.4.02	III.6.09	III.9.11	III.15.01	IV.11.05	IV.12.03	III.3.07	III.13.05	III.14.03	III.16.07	III.18.05	0.0145	5	5	0		
A033	AeB97-723.3.2	I.5.14	II.1.04	III.3.02	III.4.01	III.6.01	III.9.04	III.15.03	IV.11.05	IV.12.03	III.3.19	III.13.05	III.14.03	III.16.07	III.18.05	0.0430	9	1	0		
A078	AeA95-315.M.43.5	NA	II.1.02	III.3.02	III.4.01	III.6.02	III.9.05	III.15.17	IV.11.04	IV.12.04	III.3.03	III.13.05	III.14.04	III.16.01	III.18.05	0.0562	4	1	2		
A136	E186	I.5.09	NA	III.3.02	III.4.01	III.6.09	III.9.14	III.15.20	IV.11.02	IV.12.06	III.3.09	III.14.07	III.16.07	III.18.20	0.0620	3	1	2			
A137	ETHIOPIA-32 CGN 3255	I.5.07	NA	III.3.06	III.4.03	III.6.03	III.9.09	III.15.14	IV.11.07	IV.12.02	III.3.17	III.14.11	III.16.04	III.18.21	0.0722	2	2	2			
A130	E11	I.5.21	NA	III.3.03	III.4.01	III.6.01	III.9.04	III.15.05	IV.11.13	IV.12.02	III.3.03	III.13.07	III.14.03	III.16.03	III.18.21	0.1323	1	0	10		
A138	FRIGLIO	I.5.03	II.1.08	III.3.01	III.4.02	III.6.03	III.9.11	III.15.06	IV.11.04	IV.12.04	III.3.13	III.13.08	III.14.02	III.16.03	III.18.21	0.1850	3	3	1		
A002	0-2100	I.5.03	II.1.07	III.3.07	III.4.02	III.6.03	III.9.03	III.15.02	IV.11.03	IV.12.04	III.3.19	III.13.02	III.14.03	III.16.03	III.18.21	0.1959	4	3	1		
A074	AeD990SW- 14-8-8	I.5.03	II.1.17	III.3.02	III.4.03	III.6.02	III.9.05	III.15.17	IV.11.13	IV.12.07	III.3.03	III.13.08	III.14.03	III.16.03	III.18.09	0.2304	5	2	0		
A071	AeD990SW- 11-7-4	I.5.03	II.1.13	III.3.02	III.4.01	III.6.01	III.9.10	III.15.11	IV.11.04	IV.12.04	III.3.19	III.13.05	III.14.04	III.16.01	III.18.09	0.2399	7	1	0		
A139	GAT1259	I.5.04	II.1.03	III.3.03	III.4.02	III.6.05	III.9.15	III.15.15	IV.11.01	IV.12.02	III.3.15	III.13.15	III.14.11	III.16.02	III.18.21	0.2460	2	2	1		
MR09	BASS5.161	I.5.03	NA	III.3.02	III.4.01	III.6.01	III.9.04	III.15.12	IV.11.04	IV.12.04	III.3.19	III.13.17	III.14.11	III.16.07	III.18.24	0.2836	9	2	1		
MR04	BAP8.184	NA	NA	NA	III.4.02	NA	III.9.11	NA	NA	NA	III.3.19	III.13.17	III.14.14	III.16.08	III.18.24	0.2988	0	3	10		
A088	AeD990SW- 33-7-5	I.5.14	II.1.13	III.3.02	III.4.03	III.6.02	III.9.05	III.15.22	IV.11.13	IV.12.07	III.3.19	III.13.08	III.14.14	III.16.08	III.18.09	0.3124	4	4	0		
MR03	BAP8.128	I.5.17	NA	III.3.05	III.4.02	III.6.04	III.9.11	III.15.01	IV.11.13	IV.12.07	III.3.07	III.13.17	III.14.01	III.16.02	III.18.24	0.3876	3	4	1		
MR01	BAP8.10	I.5.11	NA	III.3.05	III.4.02	III.6.09	III.9.11	III.15.01	IV.11.13	IV.12.07	III.3.04	III.13.08	III.14.01	III.16.08	III.18.23	0.4009	2	6	3		

Line code	Name	Haplotype composition of the pea-Aphanomyces collection lines at consistent LD-blocks (a)																MFA.Dim.1 (b)	Number of favorable haplotypes (c)	Number of unfavorable haplotypes (c)	Number of missing haplotypes (d)
		I.5	II.1	III.3	III.4	III.6	III.9	III.15	III.15	IV.11	IV.12	III.3	III.3	III.13	III.14	III.16	III.18				
A135	100	I.5.03	II.1.19	III.3.01	III.4.02	III.6.06	III.9.09	III.15.09	III.15.09	IV.11.05	IV.12.04	III.3.19	III.3.19	III.14.02	III.16.03	III.18.20	0.4232	3	3	1	
MR05	BA55.8	I.5.11	NA	III.3.02	III.4.01	III.6.01	III.9.04	III.15.01	III.15.01	IV.11.05	IV.12.03	III.3.19	III.3.19	III.14.08	III.16.07	III.18.24	0.4499	6	3	1	
A030	AeB97-710.10.5	I.5.14	II.1.13	III.3.02	III.4.03	III.6.02	III.9.05	III.15.17	III.15.17	IV.11.05	IV.12.03	III.3.03	III.3.03	III.14.04	III.16.01	III.18.05	0.5501	3	1	0	
A065	AeD9905W-5-9-2	I.5.03	II.1.04	III.3.02	III.4.01	III.6.01	III.9.10	III.15.21	III.15.21	IV.11.13	IV.12.07	III.3.04	III.3.04	III.14.03	III.16.03	III.18.09	0.5813	8	0	0	
A076	AeD9905W-16-1-10	I.5.14	II.1.02	III.3.02	III.4.01	III.6.01	III.9.14	III.15.22	III.15.22	IV.11.13	IV.12.07	III.3.04	III.3.04	III.14.04	III.16.01	III.18.09	0.7230	7	1	0	
A004	O-2042	I.5.03	II.1.15	III.3.04	III.4.01	III.6.01	III.9.14	III.15.02	III.15.02	IV.11.04	IV.12.04	III.3.19	III.3.19	III.14.03	III.16.03	NA	0.7236	2	2	1	
A043	AeB97-745.2.1	I.5.11	II.1.15	III.3.02	III.4.01	III.6.01	III.9.06	III.15.02	III.15.02	IV.11.04	IV.12.04	III.3.07	III.3.07	III.14.07	III.16.07	III.18.08	0.7513	7	2	0	
A029	AeB97-709.11.1	I.5.14	II.1.13	III.3.02	III.4.03	III.6.02	III.9.05	III.15.17	III.15.17	IV.11.05	IV.12.04	III.3.03	III.3.03	III.14.04	III.16.01	III.18.05	0.7851	2	1	0	
A093	AeD9905W-38-5-4	I.5.03	II.1.06	III.3.03	III.4.03	III.6.01	III.9.10	III.15.18	III.15.18	IV.11.04	IV.12.03	III.3.19	III.3.19	III.14.13	III.16.09	III.18.09	0.7863	4	3	0	
A066	AeD9905W-6-2-3	I.5.03	II.1.17	III.3.02	III.4.01	III.6.01	III.9.10	III.15.13	III.15.13	IV.11.04	IV.12.04	III.3.19	III.3.19	III.14.11	III.16.07	III.18.09	0.7877	9	1	0	
A062	AeD9905W-2-7-8	I.5.03	II.1.17	III.3.02	III.4.01	III.6.01	III.9.10	III.15.13	III.15.13	IV.11.04	IV.12.03	III.3.19	III.3.19	III.14.11	III.16.07	III.18.09	0.8484	8	1	0	
A090	AeD9905W-35-9-2	I.5.17	II.1.13	III.3.03	III.4.03	III.6.10	III.9.11	III.15.12	III.15.12	IV.11.13	IV.12.07	III.3.19	III.3.19	III.14.11	III.16.07	III.18.09	0.8767	4	3	0	
A037	AeB97-730.10.2	I.5.14	II.1.14	III.3.03	III.4.03	III.6.10	III.9.11	III.15.01	III.15.01	IV.11.04	IV.12.04	III.3.19	III.3.19	III.14.03	III.16.03	III.18.05	0.9076	2	2	0	
A085	AeD9905W-28-8-4	I.5.14	II.1.02	III.3.03	III.4.03	III.6.10	III.9.11	III.15.13	III.15.13	IV.11.13	IV.12.07	III.3.04	III.3.04	III.14.14	III.16.08	III.18.09	0.9082	4	4	0	
A069	AeD9905W-9-9-2	I.5.14	II.1.05	III.3.02	III.4.01	III.6.01	III.9.10	III.15.03	III.15.03	IV.11.05	IV.12.04	III.3.03	III.3.03	III.14.03	III.16.03	III.18.09	0.9540	7	0	0	
A115	DSP	I.5.03	II.1.14	III.3.03	III.4.03	III.6.08	III.9.09	III.15.02	III.15.02	IV.11.04	IV.12.04	III.3.19	III.3.19	III.13.14	III.16.03	III.18.05	0.9566	5	1	1	
A112	AeD9905W-59-7-4	I.5.02	II.1.16	III.3.03	III.4.03	III.6.10	III.9.11	III.15.18	III.15.18	IV.11.13	IV.12.07	III.3.19	III.3.19	III.13.12	III.16.08	III.18.12	1.0138	2	4	1	
A086	AeD9905W-29-3-3	I.5.14	II.1.13	III.3.03	III.4.03	III.6.10	III.9.11	III.15.22	III.15.22	IV.11.13	IV.12.07	III.3.19	III.3.19	III.14.03	III.16.03	III.18.09	1.0151	3	3	0	
A087	AeD9905W-30-4-4	I.5.17	II.1.13	III.3.03	III.4.03	III.6.10	III.9.11	III.15.13	III.15.13	IV.11.13	IV.12.07	III.3.19	III.3.19	III.14.03	III.16.03	III.18.09	1.1023	3	3	0	
A122	Pawnee	I.5.19	II.1.14	III.3.01	III.4.02	III.6.10	III.9.11	III.15.09	III.15.09	IV.11.04	IV.12.04	III.3.05	III.3.05	III.14.09	III.16.07	III.18.15	1.1231	3	2	1	
A010	AeB97037.20	I.5.14	II.1.14	III.3.03	III.4.03	III.6.10	III.9.11	III.15.01	III.15.01	IV.11.04	IV.12.04	III.3.19	III.3.19	III.13.05	III.16.03	III.18.05	1.1503	2	2	1	
A036	AeB97-729.2.2	I.5.02	II.1.17	NA	III.4.03	III.6.10	III.9.11	III.15.01	III.15.01	IV.11.04	IV.12.04	III.3.05	III.3.05	III.14.04	III.16.01	III.18.05	1.1746	1	1	1	
A067	AeD9905W-7-6-7	I.5.03	II.1.17	III.3.02	III.4.01	III.6.01	III.9.10	III.15.17	III.15.17	IV.11.04	IV.12.04	III.3.19	III.3.19	III.14.03	III.16.03	III.18.09	1.2098	2	2	0	
A068	AeD9905W-8-8-1	I.5.14	II.1.13	III.3.02	III.4.01	III.6.01	III.9.03	III.15.11	III.15.11	IV.11.04	IV.12.04	III.3.19	III.3.19	III.14.04	III.16.01	III.18.09	1.2121	1	1	0	
A026	AeA95-343.M.61.5	I.5.18	II.1.17	III.3.07	III.4.02	III.6.09	III.9.05	III.15.21	III.15.21	IV.11.04	IV.12.04	III.3.19	III.3.19	III.14.03	III.16.03	III.18.05	1.2314	4	4	1	
A082	AeD9905W-25-9-4	I.5.14	II.1.17	III.3.02	III.4.03	III.6.02	III.9.05	III.15.17	III.15.17	IV.11.13	IV.12.07	III.3.19	III.3.19	III.13.08	III.16.07	III.18.09	1.2789	5	3	1	
A143	NEPAL A	I.5.10	NA	III.3.02	III.4.01	III.6.01	III.9.14	III.15.19	III.15.19	IV.11.11	IV.12.07	III.3.14	III.3.14	III.14.08	III.16.05	III.18.21	1.2795	3	1	3	
A141	LOCAL	NA	NA	III.4.01	III.6.08	III.9.07	III.9.07	III.15.01	III.15.01	IV.11.04	IV.12.04	III.3.05	III.3.05	III.14.04	III.16.01	III.18.21	1.3031	1	1	11	
A027	AeA95-359.2002.1.4	I.5.14	II.1.04	III.3.07	III.4.02	III.6.09	III.9.07	III.15.01	III.15.01	IV.11.04	IV.12.04	III.3.05	III.3.05	III.14.04	III.16.01	III.18.05	1.3414	3	3	0	
A142	MN313	I.5.03	II.1.19	III.3.03	III.4.03	III.6.08	III.9.11	III.15.11	III.15.11	IV.11.05	IV.12.03	III.3.19	III.3.19	III.14.04	III.16.06	III.18.21	1.4291	2	3	1	
A114	CAPELLA	I.5.14	II.1.13	III.3.02	III.4.03	III.6.02	III.9.05	III.15.17	III.15.17	IV.11.04	IV.12.04	III.3.03	III.3.03	III.14.04	III.16.01	III.18.13	1.4570	3	1	0	
A060	AeD9905W-24-3-2	I.5.03	II.1.15	III.3.06	III.4.02	III.6.06	III.9.07	III.15.01	III.15.01	IV.11.05	IV.12.03	III.3.11	III.3.11	III.13.05	III.16.03	III.18.09	1.5106	3	3	1	
A147	886/01	I.5.19	II.1.09	III.3.07	III.4.01	III.6.06	III.9.09	III.15.01	III.15.01	IV.11.05	IV.12.03	III.3.10	III.3.10	III.14.09	III.16.07	III.18.21	1.5418	3	3	1	
A063	AeD9905W-3-8-7	I.5.14	II.1.04	III.3.02	III.4.03	III.6.02	III.9.05	III.15.17	III.15.17	IV.11.04	IV.12.04	III.3.11	III.3.11	III.13.05	III.16.07	III.18.09	1.6164	5	1	1	
A144	NO. 9845	I.5.07	NA	III.3.07	III.4.02	III.6.03	III.9.09	III.15.23	III.15.23	IV.11.07	IV.12.02	III.3.16	III.3.16	III.14.05	III.16.04	III.18.21	1.6170	1	3	2	
A084	AeD9905W-27-7-6	I.5.03	II.1.19	III.3.02	III.4.01	III.6.01	III.9.10	III.15.17	III.15.17	IV.11.13	IV.12.07	III.3.19	III.3.19	III.13.08	III.16.08	III.18.09	1.6641	6	5	0	
A126	PUGET	I.5.03	II.1.15	III.3.06	III.4.03	III.6.09	III.9.11	III.15.02	III.15.02	IV.11.04	IV.12.04	III.3.19	III.3.19	III.14.03	III.16.03	III.18.18	1.7484	5	4	1	
A011	Mh314	I.5.03	II.1.19	III.3.02	III.4.01	III.6.01	III.9.10	III.15.13	III.15.13	IV.11.05	IV.12.03	III.3.19	III.3.19	III.13.05	III.16.03	III.18.02	1.8085	6	1	1	
A120	VICTOR	I.5.05	II.1.04	III.3.07	III.4.02	III.6.06	III.9.09	III.15.01	III.15.01	IV.11.01	IV.12.02	III.3.05	III.3.05	III.14.04	III.16.01	III.18.02	1.8750	3	3	1	
A038	AeB97-736.3.10	I.5.11	II.1.15	III.3.07	III.4.02	III.6.09	III.9.11	III.15.18	III.15.18	IV.11.04	IV.12.04	III.3.19	III.3.19	III.13.05	III.16.02	III.18.06	1.9399	2	7	0	
A057	AeD9905W-04-21-9-1	I.5.03	II.1.04	III.3.02	III.4.01	III.6.01	III.9.11	III.15.13	III.15.13	IV.11.13	IV.12.07	III.3.19	III.3.19	III.14.03	III.16.03	III.18.09	1.9896	7	3	0	
A001	90-2079	I.5.03	II.1.15	III.3.03	III.4.03	III.6.08	III.9.11	III.15.13	III.15.13	IV.11.05	IV.12.03	III.3.19	III.3.19	III.13.01	III.16.02	III.18.05	2.1220	3	3	1	
A061	AeD9905W-1-10-8	I.5.14	II.1.13	III.3.01	III.4.02	III.6.06	III.9.07	III.15.03	III.15.03	IV.11.05	IV.12.04	III.3.03	III.3.03	III.14.10	III.16.01	III.18.09	2.1523	2	0	0	
A064	AeD9905W-4-6-4	I.5.14	II.1.13	III.3.06	III.4.02	III.6.06	III.9.10	III.15.17	III.15.17	IV.11.04	IV.12.04	III.3.03	III.3.03	III.13.05	III.16.07	III.18.09	2.1875	4	3	0	
A118	Enduro	I.5.15	II.1.04	III.3.07	III.4.02	III.6.06	III.9.09	III.15.01	III.15.01	IV.11.04	IV.12.04	III.3.05	III.3.05	III.14.04	III.16.01	III.18.02	2.2039	4	3	0	
A012	P269802	I.5.24	II.1.10	III.3.07	III.4.02	III.6.03	III.9.05	III.15.23	III.15.23	IV.11.04	IV.12.04	III.3.05	III.3.05	III.13.15	III.16.04	III.18.03	2.2144	2	2	1	
A016	AeA95-304.M.4.4.5	I.5.11	II.1.15	III.3.07	III.4.02	III.6.09	III.9.11	III.15.13	III.15.13	IV.11.13	IV.12.07	III.3.07	III.3.07	III.13.05	III.16.02	III.18.05	2.2329	2	7	0	
A044	AeB97-747.3.7	I.5.12	II.1.15	III.3.03	III.4.03	III.6.10	III.9.02	III.15.13	III.15.13	IV.11.05	IV.12.04	III.3.19	III.3.19	III.14.01	III.16.02	III.18.21	2.3129	0	3	3	
A152	Belote	I.5.16	II.1.14	III.3.07	III.4.02	III.6.06	III.9.11	III.15.01	III.15.01	IV.11.08	IV.12.05	III.3.08	III.3.08	III.14.01	III.16.07	III.18.21	2.3556	1	5	1	
A006	MISTRAL	I.5.04	II.1.08	III.3.01	III.4.02	III.6.06	III.9.11	III.15.16	III.15.16	IV.11.01	IV.12.02	III.3.06	III.3.06	III.13.03	III.16.03	III.18.01	2.3759	3	3	0	
A148	DOVE HR	I.5.03	NA	NA	III.4.01	III.6.09	III.9.05	III.15.13	III.15.13	IV.11.05	IV.12.03	III.3.06	III.3.06	III.14.03	III.16.02	III.18.21	2.6955	3	1	10	
A075	AeD9905W-15-7-4	I.5.17	II.1.17	III.3.03	III.4.03	III.6.10	III.9.11	III.15.17	III.15.17	IV.11.13	IV.12.07	III.3.03	III.3.03	III.14.14	III.16.08	III.18.09	2.7435	2	5	0	
A123	ISARD	I.5.15	II.1.15	III.3.07	III.4.02	III.6.06	III.9.11	III.15.09	III.15.09	IV.11.01	IV.12.02	III.3.06	III.3.06	III.14.02	III.16.03	III.18.16	2.7530	1	5	1	
A150	Amical	I.5.11	II.1.15																		

Line code	Name	Haplotype composition of the pea-Aphanomyces collection lines at consistent LD-blocks (a)																MFA.Dim.1 (b)	Number of favorable haplotypes (c)	Number of unfavorable haplotypes (c)	Number of missing haplotypes (d)
		I.5	II.1	III.3	III.4	III.6	III.9	III.15	IV.11	IV.12	VII.3	VII.13	VII.14	VII.16	VII.18						
A149	<i>Alixor</i>	I.5.11	II.1.15	III.3.07	III.4.02	III.6.09	III.9.11	III.15.01	IV.11.07	IV.12.02	VII.3.05	VII.14.01	VII.16.02	VII.18.21	2.8012	1	7	1			
A007	<i>RUSTIC</i>	I.5.14	II.1.04	NA	III.4.02	NA	III.9.07	III.15.01	IV.11.04	IV.12.04	VII.3.11	VII.13.03	VII.14.03	NA	2.8227	3	1	3			
A059	<i>AeD99QU-04-23-8-1</i>	I.5.14	II.1.17	III.3.06	III.4.02	III.6.09	III.9.11	III.15.01	IV.11.05	IV.12.03	VII.3.11	VII.13.05	VII.14.03	VII.16.03	2.8348	3	3	0			
A155	<i>Hardy</i>	I.5.11	II.1.15	III.3.07	NA	III.6.09	III.9.11	III.15.01	IV.11.07	IV.12.02	VII.3.07	VII.14.01	VII.16.02	VII.18.22	3.0430	1	5	2			
A058	<i>AeD99QU-04-22-10-3</i>	I.5.13	II.1.13	III.3.02	III.4.01	III.6.01	III.9.10	III.15.17	IV.11.05	IV.12.03	VII.3.19	VII.13.05	VII.14.01	VII.16.02	3.0705	5	2	0			
A121	<i>Astrid</i>	I.5.23	II.1.15	III.3.07	III.4.02	III.6.06	III.9.11	III.15.01	IV.11.07	IV.12.02	VII.3.02	NA	VII.14.01	VII.16.02	3.1187	1	5	1			
A125	<i>BACCARA</i>	I.5.11	II.1.15	III.3.07	III.4.02	III.6.09	III.9.11	III.15.01	IV.11.05	IV.12.03	VII.3.07	NA	VII.14.01	VII.18.17	3.1473	1	6	2			
A151	<i>Antares</i>	I.5.14	NA	III.3.07	III.4.02	III.6.06	III.9.11	III.15.01	IV.11.07	IV.12.04	VII.3.11	NA	VII.14.01	VII.16.07	3.1755	2	5	1			
A005	<i>Odeon</i>	I.5.11	II.1.15	III.3.07	III.4.02	III.6.06	III.9.11	III.15.01	IV.11.05	IV.12.03	VII.3.02	VII.13.03	VII.14.01	VII.16.02	3.2349	1	6	0			
A124	<i>EDEN</i>	I.5.11	II.1.15	III.3.07	III.4.02	III.6.09	III.9.11	III.15.01	IV.11.05	IV.12.03	VII.3.07	NA	VII.14.01	VII.16.02	3.3301	1	6	2			
A119	<i>Lumina</i>	I.5.11	II.1.14	III.3.07	III.4.02	III.6.09	III.9.11	III.15.01	IV.11.07	IV.12.04	VII.3.07	VII.13.15	VII.14.01	VII.16.02	3.6286	1	5	1			
A153	<i>Margnan</i>	NA	NA	III.3.07	NA	III.6.09	III.9.11	NA	IV.11.05	IV.12.03	VII.3.07	NA	VII.14.01	VII.16.02	3.8938	1	4	5			
A154	<i>Safranios</i>	I.5.11	II.1.07	III.3.07	III.4.02	III.6.09	III.9.13	III.15.01	IV.11.05	IV.12.03	VII.3.07	NA	VII.14.01	VII.16.07	4.4719	2	5	1			

Additional File II.15

Annotation of markers in detected linkage disequilibrium (LD) blocks for putative candidate genes

(a) LD block were constructed with markers detected by genome-wide association mapping and markers in LD with them, LD block are named according to Additional File 11; (b) genetic positions are estimated as described in Additional File 11; For each LD block:(c) traits associated with markers in the LD block; (d) Pairwise LD between each marker and the marker(s) detected by GWA in the LD block (detected marker in bold); (e) putative protein annotation underlying each marker based on comparative Blast X data from Tayeh et al [49], see Methods section; (f) putative function of the protein in plant disease resistance based on the literature; (g) putative function of the protein in plant development based on the literature; (h) reference for protein function involved in plant disease resistance or plant development; (i) Pea cloned genes referred in literature; (j) extract of annotation data from Tayeh et al [49], Blast X of marker sequence onto *P. sativum*, *M. truncatula*, *G. max* and *A. thaliana* proteins.

LG	LD block (a)	Marker	Position (cM) (b)	Associated traits (c)	LD (r ²) with trait-associated marker (bold) (d)	Putative annotation (e)	Putative function of interest for disease resistance (f)	Putative function in plant development (g)		
I1	I1.1	PsCam045699_29329_409	38.5		0.29	not assigned	-	-		
		PsCam011022_7464_1139	38.7		0.27	UDP-glycosyltransferase	defense response (detoxification)	-		
		PsCam018890_11268_413	38.7		0.33	not assigned	-	-		
		PsCam020122_11475_329	38.7		0.33	not assigned	-	-		
		PsCam000395_347_213	38.7		0.46	Quinone reductase family protein	plant defense (oxydative burst)	-	-	
		PsCam047171_30305_572	38.7	Field_AD11_10_RI	1.00	not assigned	-	-		
		PsCam049583_32210_346	38.7		0.20	RNA binding domain protein	-	-		
		PsCam021287_11947_281	39		0.61	Thioesterase superfamily protein	-	-		
		AB56_379	73.85	Field_AD11_13_RI	1.00	not assigned	-	-		
		PsCam037161_22265_149	74.3		0.23	putative copper ion-binding/laccase	defense response (lignification of xylem tissue)	-	-	
		I3	I3.3	PsCam057507_38094_565	82.4		0.31	BTB/POZ domain-containing protein	signal transduction	-
				PsCam050808_33369_647	82.7		0.42	Lateral Organ Boundaries (LOB) domain-containing protein	signal transduction (Jasmonic Acid regulation)	-
				PsCam031033_18302_4320	83.1		0.85	Calcium-transporting ATPase 2	signal transduction	-
				PsCam037746_22815_1195	83.1		0.90	not assigned	-	-
				PsCam045028_28759_2059	83.1	Field_AD12_10_DI	1.00	Formin-like protein	defense response (oxydative burst)	-
PsCam042662_26712_871	84.3				0.81	U-box domain containing protein (E3 ubiquitin ligases) = Armadillo	signal transduction	-		
PsCam001766_1451_596	84.3				0.76	not assigned	-	-		
PsCam060131_40118_1237	84.3				0.76	BSD domain containing protein	-	-		
Af	85.6			Field_RRI_12_DI	1.00	not assigned	-	-		
PsCam025559_14651_1169	86.5				0.21	not assigned	-	-		
PsCam005297_4017_1128	87.2				0.20	not assigned	-	-		
PsCam005033_3817_61	88.2				0.83	NO APICAL MERISTEM (NAC)-like	transcription regulation	-		
PsCam027153_15841_364	88.2			Field_AD11_12_DI	1.00	syntaxin related protein (vesicule fusion)	defense response (secretion of pathogenesis related proteins)	-		
PsCam042815_26857_431	88.2				0.91	disagreeing hits	-	-		
PsCam037259_22356_1396	88.9				0.89	not assigned	-	-		
PsCam024866_14152_261	88.9		0.42	not assigned	-	-				
PsCam011366_7739_322	89.1		0.27	not assigned	-	-				
PsCam038113_23162_1036	89.1		0.39	not assigned	-	-				
PsCam044835_28587_1878	89.1	Field_AD11_All Field_AD11_13_DI Field_AD11_12_RI Field_AD12_All MFA.Dim.2	1.00	serine-threonine protein kinase, putative LRR domain	pathogen recognition/signal transduction	-				
PsCam044855_28602_1561	89.1		0.55	ACC synthase (Ethylene production)	signal transduction	-				
PsCam059295_39494_4455	89.1		0.28	Protein FAM63-like	-	-				
PsCam035316_20506_2839	89.4		0.29	RNA polymerase II largest subunit	-	-				
PsCam035496_20678_3576	89.9		0.51	RNA polymerase beta subunit (chloroplast)	-	-				
PsCam000507_444_666	90.4		0.71	disagreeing hits	-	-				
PsCam000400_351_2330	90.4		0.39	not assigned	-	-				

LG	LD block (a)	Marker	Position (cM) (b)	Associated traits (c)	LD (r^2) with trait-associated marker (bold) (d)	Putative annotation (e)	Putative function of interest for disease resistance (f)	Putative function in plant development (g)	
II	II.1	PsCam039671_24572_710	41		0.26	not assigned	-	-	
		PsCam039893_24741_949	41.7		0.55	not assigned	-	-	
		PsCam036593_21726_2984	41.8		0.50	cyclic nucleotide-gated channel 15	signal transduction	-	
		PsCam004991_3782_3995	42		0.52	not assigned	-	-	
		PsCam039641_24544_1372	42		0.54	thylakoid lumen 15.0 kDa protein (pentapeptide repeat)	-	-	
		PsCam021071_11806_2143	42.1		0.25	adenosine kinase 2-like	-	-	
		PsCam033694_19257_1741	42.1	Field_AD11_12_DI	1.00	not assigned	-	-	
				Field_AD12_11_RI					
				Field_AD11_All					
				Field_AD12_All					
				PsCam043529_27529_1349	42.1		0.55	not assigned	-
				PsCam057510_38096_1169	42.1		0.50	not assigned	-
		AB112_373	42.3		0.20	not assigned	-		
		AB112_402	42.3		0.56	not assigned	-		
		PsCam011257_7650_1194	42.7	FLO1_12_REN	1.00	bifunctional monodehydroascorbate reductase and carbonic anhydrase	-		
		PsCam039635_24538_1229	42.7		0.52	nectarin-3-like	-		
		PsCam050314_32904_1465	42.7		0.45	ribosomal protein	-		
II.2		PsCam028967_17212_242	45.2		0.20	not assigned	-	-	
		PsCam039187_24156_997	45.2		0.85	not assigned	-	-	
		PsCam059543_39701_683	45.2	FLO1_12	1.00	not assigned	-	-	
		PsCam029325_17487_1726	45.5		0.79	urease accessory protein D	-	-	
		PsCam049207_31848_271	45.5		0.91	Transmembrane amino acid transporter family protein	-	-	
		PsCam050243_32836_1352	45.5		0.81	not assigned	-	-	
		PsCam050446_33026_139	45.5		1.00	not assigned	-	-	
		PsCam059731_39835_288	45.5		1.00	not assigned	-	-	
		AB33_388	45.9	FLO1_10_REN	1.00	not assigned	-	-	
		PsCam049388_32024_1492	46.2		0.29	Serine carboxypeptidase	-	-	
		PsCam010958_7412_642	46.3		0.70	disagreeing hits	-	-	
		PsCam035387_20575_5015	46.3		0.39	ABC transporter G family member	-	plant nutrition, plant development	
		PsCam035741_20914_1106	46.3	FLO1_12_REN	1.00	exocyst complex component sec15A	-		
		PsCam043401_27414_292	46.3		0.78	disagreeing hits	-		
		PsCam050896_33451_1641	46.3		0.62	disagreeing hits	-		
		PsCam039226_24191_762	46.3		0.27	not assigned	-		
II.4		PsCam034783_20145_774	46.9		0.28	not assigned	-	-	
		PsCam037747_22816_1252	47		0.25	not assigned	-	-	
		PsCam042382_26450_3953	47.4	FLO1_10_REN	1.00	not assigned	-	-	
		PsCam056692_37460_234	47.7		0.70	not assigned	-	-	
		PsCam053665_35505_1746	67.6		0.23	paladin-like	-	-	
		PsCam036058_21208_255	68.1		0.32	not assigned	-	-	
		PsCam043114_27143_892	68.1	FLO2_10_REN	1.00	Pectinacetyltransferase family protein	-	-	
		PsCam045737_29360_2498	68.1		0.20	DHHC-type zinc finger family protein	-	shoot branching	
		PsCam038536_23560_429	101.8		0.48	not assigned	-	-	
		PsCam038159_23208_312	102.2		0.50	origin recognition complex subunit 4-like	-	-	
		PsCam038614_23632_792	102.2		0.44	fatty acyl-ACP thioesterases B	-	-	

LG	LD block (a)	Marker	Position (cM) (b)	Associated traits (c)	LD (r ²) with trait-associated marker (bold) (d)	Putative annotation (e)	Putative function of interest for disease resistance (f)	Putative function in plant development (g)		
II.7		PsCam043246_27270_1517	102.2		0.37	sec-independent protein translocase protein TATC	-	-		
		PsCam000359_318_1468	103.3		0.21	amino-acid permease	-	flower development		
		PsCam009577_6385_812	103.3	RIPE_10_REN	1.00	not assigned	-	-		
		PsCam044482_28348_1707	103.7		0.71	not assigned	-	-		
		PsCam010896_7359_1404	112.75	Field_RRI_13_DI	1.00	not assigned	-	-		
		PsCam015707_10196_516	112.9		0.26	disagreeing hits	-	-		
	III	III.1	PsCam000813_700_489	8.5		0.58	cytochrome P450 enzyme/ Abscisic acid 8' hydroxylase	-	plant defense (detoxification)	
			PsCam013963_9506_1445	8.5		0.28	bifunctional nuclease 1, partial	-	defense response (ABA-mediated callose deposition)	
			PsCam040305_25084_690	9		0.75	not assigned	-	-	
			PsCam041366_25850_304	9		0.76	not assigned	-	-	
			PsCam036018_21170_1023	9.5		0.22	Zinc finger (C3HC4-type RING finger) family protein	-	signal transduction	
			PsCam037306_22398_891	9.6	Field_AD1_12_DI	1.00	not assigned	-	-	
			PsCam050507_33085_1631	9.6		0.74	Brassinosteroid signalling positive regulator (BZR1) family protein	-	signal transduction	
			PsCam000634_552_52	10.2		0.38	disagreeing hits	-	-	
PsCam020958_11714_2207			10.2		0.51	disagreeing hits	-	-		
PsCam052721_34942_172			10.2		0.51	not assigned	-	-		
III.2				PsCam004672_3514_1959	22.4		0.57	Protein kinase superfamily protein	-	-
				PsCam017207_10682_389	22.4	RIPE_10_REN FLO1_12_REN	1.00 0.72	disagreeing hits	-	-
				PsCam030793_18164_3029	22.4		0.65	histidine kinase	-	signal transduction
				PsCam049584_32211_1847	22.4	FLO2_10_REN	0.72 1.00	3-ketoacyl-CoA synthase	-	-
	PsCam011480_7823_797	22.6			0.69	not assigned	-	-		
	PsCam034586_19967_1537	22.6			0.66	disagreeing hits	-	-		
	PsCam036107_21256_658	22.6			0.79	disagreeing hits	0.60	-		
	PsCam001506_1254_232	23.4			0.91	not assigned	0.73	-		
	AA175_282	24			0.24	not assigned	0.22	-		
	PsCam026085_15023_289	25.5			0.75	not assigned	0.75	-		
	PsCam057866_38405_252	26.4			0.90	not assigned	0.90	-		
	PsCam046091_29627_1315	27.1			0.84	RNA binding protein	0.84	-		
	PsCam006425_4782_144	27.3			0.83	not assigned	0.83	-		
	PsCam043138_27166_413	27.3		Field_AD1_12_DI FLO2_12_REN FLO1_10_REN RIPE_12_REN	27.3	1.00	disagreeing hits	-	-	
III.3		PsCam050202_32797_2026	27.3		0.85	not assigned	-	-		
		PsCam057845_38387_1581	27.6		0.78	heptahelical protein	-	signal transduction		
		PsCam038362_23399_731	27.8		0.42	cyclin D	-	plant development and flowering		
		PsCam055250_36400_562	31		0.33	not assigned	-	-		
		PsCam035590_20765_380	31.6	Field_RRI_All MFA.Dim.1 Field_RRI_11_RI Field_RRI_10_DI	1.00	Gamma-tubulin complex component	-	-		
		PsCam039309_24261_68	32.1		0.28	not assigned	-	-		
		PsCam002271_1824_370	32.6		0.25	Target SNARE coiled-coil domain protein	-	defense response (secretion of pathogenesis related proteins)		
		PsCam037297_22389_303	32.6		0.32	RING/U-box superfamily protein	-	signal transduction		
		III.4		PsCam037297_22389_303	32.6		0.32	RING/U-box superfamily protein	-	signal transduction

LG	LD block (a)	Marker	Position (cM) (b)	Associated traits (c)	LD (r ²) with trait-associated marker (bold) (d)	Putative annotation (e)	Putative function of interest for disease resistance (f)	Putative function in plant development (g)
		PsCam023109_13003_2421	32.7	Field_RRI_12_RI Field_RRI_12_DI Field_RRI_10_RI	1.00	serine/threonine-protein kinase	pathogen recognition/signal transduction	-
III.5		PsCam036264_21412_1640	32.8		0.32	disagreeing hits	-	-
		PsCam011358_7733_760	33.8	RIPE_10_REN	1.00	Ubiquitin-like modifier-activating enzyme	-	flowering
		PsCam012167_8289_560	33.8		0.31	not assigned	-	-
		PsCam057542_38126_654	33.8		0.51	not assigned	-	-
		PsCam000434_380_216	33.9		0.52	Pyridoxal phosphate (PLP)-dependent transferases superfamily protein	-	-
		PsCam051144_33677_200	33.9		0.70	not assigned	-	-
		PsCam052477_34775_186	33.9		0.22	Thylakoid luminal 17.9 kDa protein	-	-
III.6		PsCam037941_22996_137	37.3		0.30	amino acid transporter	plant defense (nitrogen burst)	-
		PsCam000088_75_1456	38.3		0.27	PRU-interacting factor G	-	-
		PsCam034645_20018_1507	38.3	Field_AD12_13_DI Field_AD11_13_DI	1.00	ABC transporter I family protein	signal transduction	-
		PsCam037415_22499_621	38.3		0.46	disagreeing hits	-	-
		PsCam001871_1531_550	38.4		0.45	RING/U-box superfamily protein	signal transduction	-
III.7		PsCam050430_33010_1078	38.7		0.51	disagreeing hits	-	-
		PsCam057273_37903_188	38.7		0.53	not assigned	-	-
		PsCam056354_37186_567	39.2	Field_AD11_10_RI	1.00	not assigned	-	-
III.8		AB92_357	43.5		0.20	not assigned	-	-
		AB92_370	43.5	Field_AD12_10_DI	1.00	not assigned	-	-
		AD57_301	43.5		0.58	not assigned	-	-
		AD57_339	43.5		0.39	not assigned	-	-
		PsCam004398_3307_281	43.7		0.61	not assigned	-	-
		PsCam017096_10628_1411	43.7		0.28	mRNA cap guanine-N7 methyltransferase	-	-
		PsCam017281_10725_414	43.7		0.28	2-like isoform X1	-	-
		PsCam022188_12477_4804	43.7		0.31	Cyclic nucleotide-gated ion channel	signal transduction	-
		PsCam030853_18210_666	43.7		0.31	Phosphomevalonate kinase	-	-
III.9		PsCam058449_38886_549	43.8		0.83	not assigned	-	-
		PsCam000125_108_427	44.5		0.21	Galactosyltransferase family protein	defense response (detoxification)	-
		PsCam045339_29025_1722	44.5		0.52	disagreeing hits	-	-
		PsCam046219_29720_864	44.5		0.23	VQ motif-containing protein (interaction with WRKY transcription factors)	transcription regulation	-
		PsCam001039_886_135	44.6		0.46	flavonol synthase/flavanone 3-hydroxylase like	disease resistance (unknown process)	-
		PsCam030789_18160_162	44.6	Field_AD11_All Field_AD11_13_RI	1.00	pectin methyltransferase	defense response (secretion of pathogenesis related proteins)	-
		PsCam056250_37103_770	44.8		0.34	Dof zinc finger DNA binding protein	signal transduction	-
		PsCam016014_10253_143	45.7		0.68	not assigned	-	-
		PsCam013984_9520_1022	45.9		0.32	alcohol dehydrogenase 1	defense response	-
		PsCam043953_27903_3736	46.15		0.59	transcriptional corepressor SEUSS-like	-	flower development (auxin responsive)
III.10		PsCam049627_32251_1328	62.5		0.40	Histone deacetylase	transcription regulation	-
		PsCam006723_4997_1323	62.9		0.78	Ras (GTPase) Leucine-rich-repeat protein	defense response (oxydative burst)	-
		PsCam035625_20800_2434	62.9		0.82	C2 and GRAM domain-containing protein	defense response (cell death)	-
		PsCam021109_11830_298	63.1		0.89	disagreeing hits	-	-
		PsCam045079_28804_1964	63.1	Field_AD11_12_RI	1.00	Elongation factor Tu	pathogen recognition	-
		PsCam034391_19785_2451	63.3		0.92	disagreeing hits	-	-
		PsCam027316_15973_1721	64		0.39	Armadillo-repeat protein	signal transduction	-

LG	LD block (a)	Marker	Position (cM) (b)	Associated traits (c)	LD (r ²) with trait-associated marker (bold) (d)	Putative annotation (e)	Putative function of interest for disease resistance (f)	Putative function in plant development (g)
		PsCam025419_14540_1296	64.1		0.25	NADH-ubiquinone oxidoreductase subunit	signal transduction	-
		PsCam056198_37060_1488	64.1		0.32	importin-5-like	-	-
		PsCam056270_37118_747	64.1		0.30	importin-5-like	-	-
III.11		PsCam045721_29348_2080	64.4		0.24	beta-1,3 glucanase	-	-
		PsCam045741_29362_2319	65.2		0.24	alpha/beta-Hydrolases superfamily protein	-	-
		PsCam055871_36784_1179	65.4		0.24	putative gibberellin signaling DELLA protein LA (plant growth)	-	plant growth
		PsCam058990_39285_415	65.7		0.49	not assigned	-	-
		PsCam012833_8640_2918	65.9	HT_10_REN	1.00	Telomeric DNA-binding protein	-	-
		PsCam037530_22609_1137	65.9		0.96	Serine/threonine protein kinase	pathogen recognition/signal transduction	-
		PsCam034012_19502_86	66.3		0.41	E3 ubiquitin-protein ligase RING1-like	signal transduction	-
		PsCam026198_15099_355	67.1		0.22	not assigned	-	-
		PsCam034850_20207_513	67.1		0.24	not assigned	-	-
		PsCam037299_22391_369	67.1		0.22	BTB/POZ/MATH-domains containing protein	signal transduction	-
III.12		PsCam052817_35000_252	70.4		0.25	not assigned	-	-
		PsCam035474_20656_922	70.7		0.82	ATPase family AAA domain-containing protein	signal transduction	-
		PsCam046081_29619_1524	70.8		0.49	General transcription factor I1H subunit	-	-
		PsCam049791_32409_803	70.8		0.95	LysM domain-containing GPI-anchored protein	pathogen recognition	-
		PsCam050434_33014_819	70.8		0.49	not assigned	-	-
		PsCam012453_8457_2872	72.1	Field_RRI_11_RI	1.00	receptor-like protein kinase like protein	pathogen recognition/signal transduction	-
		PsCam004543_3401_248	72.6		0.26	not assigned	-	-
III.13		PsCam048377_31112_1947	76.4		0.29	Cysteine-rich receptor-like protein kinase	-	growth and development
		PsCam045562_29219_567	76.8		0.94	NAD(PH)-quinone oxidoreductase	-	-
		PsCam051345_33851_551	76.8	RIPE_10_REN	1.00	50S ribosomal protein L17	-	-
		PsCam013052_8832_321	76.9		0.54	not assigned	-	-
		PsCam000915_777_561	78.1		0.65	Ethylene responsive transcription factor	-	-
III.14		PsCam037625_22700_831	91.1		0.35	not assigned	-	-
		PsCam033909_19408_223	91.4		0.28	NB-ARC domain-containing disease resistance protein-like	pathogen recognition	-
		PsCam011571_7892_294	91.5	Field_AD11_13_RI	1.00	U-box domain-containing protein (Armadillo)	signal transduction	-
		PsCam035876_21033_989	91.7		0.29	not assigned	-	-
III.15		PsCam041210_25755_211	96.5		0.35	not assigned	-	-
		PsCam051454_33947_98	96.5		0.35	HMG (high mobility group) box protein with ARID/BRIGHT DNA-binding domain	-	-
		PsCam001621_1340_484	96.6		0.29	GPI-anchored protein LORELEI	-	-
		PsCam013171_8925_143	96.7		0.34	Zinc finger-SWIM domain-containing protein	signal transduction	-
		PsCam036256_21404_2650	96.7		0.36	Endoglucanase	-	-
		PsCam000275_241_2264	96.8		0.23	disagreeing hits	-	-
		PsCam000488_427_411	96.8		0.43	Cytochrome P450	plant defense (detoxification)	-
		PsCam033941_19437_1439	96.8		0.24	paired amphipathic helix (PAHZ) protein Sin3-like	-	-
		PsCam043293_27312_1261	96.8	Field_AD11_10_RI	1.00	dehydrogenase protein	defense response (oxydative burst)	-
		PsCam046210_29713_182	96.8		0.35	NADH kinase-like	defense response (oxydative burst)	-
		PsCam051978_34391_55	96.9		0.77	not assigned	-	-

LG	LD block (a)	Marker	Position (cM) (b)	Associated traits (c)	LD (r^2) with trait-associated marker (bold) (d)	Putative annotation (e)	Putative function of interest for disease resistance (f)	Putative function in plant development (g)
		PsCam011467_7813_1433	97.1	Field_AD12_13_DI	0.30	not assigned	-	-
		PsCam033932_19429_1440	97.1		0.46	Pentatricopeptide repeat-containing protein	-	-
		PsCam036555_21689_1136	97.1		0.39	protein YLS7-like (cell wall modifications)	-	-
		PsCam039120_24097_1251	97.1		0.55	protein YLS7-like (cell wall modifications)	-	-
		PsCam013014_8801_92	97.3		0.67	Leucine-rich repeat transmembrane protein kinase	pathogen recognition/signal transduction	-
		PsCam038439_23473_1589	97.4		0.31	BES1/BZR1-like protein	signal transduction	-
		PsCam057518_38104_1138	97.4		0.21	Leucine-rich repeat transmembrane serine-threonine protein kinase	pathogen recognition/signal transduction	-
III.16		PsCam044045_27978_330	100.4		0.42	S-locus receptor kinase	-	-
		PsCam051258_33779_1408	100.4		0.44	Galactosyltransferase family protein	-	plant development
		PsCam003330_2584_574	100.6		0.42	not assigned	-	-
		PsCam034489_19875_2311	101.2		0.56	not assigned	-	-
		PsCam036194_21342_1967	101.2		0.43	RING/U-box superfamily protein	-	flower development
		PsCam036384_21526_1757	101.2		0.44	not assigned	-	-
		PsCam036878_21993_1959	101.2		0.88	not assigned	-	-
		PsCam037583_22661_232	101.2		0.46	Basic-leucine zipper (bZIP) transcription factor family protein	-	plant development
		PsCam043018_27052_1394	101.2	RIPE_10_REN	1.00	disagreeing hits	-	-
		PsCam043412_27422_489	101.2		0.44	mitochondrial import inner membrane translocase subunit TIM17-2-like	-	-
		PsCam043489_27495_1064	101.2		0.42	RNA exonuclease	-	-
		PsCam045978_29543_1362	101.2		0.42	GroES-like zinc-binding alcohol dehydrogenase family protein	-	-
		PsCam042347_26418_3657	101.6		0.23	AAA+ ATPase	-	plant growth
		PsCam012968_8762_695	103.2		0.29	not assigned	-	-
		PsCam011601_7912_909	103.7		0.34	not assigned	-	-
		PsCam031604_18516_424	103.7		0.46	not assigned	-	-
III.17		PsCam024624_13995_4847	130.6		0.75	homeobox-leucine zipper protein	-	plant development
		PsCam005215_3949_747	131		0.84	disagreeing hits	-	-
		PsCam049262_31901_1067	131		0.84	Protein disulfide isomerase Like	-	-
		PsCam037548_22627_143	131.1		0.66	not assigned	-	-
		PsCam049061_31705_537	131.1		0.84	heat shock protein hsp70	-	plant development
		PsCam038657_23672_1732	131.3		0.66	Coatomer alpha subunit-like protein	-	-
		PsCam034779_20141_1638	132		0.84	not assigned	-	-
		PsCam038309_23348_747	132	HT_12_REN	0.75	not assigned	-	-
		PsCam035617_20792_637	132	HT_10_REN	1.00	brassinosteroid receptor	-	-
		PsCam037922_22979_691	132.1		1.00	Gibberellin 3-oxidase (LE)	-	plant growth and development
		PsCam034369_19765_1448	133		0.78	not assigned	-	-
III.18		PsCam044311_28208_2067	134		0.89	formamidopyrimidine-DNA glycosylase-like	-	-
		PsCam049145_31788_282	134		0.76	Chaperone dnaJ	-	-
		PsCam050069_32675_656	134		0.66	brassinosteroid receptor	-	vascular development, flowering and senescence
		PsCam055856_36774_994	134		0.68	not assigned	-	-
		PsCam049798_32416_244	134.1		0.83	Amino acid transporter	-	flower development
		PsCam002356_1892_138	134.3		0.77	not assigned	-	-
		PsCam045932_29506_1539	134.3		0.72	F-box/LRR-repeat protein	-	plant growth and development
		PsCam049414_32050_2511	134.3		0.70	disagreeing hits	-	-
		PsCam060168_40149_2752	134.4		0.68	LRR Serine threonine protein kinase	-	plant growth and development

LG	LD block (a)	Marker	Position (cM) (b)	Associated traits (c)	LD (r ²) with trait-associated marker (bold) (d)	Putative annotation (e)	Putative function of interest for disease resistance (f)	Putative function in plant development (g)
IV	IV.1	PsCam034256_19681_301	134.6	FLO1_12_REN	1.00	not assigned	-	-
		PsCam044188_28099_1061	3.7		0.24	ADP-glucose pyrophosphorylase	-	-
		PsCam000404_355_3214	3.8	Field_RRI_11_DI	1.00	disagreeing hits	-	-
		PsCam049668_32292_713	3.8		1.00	Glucose-6-phosphate/phosphate translocator	signal transduction	-
	IV.2	PsCam010133_6813_144	5.9		0.28	not assigned	-	-
		PsCam025226_14405_465	5.9		0.87	not assigned	-	-
		PsCam050850_33409_1793	5.9	Field_AD12_12_KEN	1.00	not assigned	-	-
		PsCam029309_17475_346	7.5		0.68	not assigned	-	-
	IV.3	PsCam040341_25108_601	10.6		0.34	auxin-responsive protein IAA33-like	defense response	-
		PsCam033505_19138_1061	10.7		0.75	Replication factor A protein	-	-
PsCam035749_20920_4800		11.1		0.86	villin-4 like (root hair growth)	-	root hair growth	
PsCam034148_19622_98		11.4		0.88	mitochondrial outer membrane import complex protein METAXIN-like	-	-	
IV.4	PsCam036087_21236_2392	11.4	Field_RRI_13_DI	1.00	not assigned	-	-	
	PsCam051819_34264_514	11.4		0.93	Methyltransferase, putative	-	-	
	PsCam052249_34608_1152	11.4		0.81	not assigned	-	-	
	PsCam011134_7551_2658	11.8		0.21	Zn/Cd P(IB)-type ATPase	-	-	
	PsCam023144_13033_1388	11.8		0.27	not assigned	-	-	
	PsCam036054_21205_946	11.8		0.29	not assigned	-	-	
	PsCam043695_27683_971	11.8		0.27	Calcium-binding EF-hand family protein (helix-loop-helix)	transcription regulation	-	
	PsCam033746_19289_2936	12.3		0.25	subtilisin-like serine protease	pathogen recognition	-	
	PsCam049072_31716_1468	12.3		0.72	TBC Rab-GAP domain protein	signal transduction	-	
	PsCam050335_32924_830	12.3		0.45	Oxysterol-binding protein-related protein	transcription regulation	-	
IV.5	PsCam003906_2976_359	22.1	Field_AD12_13_DI	1.00	not assigned	-	-	
	PsCam022603_12702_205	22.1		1.00	WD40 repeat-containing protein	transcription regulation	-	
	PsCam042998_27033_1118	22.1		0.73	probable polygalacturonase-like	-	-	
	PsCam044946_28688_5469	22.1		0.73	TOPLESS-related 2 (negative regulator of jasmonate responses)	signal transduction	-	
	PsCam001099_933_92	22.3		0.24	ACC synthase (ethylene production)	signal transduction	-	
	PsCam004825_3640_1632	22.3		0.24	not assigned	-	-	
	PsCam036784_21908_580	22.3		0.70	Tetratricopeptide repeat (TPR)-like superfamily protein	signal transduction	-	
	PsCam043167_27194_1523	22.3		0.89	not assigned	-	-	
	PsCam003297_2562_75	22.7		0.49	not assigned	-	-	
	PsCam039023_24009_852	61.6	CC_RRI_09_Ae109	1.00	Prolyl 4-hydroxylase alpha-2 subunit (cell wall)	-	-	
IV.6	PsCam005728_4308_538	67.2		0.35	not assigned	-	-	
	PsCam049769_32389_581	67.2		0.26	Actin-related protein 2/3 complex subunit	-	-	
	PsCam043979_27922_2837	67.3		0.21	phototropin-like protein	-	-	
	PsCam018415_11127_149	67.6	Field_AD11_13_RI	1.00	not assigned	-	-	
IV.7	PsCam037513_22593_2935	67.9		0.70	not assigned	-	-	
	PsCam056094_36970_1559	68.7		0.39	disagreeing hits	-	-	
	PsCam004708_3544_675	68.8		0.36	Cysteine-rich receptor-like protein kinase	pathogen recognition/signal transduction	-	
PsCam037218_22318_568	68.8		0.50	High-affinity nitrate transporter	signal transduction	-		
PsCam011465_7811_255	68.9		0.47	not assigned	-	-		
PsCam000136_119_1230	69.3		0.63	BTB/POZ domain-containing protein	signal transduction	-		

LG	LD block (a)	Marker	Position (cM) (b)	Associated traits (c)	LD (r ²) with trait-associated marker (bold) (d)	Putative annotation (e)	Putative function of interest for disease resistance (f)	Putative function in plant development (g)
		PsCam001188_1003_335	69.3	Field_RRI_13_DI	1.00	GLABRA2 expression modulator (root epidermal cells)	-	root development
		PsCam001813_1486_880	69.3		0.55	not assigned	-	-
		PsCam034609_19986_173	69.3		0.52	AP-1 complex subunit mu	defense response	-
		PsCam035313_20503_2100	69.3		0.47	Nucleotide-diphospho-sugar transferase family protein	-	-
		PsCam055151_36328_103	69.3		0.97	not assigned	-	-
IV.8		PsCam027331_15987_254	76.1		0.79	BURP domain-containing protein	-	-
		PsCam000015_11_1425	76.2		0.86	Phosphoenolpyruvate carboxylase	-	-
		PsCam002207_1782_264	76.2		1.00	ribosomal protein L34 homolog	-	-
		PsCam003600_2770_552	76.2		0.86	not assigned	-	-
		PsCam025548_14643_2834	76.2		0.86	ribonucleoprotein	-	-
		PsCam026054_14999_2864	76.2		0.86	LRR Serine threonine protein kinase	pathogen recognition/signal transduction	-
		PsCam029227_17412_2455	76.2		0.86	FYVE zinc finger domain	signal transduction	-
		PsCam029388_17534_252	76.2		0.86	not assigned	-	-
		PsCam036906_22019_1424	76.2	CC_RRI_09_Ae109	1.00	Transcription factor bHLH123	transcription regulation	-
		PsCam037191_22293_279	76.2		0.93	AP2-like ethylene-responsive transcription factor	transcription regulation	-
		PsCam039135_24109_1277	76.2		1.00	disagreeing hits	-	-
		PsCam044897_28640_2303	76.2		0.86	Phospholipid-transporting ATPase	-	-
		PsCam045867_29458_1290	76.2		0.86	homogentisate 1,2-dioxygenase	-	-
		PsCam049226_31865_413	76.2		0.86	Delta-8 sphingolipid desaturase	-	-
		PsCam054935_36196_599	76.2		0.93	Vacuolar amino acid transporter	plant defense (nitrogen burst)	-
		PsCam057816_38362_1031	76.2		0.93	Vacuolar amino acid transporter	plant defense (nitrogen burst)	-
IV.9		PsCam058763_39109_322	105.7		0.21	homeobox-leucine zipper protein ATHB-15-like	defense response (cell wall modifications)	plant development
		PsCam004183_3161_311	106.1		0.40	not assigned	-	-
		PsCam049626_32250_2618	106.1		0.32	Protein kinase family protein	-	-
		PsCam056016_36903_1591	106.1		0.44	disagreeing hits	-	-
		PsCam038042_23094_246	106.2	Field_AD12_10_DI	1.00	Syringolide-induced protein B15-3-5	-	-
		PsCam050038_32646_475	106.2		0.82	x-ray induced transcript 1	-	-
		PsCam025688_14754_408	106.4		0.23	not assigned	-	-
		PsCam034462_19850_771	106.4		0.23	serine/threonine-protein kinase	pathogen recognition/signal transduction	-
IV.10		PsCam002154_1739_699	109.3		0.78	not assigned	-	-
		PsCam015423_10123_680	109.3		0.24	not assigned	-	-
		PsCam048178_30923_1956	109.3		0.24	Tuftelin-interacting-like protein	-	-
		PsCam009634_6431_1700	109.4		0.85	MLO-like protein (defense response)	transcription regulation (cell wall modification)	-
		PsCam017544_10829_660	109.5		0.24	not assigned	defense response (detoxification)	-
		PsCam055980_36875_2802	109.6		0.20	Multidrug resistance protein ABC transporter C family	-	-
		PsCam000016_12_4286	109.8	Field_AD11_12_DI	1.00	not assigned	-	-
		PsCam002975_2341_530	109.8		0.98	disagreeing hits	-	-
		PsCam004414_3317_343	109.9		0.98	not assigned	-	-
		PsCam048347_31085_1660	110		0.33	FRIGIDA-like protein (flowering time)	-	flowering time
		PsCam045032_28763_1600	110.7		0.20	not assigned	-	-
IV.11		PsCam029001_17237_215	112.05		0.23	not assigned	-	-
		PsCam010889_7353_756	112.2		0.22	disagreeing hits	-	-
		PsCam039757_24634_695	112.2		0.50	Ethylene responsive transcription factor	signal transduction (jasmonic Acid regulation)	-
		PsCam046122_29647_1454	112.2		0.55	nodulin MTN21 /EamA-like transporter family protein	-	-

LG	LD block (a)	Marker	Position (cM) (b)	Associated traits (c)	LD (r^2) with trait-associated marker (bold) (d)	Putative annotation (e)	Putative function of interest for disease resistance (f)	Putative function in plant development (g)	
IV.12	PsCam030702_18121_172	Field_RRI_12_DI	112.3	1.00	0.96	not assigned	-	-	
			PsCam030723_18124_1544	112.3	1.00	1.00	Chloride channel protein CLC-d	transcription regulation	-
			PsCam000901_767_1487	112.5	0.28	0.30	not assigned	-	-
			PsCam044074_28005_2770	112.5	0.27	0.27	Survival protein SurE-like phosphatase/nucleotidase	-	-
	PsCam058580_38979_170	Field_RRI_All	112.5	0.28	0.27	not assigned	-	-	
			PsCam005370_4073_599	112.6	0.27	0.23	Cupredoxin superfamily protein	-	-
			PsCam042808_26850_709	112.6	1.00	0.94	DNA/RNA polymerases superfamily protein	-	-
			PsCam043503_27507_1032	112.6	0.94	1.00	not assigned	-	-
			PsCam049842_32457_279	112.6	0.88	0.94	disagreeing hits	-	-
	IV.13	PsCam036805_21927_951	112.7	0.44	0.47	cysteine-rich RLK (RECEPTOR-like protein kinase)	pathogen recognition/signal transduction	-	
			PsCam004402_3309_217	112.8	0.94	not assigned	-	-	
			PsCam026907_15634_313	115.4	0.95	serine/threonine-protein kinase MHK-like isoform X1	pathogen recognition/signal transduction	-	
	V.1	PsCam053795_35577_1738	115.4	1.00	1.00	disagreeing hits	-	-	
PsCam048310_31051_2109			64.8	0.28	0.28	Cystathionine beta-synthase (CBS) family protein	-	-	
V.2	PsCam006774_5044_2706	64.9	0.41	0.41	Membrane insertion protein, OxaA/YidC with tetratricopeptide repeat domain	signal transduction	-		
		PsCam007177_5329_330	64.9	0.42	0.42	not assigned	-	-	
		PsCam022961_12908_150	64.9	0.35	0.35	not assigned	-	-	
		PsCam043555_27554_1164	64.9	0.34	0.34	disagreeing hits	-	-	
		PsCam045289_28982_363	64.9	0.42	0.42	not assigned	-	-	
		PsCam046265_29756_347	64.9	0.33	0.33	Glutaredoxin-like protein	signal transduction	-	
		PsCam050336_32925_1106	64.9	0.88	0.88	reversibly glycosylatable polypeptide	defense response (cell wall modifications)	-	
		PsCam057491_38080_3275	64.9	0.88	0.88	Heat shock protein	signal transduction	-	
		PsCam038363_23400_360	65	1.00	1.00	F-box family protein	signal transduction (hormone regulation)	-	
		PsCam012541_8505_720	65.3	0.50	0.50	WRKY transcription factor	transcription regulation	-	
		PsCam029030_17262_1220	65.3	0.24	0.24	disagreeing hits	-	-	
		PsCam049404_32040_1107	65.3	0.23	0.23	ACC synthase (Ethylene production)	signal transduction	-	
		PsCam049562_32189_262	65.3	0.49	0.49	disagreeing hits	-	-	
		PsCam002167_1750_201	75.8	0.56	0.56	not assigned	-	-	
		PsCam020963_11718_820	75.8	0.87	0.87	disagreeing hits	-	-	
		PsCam026871_15605_440	75.8	0.57	0.57	not assigned	-	-	
		PsCam027135_15827_202	75.8	0.94	0.94	disagreeing hits	-	-	
		PsCam039642_24545_120	75.8	0.43	0.43	disagreeing hits	-	-	
		PsCam043473_27479_108	75.8	0.94	0.94	isoamylase isoform 2 (starch biosynthesis)	-	-	
		V.3	PsCam048296_31038_1280	75.8	0.94	0.94	Chaperone protein dnaJ	defense response (cell death)	-
PsCam000077_4_886	75.9			1.00	1.00	putative NBS-LRR type disease resistance protein	pathogen recognition/signal transduction	-	
VI.1	PsCam039530_24449_397	75.9	0.35	0.35	disagreeing hits	-	-		
		PsCam038886_23883_1930	76.1	0.90	0.90	electron transport SCO1/SenC family protein	-	-	
		PsCam042225_26317_1763	88	1.00	1.00	DEAD-box ATP-dependent RNA helicase	defense response (oxydative burst)	-	
VI.2	PsCam057068_37736_201	88	0.20	0.20	not assigned	-	-		
		PsCam009846_6605_1377	17.4	0.22	0.22	transducin family protein / WD-40 repeat family protein	signal transduction	-	
VI.2	AC74-F66_137	17.5	1.00	1.00	not assigned	-	-		
		PsCam039048_24031_1302	39.1	0.66	0.66	not assigned	-	-	

LG	LD block (a)	Marker	Position (cM) (b)	Associated traits (c)	LD (r ²) with trait-associated marker (bold) (d)	Putative annotation (e)	Putative function of interest for disease resistance (f)	Putative function in plant development (g)
		PsCam052666_34906_475	39.1		0.72	disagreeing hits	-	-
		PsCam013085_8857_1728	39.3	Field_AD12_13_DI	1.00	not assigned	-	-
		PsCam000943_802_2706	39.4		0.47	Vesicle transport v-SNARE	defense response (secretion of pathogenesis related proteins)	-
		PsCam046918_30164_1200	39.8		0.53	not assigned	-	-
VI.3		PsCam034328_19730_5190	84.2		0.26	Protein kinase family protein with ARM repeat domain	-	plant growth and development
		PsCam036337_21482_2951	84.6		0.76	serine/threonine-protein kinase PBS1-like	-	plant growth and development
		PsCam037593_22669_918	84.6		0.26	disagreeing hits	-	-
		PsCam039509_24432_289	84.6		0.27	not assigned	-	-
		PsCam040795_25453_643	84.6		0.26	Glutaredoxin family protein	-	flower development
		PsCam042974_27010_292	84.6		0.75	Lipase	-	reproductive growth
		AA335-FG6_229	85	FLO2_10_REN	1.00	not assigned	-	-
		AA335-FG6_242	85		0.26	not assigned	-	-
		PsCam004726_3560_1453	85		0.27	Intron maturase, type II family protein	-	flower development
		PsCam030799_18169_204	85		0.53	disagreeing hits	-	-
		PsCam033628_19204_1252	85		0.73	potassium channel KAT2-like	-	-
VI.4		PsCam046316_29794_649	87.4		0.46	nodulin26-like intrinsic protein	regulation of defense response (silicon transport)	flower development
		PsCam060118_40109_376	87.4		0.74	nitrate transporter 1.2-like	-	nodulation
		PsCam001597_1323_222	87.9		0.87	Fascidin-like arabinogalactan protein	-	-
		PsCam040942_25561_669	88		0.91	not assigned	pathogen recognition	-
		PsCam039081_24060_367	88.1		0.88	not assigned	-	-
		PsCam001071_913_336	88.2		0.25	not assigned	-	-
		PsCam038183_23231_972	88.2		0.23	nodulin MtN21 /EamA-like transporter family protein	-	-
		PsCam000034_29_2253	88.3	Field_AD11_12_DI	1.00	LRR receptor-like serine/threonine-protein kinase	pathogen recognition/signal transduction	-
		PsCam036411_21551_388	88.3		0.95	not assigned	-	-
		PsCam042456_26516_2238	88.3		0.26	cellulose synthase family protein	defense response (cell wall modifications)	-
		PsCam010614_7138_199	88.4		0.92	not assigned	-	-
		PsCam035917_21074_2382	88.4		0.67	LRR receptor-like serine/threonine-protein kinase	pathogen recognition/signal transduction	-
		PsCam005430_4112_173	99.8		0.34	disagreeing hits	-	-
VI.5		PsCam039014_24000_1137	99.8		0.21	fascilin-like arabinogalactan protein	pathogen recognition	-
		PsCam009689_6479_476	100	Field_AD12_13_DI	1.00	F-box/RNI-like/FBD-like domains-containing protein	signal transduction (hormone regulation)	-
		PsCam021482_12074_4263	100		0.94	not assigned	-	-
		PsCam036513_21651_1231	100		0.81	RING/U-box superfamily protein	signal transduction	-
		PsCam039755_24632_1126	100		0.86	Transmembrane emp24 domain-containing protein	-	-
		PsCam009788_6562_2794	100.4		0.34	Potassium channel	K+ status enhances disease resistance	-
		PsCam002732_2167_92	100.7		0.31	not assigned	-	-
VII	VII.1	PsCam030182_17930_220	17.8		0.32	not assigned	-	-
		PsCam050051_32659_131	17.8		0.35	not assigned	-	-
		PsCam010078_6770_926	18.2		0.22	not assigned	-	-
		PsCam031102_18329_1418	18.2		0.27	Aminopeptidase family protein	defense response (cell death)	-
		PsCam039111_19410_3600	18.8	Field_AD12_13_DI	1.00	not assigned	-	-
		PsCam041923_26158_129	18.8		0.91	not assigned	-	-
VII.2		PsCam041431_25888_256	38.8		0.25	not assigned	-	-
		PsCam034334_19733_187	39.7		0.60	Auxin response factor	defense response	-

LG	LD block (a)	Marker	Position (cM) (b)	Associated traits (c)	LD (r^2) with trait-associated marker (bold) (d)	Putative annotation (e)	Putative function of interest for disease resistance (f)	Putative function in plant development (g)
		PsCam000133_116_1343	39.9	Field_AD12_12_RI	1.00	Inositol transporter	-	-
		PsCam001558_1292_57	39.9		0.67	not assigned	-	-
		PsCam001661_1374_958	39.9		0.67	not assigned	-	-
		PsCam021891_12310_347	39.9		0.67	not assigned	-	-
		PsCam042875_26914_2982	39.9		0.23	glycosyltransferase family (cell wall modifications)	defense response (detoxification)	-
		PsCam058482_38911_100	39.9		0.77	not assigned	-	-
		PsCam047459_30486_180	40.1		0.24	not assigned	-	-
	VII.3	PsCam038988_23977_855	42.3		0.22	senescence-related gene 1 -like	defense response (leaf senescence)	-
		PsCam003637_2797_421	42.3		0.96	not assigned	-	-
		PsCam045853_29447_1875	42.3		0.27	not assigned	-	-
		PsCam034753_20117_1972	42.3		0.64	disagreeing hits	-	-
		PsCam030917_18258_2984	42.6		0.64	disagreeing hits	-	-
		PsCam029454_47574_900	42.6		0.22	not assigned	-	-
		PsCam000907_772_3250	42.6		0.70	not assigned	-	-
		PsCam034444_19834_358	42.6	HT_10_REN	1.00	LRR serine/threonine-protein kinase	pathogen recognition/signal transduction	-
		PsCam037992_23044_806	42.6	Field_AD12_10_DI	0.65	disagreeing hits	-	-
		PsCam038332_23370_329	42.6		0.69	epsin N-terminal homology (ENTH) domain-containing protein	-	-
		PsCam039102_24080_996	42.6	HT_12_REN	0.89	senescence-related gene 1 -like	defense response (leaf senescence)	-
		PsCam046859_30131_450	42.6		0.65	not assigned	-	-
		PsCam049219_31860_2091	42.6		0.69	disagreeing hits	-	-
		PsCam048497_31223_1921	42.6		0.69	C2H2-like zinc finger protein	signal transduction	plant growth
		PsCam004854_3665_2467	42.8		0.20	Peptide transporter	signal transduction	plant growth
		PsCam023077_12977_819	42.8	Field_AD12_13_DI	0.32	Peptide transporter	signal transduction	plant growth
		PsCam037077_22184_2302	42.8		0.66	Peptide transporter	signal transduction	plant growth
		PsCam037681_22753_690	42.8		0.24	not assigned	-	-
		PsCam037946_23000_227	42.8		0.39	polyprotein	-	-
		PsCam043158_27185_1849	42.8		0.96	target of Myb protein-like 1 (TOM1)	-	-
		PsCam045791_29399_255	42.8		0.96	not assigned	-	-
		PsCam048180_30925_2357	42.8		0.72	Peptide transporter	signal transduction	plant growth
		PsCam049474_32106_2494	42.8		0.96	disagreeing hits	-	-
	VII.4	PsCam028946_17192_266	49.8		0.37	anthocyanidin reductase (PsAMR : seed coat color)	-	-
		PsCam053431_35372_133	49.8		0.78	not assigned	-	-
		PsCam056064_36944_600	49.8		0.34	F-box/FBD/LRR-repeat protein Atg78750-like isoform X1	plant growth and development	-
		PsCam046153_25672_848	50.5	FLO2_10_REN	1.00	Quinone reductase family protein	-	-
		PsCam038003_23055_1276	50.7		0.46	U-box domain-containing protein	-	-
		PsCam033982_19475_496	51.1		0.38	LRR serine/threonine-protein kinase	plant growth and development	-
		PsCam035755_20925_2095	51.1		0.37	Pentatricopeptide repeat-containing protein	-	-
		PsCam050805_33367_280	51.2		0.44	not assigned	-	-
		PsCam017692_10886_1587	51.3		0.52	not assigned	-	-
	VII.5	PsCam023431_13255_1332	57.5		0.21	Peptidyl-prolyl cis-trans isomerase CYP40	-	-
		PsCam035776_20944_3127	57.5		0.90	Agenet and bromo-adjacent homology (BAH) domain-containing protein-like protein	-	-
		PsCam049829_32444_1381	57.7	FLO2_12_REN	1.00	disagreeing hits	-	-
		PsCam000958_816_1084	58.1	FLO2_10_REN	0.24	LAG1 longevity assurance-like protein	-	-

LG	LD block (a)	Marker	Position (cM) (b)	Associated traits (c)	LD (r ²) with trait-associated marker (bold) (d)	Putative annotation (e)	Putative function of interest for disease resistance (f)	Putative function in plant development (g)
		PsCam029277_17454_150	58.1		0.23	cytochrome P450, family 707, subfamily A, polypeptide 4	-	reproductive development
		PsCam037467_22549_557	58.1		0.92	probable LRR receptor-like serine/threonine-protein kinase At1g51810-like	-	-
VII.6		PsCam011881_8108_513 PsCam050470_33049_58	59.7 59.7	Field_RRI_12_RI	1.00 0.20	not assigned glycosyltransferase-like (cell wall modifications)	defense response (detoxification)	-
VII.7		PsCam006709_4985_224 PsCam046071_29614_677	60.3 65		0.96 0.24	Electron transfer flavoprotein subunit alpha	-	-
		PsCam048495_31221_1607 PsCam022983_12922_354 PsCam035721_20895_4786 PsCam048913_31565_623 PsCam058549_38957_254	65.3 65.5 65.5 65.5 65.5	CC_RRI_09_RB84	0.28 1.00 0.41 0.79 0.40	Chaperone protein dnaJ F-box protein RMS4 (shoot branching) Trehalose synthase-like protein	defense response (cell death) signal transduction (hormone regulation) defense response	- - root growth
VII.8		PsCam008838_6114_356 PsCam027642_16235_308 PsCam032889_18913_235 PsCam050005_32614_358 PsCam011266_7658_455 PsCam027178_15859_290 PsCam058399_38849_268	67.75 67.75 68 68 68.8 69.2 69.2	CC_RRI_13_RB84	0.31 0.51 0.48 0.48 1.00 0.44 0.22	not assigned not assigned disagreeing hits disagreeing hits not assigned Transcription factor DPB-like 4-coumarate CoA ligase-like	transcription regulation plant defense (induced by pathogen attacks)	- - - - - - -
VII.9		PsCam034543_19927_2672	70.1		0.81	Wall-associated receptor kinase-like protein	-	cell expansion
		PsCam040048_24863_946	70.3		0.35	Cysteine-rich repeat receptor-like protein kinase	-	growth and development
		PsCam043617_27613_310 PsCam049131_31774_817 PsCam053743_35552_3534 PsCam054782_36090_2082	70.3 70.3 70.3 70.3	FLO2_10_REN	1.00 0.50 0.38 0.24	Ubiquitin carrier protein subtilisin-like serine protease not assigned Transducin family protein / WD-40 repeat family protein	- - - -	senescence - - -
		PsCam059143_39387_3828 PsCam000386_338_858	70.3 70.5		0.39 0.50	glycosyl hydrolase family 35 protein Senescence/dehydration-associated protein-related	- -	- -
VII.10		PsCam007194_5334_100 PsCam037149_22253_1999 PsCam000043_37_605 PsCam000852_730_514 PsCam023497_13303_2604 PsCam027009_15721_2497 PsCam023161_13047_484	70.55 70.7 70.8 70.8 70.8 70.8 72.7	CC_RRI_13_RB84	0.49 0.71 1.00 0.26 0.61 0.28 0.90	not assigned BTB/POZ domain-containing protein PWWP domain containing protein not assigned Gibberellin 2-beta-dioxigenase DNA repair protein XRCC2 homolog Sacharopine dehydrogenase family protein	signal transduction - - up-regulated through infection - defense response (oxydative burst)	- - - plant growth -
VII.11		PsCam026805_15548_347 PsCam038753_23762_1459 PsCam001904_1552_132 PsCam044935_28677_577 PsCam002863_2264_204 PsCam040861_25496_289 PsCam049618_32243_462	72.7 72.7 73.1 73.1 73.2 73.2 73.7	Field_RRI_12_RI	0.90 0.90 0.88 0.71 1.00 0.82 0.68 0.65	disagreeing hits Ubiquitin carrier protein disagreeing hits Pentatricopeptide repeat-containing protein not assigned Auxin-induced protein-like protein polyol transporter 5-like	- - signal transduction - - - defense response up-regulated through infection	- - - - - - - -

LG	LD block (a)	Marker	Position (cM) (b)	Associated traits (c)	LD (r^2) with trait-associated marker (bold) (d)	Putative annotation (e)	Putative function of interest for disease resistance (f)	Putative function in plant development (g)	
VII.12	PsCam043348_27363_200 PsCam033670_19242_2572 PsCam051731_34191_118 PsCam009638_16434_591 PsCam027856_16398_193 PsCam036623_21756_1504 PsCam038116_23165_1908 PsCam043279_27298_1255 PsCam038610_23628_1220	80.5	0.25	Cu/Zn superoxide dismutase II	-	-	-	-	
		80.9	0.30	Serine carboxypeptidase-like protein	-	-	-	-	
		80.9	1.00	disagreeing hits	-	-	-	-	-
		82.6	0.67	terpene synthase	-	-	-	defense response	-
		82.6	0.50	not assigned	-	-	-	-	-
		82.6	0.49	glycosyltransferase family	-	-	-	defense response (detoxification)	-
		82.6	0.49	not assigned	-	-	-	-	-
		82.6	0.71	not assigned	-	-	-	-	-
		82.8	0.23	Ras-related small GTP-binding family protein	-	-	-	defense response (oxydative burst)	-
		83	1.00	not assigned	-	-	-	-	-
		83	0.22	not assigned	-	-	-	-	-
		83	1.00	not assigned	-	-	-	-	-
VII.13	AA387_469 AA387_476 AA387_480 AA387_501 PsCam010536_7084_264 PsCam027452_16087_1460	83	1.00	Field_RRI_12_RI	-	-	-	-	
		83	0.22	not assigned	-	-	-	-	
		83	1.00	not assigned	-	-	-	-	
		83	1.00	Field_AD11_10_RI	-	-	-	-	
		83	1.00	RIPE_10_REN	-	-	-	-	
		83	1.00	FLO2_10_REN	-	-	-	-	
		83	0.69	not assigned	-	-	-	-	
		83	0.52	zinc finger CCHC domain-containing protein	-	-	-	signal transduction	-
		83	1.00	Homeobox-leucine zipper protein	-	-	-	defense response (cell wall modifications)	plant development
		83	0.29	not assigned	-	-	-	-	-
		83.1	0.26	Agamous-like MADS-box protein	-	-	-	signal transduction	plant development
		VII.14	AB101_390 PsCam007709_5623_138 PsCam050334_32923_389 PsCam034533_19917_1717 PsCam042921_26959_538 PsCam051717_34180_293 PsCam034407_19800_790 PsCam002799_2218_177 PsCam038378_23415_721 PsCam056236_37092_1071	83.5	1.00	Field_AD12_11_RI	-	-	-
83.5	1.00			MFA.Dim.2	-	-	-	-	
84.1	0.28			not assigned	-	-	-	-	
85.5	1.00			Proteasome assembly chaperone	-	-	-	-	
85.8	0.32			Long-chain-fatty-acid-CoA ligase	-	-	-	defense response	-
85.8	0.32			not assigned	-	-	-	-	-
85.8	0.33			ribosomal protein	-	-	-	-	-
86.2	0.59			not assigned	-	-	-	-	-
86.4	1.00			RIPE_10_REN	-	-	-	-	-
86.4	0.43			not assigned	-	-	-	-	-
86.4	1.00			not assigned	-	-	-	-	-
VII.15	PsCam039084_24063_2038 PsCam021733_12219_327 PsCam042979_27014_298 PsCam000939_798_301 PsCam040412_25162_412 PsCam042936_26973_1879 PsCam017010_10576_64 PsCam047307_30392_156 PsCam057114_37774_486 PsCam035270_20465_2660 PsCam034448_19837_782 PsCam057700_38265_886			86.9	0.23	E3 ubiquitin-protein ligase	-	-	-
		93	0.25	BTB/POZ domain-containing protein	-	-	-	-	-
		93	1.00	AT5g60050-like RNA ligase	-	-	-	-	plant development
		94.1	0.48	ribosomal protein	-	-	-	-	-
		97.4	0.68	cytochrome P450 enzyme	-	-	-	plant defense (detoxification)	-
		97.4	0.39	UDP-glucuronosyltransferase	-	-	-	defense response (detoxification)	-
		97.5	0.69	not assigned	-	-	-	-	-
		97.5	0.57	not assigned	-	-	-	-	-
		97.7	0.36	0.23	not assigned	-	-	-	-
		97.8	0.84	0.69	disagreeing hits	-	-	-	-
		97.9	1.00	disagreeing hits	-	-	-	-	-
		97.9	0.87	0.77	not assigned	-	-	-	-
VII.17	PsCam039084_24063_2038 PsCam021733_12219_327 PsCam042979_27014_298 PsCam000939_798_301 PsCam040412_25162_412 PsCam042936_26973_1879 PsCam017010_10576_64 PsCam047307_30392_156 PsCam057114_37774_486 PsCam035270_20465_2660 PsCam034448_19837_782 PsCam057700_38265_886	86.9	0.23	E3 ubiquitin-protein ligase	-	-	-	signal transduction	-
		93	0.25	BTB/POZ domain-containing protein	-	-	-	-	-
		93	1.00	AT5g60050-like RNA ligase	-	-	-	-	plant development
		94.1	0.48	ribosomal protein	-	-	-	-	-
		97.4	0.68	cytochrome P450 enzyme	-	-	-	plant defense (detoxification)	-
		97.4	0.39	UDP-glucuronosyltransferase	-	-	-	defense response (detoxification)	-
		97.5	0.69	not assigned	-	-	-	-	-
		97.5	0.57	not assigned	-	-	-	-	-
		97.7	0.36	0.23	not assigned	-	-	-	-
		97.8	0.84	0.69	disagreeing hits	-	-	-	-
		97.9	1.00	disagreeing hits	-	-	-	-	-
		97.9	0.87	0.77	not assigned	-	-	-	-

LG	LD block (a)	Marker	Position (cM) (b)	Associated traits (c)	LD (r^2) with trait-associated marker (bold) (d)	Putative annotation (e)	Putative function of interest for disease resistance (f)	Putative function in plant development (g)
		PsCam002116_1716_1030	98		0.26	not assigned	-	-
		PsCam003375_2615_454	98		0.41	not assigned	-	-
		PsCam004664_3506_3159	98		0.36	Molybdenum cofactor sulfuryase	pathogen interaction	-
		PsCam005071_3847_951	98	MFA_Dirn.1	0.65	not assigned	-	-
		PsCam021215_11903_3481	98		0.29	Dynammin-2B (response to ABA)	signal transduction	-
		PsCam021571_12130_320	98		0.38	not assigned	-	-
		PsCam037756_22824_706	98		0.36	RING/U-box superfamily protein	signal transduction	-
		PsCam027664_16252_1349	98		0.26	disagreeing hits	-	-
		PsCam040227_25014_407	98		0.36	Transcription factor GAMYB-like	-	flower development
		PsCam056030_36914_333	98		0.36	Mitochondrial import inner membrane translocase subunit TIM44	-	-
		PsCam038825_23827_1775	98		0.22	not assigned	-	-
		PsCam043396_27409_150	98		0.77	disagreeing hits	-	-
		PsCam039195_24163_318	98.4		0.74	not assigned	-	-
		PsCam030828_16189_3163	98.5	Field_AD12_10_DI FLO2_10_REN	1.00	not assigned	-	-
		PsCam049302_31940_190	98.5	FLO1_10_REN	0.63	CBL-interacting serine/threonine-protein kinase	pathogen recognition/signal transduction	-
		PsCam003105_2429_579	100.7		0.42	SPX domain-containing protein 3-like	-	-
		PsCam011584_7900_399	100.7		0.20	not assigned	-	-
					0.25		-	-

Annexe 2 : Additional files du Chapitre III

Additional file 1: Description of the Root-Collection.

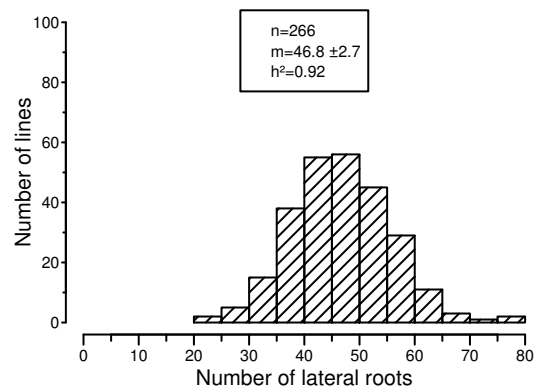
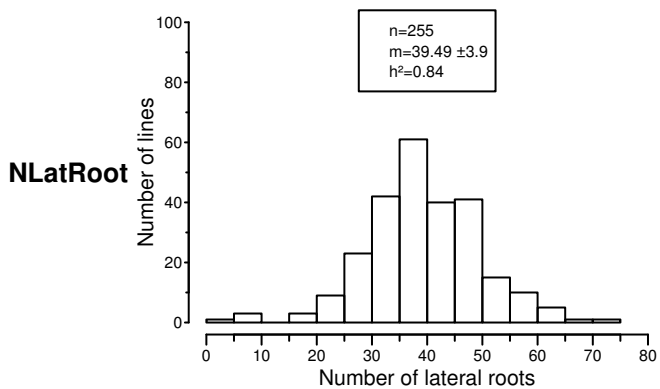
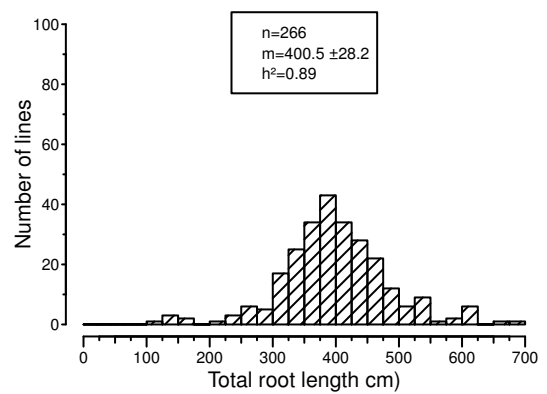
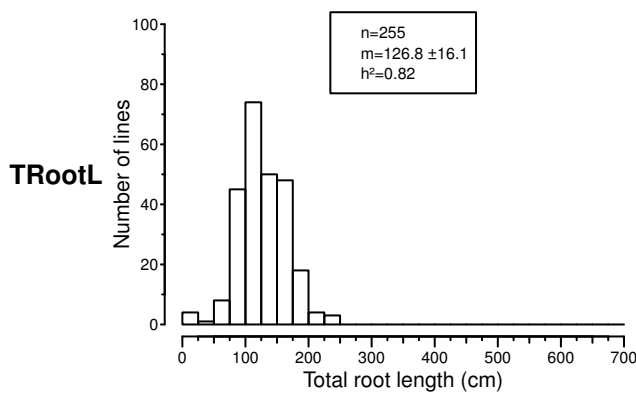
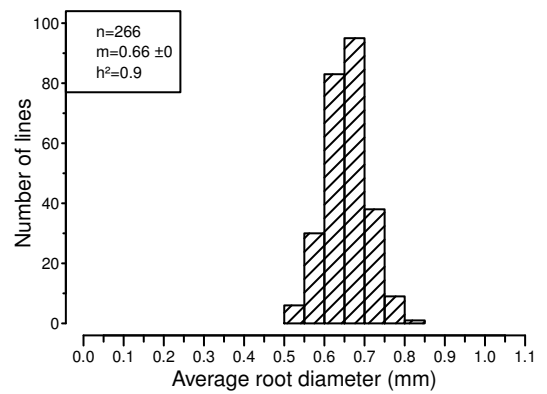
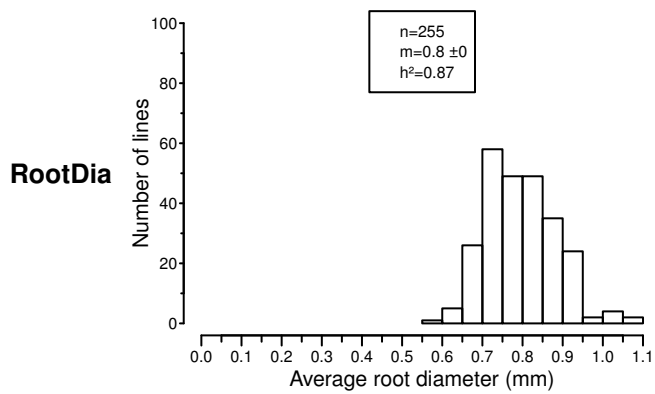
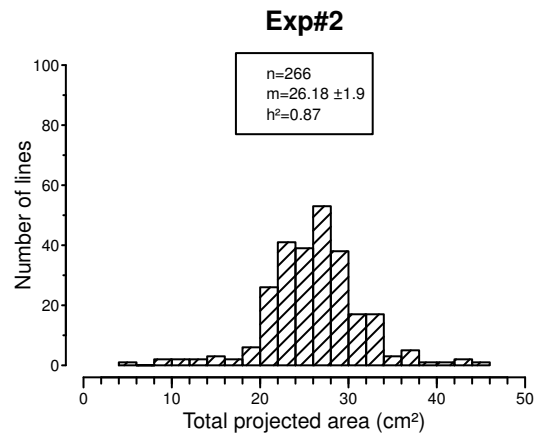
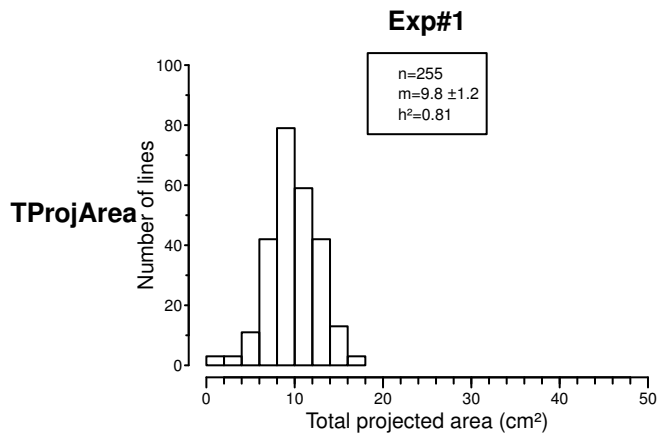
(a) Cv: Cultivar, Bl: breeding line, Gmp: germplasm, Lr: landrace, (S-)Wa: (Semi-)wild accession; (b) Fd: fodder, Gd: garden; (c) Ws: winter sown, Ss: spring sown; (d) lines included in the pea-Aphanomyces collection described in Desgroux et al (2016)

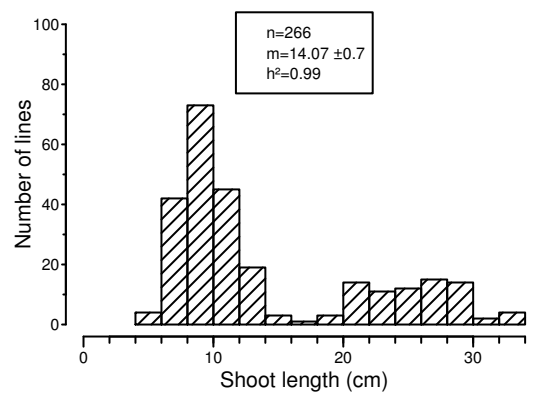
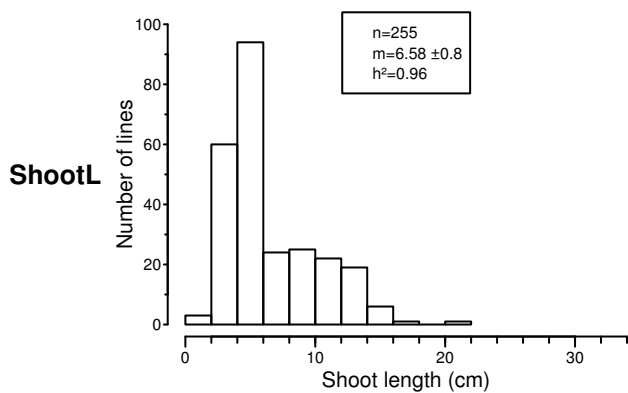
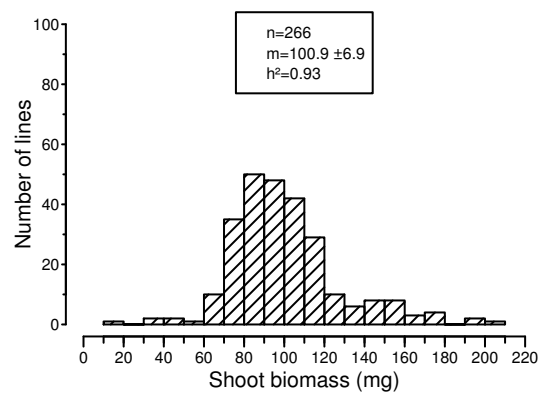
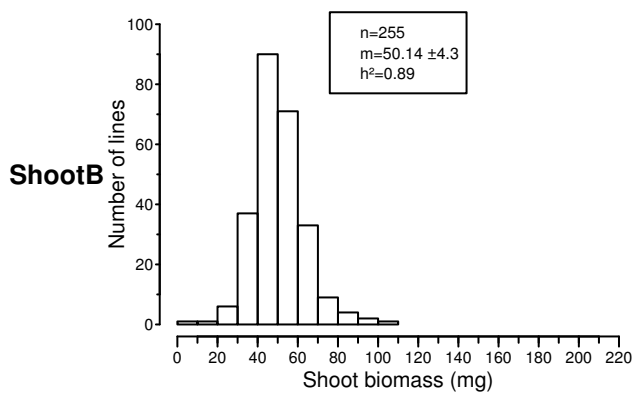
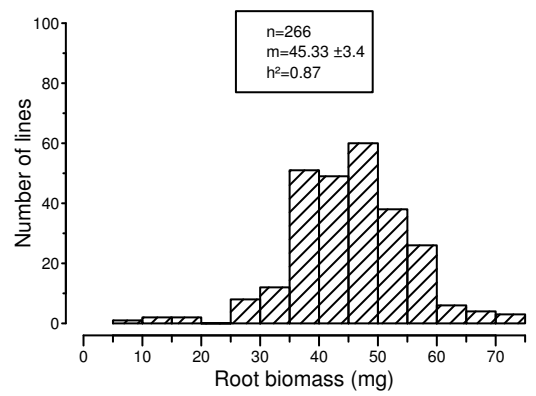
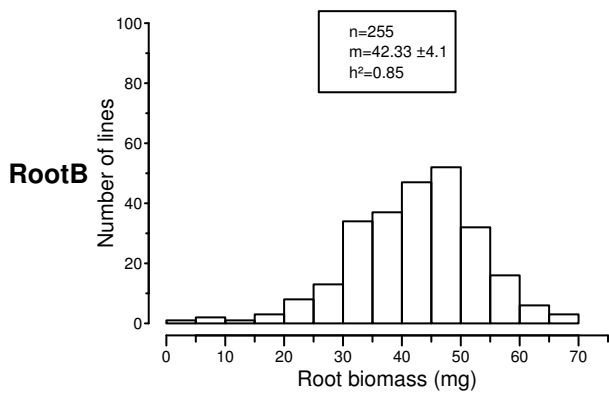
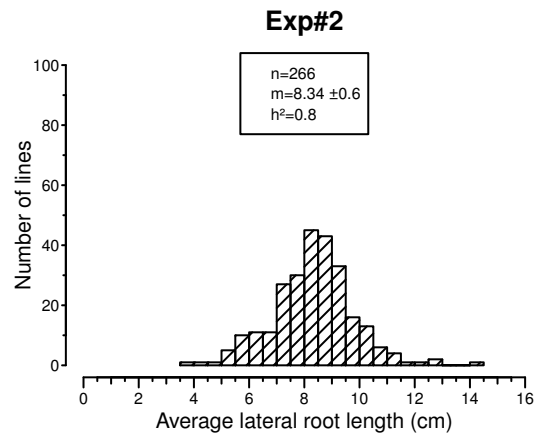
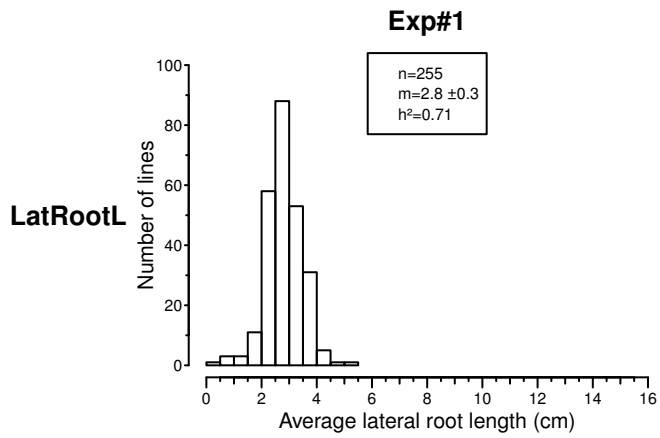
Line Name	Species/Subspecies	Cultivation status (a)	End use (b)	Sowing type (c)	Included in the pea-Aphanomyces collection (d)
KOROZA	Pisum sativum	Cv	Gd	Ss	
PISUM SATIVUM-AFGHANISTAN JI86	Pisum sativum	S-Wa/Wa			
PISUM SATIVUM-BOLIVIA JI228	Pisum sativum	S-Wa/Wa			
ALASKA	Pisum sativum	Cv	Gd	Ss	
PISUM SATIVUM-INDIA JI1267	Pisum sativum	Lr	Gd		
PESOL	Pisum sativum	Lr			
PISUM SATIVUM-MEXICO JI1844	Pisum sativum	Lr			
PISUM SATIVUM-ZAIRE JI2376	Pisum sativum	S-Wa/Wa			
PISUM SATIVUM-ZAMBIA JI2383	Pisum sativum	Lr			
PISUM ELATIUS JI1075	Pisum sativum subsp elatius	Wa			
PISUM ELATIUS JI1703	Pisum sativum subsp elatius	Wa			
PISUM HUMILE JI1794	Pisum sativum subsp humile	Wa			
PISUM ABYSSINICUM JI2202	Pisum sativum subsp abyssinicum	Lr			
PISUM FULVUM JI2473	Pisum fulvum	Wa			
PISUM FULVUM JI2523	Pisum fulvum	Wa			
PISUM TRANSCAUCASICUM JI2546	Pisum sativum	Wa			
PISUM SPECIOSUM-LIBYA JI2605	Pisum sativum	S-Wa/Wa			
AURALIA	Pisum sativum	Cv	Gd	Ss	
COTE D'OR	Pisum sativum	Lr	Fd	Ws	
MISTRAL	Pisum sativum	Cv	Dry	Ws	X
WNC 23Z SPP ARVENSE 1809	Pisum sativum	Lr	Fd	Ws	
WNC Z61Z SPP ARVENSE Z009	Pisum sativum	Lr	Fd	Ws	
PI273279	Pisum sativum subsp elatius	Lr	Fd		
PI180693	Pisum sativum	Cv	Gd		X
552	Pisum sativum	Bl	Gd	Ss	X
DARK SKIN PERFECTION	Pisum sativum	Cv	Gd	Ss	X
90-2131	Pisum sativum	Bl	Gd	Ss	X
SHRAT	Pisum sativum	Lr	Gd		
PI212112	Pisum sativum	Wa	Gd		
ENGLISH	Pisum sativum	Bl	Fd		
IREGI SARGA	Pisum sativum	Cv	Fd	Ss	
KIRIN 40	Pisum sativum	Bl	Gd	Ss	
CUZCO 1	Pisum sativum	Lr	Gd		
HAITI COLORE	Pisum sativum	Lr	Fd		
AFGHANISTAN ASIATICUM	Pisum sativum	Lr	Fd		
NEPAL A	Pisum sativum	S-Wa/Wa	Fd		X
ABYSSINICUM VAVILOVANIAM	Pisum sativum subsp abyssinicum	S-Wa/Wa	Fd		
CAPSICUM	Pisum sativum subsp elatius	Lr	Fd		
HATIVER	Pisum sativum	Cv	Gd	Ws	
TELEPHONE A RAMES	Pisum sativum	Cv	Gd	Ss	
FIN DE LA BIEVRE	Pisum sativum	Cv	Gd	Ss	
MERVEILLE D'ETAMPES	Pisum sativum	Cv	Gd	Ss	
NFG KRUPP PELUSCHKE	Pisum sativum	Cv	Fd	Ss	
PETIT PROVENCAL	Pisum sativum	Cv	Gd	Ss	
DESIREE	Pisum sativum	Cv	Fd	Ss	
LIVIOLETTA	Pisum sativum	Cv	Fd	Ss	
BINGEFORS	Pisum sativum	Bl	Fd		
MULTIRESISTANT	Pisum sativum	Bl	Gd	Ss	
AMINO	Pisum sativum	Cv	Dry	Ss	
K4626	Pisum sativum	Lr	Fd	Ss	
CAMEOR	Pisum sativum	Cv	Gd	Ss	
BALLET	Pisum sativum	Cv	Dry	Ss	

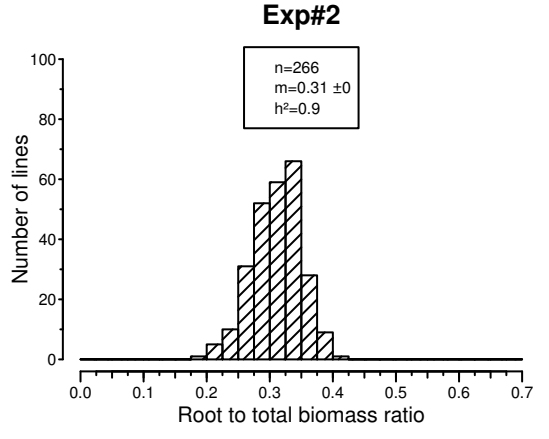
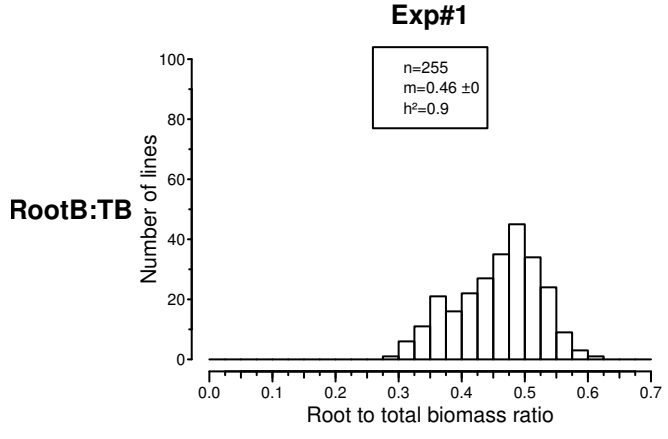
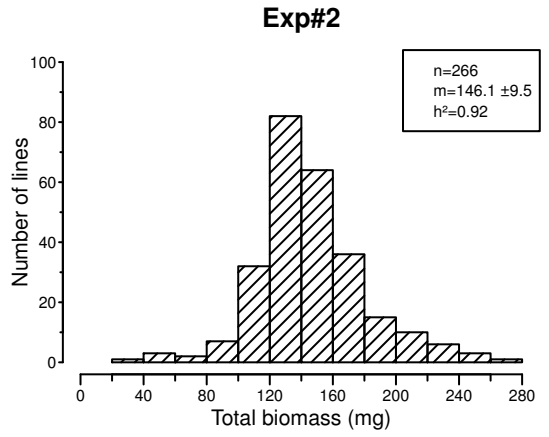
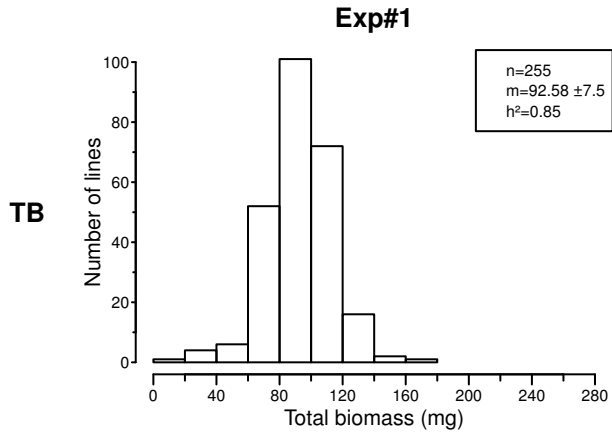
Line Name	Species/Subspecies	Cultivation status (a)	End use (b)	Sowing type (c)	Included in the pea-Aphanomyces collection (d)
CERISE-ce,CR	Pisum sativum	Gmp	Gd		
CHINA	Pisum sativum	Lr	Fd		
SOMMETTE	Pisum sativum	Cv	Gd	Ss	
MESSIRE	Pisum sativum	Cv	Dry	Ss	
ZP141	Pisum sativum	Lr	Gd	Ss	X
ZP126	Pisum sativum	Lr	Gd	Ss	
PISUM SATIVUM-ETHIOPIA JI1594	Pisum sativum	Lr	Gd		
MONGOLIA JI1345	Pisum sativum	Lr			
PISUM SATIVUM-ETHIOPIA JI1431	Pisum sativum	Lr			
COSTA RICA	Pisum sativum	Lr	Gd	Ss	
K4269	Pisum sativum	Lr	Dry	Ss	
K1666	Pisum sativum	Lr		Ss	
K8290/NORD	Pisum sativum	Cv	Dry	Ss	
K4088	Pisum sativum	Lr		Ss	
PISUM SATIVUM-HIBERNICUM JI1846 = E11	Pisum sativum	Cv	Gd		X
YANGWAN	Pisum sativum	Lr	Gd	Ss	
PISUM HUMILE JI241 = E186	Pisum sativum subsp humile	S-Wa/Wa			X
PISUM ELATIUS JI1089	Pisum sativum subsp elatius	S-Wa/Wa			
CE101=FP	Pisum sativum	Bl	Dry	Ws	
CHAMPAGNE	Pisum sativum	Lr	Fd	Ws	
DP	Pisum sativum	Gmp	Fd	Ws	
WIRAIG JI190	Pisum sativum	Lr			
PISUM ELATIUS JI261	Pisum sativum subsp elatius	S-Wa/Wa			
MELROSE	Pisum sativum	Cv	Fd	Ws	
TERESE	Pisum sativum	Cv	Dry	Ss	
BACCARA	Pisum sativum	Cv	Dry	Ss	X
CHEYENNE	Pisum sativum	Cv	Dry	Ws	
GLACIER	Pisum sativum	Cv	Dry	Ws	
HOLLY11	Pisum sativum	Lr	Fd	Ws	
KARNOBAT	Pisum sativum	Cv	Dry	Ws	
KAZAR	Pisum sativum	Cv	Dry	Ws	
TORS DAG	Pisum sativum	Cv	Gd	Ss	
WsBERGER	Pisum sativum	Lr	Fd	Ws	
CHEMIN LONG	Pisum sativum	Cv	Gd	Ss	
CLAMART HATIF	Pisum sativum	Cv	Gd	Ss	
MICHAUX DE PARIS	Pisum sativum	Cv	Gd	Ss	
CHAMPION D'ANGLETERRE	Pisum sativum	Cv	Gd	Ss	
PLEIN LE PANIER	Pisum sativum	Cv	Gd	Ss	
SERPETTE D'AUVERGNE	Pisum sativum	Cv	Gd	Ss	
CAROUBY DE MAUSSANE	Pisum sativum	Cv	Gd	Ss	
CORNE DE BELIER	Pisum sativum	Cv	Gd	Ss	
90-2079	Pisum sativum	Bl	Gd	Ss	X
FRISSON	Pisum sativum	Cv	Dry	Ws	
L1073	Pisum sativum	Bl	Gd		
AUSTIN	Pisum sativum	Cv	Dry	Ss	
271-134	Pisum sativum	S-Wa/Wa			
PISUM SATIVUM-ETHIOPIA JI281	Pisum sativum	Lr			
PUGET	Pisum sativum	Cv	Gd	Ss	X
ISARD	Pisum sativum	Cv	Dry	Ws	X
KAYANNE	Pisum sativum	Cv	Dry	Ss	
ASTRONAUTE	Pisum sativum	Cv	Dry	Ss	
TIL336/11	Pisum sativum	Bl	Dry	Ws	

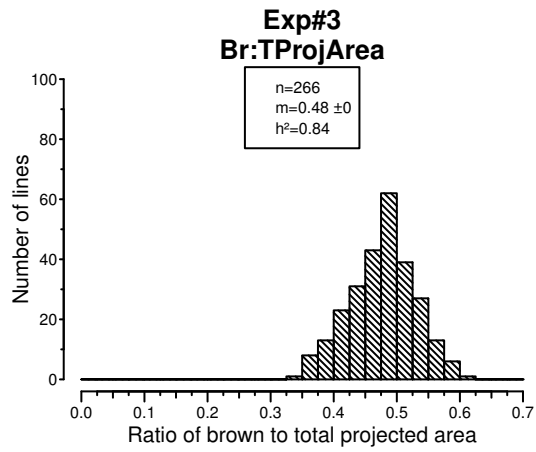
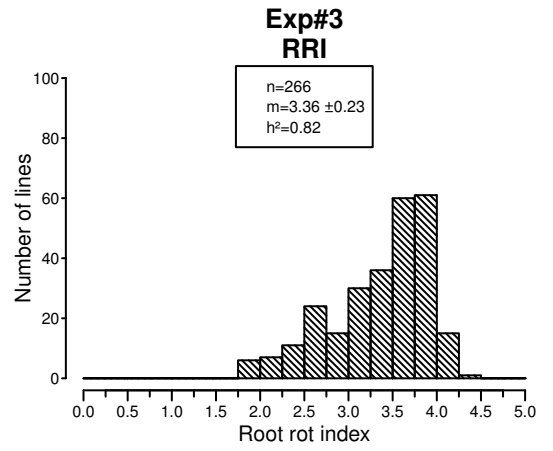
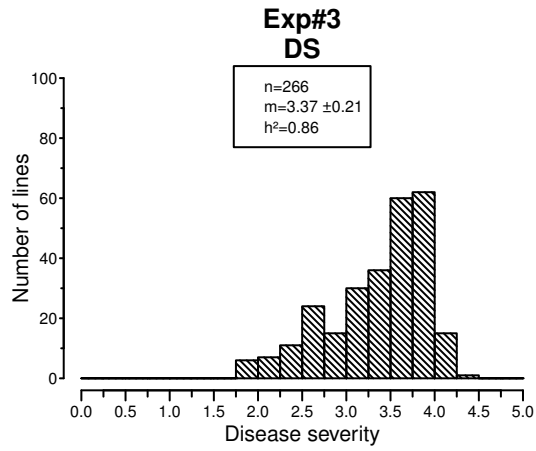
Additional file 2: Frequency distribution of least square means obtained for root, shoot and overall architecture variables, and Aphanomyces resistance in the 266-pea-line collection

Least square means were obtained from analysis of variance for six root system architecture variables (total projected area, average root diameter, total root length, number of lateral roots, average length of lateral roots, root biomass, coded as TProjArea, RootDia, TRootL, NLatRoot, LatRootL, RootB, respectively), two shoot architecture variables (shoot biomass and shoot length, coded as ShootB and ShootL, respectively), two overall plant architecture variables (total biomass and root to total biomass ratio, coded as TB and RootB:TB, respectively) and three Aphanomyces resistance variables (disease severity, root rot index and percentage of brown projected area, coded DS, RRI and Br:TProjArea, respectively). Architectural traits were assessed at 8 days in healthy green house (Exp#1) and at 14 days in healthy climatic chamber (Exp#2). Aphanomyces resistance traits were assessed a reference strain of *A. euteiches* (RB84) in controlled conditions (Exp#3). n: total number of pea lines assessed; m: mean \pm standard deviation of the 266-pea-line collection; h²: mean-based heritability.



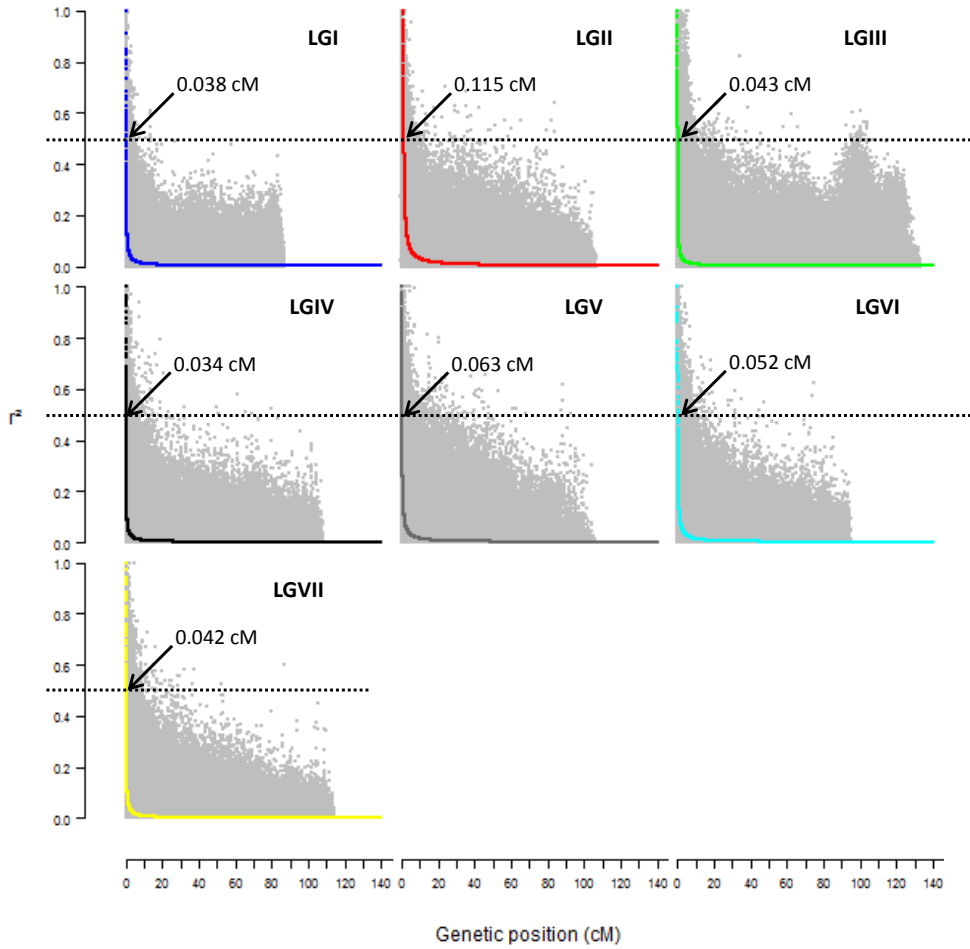






Additional file 3: Linkage disequilibrium (LD) decay in the 266-pea-line collection

Coloured curves represent the estimated LD decay for each linkage group (LG). Dashed vertical lines represent the LD decay (maximum $r^2/2$) and arrows the estimated distance (cM) to reach this LD decay rate on each LG.



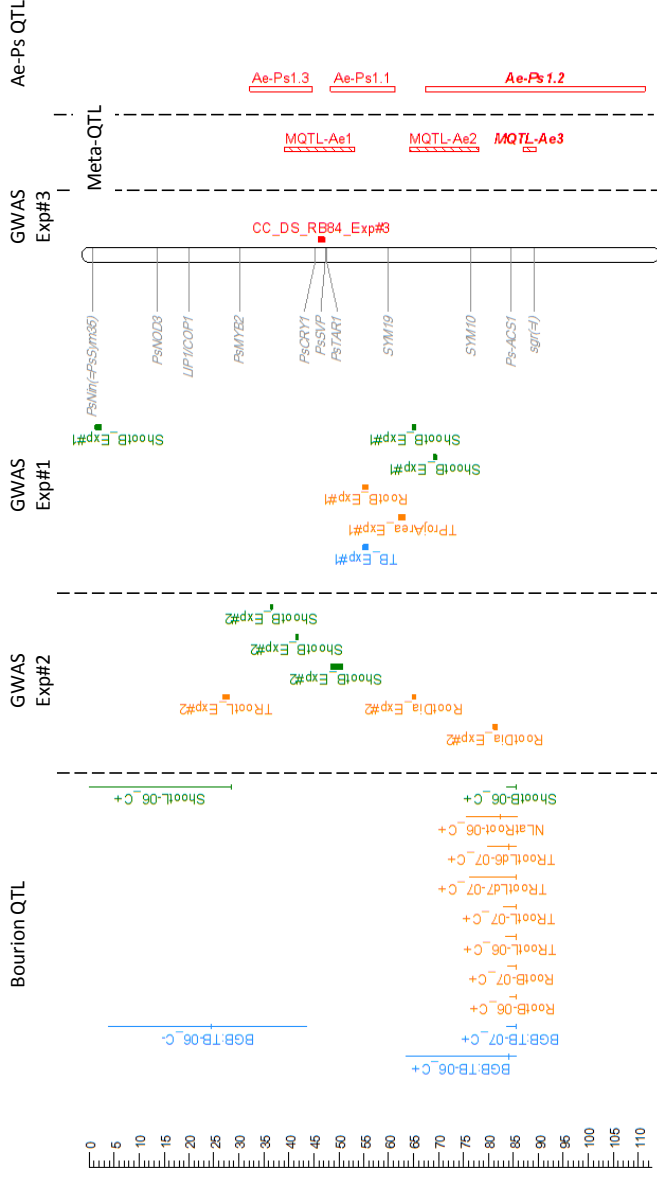
Additional file 4: Comparative genetic map of genome-wide association (GWA) and previously detected linkage quantitative trait loci (QTL) for resistance to *Aphanomyces euteiches* and plant architecture for LGI to VI

The comparative genetic map was constructed from the projection of the 953 markers from Boutet et al (submitted) onto the consensus THMap from Desgroux et al (2016). Linkage groups (LG) are named from I to VI and their size is indicated in cM Haldane. Shoot-architecture-, Root-architecture-, Overall-plant-architecture and Resistance- associated markers and QTL are indicated in green, orange, blue and red, respectively. **To the right from each LG:** Confidence intervals (CIs) around significant Resistance-associated markers, based on linkage disequilibrium (LD) value $r^2 > 0.2$, identified in this study by GWA and name of the trait are indicated (GWAS Exp#3); CIs around significant controlled-conditions-Resistance-associated markers, identified by GWA in Desgroux et al (2016), with the RB84 strain (GWAS Col175); Projected Meta-QTL (Meta-) and QTL (Ae-Ps QTL) described in Hamon et al (2010) and Hamon et al (2013), hatched bars represent Meta-QTL, while blank bars represent initial QTL before meta-analysis Hamon et al (2010) and Hamon et al (2013). The main *Aphanomyces* resistance QTL and Meta-QTL names are in bold italic. **To the left from each LG:** Genomic positions of cloned pea genes are indicated in grey; CIs around architecture-associated markers identified in this study (GWAS Exp#1 and GWAS Exp#2); Projected QTL for root, shoot and plant architecture traits described in Bourion et al (2010) (Bourion QTL).

Plant architecture

Resistance to *A. euteiches*

LGI
(113cM)



Architecture QTL

QTL linkage mapping (Bourion et al 2010)

- Overall plant architecture traits
- Root architecture traits
- Shoot architecture traits

- +
- +
- +

QTL GWA mapping (Desgroux et al in prep)

- Overall plant architecture traits
- Root architecture traits
- Shoot morphology

- +
- +
- +

Resistance QTL

QTL linkage mapping

- Ae-Ps QTL (Hamon et al 2010; Hamon et al 2013)
- Meta-QTL (Hamon et al 2013)

- +
- +

QTL GWA mapping

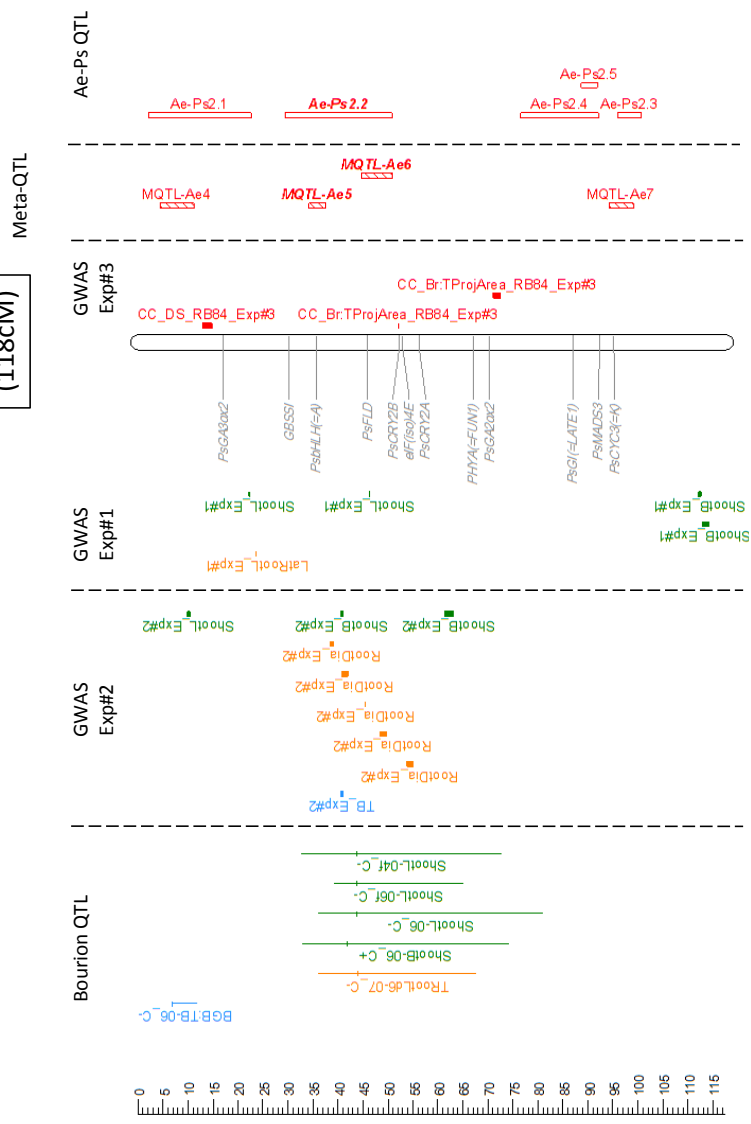
- Root rot traits (Desgroux et al 2016; Desgroux et al in prep)

- +

Plant architecture

Resistance to *A. euteiches*

LGII
(118cM)



Architecture QTL

- QTL linkage mapping** (Bourion et al 2010)
- Overall plant architecture traits
 - Root architecture traits
 - Shoot architecture traits
- QTL GWA mapping** (Desgroux et al in prep)
- Overall plant architecture traits
 - Root architecture traits
 - Shoot morphology

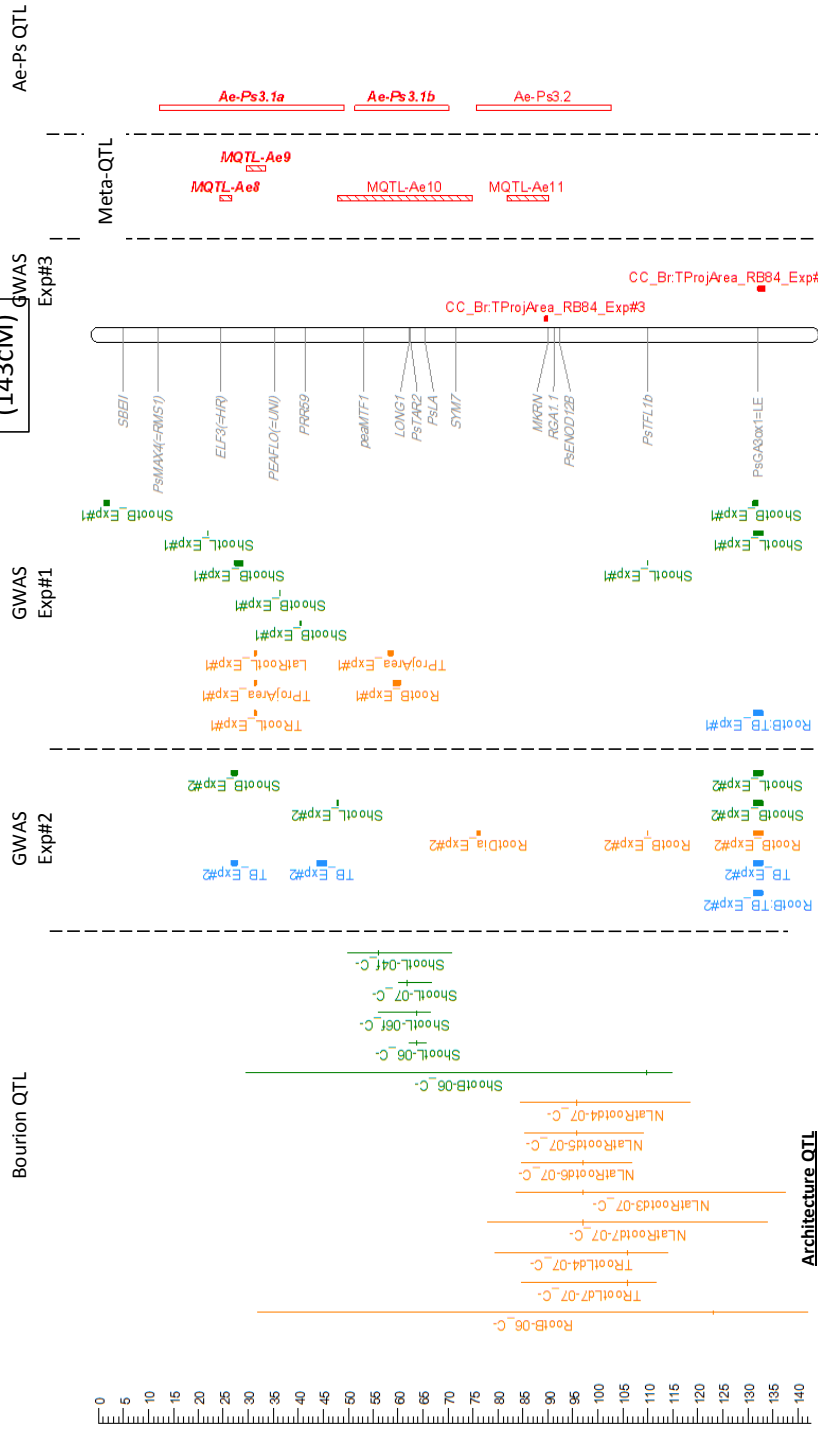
Resistance QTL

- QTL linkage mapping**
- Ae-Ps QTL (Hamon et al 2010; Hamon et al 2013)
 - Meta-QTL (Hamon et al 2013)
- QTL GWA mapping**
- Root rot traits (Desgroux et al 2016; Desgroux et al in prep)

Plant architecture

Resistance to *A. euteiches*

LGIII
(143cM)



Architecture QTL

- Overall plant architecture traits
- Root architecture traits
- Shoot architecture traits

- +
- +
- +

Resistance QTL

- Ae-Ps QTL (Hamon et al 2010; Hamon et al 2013)
- Meta-QTL (Hamon et al 2013)

- +
- +

QTL GWA mapping

- Overall plant architecture traits
- Root architecture traits
- Shoot morphology

- +
- +
- +

QTL linkage mapping

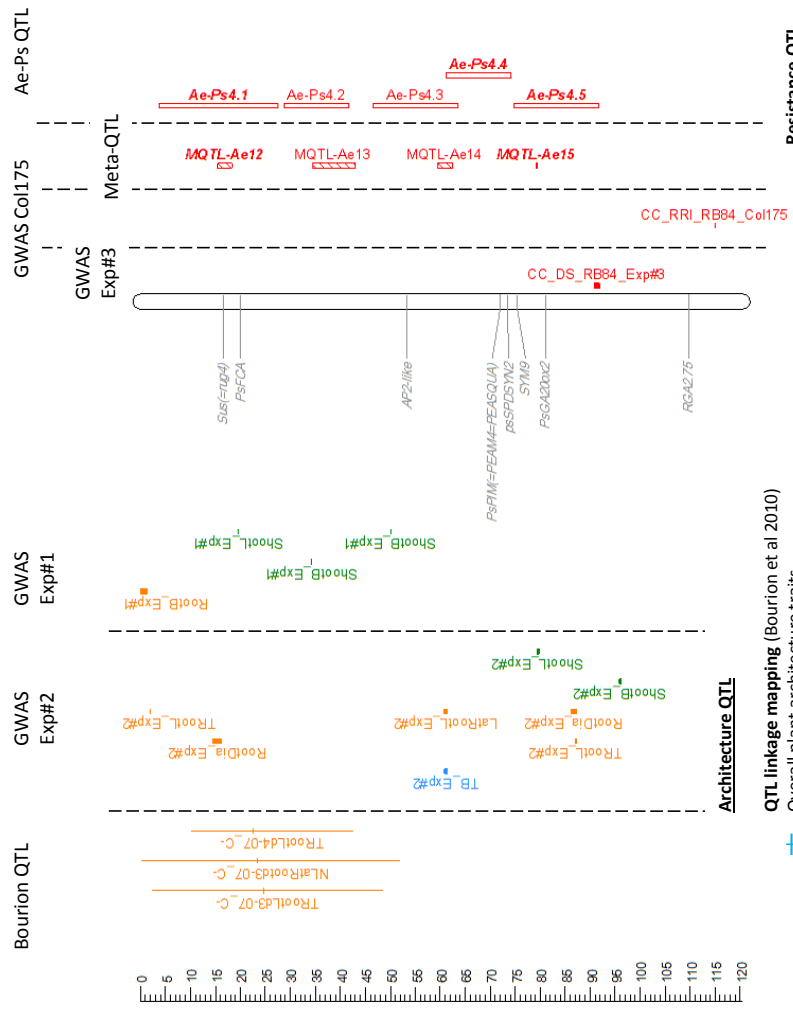
- Root rot traits (Desgroux et al 2016; Desgroux et al in prep)

- +

**LGIV
(121cM)**

Plant architecture

Resistance to *A. euteiches*



QTL linkage mapping (Bourion et al 2010)

- Overall plant architecture traits
- Root architecture traits
- Shoot architecture traits

- QTL GWA mapping** (Desgroux et al in prep)
- Overall plant architecture traits
- Root architecture traits
- Shoot morphology

Resistance QTL

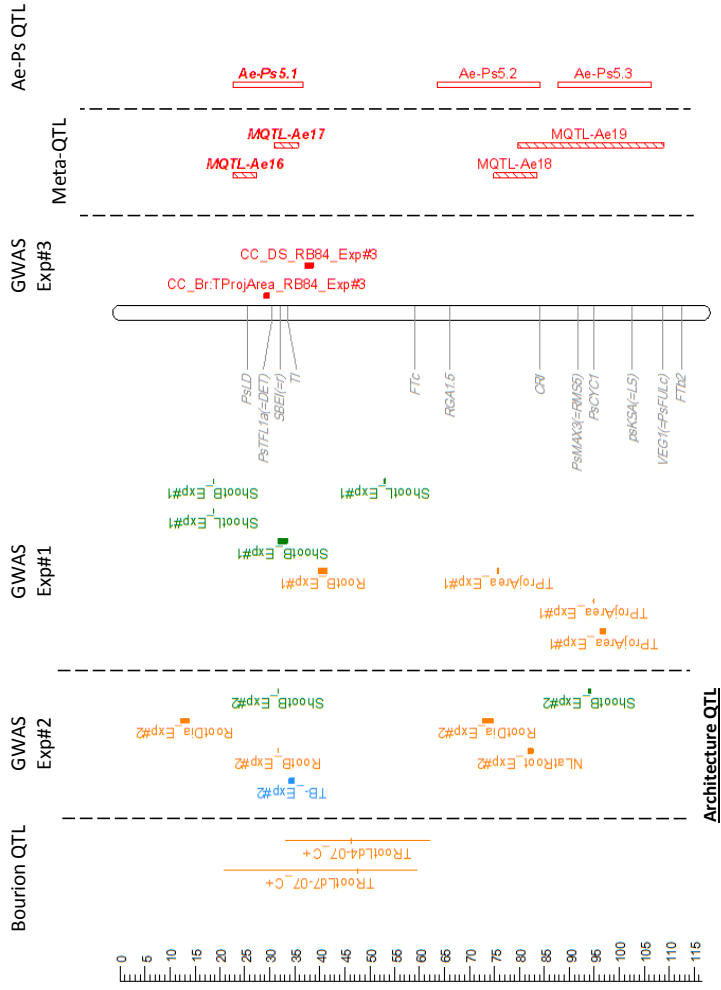
- QTL linkage mapping**
- Ae-Ps QTL (Hamon et al 2010; Hamon et al 2013)
- Meta-QTL (Hamon et al 2013)

- QTL GWA mapping**
- Root rot traits (Desgroux et al 2016; Desgroux et al in prep)

LGV
(117cM)

Plant architecture

Resistance to *A. euteiches*



Architecture QTL

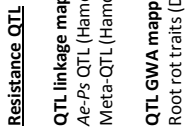
QTL linkage mapping (Bourion et al 2010)
 Overall plant architecture traits
 Root architecture traits
 Shoot architecture traits

QTL GWA mapping (Desgroux et al in prep)
 Overall plant architecture traits
 Root architecture traits
 Shoot morphology

Resistance QTL

QTL linkage mapping
 Ae-Ps QTL (Hamon et al 2010; Hamon et al 2013)
 Meta-QTL (Hamon et al 2013)

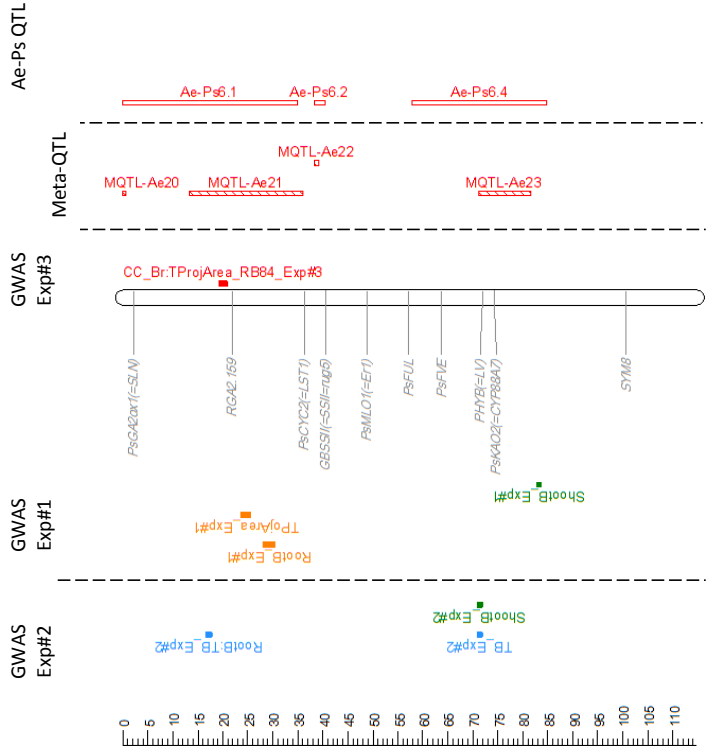
QTL GWA mapping
 Root rot traits (Desgroux et al 2016; Desgroux et al in prep)



Plant architecture

Resistance to *A. euteiches*

LGVI
(115cM)



Architecture QTL

QTL linkage mapping (Bourion et al 2010)

- + Overall plant architecture traits
- + Root architecture traits
- + Shoot architecture traits

QTL GWA mapping (Desgroux et al in prep)

- █ Overall plant architecture traits
- █ Root architecture traits
- █ Shoot morphology

Resistance QTL

QTL linkage mapping

- █ Ae-Ps QTL (Hamon et al 2010; Hamon et al 2013)
- █ Meta-QTL (Hamon et al 2013)

QTL GWA mapping

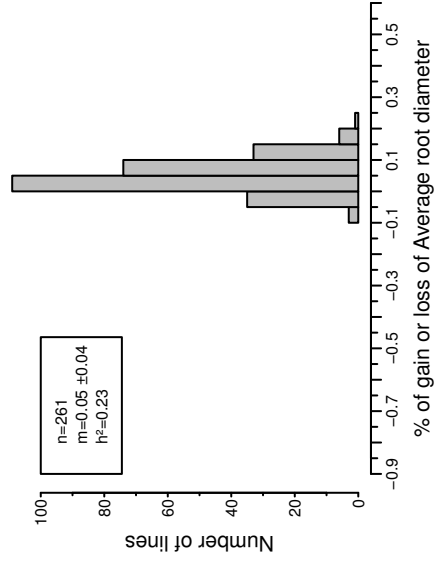
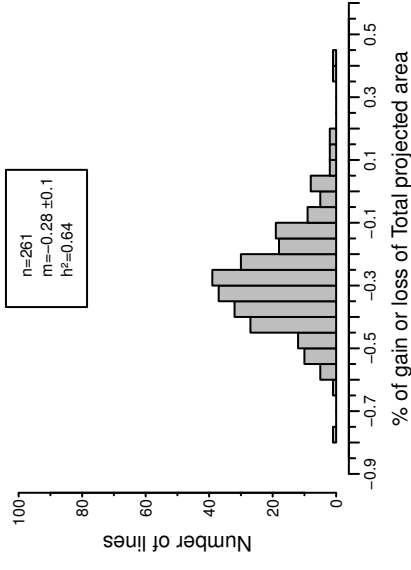
- █ Root rot traits (Desgroux et al 2016; Desgroux et al in prep)

Annexe 3 : Additional files du Chapitre IV

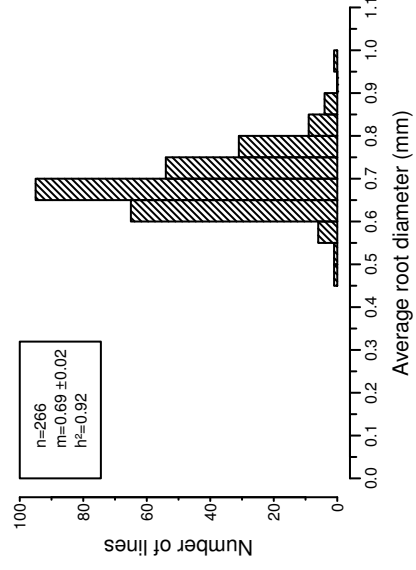
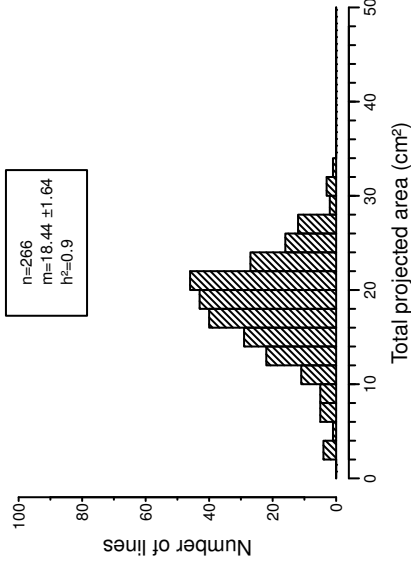
Additional file IV.1: Frequency distribution of least square means obtained for plant architecture variables in non-inoculated and inoculated controlled conditions, and disease-induced architecture traits in the 266-pea-line collection

Least square means were obtained from analysis of variance for six root architecture variables (total projected area, average root diameter, total root length, number of lateral roots, average length of lateral roots, root biomass, coded as TProjArea, RootDia, TRootL, NLatRoot, LatRootL, RootB, respectively), two shoot architecture variables (shoot biomass and shoot length, coded as ShootB and ShootL, respectively), two overall plant architecture variables (total biomass and root to total biomass ratio, coded as TB and RootB:TB, respectively) and percentage of gain or loss of all those variables (inoculated vs. non-inoculated ratios). Architecture traits in non-inoculated conditions were assessed at 14 days in a climatic chamber (Desgroux et al in prep, Exp#2). Architecture traits in inoculated conditions were assessed in the same climatic chamber, and plants were inoculated seven days post sowing with a reference strain of *A. euteiches* (RB84). Percentages of gain or loss were calculated as inoculated vs non-inoculated ratios ($X_{ratio} = (X_{I+} - X_{I-}) \times 100 / X_{I-}$, X: variable, I+: inoculated, I-: non-inoculated); n: total number of pea lines assessed; m: mean \pm standard deviation of the 266-pea-line collection; h^2 : mean-based heritability.

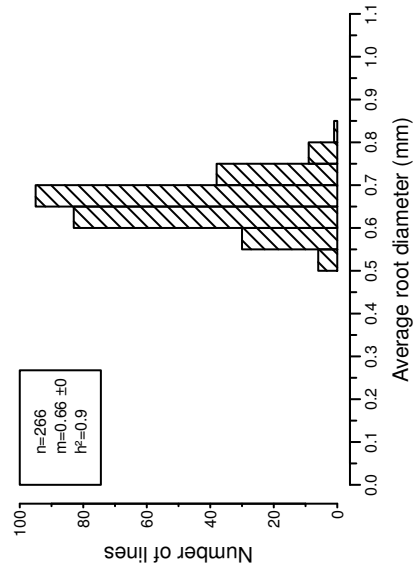
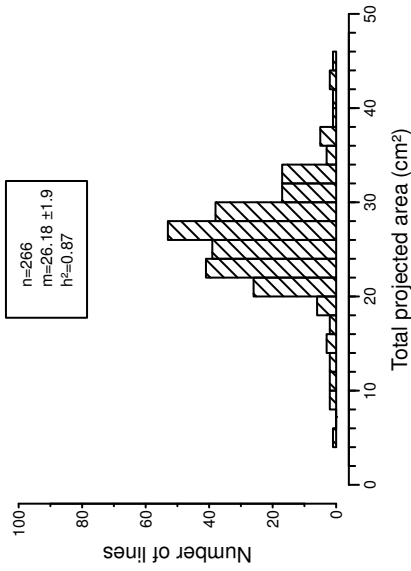
Inoculated vs non-inoculated ratios



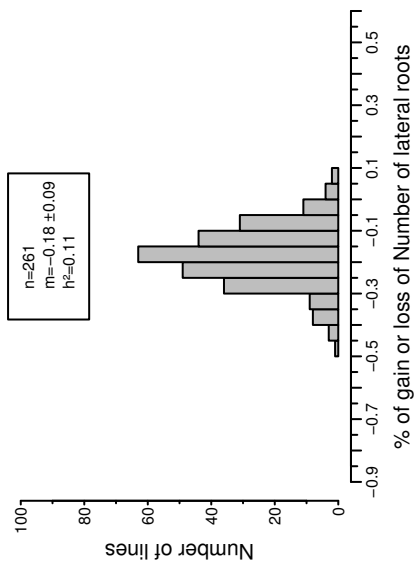
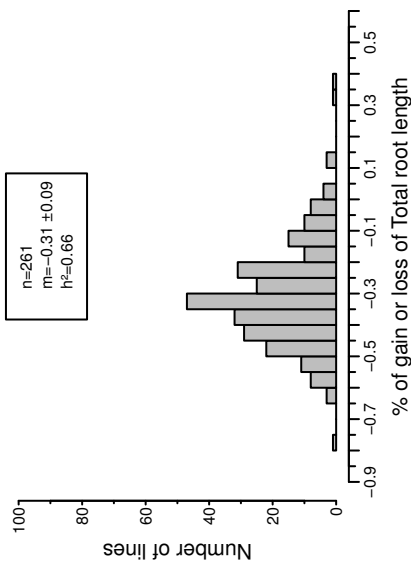
Inoculated



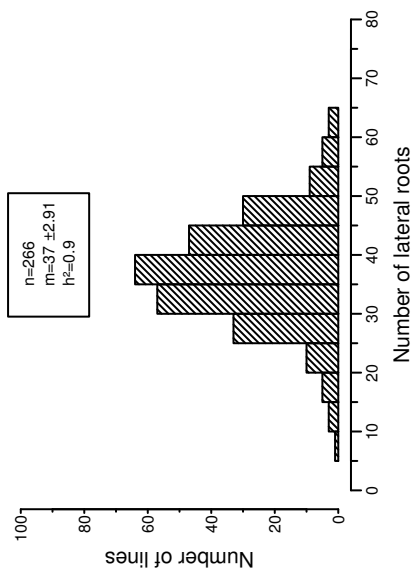
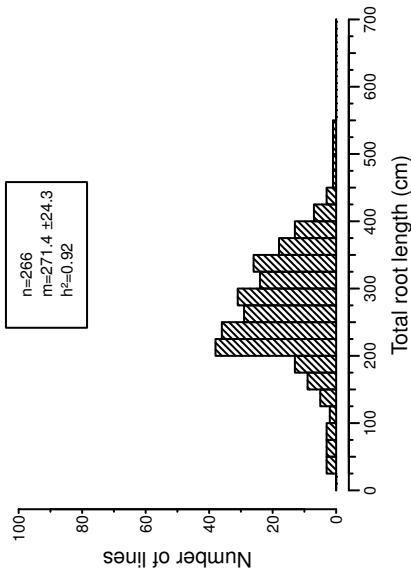
Non-inoculated



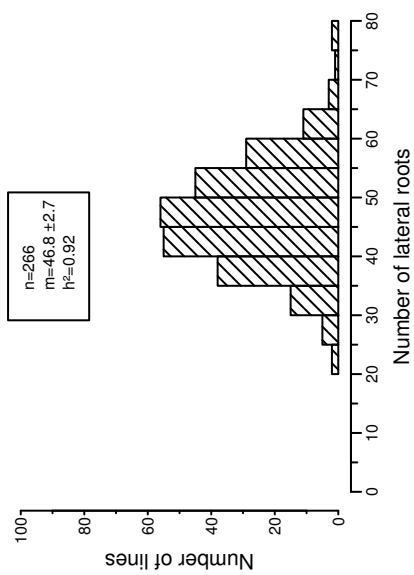
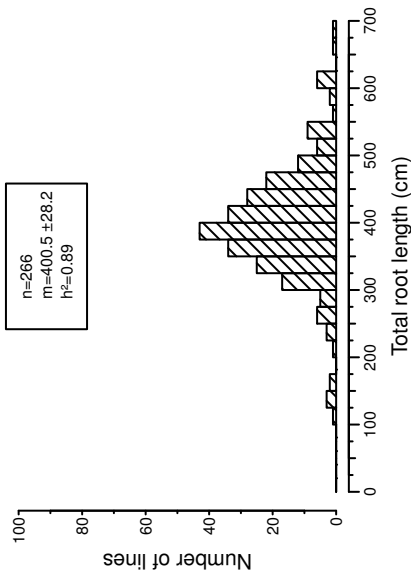
Inoculated vs non-inoculated ratios



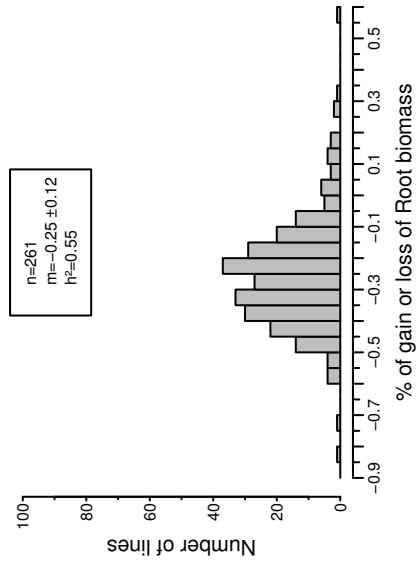
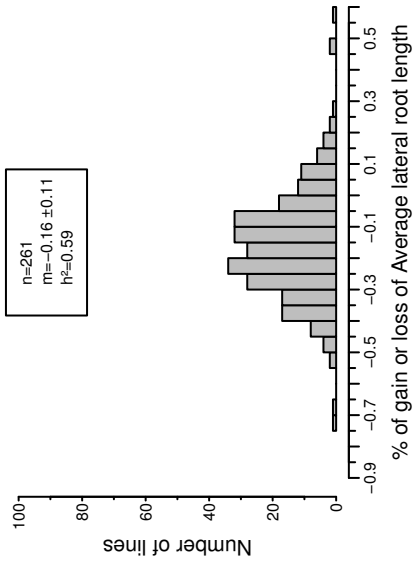
Inoculated



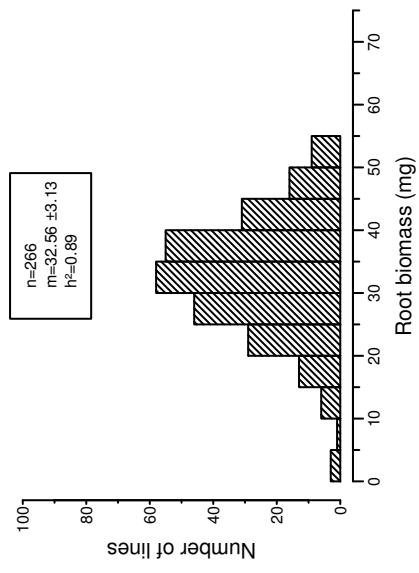
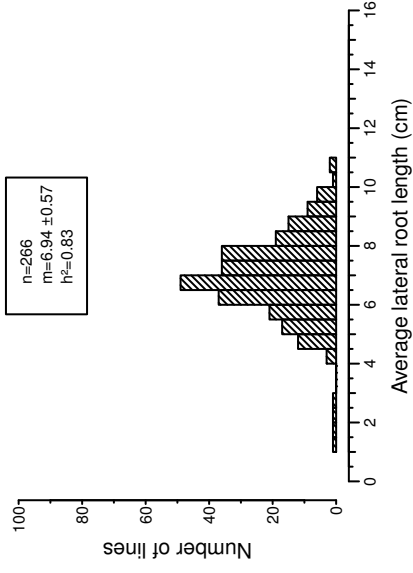
Non-inoculated



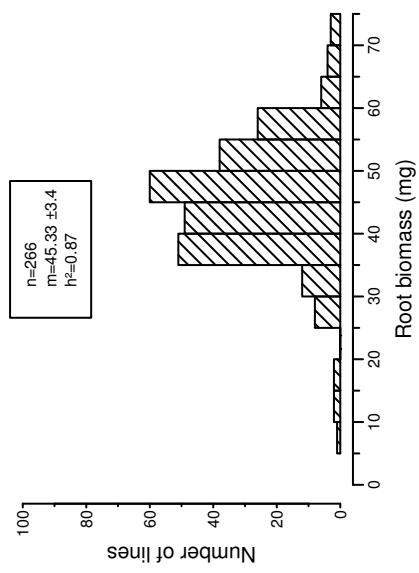
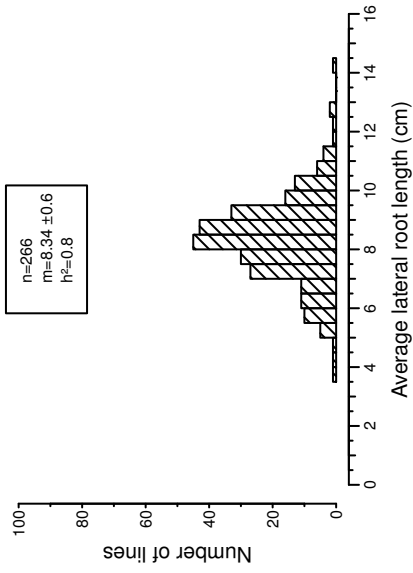
Inoculated vs non-inoculated ratios



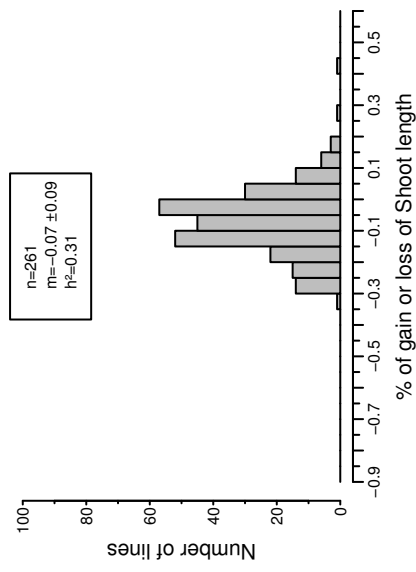
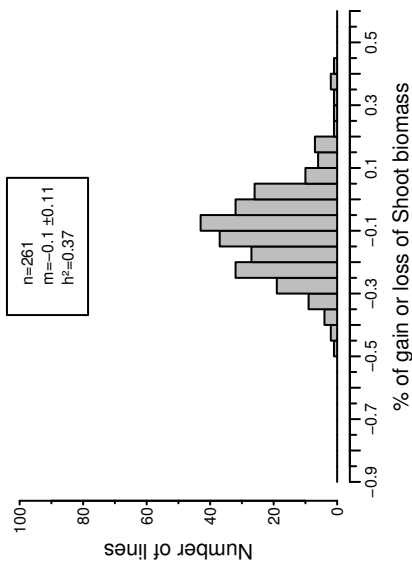
Innoculated



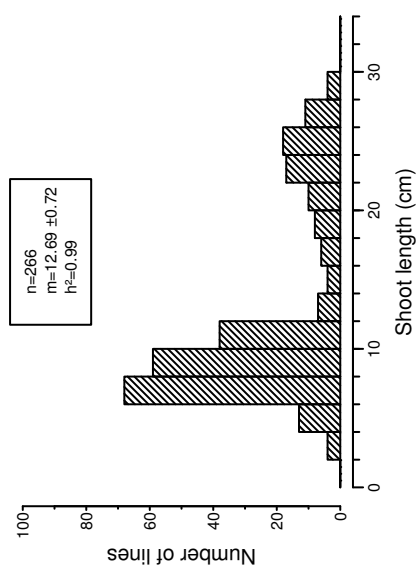
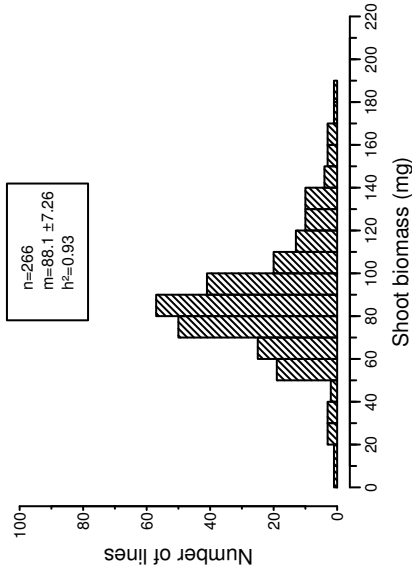
Non-inoculated



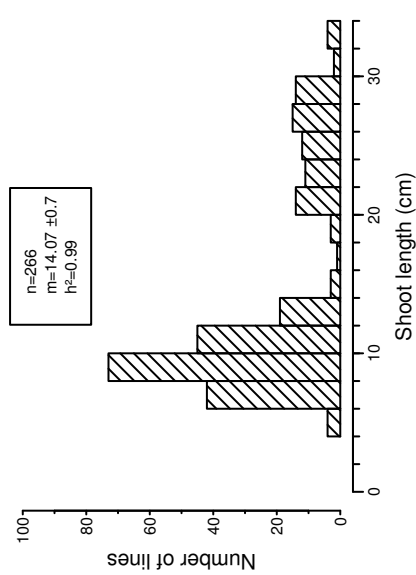
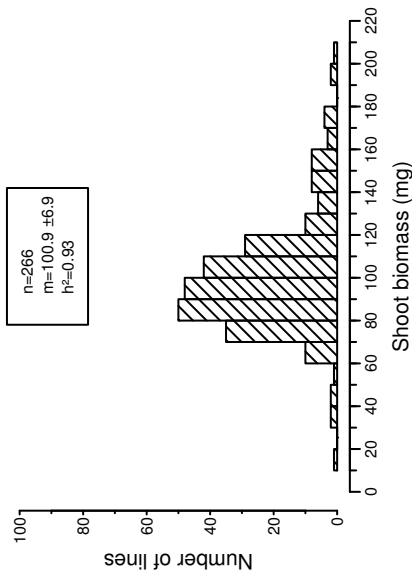
Inoculated vs non-inoculated ratios



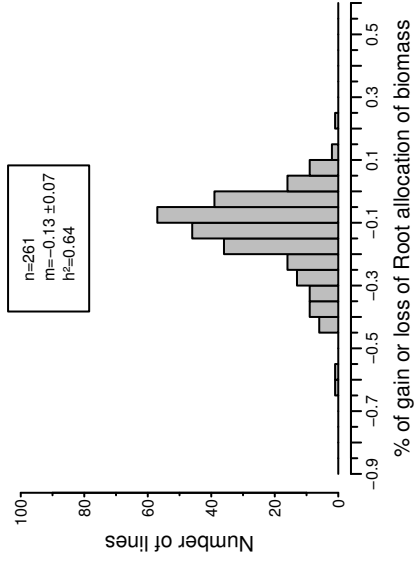
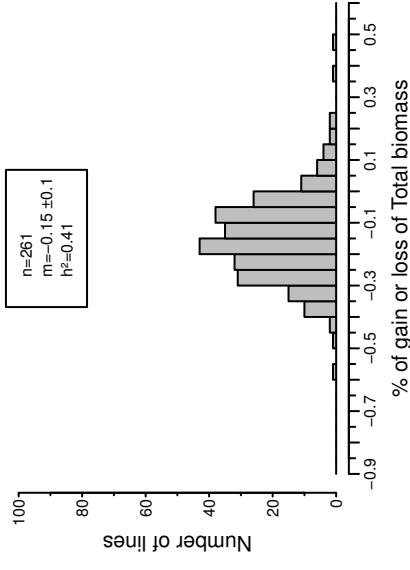
Inoculated



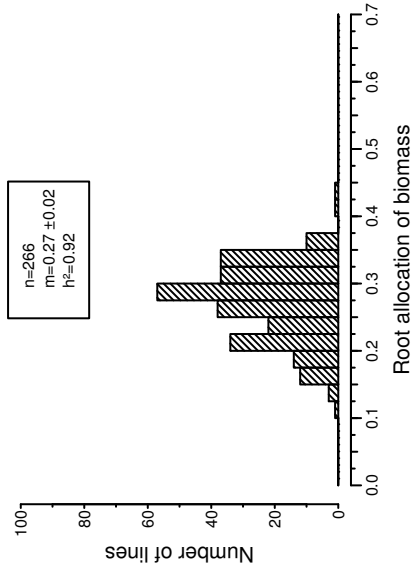
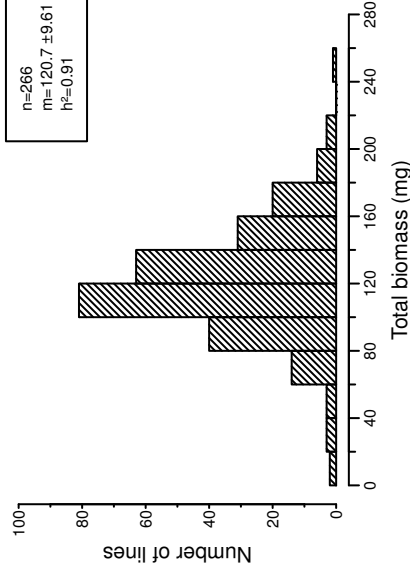
Non-inoculated



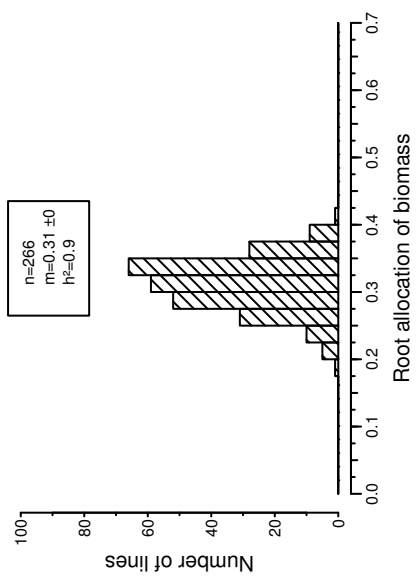
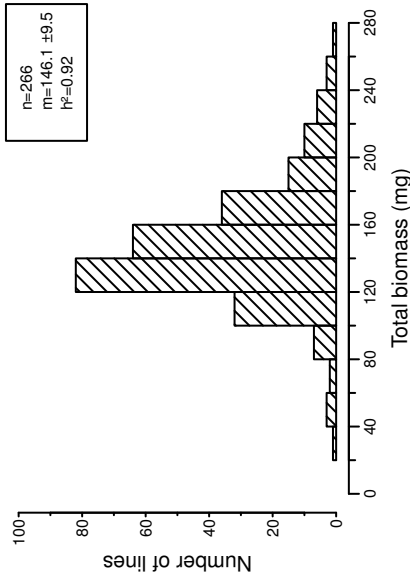
Inoculated vs non-inoculated ratios



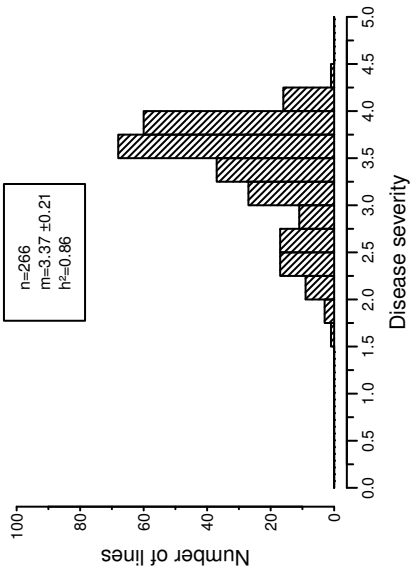
Inoculated



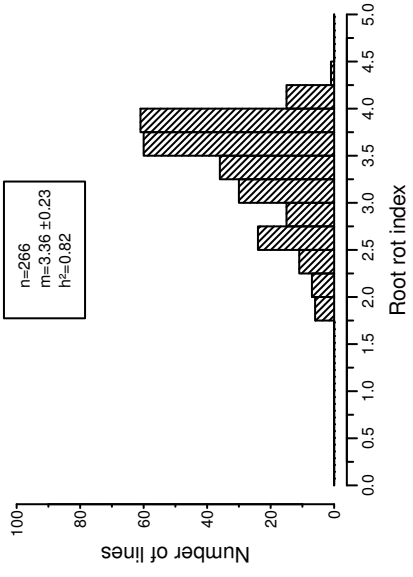
Non-inoculated



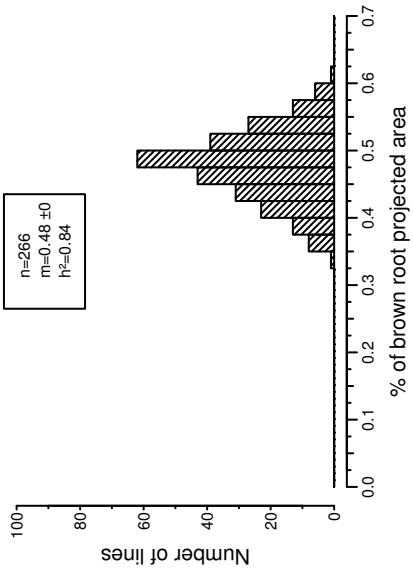
Inoculated DS



Inoculated RRI



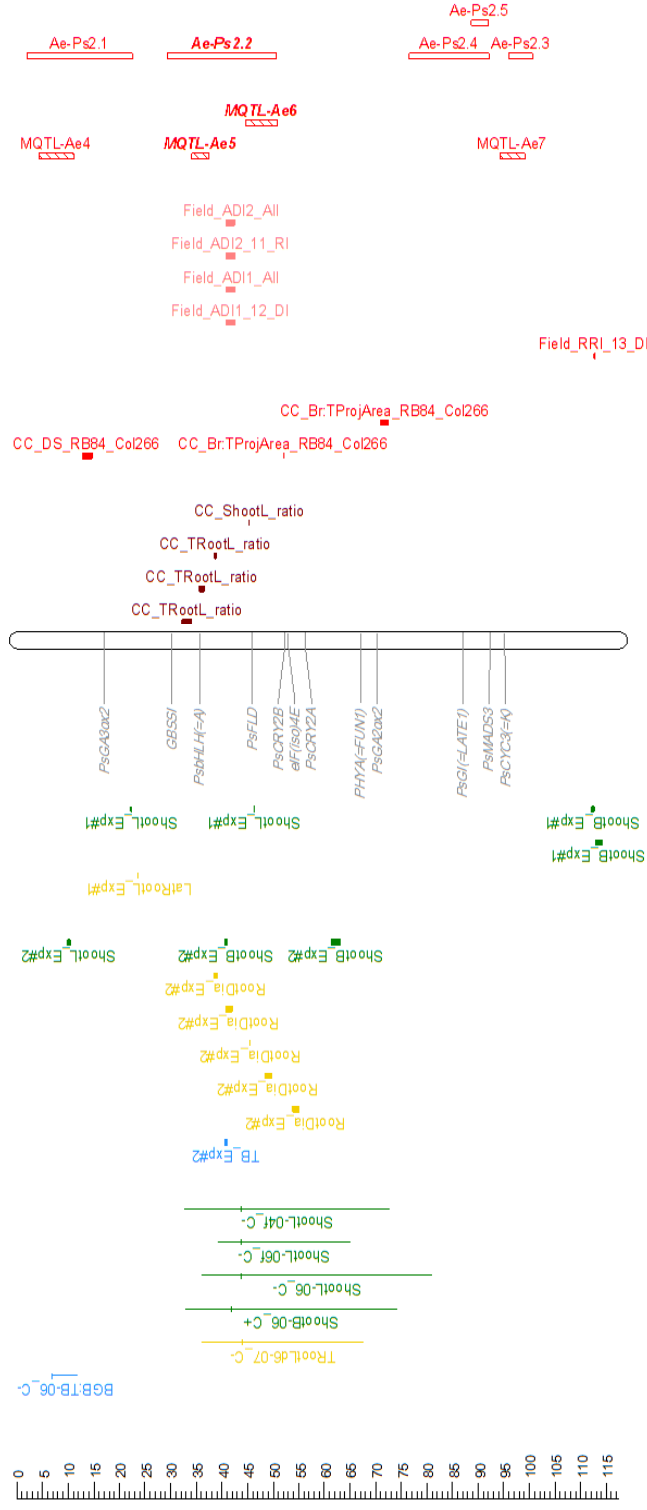
Inoculated Br:TProjArea



Additional file IV.2: Comparative genetic map of genome-wide association (GWA) and previously detected linkage quantitative trait loci (QTL) for resistance to *Aphanomyces euteiches*, constitutive plant architecture and disease-induced plant architecture for linkage groups I to VI

The comparative genetic map was constructed from the projection of the 953 markers from Boutet et al (submitted) onto the consensus THMap from Desgroux et al (2016). Only linkage groups (LG) I to VI are shown (LGVII is in Figure 1). Its size is indicated in cM Haldane. **To the right from each LG:** Confidence intervals (CIs) around significant markers associated with disease-induced architecture traits (inoculated vs non-inoculated ratios), based on linkage disequilibrium (LD) value $r^2 > 0.2$, CIs around significant Resistance-associated markers, identified by GWA in Desgroux et al (in prep); CIs around significant Resistance-associated markers, identified by GWA in Desgroux et al (2016), with the RB84 strain (GWAS Col175); Projected Meta-QTL (Meta-) and QTL (Ae-Ps QTL) described in Hamon et al (2010) and Hamon et al (2013), hatched bars represent Meta-QTL, while blank bars represent initial QTL before meta-analysis Hamon et al (2010) and Hamon et al (2013). The main *Aphanomyces* resistance QTL and Meta-QTL names are in bold italic. **To the left from each LG:** Genomic positions of cloned pea genes are indicated in grey; CIs around architecture-associated markers identified in Desgroux et al (in prep); Projected QTL for root, shoot and plant architecture traits described in Bourion et al (2010). ADI: Aerial Decline Index; BGB:TB: below ground to total biomass ratio; Br:TProjArea: % of brown root projected area; CC: Controlled conditions; Exp#1: Greenhouse non-inoculated controlled conditions at 8 days post sowing; Exp#2: Climate chamber non-inoculated controlled conditions at 14 days post sowing; MFA: Multiple factor analysis; RB84: French reference strain of *A. euteiches*; RootB: Root biomass; RootDia: Average root diameter; RRI: Root Rot Index; ShootB: Shoot biomass; ShootL: Shoot length; TB: Total biomass; TProjArea: Total root projected area; TRootL: Total root length

**LGII
(118cM)**



Resistance QTL

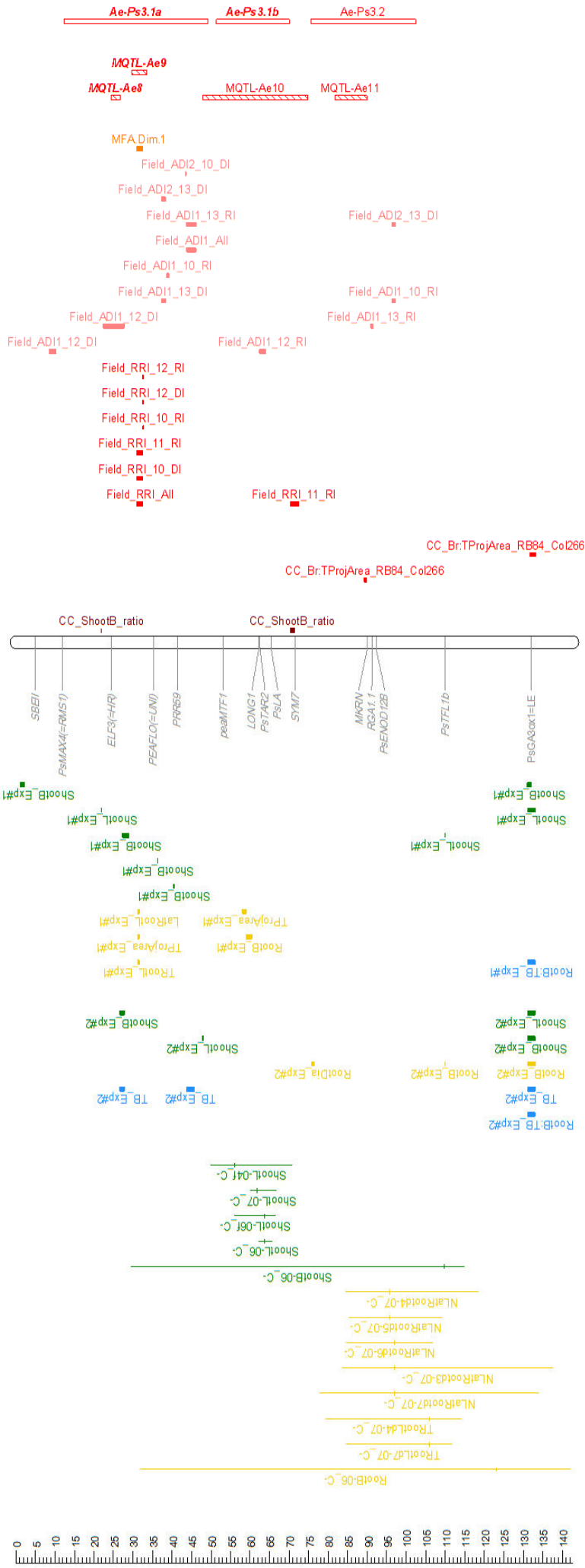
QTL linkage mapping
 Ae-Ps QTL (Hamon et al 2010; Hamon et al 2013)
 Meta-QTL (Hamon et al 2013)

Architecture QTL

QTL linkage mapping (Bourion et al 2010)
 Overall plant architecture traits
 Root architecture traits
 Shoot architecture traits

QTL GWA mapping (Desgroux et al in prep)
 Overall plant architecture traits
 Root architecture traits
 Shoot morphology

**LGIII
(143cM)**



Resistance QTL

QTL linkage mapping

Ae-Ps QTL (Hamon et al 2010; Hamon et al 2013)
Meta-QTL (Hamon et al 2013)



QTL GWA mapping

Coordinated on first and second axis of a Multiple factor analysis (Desgroux et al 2016)
Aerial decline index (Desgroux et al 2016)
Root rot traits (Desgroux et al 2016; Desgroux et al in prep)
Disease-induced architecture traits (inoculated vs non-inoculated ratios) (this study)



Architecture QTL

QTL linkage mapping (Bourion et al 2010)

Overall plant architecture traits
Root architecture traits
Shoot architecture traits

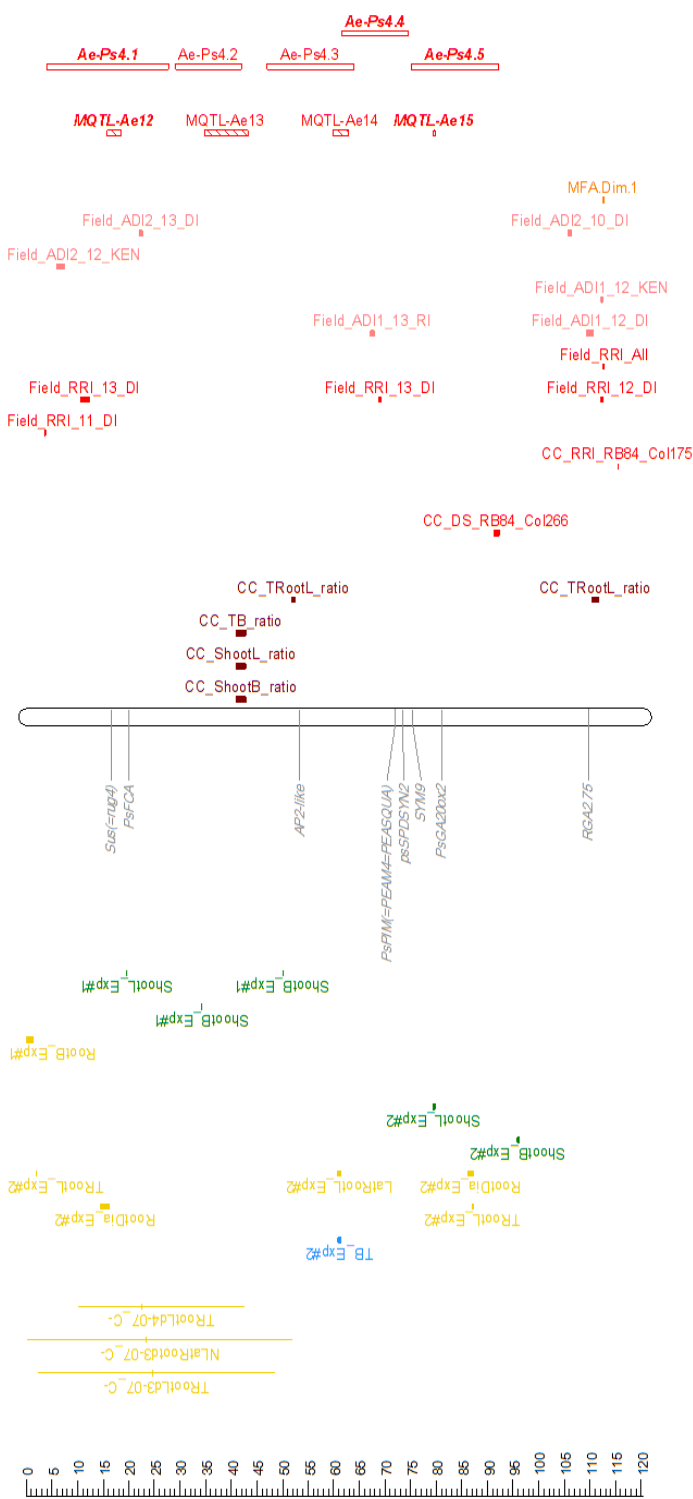


QTL GWA mapping (Desgroux et al in prep)

Overall plant architecture traits
Root architecture traits
Shoot morphology



LGIV (121cM)



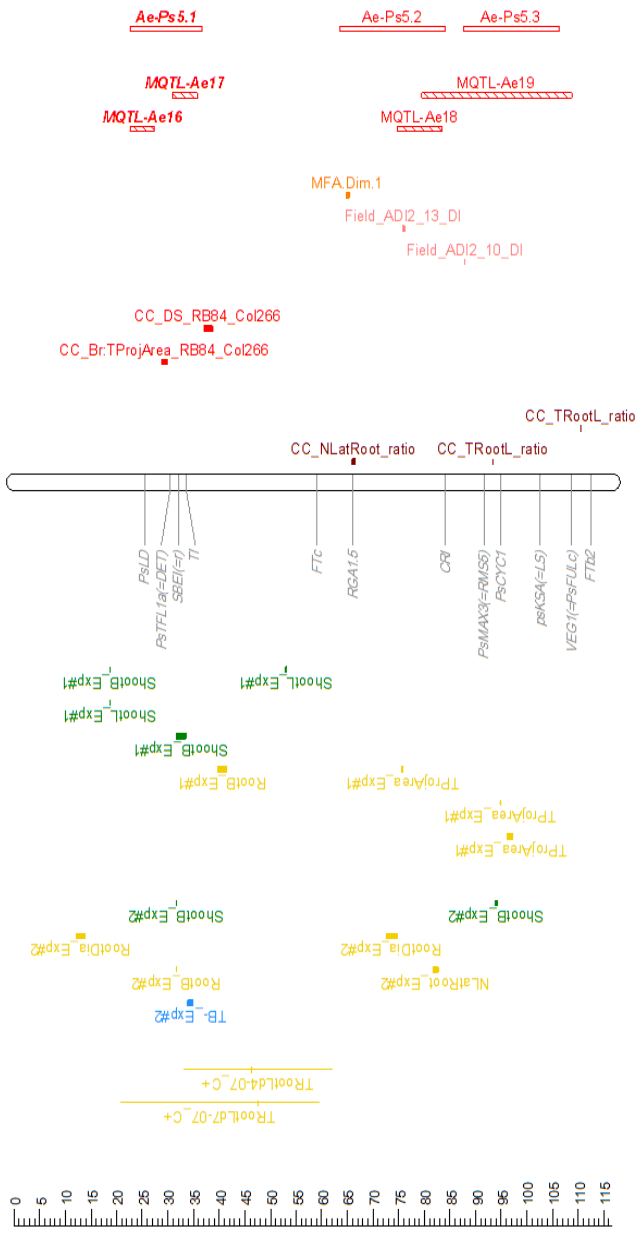
Architecture QTL

- QTL linkage mapping** (Bourion et al 2010)
 - Overall plant architecture traits
 - Root architecture traits
 - Shoot architecture traits
- QTL GWA mapping** (Desgroux et al in prep)
 - Overall plant architecture traits
 - Root architecture traits
 - Shoot morphology

Resistance QTL

- QTL linkage mapping**
 - Ae-Ps QTL (Hamon et al 2010; Hamon et al 2013)
 - Meta-QTL (Hamon et al 2013)
- QTL GWA mapping**
 - Coordinated on first and second axis of a Multiple factor analysis (Desgroux et al 2016)
 - Aerial decline index (Degroux et al 2016)
 - Root rot traits (Desgroux et al 2016; Desgroux et al in prep)
 - Disease-induced architecture traits (inoculated vs non-inoculated ratios) (this study)

LGV (117cM)



Resistance QTL

QTL linkage mapping
 Ae-Ps QTL (Hamon et al 2010; Hamon et al 2013)
 Meta-QTL (Hamon et al 2013)

QTL GWA mapping

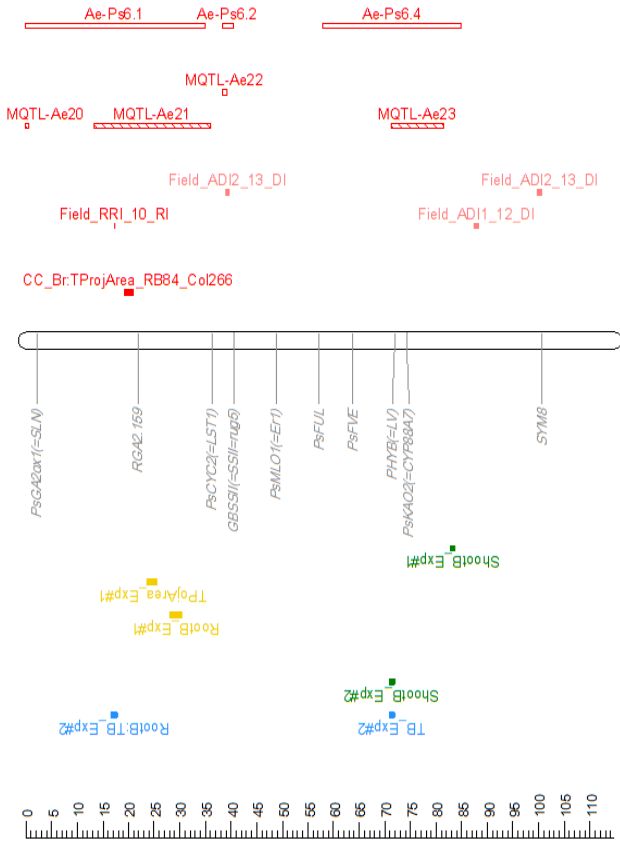
Coordinated on first and second axis of a Multiple factor analysis (Desgroux et al 2016)
 Aerial decline index (Degroux et al 2016)
 Root rot traits (Desgroux et al 2016; Desgroux et al in prep)
 Disease-induced architecture traits (inoculated vs non-inoculated ratios) (this study)

Architecture QTL

QTL linkage mapping (Bourion et al 2010)
 Overall plant architecture traits
 Root architecture traits
 Shoot architecture traits

QTL GWA mapping (Desgroux et al in prep)
 Overall plant architecture traits
 Root architecture traits
 Shoot morphology

LGVI (115cM)



Architecture QTL

- | **QTL linkage mapping** (Bourion et al 2010)
- | Overall plant architecture traits
- | Root architecture traits
- | Shoot architecture traits
- | **QTL GWA mapping** (Desgroux et al in prep)
- | Overall plant architecture traits
- | Root architecture traits
- | Shoot morphology

Resistance QTL

- | **QTL linkage mapping**
- | Ae-Ps QTL (Hamon et al 2010; Hamon et al 2013)
- | Meta-QTL (Hamon et al 2013)
- | **QTL GWA mapping**
- | Coordinated on first and second axis of a Multiple factor analysis (Desgroux et al 2016)
- | Aerial decline index (Degroux et al 2016)
- | Root rot traits (Desgroux et al 2016; Desgroux et al in prep)
- | Disease-induced architecture traits (inoculated vs non-inoculated ratios) (this study)

Rôle de l'architecture racinaire dans le contrôle génétique de la diminution des symptômes de pourriture racinaire dus à *Aphanomyces euteiches* chez le pois (*Pisum sativum*)

Dans un contexte d'agriculture durable, la création de variétés combinant des facteurs génétiques de résistance et d'architecture de la plante présente un intérêt majeur pour limiter le développement des maladies. La thèse a visé à analyser l'interdépendance des déterminants génétiques de l'architecture racinaire et de la résistance à *A. euteiches*, agent pathogène tellurique majeur du pois. Une analyse de génétique d'association conduite sur tout le génome, à partir d'une collection de 266 lignées de pois, a permis d'identifier 150 locus de résistance et d'architecture et de les comparer aux QTL précédemment étudiés. Une région commune à effet majeur a été associée à la résistance et à un caractère de développement racinaire intrinsèque. Plusieurs régions ont été détectées à la fois pour la résistance et des caractères de vigueur racinaire en réponse à l'infection. La combinaison de la résistance et de caractères d'architecture dans une sélection de lignées a été associée à une réduction des pertes de rendement au champ. Les résultats ont permis d'apporter des connaissances originales sur la génétique comparative de la résistance aux maladies et de l'architecture des plantes au niveau racinaire. Ils fournissent des outils, géniteurs et éléments de choix de QTL de résistance et d'architecture racinaire à combiner pour la sélection de futures variétés de pois résistantes à *A. euteiches*.

Mots clés : *Pisum sativum*, pourriture racinaire, résistance partielle, tolérance, vigueur racinaire, génétique d'association

Role of root architecture in the genetic control of decrease of root rot symptom caused by *Aphanomyces euteiches* in pea (*Pisum sativum*)

In a sustainable agriculture context, breeding for varieties combining genetic resistance and plant architecture is of major interest to limit diseases impacts in crops. The aim of this thesis was to analyze the interdependence of the genetic determinants of root architecture and resistance to *Aphanomyces euteiches*, a major soil born pathogen of peas. A genome wide association analysis among a 266-pea-line collection enabled us to pinpoint a total of 150 loci associated with resistance and plant architecture, and to compare them with QTL detected in previous studies. A common major locus was associated with resistance and an intrinsic root architecture trait. Several loci were detected for both resistance and disease-induced architecture. The combination of resistance and some architecture traits in selected pea lines was associated with reduced yield losses in infested fields. Results provide original knowledge on comparative genetics of disease resistance and plant architecture for root rot diseases. They provide tools (SNPs and marker haplotypes), parental lines and information for the choice of resistance and architecture QTL to combine in breeding strategies to improve resistance to *A. euteiches* in future pea varieties.

Keywords: *Pisum sativum*, root rot, partial resistance, tolerance, root vigor, association mapping

RÉSUMÉ ABSTRACT

Rôle de l'architecture racinaire dans le contrôle génétique de la diminution des symptômes de pourriture racinaire dus à *Aphanomyces euteiches* chez le pois (*Pisum sativum*)

Dans un contexte d'agriculture durable, la création de variétés combinant des facteurs génétiques de résistance et d'architecture de la plante présente un intérêt majeur pour limiter le développement des maladies. La thèse a visé à analyser l'interdépendance des déterminants génétiques de l'architecture racinaire et de la résistance à *A. euteiches*, agent pathogène tellurique majeur du pois. Une analyse de génétique d'association conduite sur tout le génome, à partir d'une collection de 266 lignées de pois, a permis d'identifier 150 locus de résistance et d'architecture et de les comparer aux QTL précédemment étudiés. Une région commune à effet majeur a été associée à la résistance et à un caractère de développement racinaire intrinsèque. Plusieurs régions ont été détectées à la fois pour la résistance et des caractères de vigueur racinaire en réponse à l'infection. La combinaison de la résistance et de caractères d'architecture dans une sélection de lignées a été associée à une réduction des pertes de rendement au champ. Les résultats ont permis d'apporter des connaissances originales sur la génétique comparative de la résistance aux maladies et de l'architecture des plantes au niveau racinaire. Ils fournissent des outils, géniteurs et éléments de choix de QTL de résistance et d'architecture racinaire à combiner pour la sélection de futures variétés de pois résistantes à *A. euteiches*.

Mots-clefs : *Pisum sativum*, pourriture racinaire, résistance partielle, tolérance, vigueur racinaire, génétique d'association

Role of root architecture in the genetic control of decrease of root rot symptom caused by Aphanomyces euteiches in pea (Pisum sativum)

In a sustainable agriculture context, breeding for varieties combining genetic resistance and plant architecture is of major interest to limit diseases impacts in crops. The aim of this thesis was to analyze the interdependence of the genetic determinants of root architecture and resistance to Aphanomyces euteiches, a major soil born pathogen of peas. A genome wide association analysis among a 266-pea-line collection enabled us to pinpoint a total of 150 loci associated with resistance and plant architecture, and to compare them with QTL detected in previous studies. A common major locus was associated with resistance and an intrinsic root architecture trait. Several loci were detected for both resistance and disease-induced architecture. The combination of resistance and some architecture traits in selected pea lines was associated with reduced yield losses in infested fields. Results provide original knowledge on comparative genetics of disease resistance and plant architecture for root rot diseases. They provide tools (SNPs and marker haplotypes), parental lines and information for the choice of resistance and architecture QTL to combine in breeding strategies to improve resistance to A. euteiches in future pea varieties.

Keywords: *Pisum sativum*, root rot, partial resistance, tolerance, root vigor, association mapping



AGROCAMPUS OUEST • Institut supérieur des sciences agronomiques, agroalimentaires, horticoles et du paysage
65 rue de Saint-Brieuc – CS84215 – F-35042 Rennes Cedex
Tél. : 02 23 48 50 00
www.agrocampus-ouest.fr

