Identification of genetic bases associated with natural variation of plant-plant interactions in arabidopsis thaliana

Cyril Libourel

THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par l'Université Toulouse 3 - Paul Sabatier

Présentée et soutenue par

Cyril LIBOUREL

Le 19 mars 2019

Identification des bases génétiques associées à la variation naturelle des interactions plante-plante chez Arabidopsis thaliana

Ecole doctorale : **SEVAB - Sciences Ecologiques, Vétérinaires, Agronomiques et Bioingenieries**

Spécialité : **Interactions plantes-microorganismes**

Unité de recherche :
LIPM - Laboratoire des Interactions Plantes-Microorganismes

Thèse dirigée par
Fabrice ROUX et Dominique ROBY

Jury

Mme Jöelle RONFORT, Rapporteur
M. Philippe REYMOND, Rapporteur
Mme Catherine RAMEAU, Examinateur
M. Christophe THEBAUD, Examinateur
M. Fabrice ROUX, Directeur de thèse
Mme Dominique ROBY, Co-directeur de thèse

Remerciements

Par ces quelques lignes, je tiens à remercier toutes les personnes qui m'ont accompagné, qui m'ont aidé, qui ont cru en moi et qui m'ont soutenu lors de ces trois années de thèse mais aussi bien avant.

Je tiens tout d'abord à remercier mon jury de thèse Christophe Thébaud, Joëlle Ronfort, Philippe Reymond et Catherine Rameau, d'avoir pris le temps de lire et d'évaluer mon travail. Je tiens également à remercier Celine Richard-Molard et Etienne Pascal Journet qui m'ont apporté conseils et bienveillance lors de mes deux comités de thèse. Enfin, je remercie le ministère de la recherche et de l'enseignement supérieur français qui a financé cette thèse, ainsi que le LIPM pour avoir financé trois mois supplémentaires indispensables pour mener à bien mes derniers travaux ainsi qu'un congrès mémorable au Brésil.

J'aimerais poursuivre ces remerciements par mes incroyables superviseurs, Fabrice et Dominique. Comme j'ai pu l'entendre un bon nombre de fois, « Avec Fabrice et Dominique en directeurs ? C'est du 5 étoiles ! ». Et en effet, je me suis rendu compte au fil des mois et des années, à quel point vous étiez présents, à l'écoute et bienveillants. Grâce à vous, j'ai eu la chance de connaitre ces fameuses réunions à « l'interface », où l'on passe plus de temps à se comprendre qu'à faire le bilan des expériences. Chacun à votre manière, vous m'avez appris à faire de la science :

- Fabrice : tu m'as fait confiance dès le début lors de mon stage et, malgré cette expérience, tu as quand même persévéré pour me proposer et m'aider à obtenir cette thèse. Et ce n'était pas gagné ! Surtout quand on pense que j'ai failli rater la date limite de soumission du projet de thèse, merci Fabrice ! Je te remercie pour l'amitié, la bonne humeur, la bienveillance et la compréhension dont tu as fait preuve durant ces 4 années. Et en particulier de m'avoir soutenu et épaulé dans la double aventure dans laquelle je me suis lancé : mener à bien une thèse à l'interface et fonder une famille ! Une banalité ! Je ne connais pas d'autre directeur de thèse conciliant aussi bien l'humour que le professionnalisme. Pour tout cela, je t'en serais toujours reconnaissant.
- Dominique : au premier abord, j'ai été impressionné voire intimidé par ta prestance et ton aura. Même si ce n'est que récemment que j'ai pu vraiment échanger et discuter avec toi, tu m'as apporté de nombreux conseils qui m'ont fait murir scientifiquement mais aussi humainement. Comme Fabrice m'a initié au GWAs, tu m'as initié à la biologie moléculaire, à la validation fonctionnelle et aux « 3 manips indépendantes ». J'ai eu beaucoup de chance de t'avoir comme directrice. Malgré les nombreuses obligations liées au LabEx TULIP et au département SPE, tu as toujours su trouver le temps de me conseiller et m'orienter. Pour tout cela, je te remercie.

Remerciements

J'ai pu affronter cette thèse grâce à la bonne ambiance de toute une équipe ! Par ordre de rencontre:

- Léa : merci pour ton accueil au sein de l'équipe, ta gentillesse et ton encadrement (et oui tu étais bien co-directrice de mon stage !). Je te remercie pour tous tes bons conseils ainsi que ton écoute, en particulier, durant les derniers jours de rédaction, cela m'a remis un petit coup de fouet pour terminer !
- Etienne : merci de m'avoir 'légué' la manip' plurispécifique lors de mon stage, j'ai pu enchaîner 3 années à travailler sur ces fascinantes interactions plante-plante.
- Claudia : sacré Claudia! Une italienne comme on en voit que trop peu. Ton dynamisme, ta culture scientifique et ta dévotion à ton travail m'ont impressionné (mais bon des fois il faut dormir Claudia). Je te remercie de m'avoir poussé à postuler au poste d'IE au sein de l'équipe...
- et toi Baptiste merci d'avoir obtenu ce poste... Je ne t'en veux qu'à moitié, après tout tu m'as quand même bien aidé pour mes débuts dans la biomol' même si toi aussi tu en as profité pour bien rigoler. Merci de m'avoir motivé (ou l'inverse je sais plus) à aller au rugby flag.
- Jaishree : 'You know what?' Thank you for your daily entertainment, every day you make me relativize about life and madness. I'm grateful to you for continuing to look for '++' interactions in the TOU-A population, I pray all the gods for you to succeed. 'Ayoyo!'.
- Tatiana : on ne s'est pas croisé longtemps mais bon on va bientôt se retrouver pour mon premier postdoc ! Je te remercie pour ton oreille attentive et ta compréhension des soucis de parents ! A très vite.
- Rémy : un sacré personnage ! Tu m'auras bien fait rire, tu as apporté un vent de fraicheur lors de cette fin de thèse. Je n'oublie pas notre projet !
- Je tiens aussi à remercier Arnaud et Kevin qui m'ont épaulé durant ma thèse ainsi que les stagiaires qui sont passés dans l'équipe ; Thomas, Mylène, Taeken, Eve et Océane.

Je tiens également à remercier toutes les personnes avec qui j'ai pu travailler durant cette thèse ;

- Merci Carine pour ton aide pour les extractions d'ARN (surtout à la fin !), les manip' en serre quand j'étais à la bourre, la récolte de graines. Grâce à toi j'ai gagné 3 à 4 mois. Je te remercie également pour ta bienveillance et ton franc parlé, ça fait du bien ! J'ai vraiment pris beaucoup de plaisir à travailler et échanger avec toi.
- Merci Marielle et Ulli' pour votre aide pour les constructions et les transfo'.
- Merci Maxime pour nos échanges sur mes scripts d'évolution moléculaire mais pas que !

Remerciements

À la bioinfo', merci :

- Jérôme Gouzy pour tes conseils qui m'ont permis de prendre du recul sur ma thèse et m'ont fait murir.
- Sébastien Carrère mais je ne sais pas trop pourquoi. À chaque fois, il faut revenir te voir pour te dire ce que tu as mal fait dans les analyses bioinfo' !
- les ludo' pour les scripts dans les différents langages et votre patience pour m'expliquer les bases.

À la serre, merci :

- Jean-Luc toujours à préparer les pots 3h avant tout le monde, Claudette toujours bienveillante et compatissante, Marine qui se la coule douce en Australie, Camille, Anaïs et bien sûr Fabrice toujours là pour les autoclaves à TPMP ou quand plus rien ne marchait.

Je tiens également à remercier de manière plus générale tout le personnel du laboratoire et la FR, notamment ;

- L'ensemble du service gestion, notamment Christophe et Audrey.
- Soon et Isabelle toujours là pour me dépanner des problèmes informatiques !
- Dominique, Christian et Florian pour votre gentillesse et la préparation du milieu MS !
- Merci également à Alissoutte et Camille, Pauline, Florent, Richard, Gaëlle, Corine, Mireille, Medhi, Rémi V, Rémi P, Sylvie, Céline, Marta, Fabienne, Narguess, Alice G, Sandra, Laurent S, Laurent D, Némo…
- Christophe et Philippe à la maintenance.
- Merci Alain, Aurélie, Cécile, Yves et Marie-Christine pour votre aide pour mes premiers pas en microscopie.

Faire une thèse, c'est également compter sur de nombreuses personnes extérieures à la science, les amis et la famille :

Merci aux amis, qui malgré les aléas de la vie sont toujours là :

- Clémentine et Romain, depuis le lycée on ne s'est plus quitté, merci d'être là avec votre fraicheur et votre bonne humeur qui vous animent.
- Les rencontres de la fac : merci Mickaël le free-rider fou et chasseur de gallinette cendrée, Jérem' Sarthe, toujours en balade aux quatre coins de la France. Marine, Bobby et Achkar je me suis vraiment régalé avec vous en M2.
- Merci aux amis par alliance : Mélanie et Philippe, Léah et Sam et Astrid. Merci aussi à la famille chat ainsi que Claire, Thomas et Livia pour les « gouter entre amis avec enfants ! » et autres moments de break. C'était un régal, pourvu que ça dure.

Remerciements

Enfin, je finirai par ma famille et les remercie de leurs soutiens à toutes épreuves depuis de longues années !

Merci Maman d'avoir toujours été là pour nous, de m'avoir permis d'aller à la fac et d'en être ici aujourd'hui. Merci Romain, malgré nos différences, tu es là et en particulier pour Mathis et Thalia.

Merci Mamie Momo, une mamie formidable d'un courage sans égal. Tu es un exemple pour tous tes petits et arrières-petits enfants, moi compris. Merci aussi à mes cousins du Tarn: Guillaume et Aurélie, Julien et Delphine et Yann et Mélanie. Merci aux cousins des Etats-Unis : Alex et Sophia, Marielle et Math', Steven et Anne. Merci à mes cousins de l'Est : Fabrice et Emilie, Sandra et Florian, Cathy et Nicolas, et ma grande grande cousine Valérie. Loin des yeux près du cœur. Je remercie aussi tous mes oncles et tantes.

Merci aussi à ma belle-famille. Merci Dany et Hervé pour vos encouragements et votre soutien sans faille depuis plus de 11 ans ! Merci tonton Tianou, t'es un super Beauf' et tu es au top avec tes neveux ! Merci Pépé et Mémé de toujours nous permettre de passer de très bons moments, à Cham' et à Dax notamment. Merci aussi aux Agenais Gisèle, Julien et Marion pour les nombreux moments à la montagne été comme hiver.

Et bien sûr Mamour, je ne te remercierais jamais assez pour toutes ces années où il t'a fallu me supporter. Tu m'as toujours poussé ou plutôt tiré vers le haut. Tu as toujours été là pour les bons comme les mauvais moments. Tu m'as permis de mener à bien cette thèse en me rassurant, m'épaulant et en étant une mère formidable. Je te remercie aussi pour ça, tout en menant à bien cette thèse nous avons aussi construit notre propre famille, Thalia puis Mathis. Ils ne portent pas ton nom mais ils sont ton portrait craché !

Thalia, Mathis vous êtes mon plus grand bonheur, ma plus grande fierté, ma plus grande fatigue mais aussi mon plus grand réconfort pour ces mois difficiles. Je vous aime.

« *Y a deux ans je comprenais pas grand-chose*
Maintenant c'est pire »

Orelsan

Sommaire

I. Introduction générale

A. Importance des interactions plante-plante 2

B. Manuscrit: The genetics underlying natural variation of plant-plant interactions, a beloved but forgotten member of the family of biotic interactions 5

C. Objectifs de la thèse 62

Arabidospsis thaliana: une espèce adaptée pour identifier les bases génétiques de la variation naturelle des interactions plante-plante ? 62

Comment identifier chez *A. thaliana* les gènes sous-jacents à la réponse à la présence de plantes voisines ? 66

D. Plan de la thèse 73

II. Chapitre 1: Identification d'une population naturelle d'*A. thaliana* adaptée à l'analyse de la variation génétique des interactions plante-plante.

A. Introduction 77

B. Manuscrit: Intermediate degrees of synergistic pleiotropy drive adaptive evolution in ecological time 81

Supporting information 122

C. Conclusion 173

III. Chapitre 2: Identification des bases génétiques sous-jacentes aux interactions mono- et pluri-spécifiques.

A. Introduction 175

B. Manuscrit: The genomic architecture of competitive response of *Arabidopsis thaliana* is highly flexible between monospecific and plurispecific neighborhoods 177

Supporting information 217

C. Conclusion 242

IV. Chapitre 3: Identification d'un récepteur kinase contrôlant la réponse compétitive à *Poa annua.*

A. Introduction 245

B. Manuscrit: An Arabidopsis receptor-like kinase mediates competitive plant-plant interactions 249

Supporting information 280

C. Conclusion 291

V. Discussion générale et perspectives

Discussion générale 295

A) Variation génétique des interactions plante-plante : spécificité de l'interaction biotique 295

B) Flexibilité de l'architecture génétique sous-jacente aux interactions plantes-plantes : maintien de la diversité génomique intra-population 298

C) Fonctions et gènes identifiés, sous-jacents aux interactions plante-plante .. 301

Perspectives 304

A) Dissection des mécanismes de perception et des voies de signalisation associées à *PERK13/RHS10* 304

B) Validation fonctionnelle d'autres QTL identifiés par GWA mapping 309

C) Valeur adaptative de *PERK13/RHS10* 311

VI. Bibliographie 313

Liste des articles scientifiques et communications orales

Articles scientifiques

A. Publiés

1. Subrahmaniam, H.J., **C. Libourel**, E.-P. Journet, J.-B. Morel, S. Muños, A. Niebel, S. Raffaele and F. Roux. 2018. The genetics underlying natural variation of plant-plant interactions: a beloved but forgotten member of the family of biotic interactions. **The Plant Journal** 93: 747-770.

2. Frachon, L.*, **C. Libourel***, R. Villoutreix, S. Carrère, C. Glorieux, C. Huard-Chauveau, M. Navascués, L. Gay, R. Vitalis, E. Baron, L. Amsellem, O. Bouchez, M. Vidal, V. Le Corre, D. Roby, J. Bergelson and F. Roux. 2017. Intermediate degrees of synergistic pleiotropy drive adaptive evolution in ecological time. **Nature Ecology & Evolution** 1: 1551-1561. *Authors contributed equally to this work.

B. En préparation

3. **C. Libourel**, E. Baron, J. Lenglet, L. Amsellem, D. Roby and F. Roux. The genomic architecture of competitive response of *Arabidopsis thaliana* is highly flexible between monospecific and plurispecific neighborhoods. Soumis.
4. **C. Libourel**, F. Roux and D. Roby. A receptor-like kinase mediates competitive plant-plant interactions in *Arabidopsis thaliana*. En préparation.

Communications Orales – Congrès & Colloques

1. **C. Libourel**, E. Baron, J. Lenglet, L. Amsellem, D. Roby and F. Roux. Septembre 2017. A Genome-Wide Association mapping approach reveals the genetic bases of monospecific and plurispecific plant-plant interactions in *Arabidopsis thaliana*. XXIIth International Congress of Genetics. Foz do Iguaçu, Brazil.

2. **C. Libourel**, H.J. Subrahmaniam, D. Roby and F. Roux. December 2017. The genetics underlying natural variation of plant-plant interactions in *Arabidopsis thaliana*. XII reunión de biología vegetal. Villarica, Chile.

3. **C. Libourel**., E. Baron, S. Carrère, J. Langlet, L. Amsellem, D. Roby and F. Roux. Mars 2017. A Genome-Wide Association mapping approach highlights the genetics of competitive ability of *Arabidopsis thaliana* to monospecific and multispecific interactions. Réunion annuelle du Réseau Ecologie des Interactions Durables. Toulouse, France.

4. L. Frachon., **C. Libourel***, R. Villoutreix, S. Carrère, C. Glorieux, C. Huard-Chauveau, M. Navascués, L. Gay, R. Vitalis, E. Baron, L. Amsellem, O. Bouchez, M. Vidal, V. Le Corre, D. Roby, J. Bergelson and F. Roux. Octobre 2016. Tracking the genetic bases of contemporary evolution in a spatially heterogeneous environment. Sfécologie-2016, International Conference of Ecological Sciences. Marseille, France.

Posters

1. H.J. Subrahmaniam, **C. Libourel**, A. Chevalier--Mairet, D. Roby and F. Roux. Septembre 2017. Genetics of intraspecific plant-plant cooperation in *Arabidopsis thaliana.* 1st Plant Adapt meeting. Banyuls-sur-mer, France.

2. **C. Libourel**, E. Baron, J. Lenglet, L. Amsellem, D. Roby and F. Roux. Septembre 2017. A Genome-Wide Association mapping approach reveals the genetic bases of monospecific and plurispecific plant-plant interactions in *Arabidopsis thaliana*. XXIInd International Congress of Genetics. Foz do Iguaçu, Brazil.

3. M. Navascués, A. Becheler, M. Julien, L. Frachon, **C. Libourel**, R. Villoutreix, J. Ronfort, F. Roux, R. Vitalis and L. Gay. October 2016. Tracking adaptation in selfing plant populations. Conférences Jacques Monod: Evolutionary genomics and systems biology: bringing together theoretical and experimental approaches. Roscoff, France.

4. L. Frachon, R. Villoutreix, **C. Libourel**, E. Baron, S. Carrière, C. Glorieux, L. Amsellem, V. Le Corre, J. Gouzy, J. Bergelson and F. Roux. Septembre 2015. Adaptive genomics to fine-grained spatial heterogeneity in a natural population of *A. thaliana*. Forum interne du laboratoire LIPM. Sorrèze, France.

5. L. Frachon, R. Villoutreix, **C. Libourel**, E. Baron, S. Carrière, C. Glorieux, L. Amsellem, V. Le Corre, J. Gouzy, J. Bergelson and F. Roux. July 2015. Adaptive genomics to fine-grained spatial heterogeneity in a natural population of *A. thaliana*. 7èmes journées des doctorants SPE. Rennes, France.

Introduction générale

Au cours de leur cycle de vie, les individus doivent faire face à une multitude de stress, qu'ils soient d'origine abiotique ou biotique. Ceci est d'autant plus vrai pour les plantes, de par leur mode de vie sessile. Dans le contexte des changements globaux actuels, nous observons une modification en profondeur de ces stress, tant au niveau de leur intensité qu'au niveau de leur identité (Vitousek *et al.* 1997, Chapin III *et al.* 2000, Sala *et al.* 2000, Millenium Ecosystem Assessment 2005). Un enjeu majeur en écologie évolutive est de comprendre et de prédire la capacité d'une espèce végétale à persister en présence de nouvelles conditions environnementales et écologiques. Trois réponses non-exclusives peuvent être adoptées par les espèces végétales pour y faire face (Hansen *et al.* 2012, Bay *et al.* 2017): (i) migration des espèces pour suivre les changements spatiaux actuels de l'environnement (Pecl *et al.* 2017), (ii) acclimatation rapide des organismes aux nouvelles conditions environnementales et écologiques *via* la plasticité phénotypique, définie comme la capacité d'un génotype donné à produire différents phénotypes quand il est exposé à différentes conditions environnementales ou écologiques (Fusco & Minelli 2010), et (iii) adaptation des espèces à de nouvelles conditions environnementales et écologiques *via* la sélection génétique (Hoffman & Sgro 2011, Bay *et al.* 2017).

Dans le dernier cas, prédire le potentiel génétique adaptatif des espèces nécessite non seulement une description de l'architecture génomique sous-jacente à l'adaptation, mais aussi la compréhension des mécanismes génétiques et moléculaires associés (Roux & Bergelson 2016). *Au niveau abiotique*, de par la disponibilité de bases de données publiques, la majorité des études chez les plantes se sont focalisées sur l'établissement d'une carte génomique de l'adaptation vis-à-vis du climat (Fournier-Level *et al.* 2011, Hancock *et al.* 2011, Bay *et al.* 2017, Frachon *et al.* 2018). Les études visant à établir une carte génomique de l'adaptation des plantes aux conditions édaphiques sont quant à elles peu nombreuses (Turner *et al.* 2010, Lasky *et al.* 2015, Pluess *et al.* 2016, Rellstab *et al.* 2016). Dans tous les cas, il est surprenant de noter que les gènes causaux sous-jacents aux Quantitative Trait Loci (QTL) n'ont été que très rarement identifiés (e.g. Hanikenne *et al.* 2008, Baxter *et al.* 2010). *Au niveau biotique*, on observe qu'un nombre bien plus conséquent de gènes causaux sous-jacents à des QTLs ont été identifiés durant les trois dernières décennies, permettant ainsi d'obtenir une vision plus ou moins complète des déterminismes génétiques et moléculaires sous-jacents à la variation naturelle de la réponse des plantes à la présence de virus, bactéries, champignons, oomycètes

ou insectes herbivores (Koornneef *et al.* 2004, Prasad *et al.* 2012, Roux *et al.* 2014, French *et al.* 2016, Roux & Bergelson 2016, Bartoli & Roux 2017, Curtin *et al.* 2017, Yang *et al.* 2017, Wang *et al.* 2017). Cependant, le nombre d'études rapportant la valeur adaptative des gènes impliqués dans la variation naturelle des interactions biotiques reste encore très limité (Brachi *et al.* 2015, Roux & Bergelson 2016, Brachi *et al.* 2017).

A) Importance des interactions plante-plante

Un type d'interaction biotique a été largement ignorée dans la description de l'architecture génétique de l'adaptation et des mécanismes génétiques et moléculaires sous-jacents. Il s'agit des interactions plante-plante. Et pourtant, comme on peut aisément l'observer dans la majorité des environnements naturels, une plante est tout au long de son cycle de vie en contact (soit directement, soit indirectement) avec d'autres plantes qui, bien souvent, correspondent à d'autres espèces. Depuis plusieurs décennies, il est reconnu que les interactions entre plantes jouent un rôle majeur dans la structure, la diversité et la dynamique des communautés végétales naturelles (Tilman 1985, Goldberg & Barton 1992, Chesson 2000). Ces interactions entre plantes peuvent aller de la compétition aux interactions positives réciproques en passant par des relations asymétriques. Parmi ces trois grands types d'interactions entre plantes, la compétition a été le plus largement étudiée et se traduit par des effets négatifs pour les deux partenaires (Figure i.1), dus notamment aux limitations en ressources telles que la disponibilité en nutriments, en eau ou en lumière (Turkington & Harper 1979, Chaney & Baucom 2014). À l'opposé, nous trouvons les interactions positives réciproques (i.e. coopération entre individus d'une même espèce ou mutualisme entre individus de différentes espèces) qui correspondent à un bénéfice pour les deux partenaires qui interagissent (Figure i.1). Bien que reconnues comme importantes par les premiers écologues (i.e. Phillips 1909, Clements 1916), ces interactions positives ont par la suite été largement ignorées au profit des interactions compétitrices. Ce n'est que très récemment que les interactions positives ont connu un regain d'intérêt (Brooker *et al.* 2008, Bukowski & Petermann 2014). Entre ces deux cas extrêmes, il existe une part non négligeable d'interactions dites asymétriques où l'un des partenaires tire avantage de l'interaction aux dépens de l'autre (Halty *et al.* 2017). Le parasitisme et l'allélopathie entrent dans cette catégorie. Les plantes parasites vivent et se développent au détriment de leur plante hôte: les

espèces du genre *Castilleja* sont connues pour parasiter pas moins de 100 hôtes différents dont les lupins, *Medicago sativa* ou des graminés comme *Lolium perenne* (Press 1998). Ces parasites réduisent la productivité de leur hôte affectant ainsi la structure et la diversité de leur communauté (Matthies 1997). Quant aux espèces allélopathiques, elles peuvent libérer *via* leurs racines des substances allélochimiques qui affectent le développement et la croissance des plantes voisines (i.e. juglone, sorgelone, etc.). De nombreuses études ont montré le rôle prépondérant de l'allélopathie dans le 'succès' de certaines espèces invasives (Callaway & Aschehoug 2000, Sakai *et al.* 2001, Bais *et al.* 2003, Hierro & Callaway 2003, Zhang *et al.* 2007, Pisula & Meiners 2010).

Cependant, les relations entre plantes sont loin d'être stables. En effet, les facteurs abiotiques peuvent fortement modifier le sens et l'intensité des interactions entre différentes espèces végétales (Callaway *et al.* 2002). De manière générale, les interactions entre plantes tendent à être positives dans des conditions de stress intense et *a contrario* elles tendent à être de nature compétitive dans des conditions moins stressantes (Callaway & Walker 1997, Pugnaire & Luque 2001, Callaway *et al.* 2002, Maestre *et al.* 2005, Maestre *et al.* 2009).

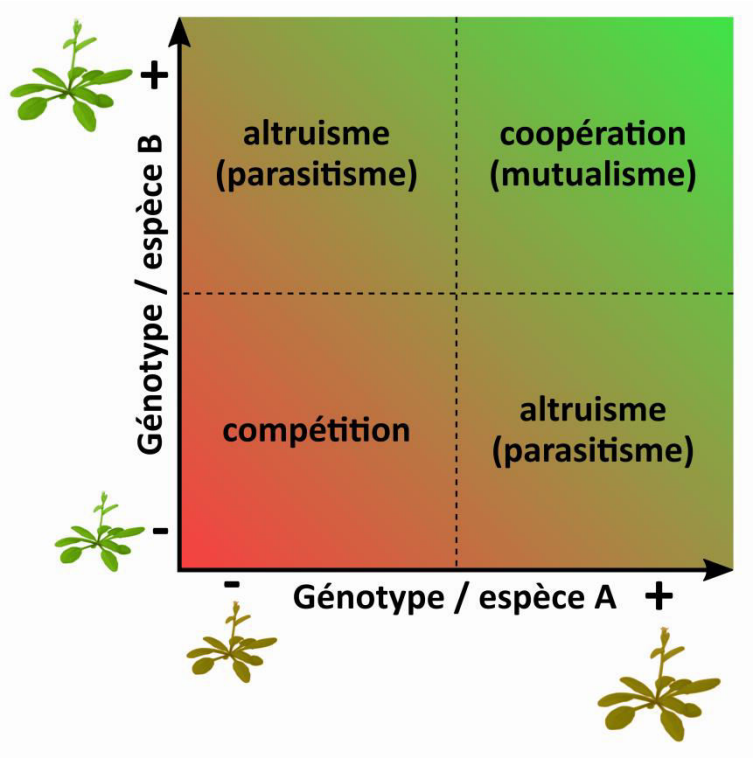

Figure i.1. Schéma représentant les types d'interactions possibles entre individus d'une même espèce ou d'espèces différentes. Les termes entre parenthèses sont utilisés dans le cadre des interactions hétérospécifiques (i.e. entre espèces différentes). D'après Subrahmaniam *et al.* 2018.

Cette diversité d'interactions plante-plante s'observe aussi au sein des agro-écosystèmes et peut donc entraîner des effets importants sur le rendement des cultures. En effet, en absence de pesticides, les plantes adventices présentent un potentiel de réduction du rendement des cultures considérablement plus élevé que pour tout autre nuisible (34% pour les plantes adventices contre 18% pour les animaux nuisibles et 16% pour les agents pathogènes; Basu *et al.* 2004, Oerke 2006, Neve *et al.* 2009). Il est donc communément admis que les interactions plante-plante sont majoritairement dominées par la compétition dans les champs cultivés. Néanmoins, comme précédemment observé au sein des communautés végétales naturelles, de plus en plus d'études empiriques mettent en avant l'importance des interactions positives réciproques dans les peuplements plurivariétaux et/ou plurispécifiques, avec des phénomènes de productivité accrue («overyielding») dans les mélanges de variétés ou d'espèces par rapport aux monocultures (Tilman *et al.* 1996, Tilman *et al.* 1997, Tilman *et al.* 2001, Hector *et al.* 1999, Loreau & Hector 2001, van Ruijven & Berendse 2003). Depuis plusieurs années, l'utilisation des propriétés allélopathiques de certaines espèces cultivées pour la gestion des mauvaises herbes a été proposée comme une alternative écologique aux produits phytosanitaires de synthèse. Cette alternative serait non seulement plus respectueuse de l'environnement mais aussi plus rentable, durable et fiable (Tesio & Ferrero 2010 Farooq *et al.* 2013, Jabran *et al.* 2015). Différentes espèces de Brassica ont montré leur potentiel allélopathique vis-à-vis de nombreuses espèces adventices présentes dans les champs cultivés et pourraient donc être considérées comme un outil durable de gestion intégrée des plantes adventices (Rehman *et al.* 2018).

Malgré cette diversité des interactions plante-plante observée aussi bien dans les écosystèmes naturels que dans les agro-écosystèmes, les déterminants génétiques et moléculaires qui sous-tendent la variation naturelle des interactions plante-plante restent encore mal connus. Dans la partie qui suit, à travers une revue, nous proposons un état des lieux des travaux portant sur la variation naturelle génétique des interactions plante-plante, l'architecture génétique associée et les déterminants génétiques identifiés. Cette revue propose aussi de futures pistes à suivre pour l'identification et l'analyse des gènes et fonctions moléculaires associée à ces interactions biotiques importantes mais encore trop peu étudiées.

B) Manuscrit: The genetics underlying natural variation of plant-plant interactions, a beloved but forgotten member of the family of biotic interactions

Harihar Jaishree Subrahmaniam[1], Cyril Libourel[1], Etienne-Pascal Journet[1,2], Jean-Benoît Morel[3], Stéphane Muños[1], Andreas Niebel[1], Sylvain Raffaele[1] and Fabrice Roux[1*]

[1] LIPM, Université de Toulouse, INRA, CNRS, Castanet-Tolosan, France

[2] AGIR, Université de Toulouse, INRA, INPT, INP-EI PURPAN, Castanet-Tolosan, France

[3] BGPI, INRA,CIRAD,SupAgro, Univ. Montpellier, Montpellier, France

* To whom correspondence should be addressed. E-mail: fabrice.roux@inra.fr

Manuscrit

"The genetics underlying natural variation of plant-plant interactions, a beloved but forgotten member of the family of biotic interactions"

Harihar Jaishree Subrahmaniam, Cyril Libourel, Etienne-Pascal Journet, Jean-Benoît Morel, Stéphane Muños, Andreas Niebel, Sylvain Raffaele and Fabrice Roux

ABSTRACT

Despite the importance of plant-plant interactions on crop yield and plant community dynamics, our understanding of the genetic and molecular bases underlying natural variation of plant-plant interactions is largely limited in comparison to other types of biotic interactions. By listing 63 QTL mapping and global gene expression studies based on plants directly challenged by other plants, we explored whether the genetic architecture and the function of the candidate genes underlying natural plant-plant interactions depend on the type of interactions between two plants (competition *vs* commensalism *vs* reciprocal helping *vs* asymmetry). The 16 transcriptomic studies are unevenly distributed between competitive interactions (n = 12) and asymmetric interactions (n = 4, all focusing on response to parasitic plants). By contrast, 17 and 30 studies were detected for competitive interactions and asymmetric interactions (either weed suppressive ability or response to parasitic plants), respectively. Surprisingly, no studies have been carried out on the identification of genetic and molecular bases underlying natural variation in positive interactions. The candidate genes underlying natural plant-plant interactions can be classified into seven categories of plant function that have been identified in artificial environments simulating plant-plant interactions either frequently (photosynthesis, hormones), only recently (cell wall modification and degradation, defense pathways against pathogens) or rarely (ABC transporters, histone modification and meristem identity / life history traits). Finally, we introduce several avenues that need to be explored in the future to obtain a thorough understanding of the genetic and molecular bases underlying plant-plant interactions within the context of realistic community complexity.

Significance statement: While plant-plant interactions are recognized as a major factor responsible for crop yield and plant community dynamics, their underlying genetic and molecular mechanisms still deserve a deeper investigation. By considering studies where plants have been directly challenged by other plants, we identified plant functions that have been rarely identified in artificial environments simulating plant-plant interactions. The next goal will be to understand the genetic and molecular mechanisms underlying positive interactions.

Keywords: competition, cooperation, altruism, mutualism, allelopathy, parasitic plant, QTL mapping, GWA mapping, gene expression, biotic diffuse interactions

INTRODUCTION

Throughout its life cycle, a plant can interact simultaneously and sequentially - directly or indirectly - with many plant neighbors, whether in crop fields or in more natural environments (Williams, 2013). In such plant networks, a large diversity of interactions can be observed both at the intraspecific and interspecific levels (Box 1). Intraspecific relationships extend from competition with conspecifics (same species) to cooperation, through altruism (Box 1) (Dudley, 2015). Interspecific relationships include competition with heterospecifics (different species), reciprocal helping (e.g. mutually beneficial interactions), commensalism (e.g. facilitation) and asymmetric interactions such as parasitism and allelopathy (Box 1).

Surprisingly, despite (i) the importance of plant-plant interactions in mediating plant community structure, diversity and dynamics (Tilman, 1985; Goldberg and Barton, 1992; Chesson, 2000); and (ii) weeds having a significantly higher average potential to reduce crop yield than any other crop pest (34% for weeds *vs* 18% for animal pests *vs* 16% for pathogens; Basu *et al.*, 2004; Oerke, 2006; Neve *et al.*, 2009), our understanding of the genetic and molecular bases underlying natural variation of plant-plant interactions is largely limited in comparison to other types of biotic interactions. For example, among the 56 genes functionally validated for being associated with natural variation in response to biotic interactions in the model plant *Arabidopsis thaliana*, more than one-third confer resistance to herbivory while the rest of the genes are more-or-less evenly distributed among interactions with viruses, bacteria, fungi and oomycetes (Roux and Bergelson, 2016). The only gene identified as involved in plant-plant interactions underlies responses to root spatial constraints (used as proxy for thigmotropic responses to other plants within the rhizosphere) and not the direct response to a neighbor plant (Joseph *et al.*, 2015). In addition, in early 2017, 35 Genome Wide Association studies (GWAS) reported the fine mapping of genomic regions associated with natural variation of plant response (either crops or natural species) to pathogen infection (Bartoli and Roux, 2017), whereas only one GWAS reported the identification of Quantitative Traits Loci (QTLs) underlying plant-plant interactions (Baron *et al.*, 2015).

Box1. Terminology of the various categories of plant-plant interactions.

According to Dudley (2015), interactions between plant individuals can be divided into various categories based on whether they occur between two species (heterospecific or interspecific interactions) or within a species (conspecific or intraspecific interactions).

Interaction	Nature	Interspecific level	Intraspecific level
Competitive	--	Competition	Competition
Commensal	+0	Facilitation	Cooperation with direct benefit
Reciprocal helping	++	Mutualism(Co adaptation)	Cooperation with reciprocal benefit
Asymmetric	+-	•Parasitism •Allelopathy	Altruism

The interaction is termed **competitive (- -)** when both interacting individuals suffer significant cost by investing in competing and therefore compromising on the benefit. In other words, the outcome of competition for both the interacting individual plants can be viewed as Benefit < 0 and Cost > 0. The terms benefit and cost are pertaining to the net effect on individual fitness of both the interacting individuals. A popular example for interspecific competition is the interaction between many crops and weeds which leads to a significant reduction in agricultural crop yield as the weeds compete for resources that would otherwise be available for the crops to use.

Commensal interactions **(+0)** are the ones where the helper plant provides benefit to another plant but does not incur any cost in the process. It can be represented as Benefit > 0, Cost = 0 for individual X, the one receiving the help and Benefit = 0, Cost = 0 for individual Y, the one providing the help. This kind of interaction is called **facilitation** when it occurs at interspecific level whereas at intraspecific level, it is called **'cooperation** with **direct benefit'**. At the interspecific level for example, epiphytes that grow on the barks of many trees purely for physical support are good examples for this type of interaction. The host tree does not incur any cost in providing anchorage to the epiphyte and the epiphyte can cling on to the host plants without being parasitic and damaging the host plant organs or functions.

Reciprocal helping **(++)** is the interaction where both the partners exchange costly help. For both the interacting individuals, the cost of providing help is significant but it is compensated for by the benefit they get in return, i.e. Benefit > Cost for both interacting individuals. This reciprocation is directed to only specific individuals that would return the favor. It is called **mutualism** when it occurs between species and **'cooperation** with **reciprocal benefit'** when it is within a species. Mutualism is thought of as a result of co-adaptation and both the interacting individuals affect the evolution of the helping trait phenotype of each other. Teste *et al.* (2014) conducted an experiment where they grew four plants species having different nutrient acquiring strategies under nutrient rich and poor conditions. They observed that under nutrient poor conditions, the focal plant *Melaleuca preissiana* (Arbuscular mycorrhizal/Ectomycorrhizal fungal network) grew better when it was grown besides *Eucalyptus marignata* (EM fungal network) and *Banksia menziesii* (cluster mining roots) in a mesh microcosm where roots were not in physical contact but only the fungal network were mingling. The plants were able to acquire nutrients and share them between neighbours depending on the nutrient acquiring

strategy of the neighbour and using the fungal network under limited soil resources. This experiment is evidence that plants can be involved in reciprocal helping but only when there is a need for them to share benefits.

Asymmetric interactions **(+-)** occur when one of the interacting partners benefits at the expense of the other (Halty *et al.*, 2017). This 'costly' help can be depicted as Benefit > 0 and Cost = 0 for the individual receiving the help and Benefit = 0 and Cost > 0 for the help provider. **Parasitism** and **allelopathy** come under this category at the interspecific level (NB: few studies also reported allelopathy at the intraspecific level such as in *Kalanchoae daigremontiana*; Groner, 1974). Parasitic plants like *Arceuthobium sp.* that derive nutrition from other plants and causing harm to the host are prime examples for this interaction at the interspecific level. Some plants release inhibitory chemicals, allelochemicals (juglone, sorgelone etc) *via* their roots that can affect the development and growth of neighboring plants. Although allelopathy includes both positive (growth promoting) and negative (growth inhibiting) effects, definitions of allelopathy often only consider negative effects (Olofsdotter *et al.*, 2002). The interaction between allelopathic plants and their neighbors is therefore considered as asymmetric. At the intraspecific level, it is related to '**altruism**' that corresponds to the preferential help given to an individual from the same population without getting any direct benefit for it. Individuals should perform actions that increase their own fitness but altruism is quite the opposite of that. Individuals that perform altruistic actions reduce their own chances of reproduction and survival in order to help another.

Altruism evolves within a population where individuals provide costly help to their relatives. Helping a relative is selfish in some sense as it increases the fitness of the altruist indirectly. Relatives that share a significant portion of genes between themselves and if the altruist decides to help a relative, that means more chance of representation of its own genes in the next generation. This nepotistic behavior of individuals within a population is called **kin recognition**. Help can also be provided if the actor recognizes a gene / set of closely linked genes and only favors the carrier of those genes. This preferential help based on genetic similarity only at some parts of the genome is called **Greenbeard effect.** But till date, there have been no report about the existence of Greenbeard genes in plants. The defining feature of kin recognition is based on the concept of inclusive fitness, a concept that has been popularly used to describe the evolution of eusociality among many animal species. However, this view has recently been debated by many theoretical and experimental studies (Nowak *e al.*, 2010; Allen *et al.*, 2013), claiming that a group can begin to cooperate even if individuals are unrelated, providing the association proves useful to both parties. This association can be a product of reciprocity or mutualistic synergism (Nowak *et al.*, 2010).

Deciphering the genetic and molecular bases underlying natural variation of plant-plant interactions can however be fundamental to propose new germplasm management strategies for maintaining sustainable provision of yield and other ecosystem services in an agro-ecological context (Litrico and Violle, 2015). For example, the identification of genetic markers usable in Marker-Assisted Selection (MAS) can largely accelerate breeding programs to address several major agro-ecological issues. Firstly, weeds are farmers' worst pests, especially in organic systems (Basu *et al.*, 2004; Neve *et al.*, 2009; Asif *et al.*, 2015). In addition, an increase of the deleterious impact of weeds on crop yield due to climate change is expected (Clements *et al.*, 2014; Peters *et al.*, 2014). QTLs associated with increased competitiveness in crops can therefore represent a durable and sustainable alternative for weed control (Worthington *et al.*, 2013). The genetics of competitiveness can be based on the detection and functional characterization of QTLs underlying enhanced crop competitive ability (such as traits linked to plant canopy establishment and nutrient acquisition capacity; Olofsdotter *et al.*, 2002), or QTLs underlying weed suppressive ability through the production of chemical defense compounds (such as allelopathy; Khanh *et al.*, 2007). Secondly, during the last decades, a particular attention from breeders has been devoted to improving yield per unit of field area by increasing plant density (Guo *et al.*, 2011), where the target of the breeding programs is population and/or community performance (i.e. group selection) rather than individual plant performance (i.e. individual selection) (Weiner *et al.*, 2010). Because the deleterious effects of a large range of abiotic (i.e. drought stress) and biotic (i.e. pathogen attack) stresses are exacerbated in high-density conditions (Gonzalo *et al.*, 2010; Ku *et al.*, 2016), there is a need for identifying the genetic basis underlying density-related stress tolerance (Gonzalo *et al.*, 2006). Thirdly, increasing species diversity and/or genotypic diversity has positive effects on plant productivity, stability of yield and ecosystem services (Tilman 1997; Tilman *et al.*, 2001; Crutsinger *et al.*, 2006; Johnson *et al.*, 2006; Isbell *et al.*, 2011; Loreau and de Mazancourt, 2013; Pietro *et al.*, 2015). Therefore, in the framework of resource-use complementarity, understanding the effects of trait combinations involved in interactions and their underlying genetics, between a focal plant and neighboring conspecific and/or heterospecific plants, may help to optimize species assemblages in cropping systems (e.g. intercropping systems) (Litrico and Violle 2015). It will undoubtedly accelerate breeding programs aimed at creating elite mixtures also called 'ideomixes' (Litrico and Violle 2015).

Identifying and characterizing the function of genes associated with natural variation of plant-plant interactions is also fundamental to predict and understand adaptive dynamics and evolutionary trajectories of natural plant communities (Pierik *et al.*, 2013). Understanding the genetic bases and modes of adaptation underlying plant-plant interactions in current plant communities is essential to accurately estimate responses of a plant species to ongoing drivers of global change (Roux and Bergelson, 2016). In particular, it can help estimate the potential of plant species to face anthropogenic modifications of plant assemblages, which may result from differences of geographic range shift among native species under climate change (Bachelet *et al.*, 2000; Gilman *et al.* 2010; Singer *et al.* 2013) or from increased plant biomass and reduced diversity under climate warming (Baldwin *et al.*, 2014). Furthermore, intraspecific diversity can largely contribute to biotic resistance to exotic invasion, as illustrated by intraspecific diversity in the dominant North American native *Pseudoroegneria spicata* improving resistance against the strong exotic invader *Centaurea stoebe* (Yang *et al.*, 2017). Identifying the genetic bases associated with natural variation of suppressive ability of invasive species may strongly help to propose management strategies, such as reinforcing invaded native populations by planting native individuals with the allelic combination that limits or suppresses invasion. Finally, because genetic diversity within plant populations can be strongly associated with species diversity in interacting communities they support (such as arthropod communities; Whitham *et al.*, 2006), this relationship may have important conservation implications. For example, the maintenance of genetic diversity of an endangered species can be dependent on the level of genetic diversity of the associated native plant species, thereby leading to the concept of minimum viable interaction population (Whitham *et al.*, 2006). Up to now, genetic diversity of plant populations has been traditionally estimated based on genetic markers that are expected to behave neutrally. Identifying the plant genetic bases associated with natural variation of associated community phenotypes may increase the power of designing appropriate management strategies to maintain endangered species.

Here, we review the genetics and molecular mechanisms underlying plant-plant interactions. We first provide an overview of the main molecular mechanisms underlying the perception of the signals related to the presence of neighboring plants and how these signals are translated into response strategies. While very informative, most of these molecular mechanisms have been initially identified in artificial environments designed to simulate plant-plant

interactions. Therefore, in a complementary way, we then list studies based on plants directly challenged by other plants. In particular, because QTL mapping and global gene expression studies are approaches well adapted to interrogate genes mediating biotic interactions in a systematic manner, we reviewed the QTL mapping studies reporting the genetics underlying natural variation of plant-plant interactions and the studies reporting global change in gene expression underlying natural interactions within and between species. We particularly explore whether the genetic architecture and the function of the candidate genes underlying natural plant-plant interactions depend on the type of interactions between two plants (conspecific *vs* heterospecific, competition *vs* commensalism *vs* reciprocal helping *vs* asymmetry). We also emphasize cases where gene functions in plant-plant interactions differ between artificial and ecologically relevant conditions. Finally, we introduce several avenues that need to be explored in the future to obtain a thorough understanding of the genetic and molecular bases underlying plant-plant interactions within the context of realistic community complexity.

NEIGHBOUR DETECTION AND RESPONSE STRATEGIES

Focal plants have the ability to perceive the nature and intensity of the interactions with neighboring plants through diverse signals, transmitted either above or below ground (reviewed in Pierik *et al.*, 2013; Gundel *et al.*, 2014). These signals can be classified as (i) indirect signals, corresponding to environmental factors modified by the presence of neighbors, such as light and nutrients, and (ii) direct signals, corresponding to molecules directly produced by neighboring plants, such as aerial volatile organic compounds (VOCs) and soluble root exudates. Recent work has led to the identification of novel genes and molecules mediating signals between plants, and improved our understanding on how the signals emitted by neighboring plants are integrated into an optimal response strategy. We review here the progress from the last three years on these points (**Figure 1**).

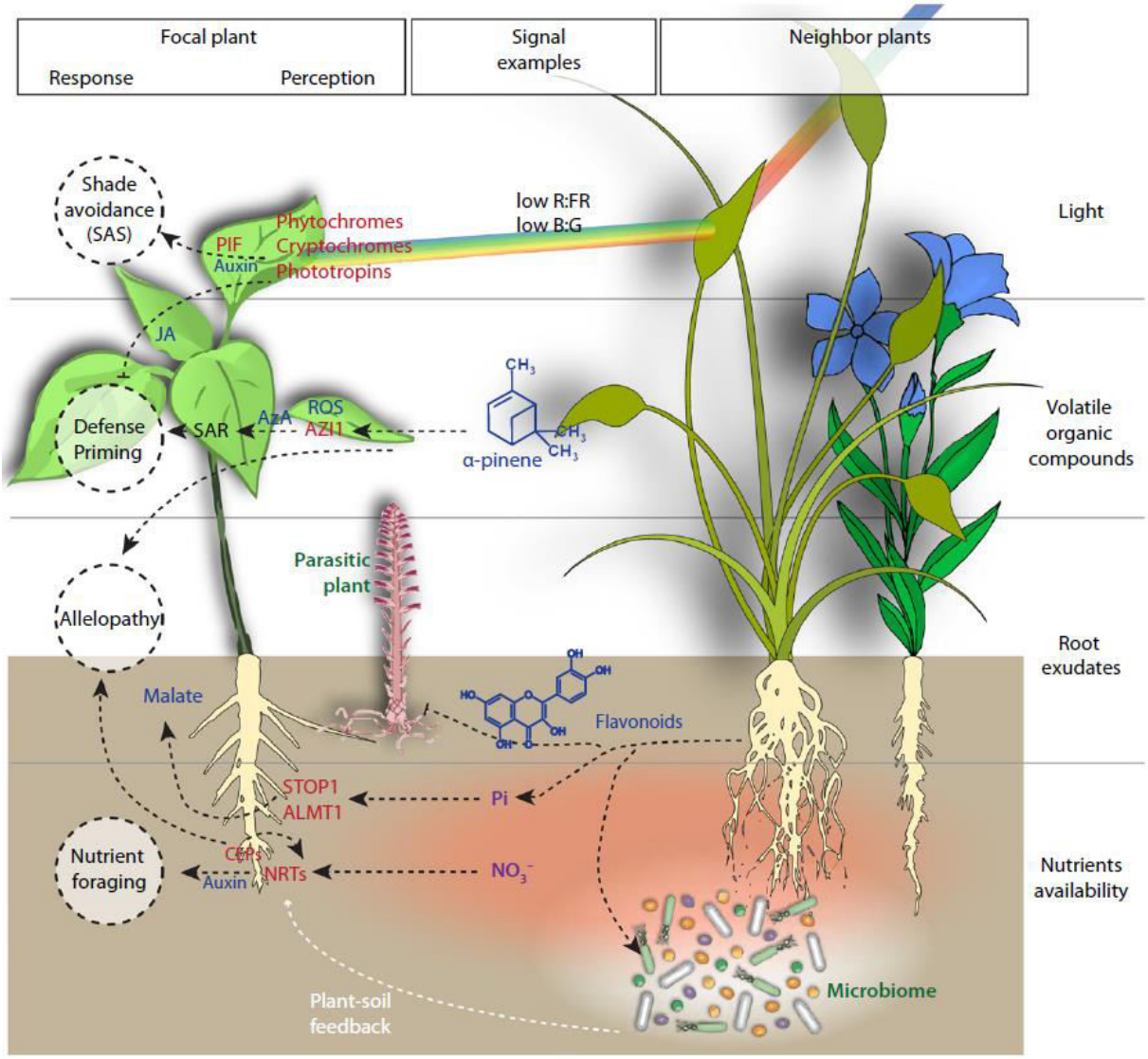

Figure 1. Neighbor detection and response strategies in plant-plant interactions. The main classes of signals and clues that mediate plant-plant communication are indicate on the right hand side of the figure: light, aerial volatile organic compounds (VOCs), root exudates and nutrient availability. Examples of signals of each class and components of the corresponding response mechanism in the focal plant discussed in the text are shown. Low ratios of red/far red light (R:FR) or blue/green light (B:G) light are signals associated with neighbor plants. Response to these signals involve notably phytochromes, cryptochromes, phototropins, PIF proteins, auxin and jasmonic acid (JA) hormones. A typical response is the shade avoidance syndrome (SAS). VOCs include for instance α- and β-pinene that mediate plant-plant interactions via AZI1 protein, reactive oxygen species (ROS) and the systemic signal azelaic acid (AzA). They trigger systemic acquired resistance (SAR) and defense priming. Root exudates include flavonoids (structure shows quercetin) and malate. Some of these compounds act on parasitic plants, on bacteria of the soil microbiome or on the availability of nutrients. Plant-plant interactions mediated by exudates are often designated as allelopathy and can result in alterations of the plant biomass allocation. The availability of nutrients such as Pi and NO_3^- is sensed by plant proteins such as STOP1, ALMT1 or NRTs, triggering the systemic movement of CEPs and release of auxin and malate notably. Plant peptides and proteins are labeled in dark red; plant metabolites are labeled in blue; black arrows show positive connections; black bar-headed lines show negative connections; some connections between elements of the figure have been omitted for clarity.

Light signal and shade avoidance syndrome

Due to absorbance of specific light wavelengths by chlorophyll, neighboring plants alter the quality and/or quantity of light perceived by the focal plant, triggering an escape strategy designated as the shade avoidance syndrome (SAS) (reviewed in Ballaré and Pierik, 2017). In crowded environments, the ratio of red/far-red light (R:FR) and blue/green light (B:G) are strongly reduced. These changes in light quality are perceived by phytochromes and cryptochromes respectively, through signaling pathways converging to the Phytochrome Interacting Factors (PIF) that integrate multiple light cues and adjust the SAS response (Fraser *et al.*, 2016; de Wit *et al.*, 2017). Interestingly, this light decoding system has been recently proposed to discriminate between kin and other neighbors, triggering altruistic or cooperative SAS (Crepy and Casal, 2015; but see Till-Botraud and de Villemereuil 2016). SAS involves the elongation of hypocotyls, stems and petioles, and the upward re-positioning of certain leaves (i.e. hyponasty) and requires biosynthesis of the auxin phytohormone (de Wit *et al.*, 2016). In *A. thaliana*, auxin fluxes have been recently shown to mediate local organ-specific responses in the focal plant to heterogeneous light signals originating from the surrounding plants (Michaud *et al.*, 2017; Pantazopoulou *et al.*, 2017). Using focal leaf illumination with low R:FR ratio, the authors showed that leaf tips, but not petioles, were sensitive to R:FR reduction, leading to a hyponastic response in the treated leaf only. In addition, global transcriptome and mutant analyses revealed the increased expression of auxin biosynthesis and auxin efflux carrier genes in response to low R:FR ratio at the leaf tip. Local exogenous auxin application and R:FR treatment on plants expressing auxin-reporter constructs indicated that auxin is transported from the leaf blade to the petiole to cause hyponasty. Mathematical leaf models showed *in silico* that perception and response to light signals in distinct leaf areas should increase growth in densely populated environments (Pantazopoulou *et al.*, 2017). Low R:FR ratio in densely planted sunflower fields leads to alternate stem inclinations, which is required for increased oil production per land area, thereby illustrating the adaptive value of neighbor-detection in cropping systems (Lopez Pereira *et al.*, 2017). Low R:FR ratio also modulates the expression of jasmonic acid (JA)-mediated immunity genes in a species-specific manner (Gommers *et al.*, 2017). JA signaling seems to be required for phytochrome B (i.e. the predominant phytochrome controlling SAS in response to low R:FR ratio) to repress plant immunity but not to trigger SAS under low R:FR ratio (Cerrudo *et al.*, 2017). Direct physical contact between leaf tips is able to induce leaf hyponasty in *A. thaliana* (de

Wit *et al.*, 2012). Such mechanosensing is also connected to defense priming via JA (Chehab *et al.*, 2012; Mbengue *et al.*, 2016), pointing towards a strong connection between physiological responses of the focal plant to its neighbor plants and pathogens.

Nutrients uptake and foraging

Uptake by roots of neighboring plants creates a heterogeneous nutrient and water environment for the focal plant that triggers morphological and physiological responses designated as foraging. These responses mainly correspond to the modulation of root distribution and architecture to increase nutrient uptake and the modulation of transport of nutrients and systemic signal across the plant (Aibara and Miwa, 2014). For example, nitrate sensing and transport by NRT1.1 leads to the repression of lateral root elongation in low nitrate conditions through the activation of the ANR1 transcription factor and modulation of auxin traffic (O'Brien *et al.*, 2017). Upon nitrogen depletion, roots secrete small C-terminally Encoded Peptides (CEPs) which are translocated to the shoot and are perceived by leucine-rich repeat receptor kinases (LRR-RKs) to activate nitrate transporters such as NRT1.1 (Sun *et al.*, 2017 for a review; Tabata *et al.*, 2014). Recent additions to the list of foraging regulators include STOP1 and ALMT1 that mediate phosphate-induced root remodeling through malate exudation (Balzergue *et al.*, 2017; Mora-Macias *et al.*, 2017). Depending on the source of phosphate available in the soil, the grass *Deschampsia cespitosa* produces more biomass when grown with a different grass species than with conspecifics, suggesting that nutrient availability regulates plant competition (Ahmad-Ramli *et al.*, 2013). In addition, growing on soil previously occupied by diverse plant genotypes increased nitrogen uptake in roots of *D. cespitosa* compared to soil conditioned by siblings (Semchenko *et al.*, 2017) and a role of soil microbiome was suggested.

Microbiome and other intermediates

The impact of soil microbiome on plant growth and responses to stress is being increasingly appreciated (Berendsen *et al.*, 2012; Lebeis *et al.*, 2015; Vandenkoornhuyse *et al.*, 2015). Microbes can also act as intermediates in plant-plant exchanges, such as in the case of

mycorrhizal networks connecting roots of several plants (reviewed in Selosse *et al.*, 2006). Nutrient exchanges through mycorrhizal networks can be highly asymmetric and may strongly favor the growth of some plant species over the others (Walder *et al.*, 2012). More generally, the impact of soil microbes on the dynamics of plant communities is designated as plant-soil feedback (PSF) (Bever *et al.*, 2011; van der Putten *et al.*, 2013). Two recent studies conducted on Mediterranean shrublands and temperate forests showed that plant diversity can be negatively impacted by soil pathogens (negative PSF) but also positively impacted by neighboring plants with distinct associations of symbiotic microbes for nutrient acquisition (positive PSF) (Bennet *et al.*, 2017; Teste *et al.*, 2017). In these studies, the strong protection against pathogens conferred by ectomycorrhizal fungi reduced plant diversity in favor of their host species, whereas arbuscular mycorrhizal fungi lead to the establishment of more diverse plant species. Reciprocally, soil suppressiveness against the fungal pathogen *Rhizoctonia solani* increased with species diversity in artificial plant communities (Latz *et al.*, 2015; Latz *et al.*, 2016). This is consistent with the information that the soils of permanent species-rich grasslands harbor a higher diverse microbiome and are more suppressive against soil-borne fungal pathogens than cultivated land (Garbeva *et al.*, 2006).

Root exudates

Nutrient and water availability, soil microbiome and small molecules released by plant roots (root exudates) form an interconnected network of belowground signals affecting plant-plant interactions (Bais *et al.*, 2006). Root exudates include a large diversity of molecules that are often species-specific (or even genotype-specific) and vary depending on the aboveground and belowground environment (reviewed in van Dam *et al.*, 2016; Massalha *et al.*, 2017). Root exudates also vary significantly at the intraspecific level. For example, some *A. thaliana* accessions lack an indolic glucosinolate hydrolysis product or a hydroxycinnamic acid conjugate, due to specific disruptive mutations affecting genes of the corresponding biosynthetic pathways (Monchgesang *et al.*, 2016). This very specific molecular signature is consistent with the idea that soluble root exudates could contribute to the ability of plants to recognize individuals of the same genotype from others (reviewed in Depuydt, 2014). For example, rice roots were shown to grow towards roots of plants from the same genotype, but away from roots of plants from different genotype (Fang *et al.*, 2013). Similarly, the growth

of *D. cespitosa* roots differed when treated with root exudates collected from plants of the same or different genotypes (Semchenko *et al.*, 2014).

Allelopathy is defined as the effect(s) of one plant on other plants through the release of chemical compounds in the environment (Rice, 1984; Olofsdotter *et al.*, 2002). Among chemical compounds, root-exuded allelochemicals, such as sorgoleone, that have negative growth effect on neighbor plants, are of primary importance to improve overall competitive ability of many crops (rice, wheat, barley, oat, sorghum…) against weeds (Olofsdotter *et al.*, 2002; Jensen *et al.*, 2008). Root-exuded allelochemicals produced by a plant can also have positive effects on other plants. Therefore, there has been a growing interest for the possible exploitation of these positive effects on plant growth in agricultural systems through intercropping (Brooker *et al.*, 2015). In maize-faba bean intercrops, maize root exudates were shown to promote flavonoid synthesis in faba bean, along with an increase of nodulation and nitrogen fixation (Li *et al.*, 2016). Reciprocally, faba bean root exudates increased maize growth *via* facilitating increased phosphorus availability in the soil (Zhang *et al.*, 2016). In this system, rhizobia that associate with faba bean root to fix nitrogen are important intermediates from the soil microbiome. Root exudates also play a key role in the interaction of crops with parasitic plants. Resistance to *Striga* parasitic plants in sorghum cultivars was found to result from a change of the dominant strigolactone 5-deoxystrigol (a highly active *Striga* germination stimulant) in root exudates to orobanchol, another strigolactone compound that does not stimulate *Striga* germination (Gobena *et al.*, 2017). *Desdemonium* plant species produce C-glycosylflavonoid in their root exudates that inhibit *Striga* parasitism on maize, making them useful intercrop species in some small-holder farms (Hooper *et al.*, 2015).

Volatile Organic Compounds

In response to endogenous or exogenous signals, plants can produce very diverse volatile organic compounds (VOCs). VOCs released in response to herbivore attack, such phenolics, alkaloids, terpenes, are well known to induce defense priming, conditioning stronger and faster subsequent defense responses (Baldwin *et al.*, 2006; Dicke and Baldwin, 2010). Owing to their long distance effect, VOCs can attract insect predators to prey-infected plants (Schnee *et al.*, 2006), and they are exploited by some plant pathogens to attract pollinators and favor

their dispersal (e.g. Roy, 1993). Recent work in *Petunia* flowers demonstrated that VOC emission can be mediated by its active transport across the plasma membrane. This could prevent toxic accumulation of VOCs in plant cells and increase the reach of emitted compounds (Adebesin *et al.*, 2017). Although various plant tissues can emit VOCs, most studies reporting plant-plant interactions mediated by VOCs involve airborne green leaf volatiles. For example, VOCs produced by damaged sagebrush plants protect neighboring *Nicotiana attenuata* plants from herbivores (Karban *et al.*, 2014). Soft mechanical stimulation also triggers VOC emission protecting plants from herbivores (Markovic *et al.*, 2016). VOCs emitted by undamaged neighboring plant can also trigger changes in biomass allocation between shoots and roots in focal plants (Ninkovic, 2003). In response to VOCs from heterospecific undamaged plants, potato plants modified the composition of their emitted VOC cocktail and were less frequently visited by aphids (Ninkovic *et al.*, 2013). By genetically manipulating VOC emissions in *N. attenuata* plants, Schuman *et al*. (2015) showed that the protective effects on the focal plant were dependent on the degree of herbivore infestation, while loss of protection in VOC-deficient plants was consistently compensated by neighboring plants. This suggests that targeted alterations in the VOC metabolism of a few plants could provide community-level protection in fields (Schuman *et al.*, 2015). By contrast, damage to a neighbor plant decreased protection against herbivores in the field for plants that were close relatives (Pearse *et al.*, 2012), highlighting the complexity of VOCs-mediated plant resistance in realistic environments (Baldwin *et al.*, 2006). During systemic acquired resistance (SAR), *A. thaliana* plants release α- and β-pinene VOCs, which in turn can elicit SAR in distal plants and protect them against the bacterial pathogen *Pseudomonas syringae* pv. *tomato*, through the induction of reactive oxygen species (ROS) accumulation (Riedlmeier *et al.*, 2017). This function requires the activity of the putative lipid-transfer protein AZI1 that stimulates the systemic movement of the SAR signal Azelaic acid (AzA) (Cecchini *et al.*, 2015; Riedlmeier *et al.*, 2017), possibly connecting plant-plant communication and the integration of plant defense signals in the focal plant.

THE GENETICS OF NATURAL VARIATION OF PLANT-PLANT INTERACTIONS

To obtain a complementary picture of the molecular bases underlying plant-plant interactions identified up to date, we examined studies where focal plants have been directly challenged by neighbor plants. Therefore, we have not considered studies performed in artificial environments designed to simulate plant-plant interactions, such as shade (Nagatani *et al.*, 1993; Reed *et al.*, 1993; Reed *et al.*, 1994) or root spatial constraint (Joseph *et al.*, 2015). Although simulated environments are highly powerful to decipher the molecular mechanisms underlying the perception of a particular signal (Gundel *et al.*, 2014; Ballaré and Pierik, 2017), it does not embrace the range and complexity of signals that are perceived by a focal plant directly interacting with a neighbor plant (Moriles *et al.*, 2012; Horvath *et al.*, 2015; **Figure 1**). Neither did we include studies focused on phenotypic traits thought to be involved in plant-plant interactions, such as improved seedling establishment and early growth measured in absence of plant-plant interactions (Addisu *et al.*, 2009). In addition, we did not cover either association studies or transcriptomic studies based on a restricted number of genes. We instead focused on studies reporting whole-genome scans. Based on these criteria, we identified a total of 63 studies reporting the identification of QTLs and/or candidate genes underlying natural plant-plant interactions in four conspecific interacting systems and 35 heterospecific interacting systems (**Table 1**).

Highlights on the nature of the data

Despite the limited number of studies that we identified, important observations have to be drawn before extracting trends on the genetic and molecular bases underlying natural plant-plant interactions. Firstly, as exemplified in *A. thaliana*, screening EMS mutants has been a widely adopted approach to start tracking down the molecular mechanisms underlying interactions with various pathogen species, in particular viruses and bacteria (Roux and Bergelson, 2016). However, although several EMS mutants initially identified in environments simulating plant-plant interactions have been subsequently tested for a role in direct interactions with neighboring plants (e.g. Schmitt *et al.*, 1995; Bates *et al.*, 2001; Cipollini, 2002; Fitter *et al.*, 2002; Cahill *et al.*, 2005), no studies reported a direct EMS mutant screen in presence of conspecifics or heterospecifics (**Table 1**). This discrepancy in

EMS mutant screens between plant-plant interactions and other types of biotic interactions may originate from the complexity of the establishment of the experiment involving interactions with neighbor plants. While screening for EMS mutants impaired in their interactions with microbial partners often requires only the spraying of a microbial solution on tens of thousands of seedlings sown at a high density, screening EMS mutants involved in plant-plant interactions would require the individual planting of the same number of seeds in presence of a neighbor plant.

Secondly, 16 of the 63 studies (~25.4%) correspond to analysis of global change in gene expression (**Table 1**). Among the remaining studies, forty-four studies (~69.8%) are based on traditional QTL mapping approaches using diverse experimental populations (F_2 populations, recombinant inbred lines, Doubled Haploid lines or back-crossed lines), while three studies (~4.7%) correspond to GWAS that have been all published in the last three years (**Table 1**). These GWAS are directly linked to the recent development of the Next-Generation Sequencing (NGS) technologies (Goodwin *et al.*, 2016; Lee *et al.*, 2016) that provide a substantial number of diverse genetic markers covering the whole genome (i.e. Single Nucleotide Polymorphisms, SNPs; copy number variation, CNV; indels, insertion-deletion), thereby allowing to fine-map genes underlying natural variation of complex traits (Bergelson and Roux, 2010). Although most GWAS in plants have been based on genetic lines collected over the entire geographic area of the studied plant species (Bartoli and Roux, 2017), we must however remember that a mapping population should be chosen according to the spatial scale at which natural variation is observed, i.e. according to the spatial scale of the ecological factors acting as selective pressures on the studied trait (Bergelson and Roux, 2010). For example, genomic regions associated with phenological variation were more significant at the regional scale than at the worldwide scale in *A. thaliana* (Brachi *et al.*, 2013). Because plants interact with neighbors over short distances, using highly genetically polymorphic local populations to fine map QTLs underlying plant-plant interactions appears to be more suitable than using worldwide genetic lines (Baron *et al.*, 2015). Therefore, in order to accelerate the identification of QTLs underlying natural variation of plant-plant interactions, we advocate development of local mapping populations that are known to interact with other plant species and are genetically diverse. While such populations can be identified within wild species (Frachon *et al.*, 2017), this may remain an important challenge in major crops.

Thirdly, the 16 transcriptomic studies are unevenly distributed between competitive interactions (n = 12 studies, equally distributed between interactions with conspecifics and interactions with heterospecifics) and asymmetric interactions (n = 4 studies, all focusing on response to parasitic plants) (**Table 1**). An opposite pattern is observed for QTL mapping studies. The 47 QTL mapping studies are unevenly distributed between competitive interactions (approximately one third), mostly testing intra-genotypic interactions (i.e. density effect), and asymmetric interactions (approximately two thirds) (**Table 1**). In the latter case, all the 30 corresponding studies were based on heterospecific interactions in the context of either allelopathy underlying weed suppressive ability (n = 8 studies) or response to parasitic plants (n = 22 studies) (**Table 1**). Surprisingly, no studies have been carried out on the identification of genetic and molecular bases underlying natural variation of positive interactions, such as facilitation and mutualism at the interspecific level and cooperation at the intraspecific level. It is a fact that negative plant-plant interactions, in particular competitive interactions, are thought of as the major factor responsible for crop yield reduction and for determining the structure of natural plant communities. However, this view has been recently challenged by several studies and the role of positive interactions (mostly facilitation) on overyielding in crop mixtures and in regulating the composition of natural plant communities has gained a lot of attention (Bertness and Callaway, 1994; Callaway, 1995; Brooker and Callaghan, 1998; Bruno *et al.*, 2003; Brooker *et al.*, 2008; Bukowski and Petermann, 2014; Li *et al.*, 2014; Wendling *et al.*, 2017). The next challenge is therefore the identification of candidate genes underlying positive interactions in various plant-plant interacting systems, which would enable testing whether some signaling pathways involved in response to neighbor presence are shared between competitive, asymmetric and reciprocal helping interactions.

Fourthly, 51 studies (~81%) involved ten crop species as focal plants, distributed across three botanical families, i.e. Asteraceae (*Helianthus annuus*), Fabaceae (*Glycine max*, *Pisum sativum*, *Vicia faba* and *Vigna unguiculata*) and Poaceae (*Hordeum vulgare*, *Oryza sativa*, *Sorghum bicolor*, *Triticum aestivum* and *Zea mays*) (**Table 1**). This major interest in crop species is consistent with the economic and environmental cost of crop weeds (Neve *et al.*, 2009) and with breeding programs for more density-related tolerant cultivars (St. Pierre *et al.*, 2011). Interestingly, while some crop species have been mainly studied for a specific type of plant-plant interactions such as *Zea mays* (i.e. response to intra-genotype competition), other

crop species have been studied for diverse types of plant-plant interactions such as *Oryza sativa* for competitive interactions with conspecifics and heterospecifics, for allelopathic effects on weeds and for response to parasitic plants (**Table 1**). The remaining focal species correspond to five wild species, i.e. *A. thaliana* (n = 8 studies), *Centaurea maculosa* (n = 1), *Medicago truncatula* (n = 1), *Solanum nigrum* (n = 1) and *Trifolium fucatum* (n = 1) (**Table 1**). In comparison to other types of biotic interactions (Roux and Bergelson, 2016; Bartoli and Roux, 2017), we observed only four QTL studies of plant-plant interactions in *A. thaliana*. This paucity of studies may stem from its status as a pioneer species; i.e. *A. thaliana* is not considered as being often challenged by other plant species in its natural habitats. However, several studies recently challenged this view (i) by revealing extensive genetic diversity in *A. thaliana* for the response to intra- and interspecific competition (Barthelheimer *et al.*, 2015), (ii) by finding that plant-plant interactions may act as selective agents on phenology in *A. thaliana* (Brachi *et al.*, 2012; Brachi *et al.*, 2013), and (iii) by demonstrating the *in situ* adaptive evolution of a highly genetically polymorphic local population of *A. thaliana* to increased interspecific competition in less than eight generations (Frachon *et al.*, 2017). Therefore, *A. thaliana* appears as a valuable model system for studying the genetics of natural variation of plant-plant interactions. On the side of neighbor species, 35 species have been used to study the genetics of plant-plant interactions (**Table 1**). This number, which is 2.3 times higher than that of the focal species, well illustrates the diversity of plant species faced by crop species in fields and wild species in natural settings (Wilson *et al.*, 2012). To summarize, we identified a total of 38 plant-plant interacting systems, corresponding to five conspecific and 33 heterospecific interacting systems (**Table 1**). Obviously, these interacting systems represent only a tiny fraction of the interactions shared between a species and its neighbors, particularly in natural environments. There is therefore an urgent need to increase the diversity of the plant-plant interacting systems studied, in particular for wild species. Such a diversity would certainly help to obtain a broader view of the pathways involved in the detection and response to the presence of neighbors.

Fifthly, the ability of a focal plant to interact with its neighbor plants results from both its competitive response (i.e. how strongly the focal plant is affected by its neighbors) and its competitive effect (i.e. how strongly a focal plant affects the performance of its neighbors) (Barthelheimer *et al.*, 2015). However, 36 QTL studies (~76.6%) reported the identification of QTLs of either one or the other component (**Table 1**). In addition, amongst these QTL

studies, the component of competitive ability under study was highly specific to the type of plant-plant interactions. Most studies on competitive interactions and asymmetric interactions reported the identification of QTLs associated with competitive response and competitive effect, respectively (**Table 1**). Interestingly, in studies reporting identification of QTLs for both components (n = 11 studies), QTLs of competitive response barely overlap with QTLs of competitive effect. While this observation suggests that competitive response may evolve independently from competitive effect (Baron *et al.*, 2015), we stress that considering simultaneously both competitive response and competitive effect would undoubtedly help to obtain an unbiased picture of the genetic and molecular mechanisms underlying the ability of a focal species to interact with its neighbor plants.

Sixthly, for competitive interactions, both above-ground (leaves) and below-ground (roots) traits have been used to study the global change of gene expression of a focal species in presence of a neighbor species, whereas all the traits measured in QTL mapping studies are exclusively above-ground (**Table 1)**. An opposite pattern was observed for asymmetric interactions. Roots have been exclusively used in transcriptomic studies on the response to parasitic plants, while both above-ground (n = 36) and below-ground (n = 27) traits have been measured in QTL mapping studies (**Table 1**). We also observed a higher number of measured traits in QTL studies on competitive interactions (mean = 8.9 traits / study, median = 8.5 traits / study) than in QTL studies on asymmetric interactions (mean = 2.1 traits / study, median = 1 trait / study) (**Table 1**). These observations are consistent with the difficulty of having access to the root compartment, especially in the case of QTL mapping studies that typically involve phenotyping several hundred or even thousands of individuals. The next frontier is therefore the development of high-throughput phenotyping for the precise root-root interactions (Mommer *et al.*, 2016). This challenge is already starting to be achieved by the development of image-analysis tools enabling quantitative analysis of root system architecture (Lobet *et al.*, 2011; Lobet *et al.*, 2013). For example, the use of a transparent gel system combined with image analysis and 3D reconstruction has allowed sophisticated analysis of the response of rice roots to another plant or physical object (Fang *et al.*, 2013). This study revealed a coordinated root system response integrating rhizophere signals into root architecture showing genotype specific root recognition *via* root tip signaling. However, experiments in controlled conditions can lead to artifactual plant responses and results need to be confirmed in less artificial conditions. Recently, novel methods have been developed to document the 3D root

system architecture within natural or field soils, using non-invasive (ground-penetrating radar for trees; Isaac *et al.*, 2013) or low-invasive tools (minirhizotrons; Johnson *et al.*, 2001).In addition, below-ground DNA-based techniques have been recently developed for quantifying species proportions in mixed root samples (Mommer *et al.*, 2008; Mommer *et al.*, 2011), thereby allowing to study below-ground species richness and rooting distributions (Kesanakurti *et al.*, 2011; Jones *et al.*, 2011) that can ultimately be linked to above-ground abundance (Frank *et al.*, 2010).

Seventhly, the environmental conditions in which phenotyping of plant-plant interactions was performed are well balanced between controlled (greenhouse/growth chamber) conditions (n = 32 studies) and field conditions (n = 26) (**Table 1**). As expected, the majority of studies on global change in gene expression were performed in controlled conditions in order to reduce variation among biological replicates. On the other hand, one transcriptomic study has taken this habit out of step, by challenging soybean plants with different weed species over three successive years in field conditions (Horvath *et al.*, 2015). This procedure allowed the authors to detect genes with consistent differential expression over the three growing seasons, thus uncovering genes underlying general soybean responses to weed presence. Four of the five remaining studies reported phenotyping experiments in both controlled and field conditions (Schmidt and Baldwin, 2006; Fondevilla *et al.*, 2010; Horvath *et al.*, 2015; Louarn *et al.*, 2016). While controlled and field conditions are complementary, natural selection acts in nature, where the neighbor plants and associated cues are numerous and complex. We therefore argue that identifying genes underlying natural variation of plant-plant interactions under natural conditions will be crucial for understanding the adaptation to the presence of neighbors, especially for wild species. Accordingly, a recent study reported for the first time (to our knowledge) a GWA mapping approach combined with an *in situ* phenotyping experiment of heterospecific interactions (Frachon *et al.*, 2017). In this study, 195 whole-genome sequenced natural accessions collected in a highly genetically polymorphic local population of *A. thaliana* were phenotyped *in situ* for 29 above-ground traits in six representative micro-habitats, consisting of the presence or absence of the bluegrass *Poa annua* (a species frequently associated with *A. thaliana* in its natural communities) crossed with three contrasting soil types. Interestingly, a minor fraction of the SNPs the most highly associated with the response to the *P. annua* was shared among the three soil types, stressing the need to consider the impact of abiotic conditions for the identification of the genetic bases

underlying competitive ability in a heterospecific neighborhood (Frachon *et al.*, 2017). Further experiments conducted in natural conditions will undoubtedly help to unravel the complexity of the molecular and genetic bases underlying natural plant-plant interactions.

Finally, in agreement with other types of biotic interactions (Bartoli and Roux, 2017), the majority of QTL mapping studies (n = 39) revealed a complex genetic architecture associated with plant-plant interactions (**Table 1**). The quantitative genetic architecture is highly diverse among plant-plant interacting systems, ranging from the identification of few medium-effect QTLs to the identification of up to tens of small-effect QTLs (Frachon *et al.* 2017). A monogenic architecture was reported for the remaining eight QTL mapping studies, all focusing on the natural variation of response to parasitic plants in three crop species, i.e. *H. annuus* (n = 3 studies), *S. bicolor* (n = 2 studies) and *V. unguiculata* (n = 3 studies). While there is a temptation to focus on cloning QTLs underlying binary traits, a polygenic architecture is more in line with theoretical works on adaptive walk to phenotypic optima (Hermisson and Pennings, 2005; Orr, 2005). Cloning medium (<30%) and small (<10%) effect QTLs rather than large-effect QTLs may therefore reveal genes involved in the adaptive response to the presence of a neighbor. Nonetheless, we should keep in mind that the functional validation of QTLs explaining less than 10% of phenotypic variation can require the phenotyping of up to thousands of plants, thereby explaining the scarcity of studies reporting the cloning of genes underlying small-effect QTLs whatever the type of biotic interactions considered (Bergelson and Roux, 2010; Roux and Bergelson, 2016).

The genetic and molecular bases underlying natural plant-plant interactions

It comes as no surprise that many more candidate genes were identified in the studies on global changes in gene expression than in the QTL mapping studies (in particular the traditional QTL mapping studies) (**Table 1**). Although the identity of the candidate genes barely overlaps between transcriptomics studies and QTL mapping studies, the functions of candidate genes are very overlapping (**Table 1**). The candidate genes can be classified into seven categories of plant function that have been identified in studies based on artificial environments designed to simulate plant-plant interactions either frequently (photosynthesis and hormones), only recently (cell wall modification and degradation, defense pathways

against pathogens) or rarely (ATP-binding cassette ABC transporters, histone modification, meristem identity / life history traits). This observation highlights the complementarity of identifying the genetic and molecular mechanisms underlying plant-plant interactions in artificial environments simulating plant-plant interactions and in environments where focal plants have been directly challenged by neighbor plants. We should also mention that (i) very few candidate genes have been identified as being involved in nutrient competition (Masclaux *et al.*, 2012), (ii) the function of the up-regulated and down-regulated genes can be highly dependent on the genotype tested, as found in barley and maize (St. Pierre *et al.*, 2011; Choe *et al.*, 2016), and (iii) several studies reported a substantial fraction of genes with unknown functions in their list of candidate genes (Horvath *et al.*, 2006; Broz *et al.*, 2008; Swarbrick *et al.*, 2008; Dita *et al.*, 2009; Bierzycki *et al.*, 2011a; Huang *et al.*, 2012; Schmid *et al.*, 2013; Baron *et al.*, 2015). The latter result suggests that some molecular mechanisms of neighbor perception and signaling pathways involved in the trigger of a response strategy remain to be identified.

Photosynthesis genes were specifically identified in presence of competitive interactions (**Table 1**). In most cases, photosynthesis genes were up-regulated in presence of conspecifics or heterospecifics, likely in connection with the shade-avoidance syndrome (SAS). Accordingly, *PHYTOCHROME B* (*PHYB*) that plays a central role in determining plant responses to changes in the R:FR ration caused by the proximity of other plants was upregulated both in barley and maize in high plant density conditions (St. Pierre *et al.*, 2011). In addition, in presence of intergenotypic competition, *PHYB* was proposed as a candidate gene for an overlapping QTL among three RIL families of *A. thaliana*. Based on transgenic analysis, further study confirmed experimentally that natural *PHYB* polymorphisms in *A. thaliana* cause differential responses in light sensitivity (Filiault *et al.*, 2008). Three studies reported a down-regulation of photosynthesis genes, all in presence of heterospecifics (Horvath *et al.*, 2006; Schmidt and Baldwin, 2006; Moriles *et al.*, 2012). A putative explanation relies on the permanent inhibition of photosynthesis that is induced when the focal plant is challenged in its early development by a neighbor plant, even if the focal plant overtops the neighbor plant later during its life cycle (Horvath *et al.*, 2006; Moriles *et al.*, 2012). The relative growth stage between two competing plants may therefore condition regulation of their photosystem genes.

The signal transduction network involved in SAS targets major physiological regulatory components such as the growth-associated hormones auxin, ethylene and gibberellins, whose biosynthesis is stimulated upon exposure to low R:FR ratios (Ballaré and Pierik, 2017). In agreement with the expression changes observed in photosystem genes, hormone-related genes were specifically detected in presence of competitive interactions (**Table 1**). In competition with conspecifics, auxin-related genes were upregulated in barley and maize (St. Pierre *et al.*, 2011; Chloe *et al.*, 2016) and a subunit (ASA1) of anthranilate synthase, which is involved in auxin synthesis, was proposed as a candidate gene underlying a QTL of response to inter-genotypic competition in *A. thaliana* (Mutic and Wolf, 2007). Ethylene-related genes were upregulated in *Trifolium fucatum* when competing with its congeneric *T. macraei* (Bowsher *et al.*, 2017). In addition, two candidates underlying two other QTLs of response to inter-genotypic competition in *A. thaliana* correspond to two polypeptides (ACS4 and ACS10) involved in the formation of 1-amino-cyclopropane-1-carboxylate synthase (ACS), which governs the rate-limiting step in ethylene formation (Mutic and Wolf, 2007). A GWAS reported the fine mapping of a genomic region associated with the length of reproductive period in *A. thaliana* in response to the presence of *Veronica arvensis* (Baron *et al.*, 2015). This genomic region of 30 kb contains the gene *AT5G66350* that codes for the SHI (for Short Internodes) protein involved in the perception of or in the response to gibberellin (Fridborg *et al.*, 1999).

Cell wall modification and degradation are important components of cell expansion, which is the driving force of organ elongation (Ballaré and Pierik, 2017). In line with the cell growth machinery being the ultimate target of the signal transduction network involved in SAS and hormone-related pathways, genes related to cell wall modification and degradation were upregulated in presence of intraspecific competition (Choe *et al.*, 2016; Bowsher *et al.*, 2017). A genomic region of less than 10kb associated with the length of reproductive period in *A. thaliana* has been fine mapped in response to the presence of *Stellaria media* (Baron *et al.*, 2015). The underlying candidate gene corresponds to the pectinacetyltransferase gene *AT5G26670*, which encodes a cell wall modification protein regulated by VOCs emitted by the rhizobacteria *Bacillus subtilis* (Zhang *et al.*, 2007). The latter case suggests that microbial-mediated below-ground communications between two plant species can ultimately lead to an above-ground adaptive response strategy.

Expression changes of numerous genes associated with defense pathways against microbial pathogens and insects have been observed in different types of plant-plant interactions (**Table 1**). Firstly, in response to parasitic plants, defense-related genes were up-regulated in incompatible interactions and down-regulated in compatible interactions. For example, the expression of *WRKY45*, a regulator of the salicylic acid / benzothiadiazole pathway, was highly induced in *Striga hermonthica*-infected rice (Mutuku *et al.*, 2015). In a study on an incompatible interaction with *Striga gesnerioides*, an up-regulation was observed in *Vignia unguiculata* for genes underlying programmed cell death and apoptosis (Huang *et al.*, 2012). In addition, the authors noticed that some genes and pathways up-regulated in *V. unguiculata* during incompatible interactions were also repressed during compatible interactions, suggesting that specific components of the host defense can be targeted and/or manipulated by *S. gesneriodes.* In line with those observations, the cloning of the first resistance gene in *V. unguiculata* to *S. gesneriodes* led to the identification of a predicted coiled-coil nucleotide-binding site leucine-rich repeat (CC-NBS-LRR) resistance protein (*R* gene) (Li and Timko, 2009). While this result suggests that similar molecular functions are shared among interactions involving microbial pathogenicity and plant parasitism, two recent studies reported the identification and functional validation of three genes conferring resistance to *Striga sp.* and having molecular functions that are distinct from *R* genes. The first study reported in rice was on the functional validation of two cytochrome P450 genes (*SBL1* and *SBL2* involved in the biosynthesis of the hormone strigolactone) as underlying a major QTL conferring resistance to the parasitic plant *S. hermonthica* (Cardoso *et al.*, 2014). The natural lines containing a deletion of *SBL1* and *SBL2* exuded lower amounts of strigolactone and had lower strigolactone content, thereby decreasing the rate of perception of the rice plants by *Striga sp.* (Cardoso *et al.*, 2014). The second study that was reported in sorghum contained the functional validation of *LGS1* (*LOW GERMINATION STIMULANT 1*) as underlying a major QTL conferring resistance to both *S. asiatica* and *S. hermonthica* (Gobena *et al.*, 2017). *LGS1* codes for an enzyme annotated as a sulfotransferase. Independent functional losses of *LGS1* in sorghum cultivars result in changes of the type of strigolactones present in the root exudates, i.e. from the dominant strigolactone 5-deoxystrigol (a highly active *Striga* germination stimulant) to orobanchol, another strigolactone compound that does not stimulate *Striga* germination (Gobena *et al.*, 2017). Secondly, for competitive interactions, none of the candidate genes proposed in QTL mapping studies are related to defense pathways (**Table 1**).

In addition, no clear pattern of regulation in defense-related genes was observed among the transcriptomic studies. For example, for conspecific interactions in *A. thaliana*, two studies reported an up-regulation of defense-related genes (Biedrzycki *et al.*, 2011a; Masclaux *et al.*, 2012), while a third study reported the opposite pattern (Geisler *et al.*, 2012). Up-regulation and down-regulation of defense-related genes were even reported within the same studies (Horvath *et al.*, 2015; Bowsher *et al.*, 2017). The growth-defense balance theory predicts that light perception by photoreceptors activates SAS and reduces the expression of defenses against microbial pathogens and insects, by a simultaneous down-regulation of jasmonate and salicylic acid signaling (Ballaré, 2014; Ballaré and Pierik, 2017). While some transcriptomic studies support this trade-off (Schmidt and Baldwin, 2006; Geisler *et al.*, 2012), other studies suggest an independent regulation of the SAS-related pathways and defense-pathways in the focal plant competing with a neighbor plant (Masclaux *et al.*, 2012; Horvath *et al.*, 2015; Bowsher *et al.*, 2017). Accordingly, some recent studies documented a reduction of disease severity in focal plants that were challenged by neighbor plants (**Table 2**). For example, a reduction of symptoms caused on soybean by the pathogenic fungus *Cylindrocladium parasiticum* was achieved in controlled conditions by growing maize in the same pot (Gao *et al.*, 2014). The direct interaction of soybean with maize roots induced, in soybean roots, the expression of most *Pathogenesis-Related* (*PR*) genes tested as well as the *Phenylalanine Ammonia Lyase* gene (*PAL*; involved in the biosynthesis of secondary metabolites). Indeed, the use of mesh or barrier separating the root systems from the two species showed that this induction of defense-related genes likely requires the diffusion of molecular signals from maize to soybean. Interestingly, exudates from maize were shown to contain salicylic acid, a potent inducer of systemic acquired resistance, which could also explain the induction of *PR* genes in soybean roots. Similar results were obtained in watermelon roots when grown together with wheat: *PAL* activity was higher in watermelon leaves and the induction of several defense-related genes was enhanced upon infection by pathogenic fungus *Fusarium oxysporum* (Xu *et al.*, 2015). In a couple of studies, exudates or purified molecules from root exudates produced by one plant species were shown to alter the expression of immunity markers in different plant species. For example, the expression of marker genes from several defense pathways was measured in shoots of healthy maize plants treated with root exudates from healthy pepper (Ding *et al.*, 2015). The induction of the AOS (Allene Oxide Synthetase) and AOC (Allene Oxide Cyclase), two genes involved in the biosynthesis of the jasmonic acid

hormone, in maize roots, was further correlated with a reduction of lesions caused by the fungal necrotrophic pathogen *Bipolaris maydis* on maize leaves (Ding *et al.*, 2015). In addition, a slight accumulation in the leaves of the secondary metabolite DIMBOA, a naturally occurring hydroxamic acid, was observed. This molecule and its major derivatives were shown to have an antimicrobial activity on *B. maydis in vitro*, suggesting that exudates from pepper roots can trigger induced systemic resistance in maize (Ding *et al.*, 2015). More recently, it was shown that p-coumaric acid secreted by rice roots could induce *PR* gene expression in watermelon and protect it against *F. oxysporum* when directly applied to watermelon (Ren *et al.*, 2016), possibly explaining the disease reduction observed when the two species are grown together (Ren *et al.*, 2008). The discrepancy between studies supporting the growth-defense balance theory and studies reporting positive effects of competitive interactions on plant immunity may originate from the diversity of signals perceived by a focal plant. While the growth-defense balance theory is mainly based on the perception of a low R:FR ratio, the perception of other signals in a more realistic environment may modify the interconnections within the network of regulatory pathways involved in the response of a focal plant to neighbor plants. Further experimental studies are clearly needed to resolve this discrepancy.

The three following categories of plant functions have only rarely been highlighted in studies on plant-plant interactions. However, since these categories have been mentioned in several studies where focal plants were directly challenged by neighbor plants, they deserve a particular attention. Firstly, ABC transporter genes were up-regulated in two studies on conspecific interactions in *A. thaliana* (Biedrzycki *et al.*, 2011a; Geisler *et al.*, 2012) and one study on response to the parasitic plant *S. hemonthica* in rice. In contrast, ABC transporter genes were down-regulated in one study on heterospecific interactions in *C. maculosa* (Broz *et al.*, 2008) (**Table 1**). Originally identified as transporters involved in detoxification processes, ABC transporters have ever since been described for being involved in a large diversity of processes, such as transport of defense-related chemicals and phytohormones (Kang *et al.*, 2011; Kretzschmar *et al.*, 2011; Hwang *et al.*, 2016). In the latter case, some ABC transporters are particularly essential to facilitate the communication between below- and above-ground structures, through the translocation of the signaling molecules cytokinins (Hwang *et al.*, 2016). An efficient communication system that coordinates the physiological and developmental processes between these two structures appears as a crucial point for a

focal plant to quickly adopting an appropriate response strategy to a neighbor plant. Interestingly, the role of three ABC transporters in the kin recognition response was confirmed in *A. thaliana* (Biedrzycki *et al.*, 2011b). In particular, their expression levels increased in the roots of plants exposed to stranger root secretions *vs* those exposed to own or kin secretions. Further functional studies are needed to establish whether ABC transporters may also be involved in recognition of heterospecific strangers.

Secondly, a plethora of histone-related genes was shown to be down-regulated in barley plant at high density (St. Pierre *et al.*, 2011) (**Table 1**). Based on the regulation of light-mediated chromatin compaction of the nuclear organizing regions (NORs) by *PHYB* and *HISTONE DEACETYLASE-6* in *A. thaliana* (Tessadori *et al.*, 2009), the authors proposed that the chromatin was more compact in plants grown at low density (i.e. with high light) than in plants grown at high density (i.e. with low light) (St. Pierre *et al.*, 2011). Furthermore, a GWAS reported the fine mapping of a genomic region associated with the number of basal branches in *A. thaliana* in response to the presence of *P. annua* (Baron *et al.*, 2015). This genomic region contains the gene *AT5G09740* that codes for the histone acetyltransferase HAM2 involved in the regulation of the expression of the well-known pleiotropic gene *FLOWERING LOCUS C* (Xiao *et al.*, 2013), a MADS-box transcription factor that regulates branching patterns in *A. thaliana* (Huang *et al.*, 2013). Those observations involving histone-related genes are intriguing and deserve in-depth investigation to validate their putative roles in competitive interactions.

Thirdly, genes related to either floral meristem identity and/or life history traits (such as flowering time and seed dispersal linked to branching patterns) were specifically identified in competitive interactions (**Table 1**). The identified candidate genes were either up-regulated in maize in heterospecific conditions (Moriles *et al.*, 2012) or proposed as underlying QTLs in five QTL mapping studies in both conspecific and heterospecific conditions (Botto and Collucio, 2007; Asif *et al.*, 2015; Granberry *et al.*, 2016; Frachon *et al.*, 2017; Kikuchi *et al.*, 2017) (**Table 1**). In competitive environments, such candidate genes may trigger an adaptive escape strategy that would correspond to an increased reproductive efficiency, mediated by a shortening of the life cycle and a faster reallocation of vegetative resources to reproductive structures, which is itself facilitated by an increased number of branches (Bonser *et al.*, 2013). In agreement with this hypothesis, in a natural plant community dominated by grasses, an

adaptive evolution towards an escape strategy was observed in *A. thaliana* in less than eight generations (Frachon *et al.*, 2017). This response to increased interspecific competition was mediated in part by the meristem identity gene *TWIN SISTER OF FT* (*TSF*) found associated with bolting time, the length of reproductive period and the number of branches on the main stem (Frachon *et al.*, 2017). This result suggests that phenotyping life history traits can help to obtain a better understanding of the genetic and molecular mechanisms underlying natural variation in plant-plant interactions, especially in wild plant species.

Although many studies proposed candidate genes involved in natural plant-plant interactions, only four of these studies (~6.3%) have been followed up by studies aiming at functionally validating the causal genes (**Table 1**). In QTL mapping studies, functional validation of candidate genes is however a pre-requisite to analyze the transcriptional and/or posttranscriptional regulation of the causal gene and to search for proteins directly interacting with the causal gene, that will in turn facilitate the identification of the downstream signaling pathways. Such complementary studies may then provide new candidate genes for breeding programs based on MAS.

FUTURE AVENUES

Here, we introduce several avenues that need to be explored in the future to obtain a thorough understanding of the genetic and molecular bases underlying plant-plant interactions within the context of realistic community complexity.

Identifying the genetic and molecular bases underlying natural variation of mutualism

As previously mentioned, studies reporting the genetic and molecular bases underlying natural variation of reciprocal helping are scarce (not to say absent) despite the role of positive plant-plant interactions on overyielding in crop mixtures and in regulating the composition of natural plant communities (Brooker *et al.*, 2008; Li *et al.*, 2014; Wendling *et al.*, 2017). Based on an innovative strategy recently developed for global genome-to-genome analysis and employed in the human-HIV pathosystem (Bartha *et al.*, 2013), we propose an ecological genomics strategy of GWA mapping to identify natural genetic variants underlying

mutualism between two plant species, without the need to obtain large phenotypic data sets. The strategy is composed of four steps (**Figure 2a**). The first step consists in collecting a substantial number (> 100) of paired genotypes (one per species) across a geographic area. The sampling area should be defined according to the same factors mentioned earlier for altruism (i.e., degree of coexistence, repeated interactions…). The second step would be testing for mutualism based on a small number of paired genotypes, i.e. whether the genotype from species A sampled in community X has a better performance in presence of the genotype from species B sampled from the same community than when growing alone or when growing with other genotypes from species B sampled in other communities. If mutualism between the two species is confirmed, the third step would include generating paired plant-plant genomic data, which will be facilitated by ever-cheaper genome-sequencing technologies. The fourth step would then comprise of performing joint association mapping analysis using both plant genomes in order to identify genetic markers in strong Linkage Disequilibrium across the two genomes. Based on co-evolutionary processes, this strategy of joint GWA mapping should allow description of the adaptive molecular landscape underlying mutualism between two plant species.

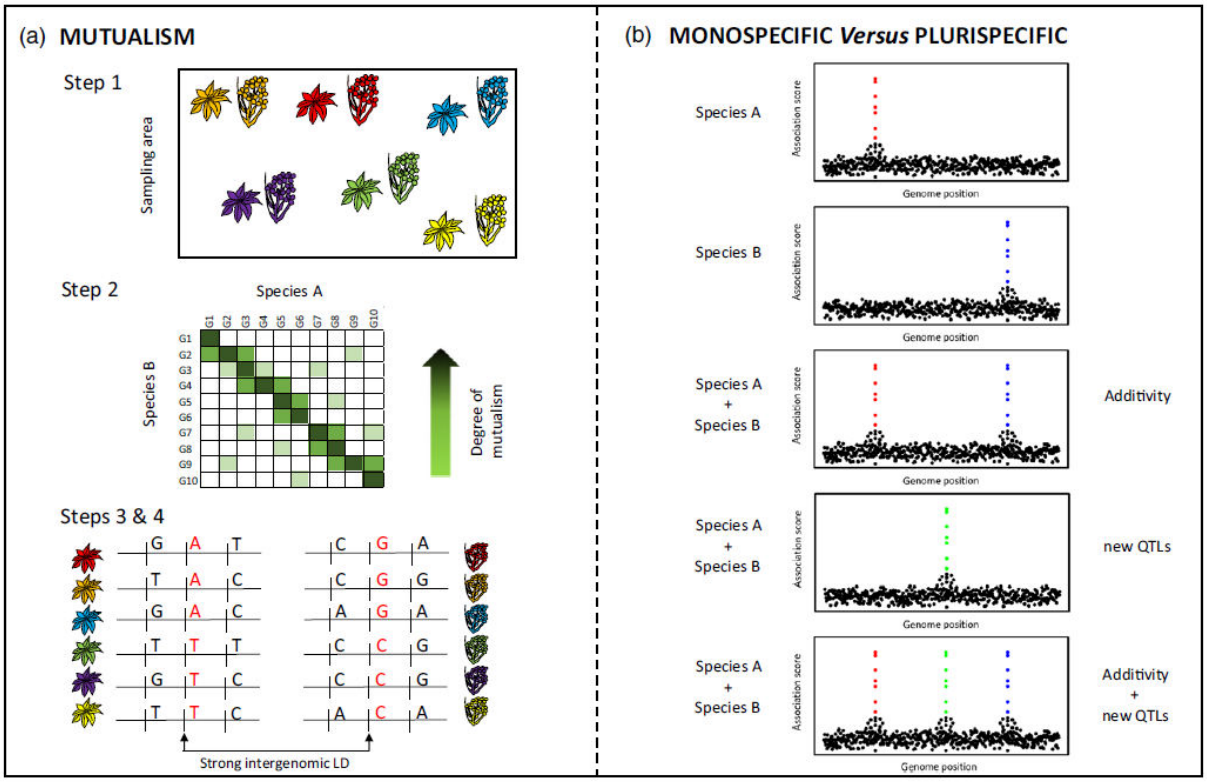

Figure 2. Future avenues on the genetics underlying natural variation of plant-plant interactions. (**a**) *Mutualism*. Step 1: paired sampling of genotypes from species A and genotypes from species B. Step 2: Testing for mutualism based a small number of paired genotypes. Steps 3 and 4: whole-genome sequencing of both plants species and genome-to-genome statistical analysis. (**b**) *Monospecific vs plurispecific heterospecific interactions*. Hypothetical genetic architectures expected in a focal species in the context of plurispecific competition with two species A and B, as illustrated by Manhattan plots based on GWAS.

Monospecific interactions vs *plurispecific interactions* vs *biotic diffuse interactions within plant communities*

Most studies reporting the genetic and molecular mechanisms of natural plant-plant interactions are based on monospecific heterospecific interactions (**Table 1**). However, a focal plant rarely interacts with only one neighbor species, either in crop fields or in more natural environments. Instead, a focal plant interacts simultaneously with multiple plant partners belonging to several species. This strengthens the need to study plant-plant interactions in a community context. One may wonder whether the phenotypic response of a focal plant to plurispecific interaction results from the additivity of the individual phenotypic responses to monospecific interactions. Similarly, it remains to be tested whether QTLs of competitive responses of a focal plant in a plurispecific neighborhood correspond to the sum of QTLs that

are specific to a neighbor species and/or to the emergence of new QTLs (**Figure 2b**). The approach for evaluating the operation of plurispecific interactions between more than two plant species will require the evaluation of the performance of numerous genetic lines in all two-way, three-way and so on combinations. For the experiment to be feasible, such an evaluation can only be achieved by considering a reasonable number of interacting species (i.e. 3-4 species), which will however still be less than the number of plant species that a focal species may encounter in its natural communities (up to 89 species; Wilson *et al.*, 2012). To resolve this issue, a Genome Environment Analysis (GEA) approach can be used to identify the genetic and molecular basis associated with the interaction of a focal species with multiple and simultaneous interactors in plant communities. This approach will require (i) the characterization of the plant communities associated with a given focal species, (ii) the genome sequencing of the focal species within each plant community, and (iii) statistical analyses aimed at identifying genomic regions of the focal species associated with descriptors of plant communities such as richness, alpha-diversity and composition. While GEA has been proved to successfully identify the genetic bases of adaptation to climate in plants (Hancock *et al.*, 2011), its power to identify the adaptive genetic bases to biotic diffuse and complex conditions remains to be tested. Nevertheless, by exploring diffuse biotic plant-plant interactions, the lofty goal of identifying adaptive QTLs associated with plant community descriptors can help to understand the role of community-wide selection.

ACKNOWLEDGMENTS

This work was funded by a PhD fellowship from the LABEX TULIP to HJS (ANR-10-LABX-41, ANR-11-IDEX-0002-02) and a PhD fellowship from the University of Toulouse to CL. This work was also supported by the Région Midi-Pyrénées (CLIMARES project) and the INRA Santé des Plantes et Environnement department (RESURRECTION project). SR is supported by a Starting Grant of the European Research Council (ERC-StG-336808). The authors declare no conflict of interest.

REFERENCES

Addisu, M., Snape, J.W., Simmonds, J.R. and Gooding, M.J. (2010) Effects of reduced height (*Rht*) and photoperiod insensitivity (*Ppd*) alleles on yield of wheat in contrasting production systems. *Euphytica*, *172*(2), 169–181.

Adebesin, F., Widhalm, J.R., Boachon, B. *et al.* (2017) Emission of volatile organic compounds from petunia flowers is facilitated by an ABC transporter. *Science*, *356*(6345), 1386-1388.

Ahmad-Ramli, M.F., Cornulier, T. and Johnson, D. (2013) Partitioning of soil phosphorus regulates competition between *Vaccinium vitis-idaea* and *Deschampsia cespitosa*. *Ecology and Evolution*, *3*(12), 4243–4252.

Aibara, I. and Miwa, K. (2014) Strategies for optimization of mineral nutrient transport in plants: multilevel regulation of nutrient-dependent dynamics of root architecture and transporter activity. *Plant and Cell Physiology*, *55*(12), 2027–2036.

Allen, B., Nowak, M.A. and Wilson, E.O. (2013) Limitations of inclusive fitness. *Proceedings of the National Academy of Sciences*, *110*(50), 20135–20139.

Asif, M., Yang, R.C., Navabi, A. *et al.* (2015) Mapping QTL, selection differentials, and the effect of *Rht-B1* under organic and conventionally managed systems in the Attila × CDC Go spring wheat mapping population. *Crop Science*, *55*(3), 1129–1142.

Bachelet, D., Neilson, R.P., Lenihan, J.M. and Drapek, R.J. (2001) Climate change effects on vegetation distribution and carbon budget in the United States. *Ecosystems.* 4, 164–185.

Bais, H.P., Weir, T.L., Perry, L.G., Gilroy, S. and Vivanco, J.M. (2006) The role of root exudates in rhizosphere interactions with plants and other organisms. *Annual Review of Plant Biology*, *57*(1), 233–266.

Baldwin, A.H., Jensen, K. and Schönfeldt, M. (2014) Warming increases plant biomass and reduces diversity across continents, latitudes, and species migration scenarios in experimental wetland communities. *Global Change Biology*, *20*(3), 835–850.

Baldwin, I.T., Halitschke, R., Paschold, A., von Dahl, C.C. and Preston, C.A. (2006) Volatile signaling in plant-plant interactions: "talking trees" in the genomics era. *Science*, *311*(5762), 812-815.

Ballaré, C.L. and Pierik, R. (2017) The shade-avoidance syndrome: multiple signals and ecological consequences. *Plant Cell and Environment*.

Ballaré, C.L. (2014) Light regulation of plant defense. *Annual Review of Plant Biology*, *65*(1), 335–363.

Balzergue, C., Dartevelle, T., Godon, C. *et al.* (2017) Low phosphate activates STOP1-ALMT1 to rapidly inhibit root cell elongation. *Nature Communications*, 15300.

Baron, E., Richirt, J., Villoutreix, R., Amsellem, L. and Roux, F. (2015) The genetics of intra- and interspecific competitive response and effect in a local population of an annual plant species. *Functional Ecology*, *29*(10), 1361–1370.

Bartha, I., Carlson, J. M., Brumme, C.J. *et al.* (2013) A genome-to-genome analysis of associations between human genetic variation, HIV-1 sequence diversity, and viral control. *eLife*, *2013*(2), 1–16.

Bartelheimer, M., Schmid, C., Storf, J., Hell, K. and Bauer, S. (2015) Interspecific competition in *Arabidopsis thaliana*: A knowledge gap is starting to close. In U. Lüttge & W. Beyschlag (Eds.), Vol. 76. *Progress in Botany* (pp. 303–319). Berlin: Springer.

Bartoli, C. and Roux, F. (2017) Genome-Wide Association studies in plant pathosystems: toward an ecological genomics approach. *Frontiers in Plant Science*, *8,* article 763.

Basu, C., Halfhill, M.D., Mueller, T.C. and Stewart, C.N. (2004) Weed genomics: new tools to understand weed biology. *Trends in Plant Science*, *9*(8), 391–398.

Bates, T. R. and Lynch, J. P. (2001) Root hairs confer a competitive advantage under low phosphorus availability. *Plant and Soil*, *236*(2), 243–250.

Bennett, J.A., Maherali, H., Reinhart, K.O., Lekberg, Y., Hart, M.M. and Klironomos, J. (2017) Plant-soil feedbacks and mycorrhizal type influence temperate forest population dynamics. *Science*, *355*(6321), 181-184.

Berendsen, R.L., Pieterse, C.M.J. and Bakker, P.A.H.M. (2012) The rhizosphere microbiome and plant health. *Trends in Plant Science*, *17*(8), 478–486.

Bergelson, J. and Roux, F. (2010) Towards identifying genes underlying ecologically relevant traits in *Arabidopsis thaliana*. *Nature Reviews Genetics*, *11*(12), 867–879.

Bertness, M.D. and Callaway, R. (1994) Positive interactions in communities. *Trends in Ecology and Evolution*, *9*(5), 187–191.

Bever, J.D., Dickie, I.A., Facelli, E. *et al.* (2011) Rooting theories of plant community ecology in microbial interactions. *Trends in Ecology & Evolution*, *2010*(8), 468–478.

Biedrzycki, M.L., Lakshmannan, V. and Bais, H.P. (2011a) Transcriptome analysis of *Arabidopsis thaliana* plants in response to kin and stranger recognition. *Plant Signaling & Behavior*, *6*(10), 1515–1524.

Biedrzycki, M. L., Lakshmannan, V. and Bais, H.P. (2011b) The role of ABC transporters in kin recognition in *Arabidopsis thaliana*. *Plant Signaling & Behavior*, *6*(8), 1154–61.

Bonser, S.P. (2013) High reproductive efficiency as an adaptive strategy in competitive environments. *Functional Ecology*, *27*(4), 876–885.

Botto, J.F. and Coluccio, M.P. (2007) Seasonal and plant-density dependency for quantitative trait loci affecting flowering time in multiple populations of *Arabidopsis thaliana*. *Plant, Cell and Environment*, *30*(11), 1465–1479.

Boukar, O., Kong, L., Singh, B.B., Murdock, L. and Ohm, H.W. (2004) AFLP and AFLP-derived SCAR markers associated with *Striga gesnerioides* resistance in cowpea. *Crop Science*, *44*(4), 1259–1264.

Bowsher, A.W., Shetty, P., Anacker, B.L., Siefert, A., Strauss, S.Y. and Friesen, M.L. (2017) Transcriptomic responses to conspecific and congeneric competition in co-occurring *Trifolium*. *Journal of Ecology*, *105*(3), 602–615.

Brachi, B., Aimé, C., Glorieux, C., Cuguen, J. and Roux, F. (2012) Adaptive value of phenological traits in stressful environments: predictions based on seed production and laboratory natural selection. *PLoS ONE*, *7*(3).

Brachi, B., Villoutreix, R., Faure, N. *et al.* (2013) Investigation of the geographical scale of adaptive phenological variation and its underlying genetics in *Arabidopsis thaliana*. *Molecular Ecology*, *22*(16), 4222–4240.

Brooker, R. W., Bennett, A. E., Cong, W.F. *et al.* (2015) Improving intercropping: a synthesis of research in agronomy, plant physiology and ecology. *New Phytologist*, *206*(1), 107–117.

Brooker, R.W. and Callaghan, T.V. (1998) The Balance between positive and negative plant interactions and its relationship to environmental gradients: a model. *Oikos*, *81*(1), 196–207.

Brooker, R.W., Maestre, F.T., Callaway, R.M. *et al.* (2008) Facilitation in plant communities: the past, the present, and the future. *Journal of Ecology*, *96*(1), 18–34.

Broz, A.K., Manter, D.K., Callaway, R.M., Paschke, M.W. and Vivanco, J.M. (2008) A molecular approach to understanding plant-plant interactions in the context of invasion biology. *Functional Plant Biology*, *35*(11), 1123–1134.

Bruno, J.F., Stachowicz, J.J. & Bertness, M.D. (2003) Inclusion of facilitation into ecological theory. *Trends in Ecology and Evolution*, **18**, 119–125.

Bukowski, A.R. and Petermann, J.S. (2014) Intraspecific plant-soil feedback and intraspecific overyielding in *Arabidopsis thaliana*. *Ecology and Evolution*, *4*(12), 2533–2545.

Cahill Jr., J.F., Kembel, S.W. and Gustafson, D.J. (2005) Differential genetic influences on competitive effect and response in *Arabidopsis thaliana*. *Journal of Ecology*, *93*(5), 958–967.

Callaway, R.M. (1995) Positive interactions among plants. *The Botanical Review*, *61*(4), 306–349.

Cardoso, C., Zhang, Y., Jamil, M. *et al.* (2014) Natural variation of rice strigolactone biosynthesis is associated with the deletion of two *MAX1* orthologs. *Proceedings of the National Academy of Sciences*, *111*(6), 2379–2384.

Cecchini, N. ., Steffes, K., Schläppi, M.R., Gifford, A.N. and Greenberg, J.T. (2015) *Arabidopsis* AZI1 family proteins mediate signal mobilization for systemic defence priming, *6*, 7658.

Cerrudo, I., Caliri-Ortiz, M.E., Keller, M.M., Degano, M.E., Demkura, P.V. and Ballaré, C.L. (2017) Exploring growth-defence trade-offs in *Arabidopsis*: phytochrome B inactivation requires JAZ10 to suppress plant immunity but not to trigger shade-avoidance responses. Plant, Cell & Environment, 40, 635–644.

Chehab, E.W., Yao, C., Henderson, Z., Kim, S. and Braam, J. (2012) *Arabidopsis* touch-induced morphogenesis is jasmonate mediated and protects against pests. *Current Biology*, *22*(8), 701–706.

Chesson, P. (2000) Mechanism of maintenance of species diversity. *Annual Review of Ecology and Systematics*, *31*, 343–66.

Choe, E., Drnevich, J. and Williams, M.M. (2016) Identification of crowding stress tolerance co-expression networks involved in sweet corn yield. *PLoS ONE*, *11*(1), 1–20.

Cipollini, D.F. (2002) Does competition magnify the fitness costs of induced responses in *Arabidopsis thaliana*? A manipulative approach. *Oecologia*, *131*(4), 514–520.

Clements, D.R., Ditommaso, A. and Hyvönen, T. (2014) Recent advances in weed management. 13–38.

Coleman, R., Gill, G. and Rebetzke, G. (2001) Identification of quantitative trait loci for traits conferring weed competitiveness in wheat *(Triticum aestivum L.)*. *Crop and Pasture Science* (Vol. 52).

Crepy, M.A. and Casal, J.J. (2015) Photoreceptor-mediated kin recognition in plants. *New Phytologist*, *205*(1), 329–338.

Crutsinger, G.M., Collins, M. D., Fordyce, J.A., Gompert, Z., Nice, C.C. and Sanders, N.J. (2006) Plant genotypic diversity predicts community structure and governs an ecosystem process. *Science*, *313*(5789), 966–968.

Depuydt, S. (2014). Arguments for and against self and non-self root recognition in plants. *Frontiers in Plant Science*, *5*, 1–7.

de Wit, M., Kegge, W., Evers, J.B., Vergeer-van Eijk, M.H., Gankema, P., Voesenek, L.A.C.J. and Pierik, R. (2012) Plant neighbor detection through touching leaf tips

precedes phytochrome signals. *Proceedings of the National Academy of Sciences*, *109*(36), 14705–14710.

de Wit, M., Keuskamp, D.H., Bongers, F.J. *et al.* (2017) Integration of phytochrome and cryptochrome signals determines plant growth during competition for light. *Current Biology*, *26*(24), 3320–3326.

de Wit, M., Galvão, V.C. and Fankhauser, C. (2016) Light-mediated hormonal regulation of plant growth and development. *Annual Review of Plant Biology*, *67*(1), 513–537.

Diaz-Ruiz, R., Torres, A., Gutierrez, M.V. *et al.* (2009) Mapping of quantitative trait loci controlling *Orobanche foetida* Poir. resistance in faba bean (*Vicia faba* L.). *African Journal of Biotechnology*, *8*(12), 2718–2724.

Díaz-Ruiz, R., Torres, A.M., Satovic, Z., Gutierrez, M.V., Cubero, J.I. and Román, B. (2010) Validation of QTLs for *Orobanche crenata* resistance in faba bean (*Vicia faba* L.) across environments and generations. *Theoretical and Applied Genetics*, *120*(5), 909–919.

Dicke, M. and Baldwin, I.T. (2010) The evolutionary context for herbivore-induced plant volatiles: beyond 'cry for help'. *Trends in Plant Science*, *15*(3), 167–175.

Ding, X., Yang, M., Huang, H. *et al.* (2015) Priming maize resistance by its neighbors: activating 1,4-benzoxazine-3-ones synthesis and defense gene expression to alleviate leaf disease. *Frontiers in Plant Science*, *6*(October), 1–11.

Dita, M.A., Die, J.V., RomÁn, B. *et al.* (2009) Gene expression profiling of *Medicago truncatula* roots in response to the parasitic plant *Orobanche crenata*. *Weed Research*, *49*(SUPPL. 1), 66–80.

Dudley, S.A. (2015) Plant cooperation. *AoB Plants*, *7*, plv113.

Ebana, K., Yan, W., Dilday, R.H., Namai, H. and Okuno, K. (2001) Analysis of QTL associated with the allelopathic effect of rice using water-soluble extracts. *Breeding Science*.

Fang, S., Clark, R.T., Zheng, Y. *et al.* (2013) Genotypic recognition and spatial responses by rice roots. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(7), 2670–5.

Filiault, D.L., Wessinger, C.A., Dinneny, J.R. *et al.* (2008) Amino acid polymorphisms in *Arabidopsis* phytochrome B cause differential responses to light. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(8), 3157–62.

Fitter, A., Williamson, L., Linkohr, B. and Leyser, O. (2002) Root system architecture determines fitness in an *Arabidopsi*s mutant in competition for immobile phosphate ions but not for nitrate ions. *Hungarian Quarterly*, *49*(191), 2017–2022.

Fondevilla, S., Fernández-Aparicio, M., Satovic, Z., Emeran, A.A., Torres, A.M., Moreno, M.T. and Rubiales, D. (2010) Identification of quantitative trait loci for specific mechanisms of resistance to *Orobanche crenata* Forsk. in pea *(Pisum sativum* L.). *Molecular Breeding*, *25*(2), 259–272.

Frachon, L., Libourel, C., Villoutreix, R. *et al.* (2017) Intermediate degrees of synergistic pleiotropy drive adaptive evolution in ecological time. *Nature Ecology & Evolution*. **1**, 1551–1561.

Frank, D.A., Pontes, A.W., Maine, E.M., Caruana, J., Raina, R., Raina, S. and Fridley, J.D. (2010) Grassland root communities: Species distributions and how they are linked to aboveground abundance. *Ecology*, *91*(11), 3201–3209.

Fraser, D.P., Hayes, S. and Franklin, K.A. (2016) Photoreceptor crosstalk in shade avoidance. *Current Opinion in Plant Biology*, *33*(Supplement C), 1–7.

Fridborg, I., Kuusk, S., Moritz, T. and Sundberg, E. (1999) The *Arabidopsis* dwarf mutant shi exhibits reduced gibberellin responses conferred by overexpression of a new putative zinc finger protein. *The Plant Cell*, *11*(6), 1019–1032.

Fu, X., Wu, X., Zhou, X., Liu, S., Shen, Y. and Wu, F. (2015) Companion cropping with potato onion enhances the disease resistance of tomato against *Verticillium dahliae*. *Frontiers in Plant Science*, *6*(September), 1–13.

Gao, X., Wu, M., Xu, R., Wang, X., Pan, R., Kim, H.-J. and Liao, H. (2014) Root interactions in a maize/soybean intercropping system control soybean soil-borne disease, Red Crown Rot. *PLOS ONE*, *9*(5), e95031.

Garbeva, P., Postma, J., Van Veen, J.A. and Van Elsas, J.D. (2006) Effect of above-ground plant species on soil microbial community structure and its impact on suppression of *Rhizoctonia solani* AG3. *Environmental Microbiology*, *8*(2), 233–246.

Geisler, M., Gibson, D.J., Lindsey, K.J., Millar, K. and Wood, A.J. (2012) Upregulation of photosynthesis genes, and downregulation of stress defense genes, is the response of *Arabidopsis thaliana* shoots to intraspecific competition. *Botanical Studies*, *53*(1), 85–96.

Gilman, S.E., Urban, M.C., Tewksbury, J., Gilchrist, G.W. and Holt, R.D. (2010) A framework for community interactions under climate change. *Trends in Ecology and Evolution*, *25*(6), 325–331.

Gobena, D., Shimels, M., Rich, P.J. *et al.* (2017) Mutation in sorghum *LOW GERMINATION STIMULANT 1* alters strigolactones and causes *Striga* resistance. *Proceedings of the National Academy of Sciences*, *114*(17), 4471–4476.

Goldberg, D.E. and Barton, A.M. (1992) Patterns and consequences of interspecific competition in natural communities - a review of field experiments with plants. *American Naturalist*, *139*(4), 771–801.

Gommers, C.M.M., Keuskamp, D.H., Buti, S. *et al.* (2017) Molecular profiles of contrasting shade response strategies in wild plants: differential control of immunity and shoot elongation. *The Plant Cell*, *29*(2), 331–344.

Gonzalo, M., Holland, J.B., Vyn, T.J. and McIntyre, L.M. (2010) Direct mapping of density response in a population of B73 × Mo17 recombinant inbred lines of maize (*Zea Mays* L.). *Heredity*, *104*(6), 583–599.

Gonzalo, M., Vyn, T.J., Holland, J.B. and McIntyre, L.M. (2006) Mapping density response in maize: a direct approach for testing genotype and treatment interactions. *Genetics*, *173*(1), 331–348.

Goodwin, S., McPherson, J.D. and McCombie, W.R. (2016) Coming of age: ten years of next-generation sequencing technologies. *Nature Reviews Genetics*, *17*(6), 333–351.

Granberry, M.C. (2016) Identifying Quantitative Trait Loci (QTL) associated with suppression of Italian Ryegrass in an AGS 2000 x NC-Neuse Soft Red Winter wheat population. PhD Thesis.

Groner, M.G. (1974) Intraspecific allelopathy in *Kalanchoe daigremontiana*. *Botanical Gazette*, *135*(1), 73–79.

Gundel, P.E., Pierik, R., Mommer, L. and Ballaré, C.L. (2014) Competing neighbors: light perception and root function. *Oecologia*, *176*(1), 1–10.

Guo, J., Chen, Z., Liu, Z. *et al.* (2011) Identification of genetic factors affecting plant density response through QTL mapping of yield component traits in maize (*Zea mays* L.). *Euphytica*, *182*(3), 409–422.

Gurney, A.L., Slate, J., Press, M.C. and Scholes, J.D. (2006) A novel form of resistance in rice to the angiosperm parasite *Striga hermonthica*. *New Phytologist*, *169*(1), 199–208.

Gutiérrez, N., Palomino, C., Satovic, Z. *et al.* (2013) QTLs for *Orobanche spp*. resistance in faba bean: Identification and validation across different environments. *Molecular Breeding*, *32*(4), 909–922.

Halty, V., Valdés, M., Tejera, M., Picasso, V. and Fort, H. (2017) Modelling plant interspecific interactions from experiments of perennial crop mixtures to predict optimal combinations. *Ecological Applications* (in press).

Hancock, A.M., Brachi, B., Faure, N. *et al.* (2011) *Arabidopsis thaliana* Genome. *Science*, *334*, 83–86.

Haussmann, B.I.G., Hess, D.E., Omanya, G.O. *et al.* (2004) Genomic regions influencing resistance to the parasitic weed *Striga hermonthica* in two recombinant inbred populations of sorghum. *Theoretical and Applied Genetics*, *109*(5), 1005–1016.

Haussmann, B.I.G., Hess,D. E., Welz, H.G. and Geiger, H.H. (2000) Improved methodologies for breeding striga-resistant sorghums. *Field Crops Research*, *66*(3), 195–211.

Hermisson, J. and Pennings, P.S. (2005) Soft sweeps: molecular population genetics of adaptation from standing genetic variation. *Genetics, 169,* 2335-2352.

Hooper, A.M., Caulfield, J.C., Hao, B., Pickett, J.A., Midega, C.A.O. and Khan, Z.R. (2015) Isolation and identification of *Desmodium* root exudates from drought tolerant species used as intercrops against *Striga hermonthica*. *Phytochemistry*, *117*, 380–387.

Horvath, D.P., Hansen, S.A., Moriles-Miller, J.P. *et al.* (2015) RNAseq reveals weed-induced PIF3-like as a candidate target to manipulate weed stress response in soybean. *New Phytologist*, *207*(1), 196–210.

Horvath, D.P., Gulden, R. and Clay, S.A. (2006) Microarray analysis of late- season velvetleaf (*Abutilon theophrasti*) impact on corn. *Weed Science,* 54, 983–994

Huang, X., Ding, J., Effgen, S., Turck, F. and Koornneef, M. (2013) Multiple loci and genetic interactions involving flowering time genes regulate stem branching among natural variants of Arabidopsis. *New Phytologist*, *199*(3), 843–857.

Huang, K., Mellor, K.E., Paul, S.N., Lawson, M.J., Mackey, A.J. and Timko, M.P. (2012) Global changes in gene expression during compatible and incompatible interactions of cowpea (*Vigna unguiculata* L.) with the root parasitic angiosperm *Striga gesnerioides*. *BMC Genomics*, *13*(1), 402.

Hwang, J.U., Song, W.Y., Hong, D. *et al.* (2016) Plant ABC transporters enable many unique aspects of a terrestrial plant's lifestyle. *Molecular Plant*, *9*(3), 338–355.

Imerovski, I., Dimitrijević, A., Miladinović, D., Dedić, B., Jocić, S., Tubić, N.K. and Cvejić, S. (2016) Mapping of a new gene for resistance to broomrape races higher than F. *Euphytica*, *209*(2), 281–289.

Isaac, M. and Anglaaere L.C.N. (2013) An *in situ* approach to detect tree root ecology: linking ground-penetrating radar imaging to isotope-derived water acquisition zones. *Ecology and Evolution, 3*, 1330–1339.

Isbell, F., Calcagno, V., Hector, A. *et al.* (2011) High plant diversity is needed to maintain ecosystem services. *Nature*, *477*(7363), 199–202.

Jensen, L.B., Courtois, B. and Olofsdotter, M. (2008) Quantitative trait loci analysis of allelopathy in rice. *Crop Science*, *48*(4), 1459–1469.

Jensen, L.B., Courtois, B., Shen, L., Li, Z., Olofsdotter, M. and Mauleon, R.P. (2001) Locating genes controlling allelopathic effects against barnyardgrass in upland rice. *Agronomy Journal*, *93*, 21–26.

Johnson, M.G., Tingey, D.T., Phillips, D.L. and Storm, M.J. (2001) Advancing fine root research with minirhizotrons. *Environmental and Experimental Botany, 45,* 263–289.

Johnson, M.T.J., Lajeunesse, M. . and Agrawal, A.A. (2006) Additive and interactive effects of plant genotypic diversity on arthropod communities and plant fitness. *Ecology Letters*, *9*(1), 24–34.

Jones, F.A., Erickson, D.L., Bernal, M.A. *et al.* (2011) The roots of diversity: below ground species richness and rooting distributions in a tropical forest revealed by DNA barcodes and inverse modeling. *PLoS ONE*, *6*(9). e24506.

Joseph, B., Lau, L. and Kliebenstein, D.J. (2015) Quantitative variation in responses to root spatial constraint within *Arabidopsis thaliana*. *The Plant Cell*, *27*(8), 2227–43.

Kaewchumnong, K. and Price, A.H. (2008) A study on the susceptibility of rice cultivars to *Striga hermonthica* and mapping of *Striga* tolerance quantitative trait loci in rice. *New Phytologist*, *180*(1), 206–216.

Kang, J., Park, J., Choi, H., Burla, B., Kretzschmar, T., Lee, Y. and Martinoia, E. (2011) Plant ABC transporters. *The Arabidopsis Book*, *9*, e0153.

Karban, R., Yang, L.H. and Edwards, K.F. (2014) Volatile communication between plants that affects herbivory: a meta-analysis. *Ecology Letters*, *17*(1), 44–52.

Kesanakurti, P.R., Fazekas, A.J., Burgess, K.S. *et al.* (2011) Spatial patterns of plant diversity below-ground as revealed by DNA barcoding. *Molecular Ecology*, *20*(6), 1289–1302.

Khanh, T.D., Xuan, T.D. and Chung, I.M. (2007) Rice allelopathy and the possibility for weed management. *Annals of Applied Biology*, *151*(3), 325–339.

Kikuchi, S., Bheemanahalli, R., Jagadish, K.S.V. *et al.* (2017) Genome-wide association mapping for phenotypic plasticity in rice. *Plant, Cell & Environment*, *40*(8), 1565–1575.

Kretzschmar, T., Burla, B., Lee, Y., Martinoia, E. and Nagy, R. (2011) Functions of ABC transporters in plants. *Essays In Biochemistry*, *50*, 145–160.

Ku, L., Ren, Z., Chen, X. *et al.* (2016) Genetic analysis of leaf morphology underlying the plant density response by QTL mapping in maize (*Zea mays* L.). *Molecular Breeding*, *36*(5), 1–16.

Ku, L., Zhang, L., Tian, Z. *et al.* (2015) Dissection of the genetic architecture underlying the plant density response by mapping plant height-related traits in maize (*Zea mays* L.). *Molecular Genetics and Genomics*, *290*(4), 1223–1233.

Latz, E., Eisenhauer, N., Rall, B.C., Scheu, S. and Jousset, A. (2016) Unravelling linkages between plant community composition and the pathogen-suppressive potential of soils. *Scientific Reports*, *6*(1), 23584.

Latz, E., Eisenhauer, N., Scheu, S. and Jousset, A. (2015) Plant identity drives the expression of biocontrol factors in a rhizosphere bacterium across a plant diversity gradient. *Functional Ecology*, *29*(9), 1225–1234.

Lebeis, S.L., Paredes, S.H., Lundberg, D.S. *et al.* (2015) Salicylic acid modulates colonization of the root microbiome by specific bacterial taxa. *Science*, *349*(6250), 860-864.

Lee, H., Gurtowski, J., Yoo, S. *et al.* (2016) Third-generation sequencing and the future of genomics. *bioRxiv*. 48603.

Lee, S., Seo, K.I., Koo, J.H., Hur, H.S. and Shin, J.C. (2005) QTLs and molecular markers associated with rice allelopathy. 4th World Congress on Allelopathy. *The Regional Institute Ltd*, p. 21-26.

Li, B., Li, Y.-Y., Wu, H.-M.*et al.* (2016) Root exudates drive interspecific facilitation by enhancing nodulation and N_2 fixation. *Proceedings of the National Academy of Sciences*, *113*(23), 6496–6501.

Li, J. and Timko, M.P. (2009) Gene-for-Gene resistance in *Striga*-Cowpea associations. *Science*, *325*(5944), 1094–1094.

Li, L., Tilman, D., Lambers, H. and Zhang, F.S. (2014) Plant diversity and overyielding: Insights from belowground facilitation of intercropping in agriculture. *New Phytologist*, *203*(1), 63–69.

Litrico, I. and Violle, C. (2015) Diversity in plant breeding: a new conceptual framework. *Trends in Plant Science*, *20*(10), 604–613.

Lobet, G., Pagès, L. and Draye, X. (2011) A novel image-analysis toolbox enabling quantitative analysis of root system architecture. *Plant Physiology*, *157*(1), 29–39.

Lobet, G., Draye, X. and Périlleux, C. (2013) An online database for plant image analysis software tools. *Plant Methods, 9*, 38.

López Pereira, M., Sadras V.O., Batista W., Casal, J.J. and Hall, A.J. (2017) Light-mediated self-organization of sunflower stands increases oil yield in the field. *Proceedings of the National Academy of Sciences*. 114(30),7975-7980.

Loreau, M. and de Mazancourt, C. (2013) Biodiversity and ecosystem stability: a synthesis of underlying mechanisms. *Ecology Letters*. 16, 106–115

Louarn, J., Boniface, M.-C., Pouilly, N., Velasco, L., Pérez-Vich, B., Vincourt, P. and Muños, S. (2016) Sunflower resistance to broomrape (*Orobanche cumana*) is controlled by specific QTLs for different parasitism stages. *Frontiers in Plant Science*, *7*, 1–14.

Lu, Y.H., Melero-Vara, J.M., García-Tejada, J.A. and Blanchard, P. (2000) Development of SCAR markers linked to the gene Or5 conferring resistance to broomrape (*Orobanche cumana* Wallr.) in sunflower. *Theoretical and Applied Genetics*, *100*(3), 625–632.

Markovic, D., Nikolic, N., Glinwood, R., Seisenbaeva, G. and Ninkovic, V. (2016) Plant responses to brief touching: a mechanism for early neighbour detection? *PLOS ONE*, *11*(11), e0165742.

Masclaux, F.G., Bruessow, F., Schweizer, F., Gouhier-Darimont, C., Keller, L. and Reymond, P. (2012) Transcriptome analysis of intraspecific competition in *Arabidopsis thaliana* reveals organ-specific signatures related to nutrient acquisition and general stress response pathways. *BMC Plant Biology*, *12*(1), 227.

Massalha, H., Korenblum, E., Tholl, D. and Aharoni, A. (2017) Small molecules below-ground: the role of specialized metabolites in the rhizosphere. *The Plant Journal*, 90: 788–807.

Mbengue, M., Navaud, O., Peyraud, R. *et al.* (2016) Emerging trends in molecular interactions between plants and the broad host range fungal pathogens *Botrytis cinerea* and *Sclerotinia sclerotiorum* . *Frontiers in Plant Science*, 31, article 422.

Michaud, O., Fiorucci, A., Xenarios, I. and Fankhauser, C. (2017) Local auxin production underlies a spatially restricted neighbor-detection response in *Arabidopsis*. *Proceedings of the National Academy of Sciences*. 114 (28) 7444-7449.

Mommer, L., Wagemaker, C.A.M., De Kroon, H. and Ouborg, N.J. (2008) Unravelling below-ground plant distributions: a real-time polymerase chain reaction method for quantifying species proportions in mixed root samples. *Molecular Ecology Resources*, *8*(5), 947–953.

Mommer, L., Dumbrell, A.J., Wagemaker, C.A.M. and Ouborg, N.J. (2011) Belowground DNA-based techniques: untangling the network of plant root interactions. *Plant and Soil*, *348*(1–2), 115–121.

Mommer, L., Kirkegaard, J. and van Ruijven, J. (2016) Root-Root interactions: towards a rhizosphere framework. *Trends in Plant Science*, *21*(3), 209–217.

Moncada, P., Martínez, C.P., Borrero, J. *et al.* (2001). Quantitative trait loci for yield and yield components in *an Oryza sativa* × *Oryza rufipogon* BC 2 F 2 population evaluated in an upland environment. *TAG Theoretical and Applied Genetics*, *102*(1), 41–52.

Mönchgesang, S., Strehmel, N., Schmidt, S. *et al.* (2016) Natural variation of root exudates in *Arabidopsis thaliana*-linking metabolomic and genomic data. *Scientific Reports*, *6*(1), 29033.

Mora-Macías, J., Ojeda-Rivera, J.O., Gutiérrez-Alanís, D. *et al.* (2017) Malate-dependent Fe accumulation is a critical checkpoint in the root developmental response to low phosphate. *Proceedings of the National Academy of Sciences*, *114*(17), E3563–E3572.

Moriles, J., Hansen, S., Horvath, D.P., Reicks, G., Clay, D.E. and Clay, S.A. (2012) Microarray and growth analyses identify differences and similarities of early corn response to weeds, shade, and nitrogen stress. *Weed Science*, *60*(2), 158–166.

Mutic, J.J. and Wolf, J.B. (2007) Indirect genetic effects from ecological interactions in *Arabidopsis thaliana*. *Molecular Ecology*, *16*(11), 2371–2381.

Mutuku, J.M., Yoshida, S., Shimizu, T. *et al.* (2015) The WRKY45-dependent signaling pathway is required for resistance against *Striga hermonthica* parasitism. *Plant Physiology* (Vol. 168).

Nagatani, A., Reed, J.W. and Chory, J. (1993) Isolation and initial characterization of *Arabidopsis* mutants that are deficient in Phytochrome A. *Plant Physiology*, *102*(1), 269–277.

Neve, P., Vila-aiub, M. and Roux, F. (2009) Evolutionary-thinking in agricultural weed managment. *Wiley*, *184*(4), 783–793.

Ninkovic, V. (2003) Volatile communication between barley plants affects biomass allocation. *Journal of Experimental Botany*, *54*(389), 1931–1939.

Ninkovic, V., Dahlin, I., Vucetic, A., Petrovic-Obradovic, O., Glinwood, R. and Webster, B. (2013) Volatile exchange between undamaged plants - a new mechanism affecting insect orientation in intercropping. *PLOS ONE*, *8*(7), e69431.

Nowak, M.A., Tarnita, C.E. and Wilson, E.O. (2010) The evolution of eusociality. *Nature, 466*, 1057-1062.

O'Brien, J.A., Vega, A., Bouguyon, E., Krouk, G., Gojon, A., Coruzzi, G. and Gutiérrez, R.A. (2017) Nitrate transport, sensing, and responses in plants. *Molecular Plant*, *9*(6), 837–856.

Oerke, E.C. (2006) Crop losses to pests. *Journal of Agricultural Science* 144, 31–43.

Okuno, K. and Ebana, K. (2003) Identification of QTL controlling allelopathic effects in rice: Genetic approaches to biological control of weeds. *Japan Agricultural Research Quarterly*, *37*(2), 77–81.

Olofsdotter, M., Jensen, L.B. and Courtois, B. (2002) Improving crop competitive ability using allelopathy - An example from rice. *Plant Breeding*, *121*(1), 1–9.

Orr, H.A. (2005) The genetic theory of adaptation: a brief history. *Nature Reviews Genetics, 6*, 119-127.

Ouédraogo, J.T., Maheshwari, V., Berner, D.K., St-Pierre, C.A., Belzile,F. and Timko, M. P. (2001) Identification of AFLP markers linked to resistance of cowpea (*Vigna unguiculata* L.) to parasitism by *Striga gesnerioides*. *Theoretical and Applied Genetics*, *102*(6–7), 1029–1036.

Ouédraogo, J.T., Tignegre, J.-B., Timko, M.P. and Beizile, F.J. (2002) AFLP markers linked to resistance against *Striga gesnerioides* race 1 in cowpea (*Vigna unguiculata*). *Genome, 45*(5), 787–793.

Pantazopoulou, C.K., Bongers, F.J., Küpers, J.J. *et al.* (2017) Neighbor detection at the leaf tip adaptively regulates upward leaf movement through spatial auxin dynamics. *Proceedings of the National Academy of Sciences*, *114*(28), 7450–7455.

Pearse, I.S., Porensky, L.M., Yang, L.H. *et al.* (2012) Complex consequences of herbivory and interplant cues in three annual plants. *PLoS One, 7*(5), e38105.

Peters, K., Breitsameter, L. and Gerowitt, B. (2014) Impact of climate change on weeds in agriculture: A review. *Agronomy for Sustainable Development*, *34*(4), 707–721.

Pierik, R., Mommer, L. and Voesenek, L.A. (2013) Molecular mechanisms of plant competition: Neighbour detection and response strategies. *Functional Ecology*, *27*(4), 841–853.

Prieto, I., Violle, C., Barre, P., Durand, J.-L., Ghesquiere, M. and Litrico, I. (2015) Complementary effects of species and genetic diversity on productivity and stability of sown grasslands. *Nature Plants*, *1*(4), 15033.

Reed, J. W., Nagpal, P., Poole, D.S., Furuya, M. and Chory, J. (1993) Mutations in the gene for the red/far-red light receptor phytochrome B alter cell elongation and physiological responses throughout *Arabidopsis* development. *The Plant Cell*, *5*(2), 147–157.

Reed, J. W., Nagatani, A., Elich, T.D., Fagan, M. and Chory, J. (1994) Phytochrome A and Phytochrome B have overlapping but distinct functions in Arabidopsis development. *Plant Physiology*, *104*(4), 1139–1149.

Ren, L., Huo, H., Zhang, F., Hao, W., Xiao, L., Dong, C. and Xu, G. (2016) The components of rice and watermelon root exudates and their effects on pathogenic fungus and watermelon defense. *Plant Signaling and Behavior*, *11*(6), 1–9.

Ren, L., Su, S., Yang, X., Xu, Y., Huang, Q. and Shen, Q. (2008) Intercropping with aerobic rice suppressed *Fusarium* wilt in watermelon. *Soil Biology and Biochemistry*, *40*(3), 834–844.

Rice, E.L. (1984) Allelopathy, 2nd edn. Academic Press, Orlando.

Riedlmeier, M., Ghirardo, A., Wenig, M. *et al.* (2017) Monoterpenes support systemic acquired resistance within and between plants. *The Plant Cell*, *29*(6), 1440 LP-1459.

Román, B., Torres, A.M.A., Rubiales, D., Cubero, J.I., Satovic, Z. and Román, B. (2002) Mapping of quantitative trait loci controlling broomrape (*Orobanche crenata* Forsk.) resistance in faba bean (*Vicia faba* L.). *Genome*, *45*, 1057–1063.

Roux, F. and Bergelson, J. (2016) The genetics underlying natural variation in the biotic interactions of *Arabidopsis thaliana*: the challenges of linking evolutionary genetics and community ecology. *Current Topics in Developmental Biology* (1st ed., Vol. 119). Elsevier Inc.

Roy, B.A. (1993) Floral mimicry by a plant pathogen. *Nature*, *362*(6415), 56–58.

Satish, K., Gutema, Z., Grenier, C., Rich, P.J. and Ejeta, G. (2012) Molecular tagging and validation of microsatellite markers linked to the low germination stimulant gene (lgs) for *Striga* resistance in sorghum [*Sorghum bicolor* (L.) Moench]. *Theoretical and Applied Genetics*, *124*(6), 989–1003.

Schmid, C., Bauer, S., Müller, B. and Bartelheimer, M. (2013) Belowground neighbor perception in *Arabidopsis thaliana* studied by transcriptome analysis: roots of *Hieracium pilosella* cause biotic stress. *Frontiers in Plant Science*, *4*(August), 1–17.

Schmidt, D.D. and Baldwin, I.T. (2006) Transcriptional responses of *Solanum nigrum* to methyl jasmonate and competition: A glasshouse and field study. *Functional Ecology*, *20*(3), 500–508.

Schmitt, J., McCormac, A.C. and Smith, H. (1995) A test of the adaptive plasticity hypothesis using transgenic and mutant plants disabled in phytochrome-mediated elongation responses to neighbors. *The American Naturalist*, *146*(6), 937.

Schnee, C., Kollner, T.G., Held, M., Turlings, T.C.J., Gershenzon, J. and Degenhardt, J. (2006) The products of a single maize sesquiterpene synthase form a volatile defense signal that attracts natural enemies of maize herbivores. *Proceedings of the National Academy of Sciences*, *103*(4), 1129–1134.

Schuman, M.C., Allmann, S. and Baldwin, I.T. (2015) Plant defense phenotypes determine the consequences of volatile emission for individuals and neighbors. *eLife*, *4*(2008), 1–43.

Selosse, M.-A., Richard, F., He, X. and Simard, S. W. (2017) Mycorrhizal networks: des liaisons dangereuses? *Trends in Ecology & Evolution*, *21*(11), 621–628.

Semchenko, M., Saar, S. and Lepik, A. (2017) Intraspecific genetic diversity modulates plant–soil feedback and nutrient cycling. *New Phytologist* 216, 90–98.

Semchenko, M., Saar, S. and Lepik, A. (2014) Plant root exudates mediate neighbour recognition and trigger complex behavioural changes. *New Phytologist*, *204*(3), 631–637.

Shimizu, H. (2012) Developmental change of vegetative plant architecture of annual-form-wild rice (*Oryza rufipogon* Griff.) elevates competitive ability during the late development under a dense condition. *American Journal of Plant Sciences*, *3*(5), 670–687.

Shimizu, H., Maruoka, M., Ichikawa, N., Baruah, A. R., Uwatoko, N., Sano, Y. and Onishi, K. (2010) Genetic control of phenotypic plasticity in Asian cultivated and wild rice in response to nutrient and density changes. *Genome.* 53(3), 211–23.

Singer, A., Travis, J.M.J. and Johst, K. (2013) Interspecific interactions affect species and community responses to climate shifts. *Oikos*, *122*(3), 358–366.

St. Pierre, S., Springer, N.M. and Muehlbauer, G.J. (2011) Density stress has minimal impacts on the barley or maize seedling transcriptome. *The Plane Genome*, *4*(1), 47–54.

Sun, C.-H., Yu, J.-Q. and Hu, D.-G. (2017) Nitrate: a crucial signal during lateral roots development. *Frontiers in Plant Science*, *8*, 1–9.

Swarbrick, P.J., Scholesa, J.D., Press, M.C. and Slatea, J. (2009) A major QTL for resistance of rice to the parasitic plant *Striga hermonthica* is not dependent on genetic background. *Pest Management Science*, *65*(5), 528–532.

Swarbrick, P., Huang, K., Liu, G., Slate, J., Press, M. and Scholes, J. (2008) Global patterns of gene expression in rice cultivars undergoing a susceptible or resistant interaction with the parasitic plant *Striga hermonthica*. *New Phytol*, *179*, 515–529.

Tabata, R., Sumida, K., Yoshii, T., Ohyama, K., Shinohara, H. and Matsubayashi, Y. (2014) Perception of root-derived peptides by shoot LRR-RKs mediates systemic N-demand signaling. *Science*, *346*(6207), 343-346.

Tang, S., Heesacker, A., Kishore, V.K., Fernandez, A., Sadik, E.S., Cole, G. and Knapp, S.J. (2003) Genetic mapping of the gene for resistance to race E in sunflower. *Crop Science*, *43*(3), 1021–1028.

Tessadori, F., van Zanten, M., Pavlova, P. *et al.* (2009) PHYTOCHROME B and HISTONE DEACETYLASE 6 control light-induced chromatin compaction in *Arabidopsis thaliana*. *PLOS Genetics*, *5*(9), e1000638.

Teste, F.P., Kardol, P., Turner, B.L., Wardle, D.A., Zemunik, G., Renton, M. and Laliberté, E. (2017) Plant-soil feedback and the maintenance of diversity in Mediterranean-climate shrublands. *Science*, *355*(6321), 173-176.

Teste, F.P., Veneklaas, E.J., Dixon, K.W. and Lambers, H. (2014) Complementary plant nutrient-acquisition strategies promote growth of neighbour species. *Functional Ecology*, *28*(4), 819–828.

Till-Bottraud, I. and de Villemereuil, P. (2016) Kin recognition or phenotype matching? *New Phytologist*, *209*(1), 13–14.

Tilman, D. (1985) The resource ratio hypothesis of succession. *The American Naturalist*, 125, 827-852.

Valderrama, M.R., Roman, B., Satovic, Z., Rubiales, D., Cubero, J.I. and Torres, A.M. (2004) Locating quantitative trait loci associated with *Orobanche crenata* resistance in pea. *Weed Research*, *44*(4), 323–328.

van Dam, N.M. and Bouwmeester, H.J. (2016) Metabolomics in the rhizosphere: tapping into belowground chemical communication. *Trends in Plant Science*, *21*(3), 256–265.

Vandenkoornhuyse, P., Quaiser, A., Duhamel, M., Le Van, A. and Dufresne, A. (2015) The importance of the microbiome of the plant holobiont. *New Phytologist*. 206, 1196–1206.

Van de Putten, W.H., Bardgett, R.D., Bever, J.D. *et al.* (2013). Plant-soil feedback: the past, the present and future challenges. *Journal of Ecology*, *101*(2), 265-276.

Walder, F., Niemann, H., Natarajan, M., Lehmann, M.F., Boller, T. and Wiemken, A. (2012) Mycorrhizal networks: common goods of plants shared under unequal terms of trade. *Plant Physiology*, *159*(2), 789–797.

Weiner, J., Andersen, S.B., Wille, W.K.M., Griepentrog, H.W. and Olsen, J.M. (2010) Evolutionary Agroecology: The potential for cooperative, high density, weed-suppressing cereals. *Evolutionary Applications*, *3*(5–6), 473–479.

Wendling, M., Büchi, L., Amossé, C., Jeangros, B., Walter, A. and Charles, R. (2017) Specific interactions leading to transgressive overyielding in cover crop mixtures. *Agriculture, Ecosystems and Environment*, *241*, 88–99.

Whitham, T.G., Bailey, J.K., Schweitzer, J.A. *et al.* (2006) A framework for community and ecosystem genetics: from genes to ecosystems. *Nature Reviews Genetics*, *7*(7), 510–523.

Williams, M.E. (2013) Plants are not alone. teaching tools in plant biology. *The Plant Cell (online)*.

Wilson, J.B., Peet, R.K., Dengler, J. and Pärtel, M. (2012) Plant species richness: The world records. *Journal of Vegetation Science*, *23*(4), 796–802.

Worthington, M. and Reberg-Horton, C. (2013) Breeding cereal crops for enhanced weed suppression: optimizing allelopathy and competitive ability. *Journal of Chemical Ecology*, *39*(2), 213–231.

Wu, H., Pratley, J., Ma, W. and Haig, T. (2003) Quantitative trait loci and molecular markers associated with wheat allelopathy. *Theoretical and Applied Genetics*, *107*(8), 1477–1481.

Xiao, J., Zhang, H., Xing, L., Xu, S., Liu, H., Chong, K. and Xu, Y. (2013) Requirement of histone acetyltransferases HAM1 and HAM2 for epigenetic modification of *FLC* in regulating flowering in *Arabidopsis*. *Journal of Plant Physiology*, *170*(4), 444–451.

Xu, W., Liu, D., Wu, F. and Liu, S. (2015) Root exudates of wheat are involved in suppression of *Fusarium* wilt in watermelon in watermelon-wheat companion cropping. *European Journal of Plant Pathology*, *141*(1), 209–216.

Yang, L., Callaway, R.M. and Atwater, D.Z. (2017) Ecotypic diversity of a dominant grassland species resists exotic invasion. *Biological Invasions*, *19*(5), 1483–1493.

Zeng, D., Qian, Q., Teng, S., Dong, G., Fujimoto, H., Yasufumi, K. and Zhu, L. (2003) Genetic analysis of rice allelopathy. *Chinese Science Bulletin*, *48*(3), 265–268.

Zhang, H., Kim, M.S., Krishnamachari, V.*et al.* (2007) Rhizobacterial volatile emissions regulate auxin homeostasis and cell expansion in *Arabidopsis*. *Planta*, *226*(4), 839–851.

Zhang, D., Zhang, C., Tang, X., Li, H., Zhang, F., Rengel, Z., Whalley, W. R., Davies, W. J. and Shen, J. (2016) Increased soil phosphorus availability induced by faba bean root exudation stimulates root growth and phosphorus uptake in neighbouring maize. New Phytol, 209, 823–831.

Zuo, S., Liu, G. and Li, M. (2012) Genetic basis of allelopathic potential of winter wheat based on the perspective of quantitative trait locus. *Field Crops Research*, *135*, 67–73.

Table 1. Studies on global change in gene expression and QTL mapping studies reporting the genetics of natural plant-plant interactions.

Approach[a]	Type of interaction	Focal species[b]	Neighbor species / Treatment	CR[c]	CE[c]	Plant organ[d]	Growth cond.[e]	Gen. Arch.[f]	Candidate genes[g]	References[h]
Global change in gene expression	competition conspecific	*Arabidopsis thaliana*	intra-genotypic competition *vs* inter-genotypic competition	ne	ne	roots	C	ne	UP: ABC transporter, pathogen response, metabolism, cytochrome	Biedrzycki *et al.* , 2011
		Arabidopsis thaliana	intra-genotypic competition (four density conditions)	ne	ne	shoot and leaves	C	ne	UP: photosynthesis, ABC transporter, metal ion transporter DOWN: regulation of active oxygen species, defense response DOWN in low density but UP in high density : biotic and abiotic stress perception	Geisler *et al.* , 2012
		Arabidopsis thaliana	absence and presence of intra-genotypic competition at multiple densities	ne	ne	rosette leaves and roots	C	ne	UP (leaves + roots): nutrient starvation, biotic and abiotic stress response, defense response UP (leaves): genes expressed during low R:FR ratio	Masclaux *et al.* , 2012
		Hordeum vulgare (4 genotypes)	intra-genotypic competition (two density conditions)	ne	ne	leaves	C	ne	UP: photosystem, auxin, DOWN: histones	St. Pierre *et al.* , 2011
		Trifolium fucatum	absence and presence of inter-genotypic competition	ne	ne	roots	C	ne	UP: disease resistance, cell wall modification and flavanoid pathway DOWN: disease resistance	Bowsher *et al.* , 2017
		Zea mays (5 genotypes)	intra-genotypic competition (two density conditions)	ne	ne	leaves	C	ne	UP: auxin, photosystem	St. Pierre *et al.* , 2011
		Zea mays (3 high-yielding hybrids + 3 low-yielding hybrids)	intra-genotypic competition	ne	ne	ear leaves	F	ne	UP: amino acid degradation/polyamine metabolism, cell wall degradation, glycolysis cytosolic branch, hormone metabolism (auxin), RNA transcription regulation, signaling receptor kinases and tetrapyrrole synthesis DOWN: flavonoids secondary metabolism, miscellaneous enzyme families, protein folding, protein post-translational modification and development	Choe *et al.* , 2016
	heterospecific	*Arabidopsis thaliana*	*Hieracium pilosella*	ne	ne	roots	C	ne	UP: photosynthesis, biotic stress response ,transcription factors, signaling lipids DOWN: tricarboxylic acid, amino acid metabolism, glucosinolate	Schmid *et al.*, 2013
		Centaurea maculosa	*Gaillardia aristata, Festuca idahoensis*	ne	ne	leaves and roots	C	ne	UP (in presence of each species): ADP ribosylation DOWN (in presence of each species): ABC transporter	Broz *et al.* , 2008
		Glycine max	*Abutilon thoephrasti* , *Helianthus annuus* , *Linum usitatissimum* , *Polygonum convolvulus* , *Brassica napus*	ne	ne	leaves	F + C	ne	UP: phytochrome signaling, photosynthesis, shade avoidance, oxidative stress responses, salicylic acid DOWN: heat shock response, protein synthesis, jasmonate acid signalling	Horvath *et al.* , 2015
		Solanum nigrum	*Plantago lanceolata, Lolium perenne, Trifolium pratense*	ne	ne	leaves	F + C	ne	UP (in presence of each species): signal response, defense response DOWN: photosynthesis	Schmidt and Baldwin, 2006
		Trifolium fucatum	*Trifolium macraei*	ne	ne	roots	C	ne	UP: abiotic and biotic stress, ethylene responses, protein synthesis and photosynthesis, disease resistance, heat shock response DOWN: disease resistance	Bowsher *et al.* , 2017
		Zea mays	*Abutilon theophrasti*	ne	ne	leaves	F	ne	DOWN: carbon and nitrogen utilization, photosynthesis, oxidative stress, signal transduction, responses to auxin and ethylene, zinc transport	Horvath *et al.* , 2006
		Zea mays	*Abutilon thoephrasti* , *Brassica napus*	ne	ne	leaves	F	ne	UP: 13-LOX, 13-HPL, nutrient reservoir activity, meristem identity DOWN: light/photosynthesis, energy conversion, transcription activity, protein kinase activity	Moriles *et al.* , 2012

Table 1. Continued

Approach[a]	Type of interaction	Focal species[b]	Neighbor species / Treatment	CR[c]	CE[c]	Plant organ[d]	Growth cond.[e]	Gen. Arch.[f]	Candidate genes[g]	References[h]
Global change in gene expression	asymmetric parasitism									
	heterospecific	*Medicago truncatula*	*Orobanche crenata*	ne	ne	roots	C	ne	UP: cell wall degradation and modification, primary metabolism, defense response DOWN: pirmary metabolism, defense response and cell rescue	Dita *et al.*, 2009
		Oryza sativa (resistant cutlivar)	*Striga hermonthica*	ne	ne	roots	C	ne	UP: defense response, ABC transporters, phenylpropanaoid metabolism, WRKY transciptions factors	Swarbrick *et al.*, 2008
		Oryza sativa (susceptible cultivar)	*Striga hermonthica*	ne	ne	roots	C	ne	UP: nutrient transporters, enzymes of amino acid metabolism DOWN: plant growth regulator signalling and metabolism, biogenesis of cellular components and cell division	Swarbrick *et al.*, 2008
		Oryza sativa	*Striga hermonthica*	ne	ne	roots	C	ne	UP: defense response, jasmonic acid, salicycilic acid ***WRKY45***	Mutuku *et al.*, 2015
		Vigna unguiculata	*Striga gesnerioides* - incompatible interaction	ne	ne	roots	C	ne	UP: signal transduction, programmed cell death and apoptosis, components of lignification, secondary wall formation DOWN: cell cycle regulators, cellular transporters	Huang *et al.*, 2012
		Vigna unguiculata	*Striga gesnerioides* - compatible interaction	ne	ne	roots	C	ne	UP: cellular transport processes for nitrogen and sulfur DOWN: defense pathways, lignin biosynthesis, secondary wall modifications, plant growth regulators	Huang *et al.*, 2012
QTL mapping Traditional QTL mapping	competition conspecific	*Arabidopsis thaliana* (411 RILs)	only presence of inter-genotypic competition	X	X	above-ground (n = 9 + 9)	C	P	*ACS4*, *ACS10*, *ASA1*	Mutic and Wolf, 2007
		Arabidopsis thaliana (three RIL populations, 340 RILs)	absence and presence of inter-genotypic competition	X		above-ground (n = 2)	C	P	*PHYB*, *FLC*, *FRI*	Botto and Coluccio, 2007
		Oryza sativa x *Oryza rufipogon* (35 F_7 RILs)	intra-genotypic competition (two density conditions)	X	X	above-ground (n = 2 + 1)	F	P	nr	Shimizu and Itoh, 2012
		Oryza sativa x *Oryza rufipogon* (89 F_7 RILs)	intra-genotypic competition (three density conditions)	X		above-ground (n = 4)	C	P	nr	Shimizu *et al.*, 2010
		Zea mays (8 SILs)	intra-genotypic competition (two density conditions)	X		above-ground (n = 12)	F	P	nr	Gonzalo *et al.*, 2006
		Zea mays (186 RILs)	intra-genotypic competition (two density conditions)	X		above-ground (n = 8)	F	P	nr	Gonzalo *et al.*, 2010
		Zea mays (231 $F_{2:3}$ families)	intra-genotypic competition (two density conditions)	X		above-ground (n = 12)	F	P	nr	Guo *et al.*, 2011
		Zea mays (four RIL families, 840 RILs)	intra-genotypic competition (two density conditions)	X		above-ground (n = 3)	F	P	nr	Ku *et al.*, 2015
		Zea mays (four RIL families, 858 RILs)	intra-genotypic competition (two density conditions)	X		above-ground (n = 6)	F	P	nr	Ku *et al.*, 2016
		Zea mays (220 F_{10} RILs)	intra-genotypic competition (three density conditions)	X		above-ground (n = 3)	F	P	nr	Wang *et al.*, 2017
	heterospecific	*Oryza sativa* x *Oryza rufipogon* (274 BC_2F_2 families)	*Brachiara brizanta*	X		above-ground (n = 7)	F	P	nr	Moncada *et al.*, 2001
		Triticum aestivum (161 Doubled Haploid lines)	*Lolium rigidum*	X	X	above-ground (n = 9 + 1)	F	P	nr	Coleman *et al.*, 2001
		Triticum aestivum (163 F_7-derived RILs)	multiple weed species (conventional *vs* organic plots)	X	X	above-ground (n = 9 + 1)	F	P	*Rht-B1*	Asif *et al.*, 2015
		Triticum aestivum (140 F_5-derived RILs)	*Lolium perenne*	X	X	above-ground (n = 6 + 1)	F	P	*Ppd1-B1*, *Vrn-B1*, *Rht*, *Vrn3*	Granberry, 2016

Table 1. Continued

Approach[a]	Type of interaction	Focal species[b]	Neighbor species / Treatment	CR[c]	CE[c]	Plant organ[d]	Growth cond.[e]	Gen. Arch.[f]	Candidate genes[g]	References[h]
QTL mapping										
Traditional QTL mapping	asymmetric allelopathy									
	heterospecific	*Oryza sativa* (121 RILs)	*Echinochloa crus-galli*		X	below-ground (n = 1)	C	P	nr	Jensen *et al.*, 2001
		Oryza sativa (192 F_3 families)	*Lactuca sativa*		X	below-ground (n = 1)	C	P	nr	Ebana et *al.*, 2001 Okuno and Ebana, 2003
		Oryza sativa ($F_{2:3}$ families)	*Echinochloa crus-galli*		X	above-ground (n = 1)	C	P	nr	Lee *et al.*, 2005
		Oryza sativa (123 Doubled Haploid lines)	*Lactuca sativa*		X	below-ground (n = 1)	C	P	nr	Zeng *et al.*, 2003
		Oryza sativa (150 RILs)	*Echinochloa crus-galli*		X	above-ground (n = 2) + below-ground (n = 2)	C	P	nr	Jensen *et al.*, 2008
		Triticum aestivum (271 Doubled Haploid lines)	*Lolium rigidum*		X	below-ground (n = 1)	C	P	nr	Wu *et al.*, 2003
		Triticum aestivum (277 $F_{2:3}$ families)	weeds + *Zea mays* + *Lolium rigidum*	X	X	above-ground (n = 1 + 4) + below-ground (n = 0 + 1)	F	P	nr	Zuo *et al.*, 2012
	parasitism heterospecific	*Helianthus annuus* (230 F_2 families + 204 F_2 families)	*Orobanche cumana*	X		above-ground (n = 1)	C	M	nr	Lu *et al.*, 2000
		Helianthus annuus (113 F_2 families)	*Orobanche cumana*		X	above-ground (n = 1) + below-ground (n = 1)	F	P	nr	Pérez-Vich *et al.*, 2002
		Helianthus annuus (262 RILs)	*Orobanche cumana*		X	above-ground (n = 1) + below-ground (n = 1)	C	M	nr	Tang *et al.*, 2003
		Helianthus annuus (101 RILs)	*Orobanche cumana*		X	above-ground (n = 1) + below-ground (n = 2)	F + C	P	nr	Louarn *et al.*, 2015
		Helianthus annuus (96 $F_{2:3}$ families)	*Orobanche cumana*		X	below-ground (n = 1)	C	M	nr	Imerovski *et al.*, 2016
		Oryza sativa (98 BILs)	*Striga hermonthica*		X	below-ground (n = 1)	C	P	nr	Gurney *et al.*, 2006
		Oryza sativa (141 BILs)	*Striga hermonthica*		X	below-ground (n = 1)	C	P	nr	Swarbrick *et al.*, 2009
		Oryza sativa (115 F_6 RILs)	*Striga hermonthica*	X	X	above-ground (n = 6 + 2)	C	P	nr	Kaewchumnong and Price, 2008
		Oryza sativa (115 F_6 RILs)	*Striga hermonthica*	X	X	above-ground (n = 2 + 0) + below-ground (n = 2 + 1)	C	P	**SLB1 & SBL2** (two cytochrome P450 genes)	Cardoso *et al.*, 2014
		Pisum sativum (115 RILs)	*Orobanche crenata*		X	above-ground (n = 1)	F	P	nr	Valderrama *et al.*, 2004
		Pisum sativum (111 $F_{6:7}$ RILs)	*Orobanche crenata*	X	X	above-ground (n = 0 + 1) + below-ground (n = 1 + 3)	F + C	P	nr	Fondevilla *et al.*, 2010
		Sorghum bicolor (226 RIPs + 226 RIPs)	*Striga hermonthica*		X	above-ground (n = 1)	F	P	nr	Haussmann *et al.*, 2004
		Sorghum bicolor (354 RILs)	*Striga asiatica* + *Striga hermonthica*		X	below-ground (n=1)	C	M	nr	Satish *et al.*, 2012
		Sorghum bicolor (selected RILs from 354 RILs)	*Striga asiatica* + *Striga hermonthica*		X	below-ground (n=1+1)	C	M	***LGS1*** (sulfotransferase)	Gobena *et al.*, 2017

Table 1. Continued

Approach[a]	Type of interaction	Focal species[b]	Neighbor species / Treatment	CR[c]	CE[c]	Plant organ[d]	Growth cond.[e]	Gen. Arch.[f]	Candidate genes[g]	References[h]
QTL mapping										
Traditional QTL mapping	asymmetric parasitism									
	heterospecific	*Vicia faba* (196 F_2 families)	*Orobanche crenata*		X	above-ground (n = 1)	F	P	nr	Roman *et al.* , 2002
		Vicia faba (144 RILs)	*Orobanche foetida*		X	above-ground (n = 1)	F	P	nr	Díaz-Ruiz *et al.* , 2009
		Vicia faba (156 F_6 RILs)	*Orobanche crenata*		X	above-ground (n = 1)	F	P	nr	Díaz-Ruiz *et al.* , 2010
		Vicia faba (119 F_7-F_8 RILs)	*Orobanche crenata*		X	above-ground (n = 1)	F	P	nr	Gutiérrez *et al.* , 2013
		Vicia faba (119 F_7-F_8 RILs)	*Orobanche foetida*		X	above-ground (n = 1)	F	P	nr	Gutiérrez *et al.* , 2013
		Vigna unguiculata (four 150 F_2 families + 153 F_3 lines)	*Striga gesnerioides*		X	above-ground (n = 1) + below-ground (n = 1)	F	M	nr	Ouédraogo *et al.* , 2001
		Vigna unguiculata (two 150 F_2 families)	*Striga gesnerioides*		X	above-ground (n = 1)	F	P	nr	Ouédraogo *et al.* , 2002
		Vigna unguiculata (150 F_9 RILs)	*Striga gesnerioides*		X	above-ground (n = 1)	C	M	***Rsg3-301*** (CC-NBS-LRR)	Li and Timko, 2009
		Vigna unguiculata (62 F_2 families + 35 F_2 families)	*Striga gesnerioides*		X	below-ground (n=1)	C	M	nr	Boukar *et al.* , 2004
GWA mapping	competition conspecific	*Arabidopsis thaliana* (48 local accessions)	absence and presence of intra-genotypic competition	X	X	above-ground (n = 9)	F	P	nr	Baron *et al.* , 2015
		Oryza sativa (301 accessions + 151 MAGIC lines)	intra-genotypic competition (two density conditions)	X		above-ground (n = 9)	F	P	*Os03g0275400*	Kikuchi *et al.* , 2017
	heterospecific	*Arabidopsis thaliana* (48 local accessions)	*Poa annua* , *Stellaria media* , *Trifolium repens* , *Veronica arvensis*	X	X	above-ground (n = 8 + 1)	F	P	*HAM2* , *AT5G26670* , *AT5G66530*	Baron *et al.* , 2015
		Arabidopsis thaliana (195 local accessions)	*Poa annua*	X		above-ground (n = 29)	IS	P	*FLC* , *TSF*	Frachon *et al.* , 2017

[a] QTL mapping: Quantitative Trait Loci, GWA: Genome-Wide Association. [b] RILs : Recombinant Inbred Lines, SILs : Segmental Introgressions Lines, BILs: Backcross Inbred Lines, RIPs: Recombinant Inbred Populations. [c] CR: competitive response, CE: competitive effect. ne: not estimated. [d] Numbers in brackets indicate the number of measured above-ground/below-ground traits. One number stands for the number of traits measured either for competitive response or for competitive effect, whereas two numbers stand for the number of traits measured for competitive response and competitive effect. The measured traits are listed in Supplementary Data Set 1. [e] Growth conditions. C: controlled conditions (greenhouse and growth chambers), F: field conditions, IS: *in situ*. [f] Genetic architecture. M: monogenic, P: polygenic. ne: not estimated. [g] UP and DOWN correspond to an up-regulation and down-regulation of genes associated with a specific functional category. nr: not reported. Candidate genes highlighted in red correspond to functionally validated genes.

Table 2. Evidences for modification of immunity by the plant neighborhood.

Plant combination[1]	Pathogen species	Symptom / damage fold reduction	Molecular immunity-related phenotype observed on focal plant in conspecific *vs* heterospecific [3]	References
Watermelon/Rice	*Fusarium oxysporum* f. sp. *niveum*	∞ [2]	several defense-related activities reduced (before infection)	Ren *et al.*, 2008
Soybean/Maize	*Cylindrocladium parasiticum*	1.7	*PR*, *PAL* and *PPO* gene induction enhanced (after infection)	Gao *et al.*, 2014
Watermelon/Wheat	*Fusarium oxysporum* f. sp. *niveum*	3.9	Induction of *AOS*, *PAL* and other genes enhanced (after infection but not before)	Xu *et al.*, 2015
Tomato/Onion	*Verticillium dalhiae*	1.4	many genes involved in biotic response enhanced (RNASeq after infection)	Fu *et al.*, 2015

[1] the first species is the focal plant on which measures (such as disease and immunity) were scored, whereas the second species corresponds to the identity of the neighbor species inducing changes. [2] 0% in interspecific versus 66% in conspecific. [3] AOS: Allene Oxide Synthase; PAL: Phenyl Ammonia Lyase; PDF: Plant Defensin; PPO: polyphenol oxidase; PR: Pathogenesis-related. n.t. not tested.

C) Objectifs de la thèse

Comme nous avons pu le voir précédemment, malgré l'importance des interactions plante-plante aussi bien dans les communautés végétales naturelles que dans les champs cultivés, les bases génétiques et moléculaires sous-jacentes à la variation naturelle de ce type d'interactions restent encore largement méconnues, en comparaison avec la résistance à des espèces pathogènes par exemple (Roux & Bergelson 2016). A l'ère de la génomique écologique, l'identification des bases génétiques sous-jacentes à la variation naturelle des interactions plante-plante doit reposer sur des approches innovantes et situées à l'interface entre écologie, génomique et biologie moléculaire.

Etant donné l'importance des interactions compétitrices dans le fonctionnement des communautés végétales et l'absence d'identification à ce jour de gène impliqué dans ce type d'interactions, l'objectif principal de ma thèse est d'identifier et valider fonctionnellement un/des gène/s impliqué/s dans la variation naturelle de la réponse à la présence d'une espèce compétitrice. Pour cela, différents points doivent être abordés, que l'on peut formuler à travers quatre questions principales :

i. Quelle population de cartographie utiliser pour l'identification des bases génétiques associées à la variation naturelles des interactions plante-plante?
ii. Quelle est le niveau de variation génétique des interactions plante-plante ?
iii. Quelle est l'architecture génétique sous-jacente à cette variation ?
iv. Quels sont les mécanismes moléculaires sous-jacents ?

Avant toute chose, il est nécessaire de déterminer l'espèce la plus adaptée pour répondre à ces questions, mais aussi pour laquelle nous sommes capables de mettre en place une approche interdisciplinaire durant la durée de ma thèse.

Arabidospsis thaliana : une espèce adaptée pour identifier les bases génétiques de la variation naturelle des interactions plante-plante ?

Une espèce porte-drapeau en génomique végétale

A. thaliana (L.) est une plante annuelle de la famille des Brassicaceae. Dans le domaine de la recherche en sciences végétales, elle est à l'heure actuelle toujours considérée comme l'espèce modèle en génomique fonctionnelle, biologie moléculaire et physiologie en raison de

sa facilité de culture, son cycle de vie court (en conditions contrôlées), ainsi que sa capacité à s'autoféconder, permettant ainsi de maintenir des lignées homozygotes et de les phénotyper un nombre infini de fois (Weigel & Nordborg 2005). Ces caractéristiques, combinées à la petite taille de son génome (5 chromosomes, ~119 Mb), ont conduit à son séquençage complet — le premier chez les plantes supérieures — achevé en 2000 (accession Col-0, The *Arabidopsis* Genome Initiative 2000). Actuellement, la base de données TAIR 10 (The *Arabidopsis* Information Resource ; http://www.arabidopsis.org) compte 33 602 gènes dont 27 416 codant pour des protéines. Par ailleurs, d'importantes ressources génétiques pour l'analyse des bases génétiques de la variation phénotypique et la dissection des mécanismes moléculaires sous-jacents sont publiquement disponibles (altération ou perturbation aléatoire des gènes : mutagénèse EMS, mutants T-DNA; altération ou perturbation spécifique de gènes : 'gene silencing' par microARN artificiel, amiRNA) (Alonso & Ecker 2006, O'Malley & Ecker 2010), permettant la validation fonctionnelle par complémentation quantitative ou 'silencing' quantitatif.

Une espèce modèle en écologie évolutive

Depuis quelques années, *A. thaliana* est aussi considérée comme une espèce modèle en écologie évolutive (Gaut 2012). Native d'Eurasie, elle a aujourd'hui une répartition mondiale et est rencontrée dans des habitats très contrastés aussi bien en termes d'environnement abiotique que biotique (Mitchell-Olds & Schmitt 2006, Shindo *et al.* 2007). Le taux d'autogamie de 98% décrit dans les années 1970 a longtemps laissé penser que les populations naturelles d'*A. thaliana* étaient majoritairement monomorphes. Or, il a été mis en évidence récemment que des populations peuvent être très polymorphes aussi bien d'un point de vue génétique que phénotypique (Le Corre 2005, Kronholm *et al.* 2012, Brachi *et al.* 2013). Les taux d'allogamie calculés au sein des populations naturelles d'*A. thaliana* sont en moyenne de l'ordre de 2% mais peuvent atteindre jusqu'à 20 % dans certaines populations (Bomblies *et al.* 2010, Platt *et al.* 2010). Il est donc possible de s'intéresser aux patrons d'évolution de traits phénotypiques et de leurs bases génétiques à l'échelle de l'aire de distribution de l'espèce, mais aussi à une échelle très locale.

Une espèce modèle pour étudier la génétique de la variation naturelle des interactions plante-plante ?

Alors qu'*A. thaliana* est généralement décrite comme une espèce pionnière souvent trouvée dans des milieux pauvres ou perturbés, rarement en compétition avec d'autres espèces, de récentes études et observations sur le terrain, menées au sein de l'équipe, semblent indiquer le contraire :

- Le nombre d'espèces cohabitant avec *A. thaliana* dans les communautés végétales peut être très important. En effet, au cours d'une campagne de prospection de populations naturelles d'*A. thaliana* dans la région Midi-Pyrénées (sud-ouest de la France) au printemps 2014, 168 populations ont été identifiées (Bartoli *et al.* 2018, Frachon *et al.* 2018). La caractérisation des communautés végétales associées à ces 168 populations a permis de montrer qu'*A. thaliana* pouvait cohabiter avec des communautés végétales (i) très différentes au niveau de leur composition, et (ii) pouvant contenir jusqu'à 28 espèces végétales (moyenne = 12.1, Figure i.2) (Frachon *et al.* 2019).

Figure i.2. Photos d'*A. thaliana* dans des milieux présentant de fortes interactions plante-plante dans la région Midi-Pyrénées. Les flèches rouges indiquent *A. thaliana* (d'après la thèse de Léa Frachon).

- À partir de 49 populations naturelles françaises d'*A. thaliana* provenant de quatre régions Françaises (Bourgogne, Bretagne, Languedoc et Nord) et caractérisées au niveau (i) phénotypique en conditions contrôlées (serre) et (ii) écologique (climat, sol et intensité des interactions plante-plante au niveau intra- et interspécifique), une étude corrélative menée au sein de l'équipe a suggéré que les interactions plante-plante au niveau

interspécifique pouvaient constituer une importante pression de sélection sur des traits phénologiques, telle que la date de floraison (Brachi *et al.* 2013). En utilisant une population de 160 lignées recombinantes consanguines (Recombinant Inbred Lines, RILs), une expérience d'évolution expérimentale réalisée sur 4 générations en conditions contrôlées (serre) a permis de confirmer que la présence d'une autre espèce végétale (i.e. le pâturin annuel, *Poa annua*) pouvait être un agent sélectif majeur de la phénologie chez *A. thaliana* (Brachi *et al.* 2012).

- À une échelle locale, un suivi sur plus de dix ans d'une population naturelle d'*A. thaliana* au sein d'une prairie permanente a permis de mettre en évidence une augmentation de la taille de la population (données non publiées, Fabrice Roux), démontrant qu'*A. thaliana* peut se maintenir sur de nombreuses générations dans un milieu où les interactions plante – plante sont prépondérantes.
- Toujours à une échelle locale, Etienne Baron (ancien doctorant au sein de l'équipe) a pu détecter lors d'une expérience réalisée sur un terrain expérimental, une variation génétique importante de la réponse à la compétition chez *A. thaliana* aussi bien dans un contexte d'interaction intraspécifique que dans un contexte d'interaction interspécifique (Baron *et al.* 2015).

Ainsi, (i) les connaissances sur le développement, la génétique et la physiologie d'*A. thaliana*, (ii) la diversité des habitats rencontrés par *A. thaliana*, (iii) la disponibilité de ressources génétiques artificielles et naturelles, et (iv) le développement des technologies NGS couplé au développement de méthodes d'analyses statistiques puissantes, font d'*A. thaliana* une espèce de choix pour aborder des questions en écologie et en biologie évolutive (Koornneef *et al.* 2004, Mitchell-Olds & Schmitt 2006), et notamment l'identification des bases génétiques sous-jacentes à la variation naturelle des interactions plante-plante.

Comment identifier chez A. thaliana les gènes sous-jacents à la réponse à la présence de plantes voisines ?

Comme nous avons pu le mentionner dans la revue, deux types de méthodes non-exclusives peuvent être utilisés chez *A. thaliana* pour identifier les gènes sous-jacents à la réponse à la présence de plantes voisines. Le premier type de méthodes repose sur l'utilisation d'un seul fond génétique :

i. soit par analyse de mutants pour lesquels le phénotype d'une lignée sauvage et de lignées mutantes issues de cette lignée sauvage est mesuré dans un environnement où un stress biotique ou abiotique est appliqué. Parmi ces mutants, il existe des lignées mutantes obtenues par mutagénèse insertionnelle qui permet des approches sans *a priori* mais également des approches *a priori*, où des lignées mutantes pour des fonctions moléculaires connues sont testées pour leur phénotype dans des environnements choisis. L'utilisation de tels mutants a été utilisée pour identifier des mécanismes supposés de détection et/ou de réponse à des plantes voisines (e.g. qualité de la lumière, nutriments, exsudats racinaires et composés organiques volatiles, Pierik *et al.* 2013). Cependant, cette méthode repose sur des présupposés forts qui ne présentent pas forcément de point commun avec les bases génétiques associées à la variation des interactions plante-plante observables dans les populations naturelles. De façon complémentaire, il existe des mutants de type EMS (ethyl mehtyl sulfonate), composé chimique qui permet de générer des mutations ponctuelles, créant ainsi de la diversité 'artificielle' dans les descendants qui seront ensuite phénotypés sans *a priori*. Bien que largement adoptée pour identifier des mécanismes génétiques et moléculaires impliqués dans les interactions avec des espèces pathogènes, cette approche n'a pas été utilisée à notre connaissance dans le cadre des interactions plante-plante, certainement en raison de la complexité des expériences à mettre en place, où chacune des dizaines de milliers de lignées mutantes doit être mise en présence d'une plante voisine.
ii. soit par une approche de transcriptomique où l'expression de la majorité des gènes présents dans le génome est mesurée dans une lignée génétique dans un environnement contrôle et dans chacun des environnements d'interaction testés. Bien que très intéressante et utilisée de manière assez régulière dans le cadre des

interactions plante-plante, cette méthode peut s'avérer coûteuse et implique donc souvent de limiter les analyses à quelques environnements, organes et points de cinétique. Cependant, la baisse constante des coûts liés à l'utilisation de technologies NGS permet d'envisager dans un futur proche d'étudier la dynamique de réponse du transcriptome dans différents organes d'une plante soumise aux stress multiples imposés par les plantes avoisinantes. Cette méthode peut aussi s'appliquer entre une lignée sauvage et un mutant afin de détecter l'ensemble des gènes présentant une régulation différentielle : cette approche est plus ciblée et plus puissante. Malgré tout, les approches de transcriptomique mettent en avant un nombre important de gènes dérégulés qui ne sont pas forcément directement impliqués dans le processus étudié.

Bien que très pertinentes, ces méthodes ne nous renseignent malheureusement en rien (i) sur l'architecture génétique (le nombre, l'effet et la position des régions associées aux traits phénotypiques, i.e. QTL), et (ii) sur les variants génétiques naturels associés à la variation phénotypique observée dans les populations naturelles, et donc potentiellement retenus par la sélection naturelle.

Le deuxième type de méthodes vise donc à identifier les QTL associés à la variation naturelle de la réponse phénotypique aux interactions plante-plante. Deux approches peuvent être mentionnées:

i. L'approche par cartographie QTL traditionnelle basée sur des populations artificielles issues de croisements entre accessions. Plusieurs types de populations de QTL mapping traditionnel peuvent être utilisés pour cartographier les marqueurs génétiques associés à la variation phénotypique naturelle (Figure i.3). Bien que largement utilisée chez *A. thaliana*, ces méthodes de cartographie traditionnelle restent peu résolutives avec des régions QTL recouvrant encore des centaines de gènes (Bergelson & Roux 2010). Ce nombre important de gènes rend laborieux l'identification des gènes sous-jacents et donc leur validation fonctionnelle.

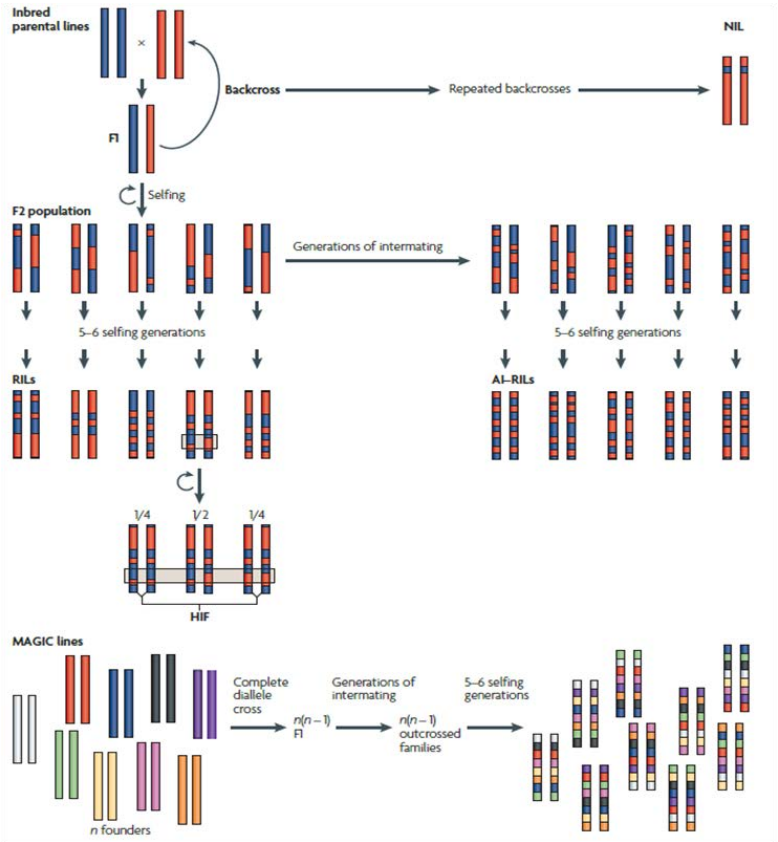

Figure i.3. Illustration des différentes populations de QTL mapping traditionnel pouvant être utilisées pour cartographier les marqueurs génétiques associés à la variation naturelle phénotypique. **RILs**: Recombinant inbred line, **AI-RILS**: Advanced intercross-recombinant inbred lines, **HIF**: Heterogeneous inbred family, **MAGIC**: multiparent advanced generation inter-cross lines, **NIL**: near-isogenic line. D'après Bergelson & Roux 2010.

ii. Une alternative pour cartographier finement les régions QTL correspond à la cartographie par association pangénomique (Genome-Wide Association mapping, GWA, Mitchell-Olds & Schmitt 2006), méthode développée chez *A. thaliana* depuis plus d'une décennie (Aranzana *et al.* 2005, Zhao *et al.* 2007, Atwell *et al.* 2010). Cette méthode repose sur le déséquilibre de liaison (Linkage Disequilibrium, LD) présent dans les populations naturelles d'*A. thaliana* pour identifier les polymorphismes génétiques associés à la variation phénotypique naturelle. Il est donc nécessaire que la population utilisée présente un niveau suffisant de diversité génétique ainsi qu'un LD moyen (r^2~0.5) le plus court possible, ceci afin d'être le plus précis dans la cartographie des régions génomiques associées. Dans le cas *d'A. thaliana*, il a été estimé que le LD moyen était en moyenne de 10kb, bien que cette estimation puisse être très variable le long du génome (entre 50bp et 200kb) (Kim *et al.* 2007).

Bien que très puissante, la méthode de GWA mapping présente deux inconvénients majeurs, à savoir la présence de faux positifs et l'hétérogénéité génétique et/ou allélique (Figure i.4).

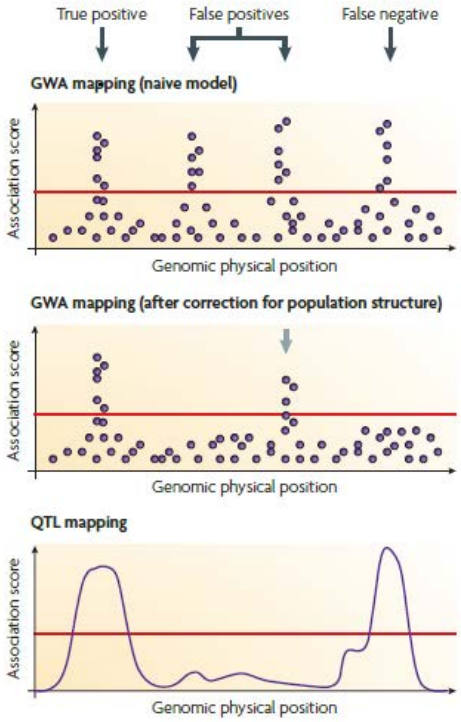

Figure i.4. Illustration des faux positifs et faux négatifs dans les études de GWA mapping. D'après Bergelson & Roux 2010.

Les faux positifs correspondent à de fausses associations génotype-phénotype qui résultent de l'effet de l'histoire démographique de l'espèce (Box 2). Ce phénomène chez *A. thaliana* est notamment observé à l'échelle européenne, où la colonisation de l'Europe Centrale et du Nord s'est en partie faite à partir de différents refuges situés sur le pourtour méditerranéen lors de la dernière période glaciaire et entre lesquels une différenciation à l'échelle du génome s'est mise en place durant cette période (Nordborg *et al.* 2005).

Box2. Illustration de l'obtention de faux positifs dans les études de GWA mapping.

Prenons l'exemple théorique de deux populations i et j. De par leur isolement géographique (par exemple, dans des refuges glaciaires), les populations i et j ont divergé au niveau génomique, ce qui est illustré par des génomes de couleur différente sur la Figure i.5. Une mutation récente conduisant à la résistance à un agent pathogène, représentée par un triangle rouge, apparait dans la population j. Dans ce cas, cette mutation est en déséquilibre de liaison (DL) complet avec l'ensemble des polymorphismes le long du génome qui différencient la population j de la population i. Dans une analyse de GWA mapping, le phénotype de résistance se révèle donc non seulement associé à la mutation causale mais aussi à tous les autres polymorphismes en DL, ces derniers correspondant à des faux positifs (Figure i.5A).

Heureusement, une situation si extrême est rarement rencontrée dans la nature. En effet, des flux géniques sont observés entre les populations. Plus le taux de migration est important, plus le taux de recombinaison efficace augmente au sein de chacune des populations, plus le DL est court, plus le taux de faux positifs est réduit (Figure i.5B et i.5C).

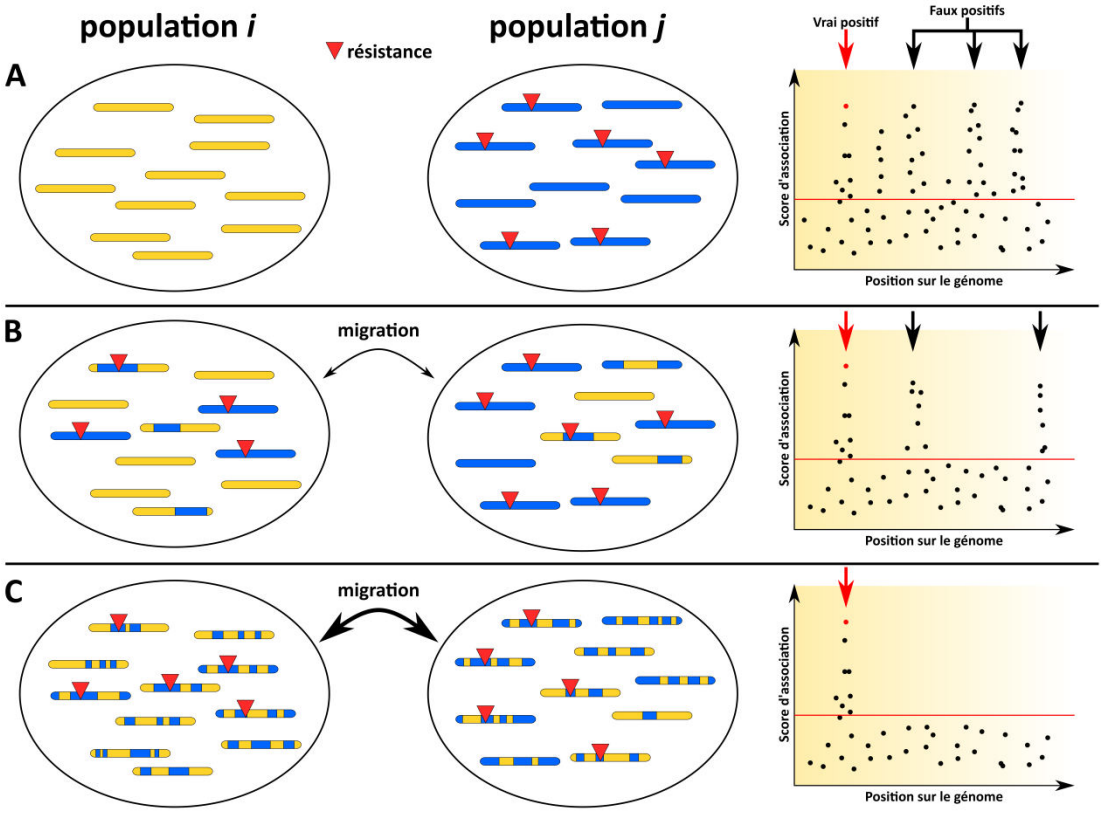

Figure i.5. Illustration de l'effet de l'histoire démographique d'une espèce sur le taux de fausses associations entre génotype et phénotype. A | Absence de migration entre populations provoquant de fausses associations entre le phénotype de résistance et l'ensemble des polymorphismes en DL significatif avec la mutation causale. B | Cas d'un flux génique faible entre populations. C | Cas d'un flux génique fort entre populations. Triangle rouge : allèle procurant la résistance à une espèce pathogène.

Plusieurs méthodes statistiques ont été développées pour tenter de corriger ces faux positifs, comme par exemple l'intégration d'une matrice d'apparentement génomique entre accessions dans les modèles statistiques (Kang *et al.* 2010). Bien que très performantes, ces méthodes statistiques entraînent aussi l'apparition de faux négatifs (Figure i.4), c'est-à-dire des marqueurs génétiques réellement associés à la variation phénotypique naturelle (i.e. marqueurs causaux) mais qui sont perdus après correction pour l'effet de l'histoire démographique de l'espèce (Bergelson & Roux 2010). Une solution pour identifier les faux négatifs est de combiner l'approche de GWA mapping à des approches de QTL mapping traditionnel (Figure i.4) (Brachi *et al.* 2010).

L'hétérogénéité génétique et allélique provient du fait qu'une même valeur phénotypique observée au sein d'une population de cartographie peut être le résultat de différents QTL ou de différents allèles à un même QTL, respectivement (Figure i.6). Ces hétérogénéités diminuent la puissance statistique de détection des QTL car elles augmentent le niveau d'asymétrie entre les variances intra-alléliques à un marqueur génétique donné (Figure i.6). Pour limiter les effets de de l'hétérogénéité génétique/allélique, il a été proposé de travailler à une échelle géographique restreinte (échelle régionale par exemple) afin de limiter le nombre d'allèles rares tout en conservant une diversité génétique relativement importante en comparaison de la diversité génétique observée à une échelle mondiale (Bergelson & Roux 2010).

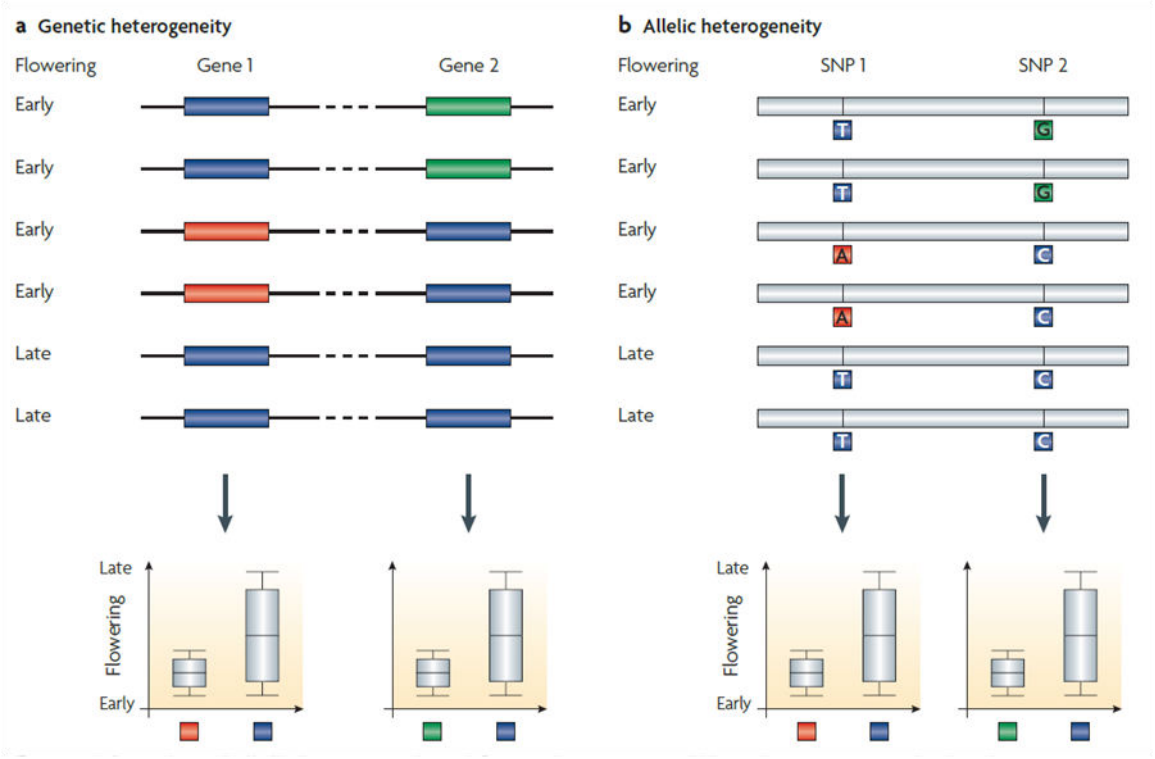

Figure i.6. Illustration de l'effet de l'hétérogénéité génétique et de l'hétérogénéité allélique sur la détection de QTL dans les études de GWA mapping. Cas de la date de floraison. D'après Bergelson & Roux (2010).

Depuis plusieurs années, l'équipe a développé des compétences et une expertise en GWA mapping. Outre le développement de nouvelles approches statistiques en GWA mapping (Brachi *et al.* 2010, Brachi *et al.* 2013, Huard-Chauveau *et al.* 2013, Wang *et al.* 2018), elle a mis en avant la nécessité de choisir la population de cartographie en fonction de l'échelle géographique à partir de laquelle la variation génétique du trait phénotypique étudié ne change plus ; ce qui dépend donc directement de l'échelle géographique à laquelle s'exerce les pressions de sélection correspondantes (Bergelson & Roux 2010). Par exemple, pour identifier chez *A. thaliana* les bases génétiques associées à la tolérance au froid, il est préférable d'utiliser une population de GWA mapping basée sur des accessions mondiales, ceci afin de recouvrir le large gradient climatique rencontré par *A. thaliana* sur son aire de distribution (Horton *et al.* 2012). À l'opposé, pour des interactions hôte-pathogène qui correspondent plus à des processus régionaux ou locaux, il serait préférable d'utiliser une population régionale ou locale de GWA mapping. Par exemple, le même niveau de diversité génétique pour la résistance quantitative à la bactérie phytopathogène *Xanthomonas campestris* a été observé à différentes échelles géographiques (Huard-Chauveau *et al.* 2013, Debieu *et al.* 2016). En accord avec ces observations, le principal gène de résistance

quantitative à *X. campestris* (i.e. *RKS1*) a été détecté aussi bien au niveau local qu'au niveau mondial (Huard-Chauveau *et al.* 2013).

D) Plan de la thèse

En adoptant une approche interdisciplinaire à l'interface entre génomique, génétique d'association et biologie moléculaire, je me suis intéressé lors de ma thèse (i) à caractériser aux niveaux phénotypique et génomique une population locale d'*A. thaliana*, (ii) permettant ainsi d'adopter une approche de GWA mapping pour décrire l'architecture génétique associée aux interactions plante-plante dans différents contextes de compétition, (iii) ceci dans le but de cloner le premier QTL (à ma connaissance) associé à la variation génétique naturelle de la réponse compétitive d'*A. thaliana* à la présence d'une autre espèce végétale. Ces objectifs sont articulés selon 3 chapitres (Figure i.7).

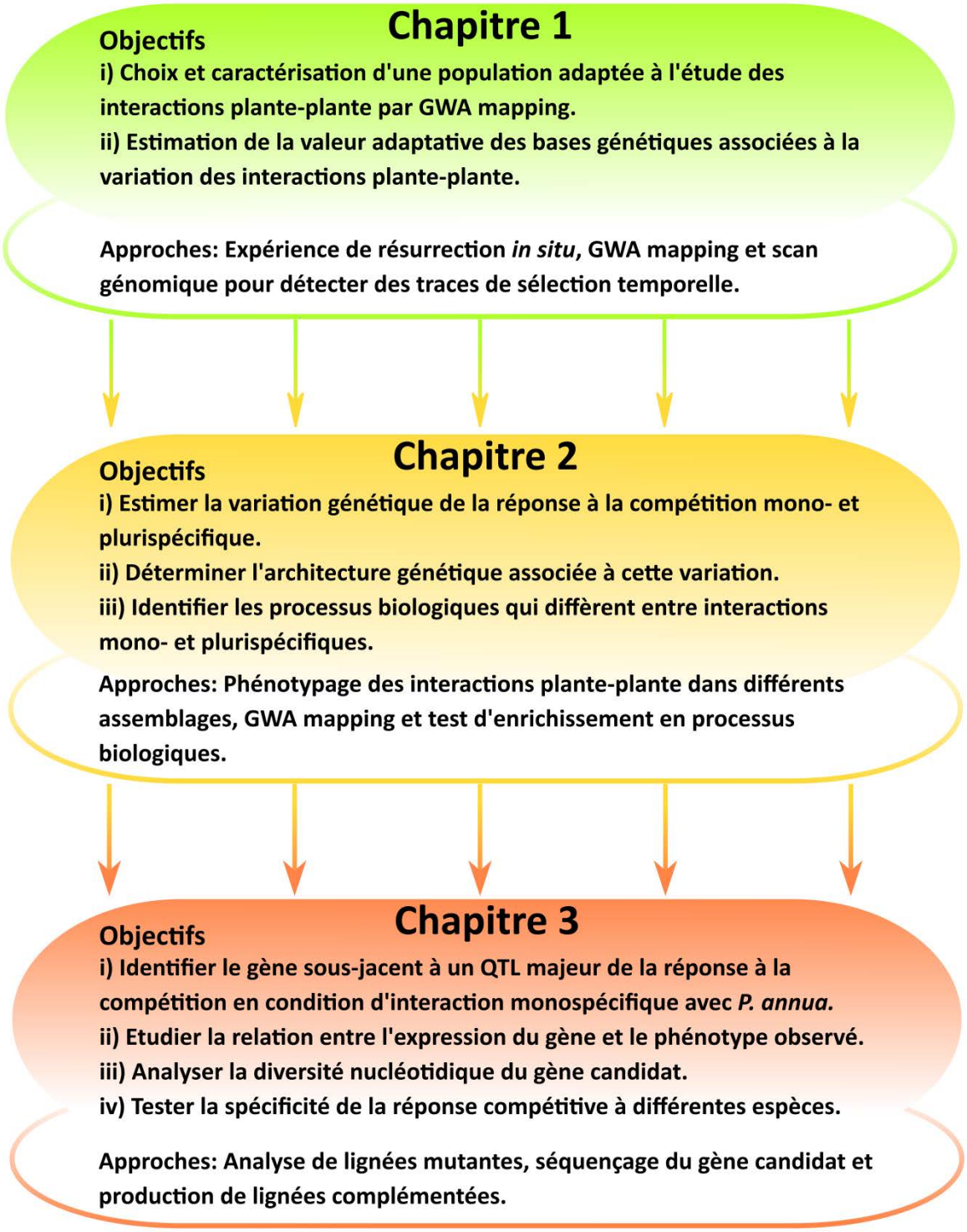

Figure i.7. Schéma représentant les principaux objectifs et approches retenus dans le cadre des 3 chapitres de ma thèse.

Les espèces végétales sont susceptibles de développer des mécanismes de réponse à la compétition qui sont fortement dépendants des autres espèces végétales présentes localement au sein de leur communauté. Il semble donc nécessaire d'utiliser une population locale de GWA mapping (i) qui évolue dans un habitat où les interactions plante-plante sont prédominantes, (ii) qui présente une variation naturelle adaptative de la réponse à la compétition interspécifique, (iii) une diversité génomique importante et (iv) un déséquilibre de liaison le plus court possible. En d'autres termes, la population parfaite! Une population locale d'*A. thaliana* ayant été collectée au sein de l'équipe et ayant fait l'objet de plusieurs caractérisations aux niveaux génétique et phénotypique, semble réunir un nombre important des critères cités précédemment. Dans le premier chapitre, je me suis donc intéressé à caractériser cette population française locale TOU-A d'*A. thaliana* au niveau génomique mais aussi au niveau phénotypique, afin de déterminer si elle pouvait être un outil puissant pour l'identification des bases génétiques adaptatives sous-jacentes aux interactions plante-plante. Plus précisément, j'ai cherché dans ce premier chapitre à répondre aux trois questions suivantes:

i. Observe-t-on une variation importante entre les 195 accessions locales en réponse à la présence d'une autre espèce végétale?
ii. Cette population présente-elle une diversité génomique suffisante et un LD court compatibles avec des analyses d'association pangénomique?
iii. Les bases génétiques de cette variation sont-elles adaptatives ?

Dans la nature, les plantes interagissent de manière simultanée ou séquentielle avec plusieurs espèces végétales. Cependant, les études qui ont cherché à identifier les bases génétiques sous-jacentes aux interactions interspécifiques se sont focalisées la plupart du temps sur des interactions entre seulement deux espèces végétales. Durant le second chapitre, j'ai donc cherché, par une approche expérimentale, à caractériser au sein de la population TOU-A l'architecture génétique de la réponse compétitive d'*A. thaliana* en conditions d'interactions monospécifique et plurispécifique. À travers ce second chapitre, je me suis intéressé à répondre à plusieurs questions :

i. Quelle est l'étendue de la variation génétique au sein de cette population locale d'*A. thaliana* dans différents contextes d'interactions plante-plante mono- et pluri-spécifiques?
ii. Peut-on prédire l'architecture génétique en conditions d'interaction plurispécifique à partir des architectures génétiques observées en conditions d'interaction monospécifique (hypothèse d'additivité)? Ou observe-t-on l'émergence de nouveaux QTLs?
iii. Les principaux processus biologiques sous-jacents à la réponse à la compétition diffèrent-ils entre interactions monospécifiques et interactions plurispécifiques?

Comme déjà mentionné, aucun gène impliqué dans la réponse à la présence d'une espèce compétitrice n'a été cloné et validé fonctionnellement à l'heure actuelle. Dans le troisième chapitre, je me suis donc attaché à identifier le gène causal sous-jacent à un QTL majeur identifié dans le second chapitre en conditions d'interaction monospécifique avec *P. annua*. En combinant (i) analyse de mutants insertionnels, (ii) étude de la diversité nucléotidique du gène candidat au sein de la population locale TOU-A et (iii) production de lignées complémentées par différents haplotypes naturels, j'ai cherché à répondre à trois questions :

i. Quel est le gène causal sous-jacent au QTL sélectionné ?
ii. Quelle est la fonction codée par ce gène ?
iii. Ce gène confère-t-il une réponse compétitive à d'autres espèces que *P. annua* ?

Chapitre 1

Identification d'une population naturelle d'*A. thaliana* adaptée à l'analyse de la variation génétique des interactions plante-plante

A) Introduction

Sachant que les plantes interagissent entre elles sur de courtes distances, les mécanismes mis en jeux lors de ces interactions doivent être dépendants de la composition de la communauté végétale au sein de laquelle les plantes interagissent. En d'autres termes, les gènes de perception et de réponse impliqués dans la variation naturelle des interactions plante-plante pourraient être très différents entre populations naturelles d'*A. thaliana*. Afin d'identifier ces gènes, il semblait donc préférable dans le cadre de ma thèse de se focaliser sur une population locale d'*A. thaliana* plutôt que sur un jeu d'accessions mondiales. En effet, afin d'obtenir le maximum de puissance pour cartographier finement des régions génomiques associées à de la variation phénotypique naturelle, il est recommandé d'ajuster l'échelle géographique d'une population de GWA mapping au grain environnemental de la pression de sélection étudiée (Bergelson & Roux 2010).

Dans un premier temps, il m'a donc fallu identifier et caractériser une population locale d'*A. thaliana* adaptée à l'utilisation d'une approche de GWA mapping dans le cadre des interactions plante-plante. Pour autant, il est nécessaire que cette population locale présente une variation phénotypique non-négligeable en réponse aux interactions plante-plante et si telle est le cas, déterminer si cette variation est adaptative. Dans le but de cartographier de façon fine les régions génomiques sous-jacentes à ces mécanismes par une approche de GWA mapping, il est aussi indispensable que cette population présente une diversité génomique suffisante qui s'accompagne d'un déséquilibre de liaison (LD) le plus court possible. Autant chercher une aiguille dans une botte de foin ! En effet, dans le cas d'une population locale, et en particulier pour une espèce principalement autogame comme *A. thaliana* (Platt *et al.* 2010), il est attendu que la diversité génétique intra-populationnelle soit faible et s'accompagne d'un LD relativement long en raison d'une forte dérive génétique dû à un effectif efficace faible.

Pourtant, au sein de l'équipe, une population locale d'*A. thaliana* a été collectée et a fait l'objet au cours de ces dernières années de plusieurs caractérisations génétiques et phénotypiques qui semblent indiquer que cette population pourrait respecter les pré-requis mentionnés précédemment. Cette population, nommée TOU-A, a été observée pour la première fois en avril 2001 et est située dans le village de Toulon-sur-Arroux (Saône-et-Loire, Bourgogne, France) sous une clôture électrique de 350m délimitant deux prairies permanentes, Figure 1.1). Elle est donc située dans un habitat présentant un niveau de compétition

interspécifique important. Par ailleurs, cette population naturelle a expérimenté une augmentation de la température annuelle de plus de 1°C durant les 30 dernières années. Les graines de 80 plantes (population TOU-A1), 115 plantes (population TOU-A5) ainsi que 115 plantes (population TOU-A6) situées le long de la clôture électrique ont été collectées en 2002, 2007 et 2010, respectivement.

Figure 1.1. Habitat naturel de la population locale TOU-A située sous une clôture électrique de 350m.

Au niveau de la diversité phénotypique, différentes études menées au sein de l'équipe ont pu montrer une forte variation génétique entre 48 accessions collectées en 2002 (population TOU-A1) pour une large gamme de traits phénotypiques, tels que la date de floraison (Brachi *et al.* 2013), la résistance quantitative à la bactérie phytopathogène *Xanthomonas campestris* (Huard-Chauveau *et al.* 2013, Debieu *et al.* 2016), la résistance quantitative à une souche du *Turnip Mosaic Virus* (Rubio *et al.* 2018), mais aussi pour la réponse à la compétition interspécifique (Baron *et al.* 2015). Dans cette dernière expérience réalisée sur un terrain expérimental, Etienne Baron (ancien doctorant au sein de l'équipe) a fait pousser 48 accessions de la population TOU-A1 selon six traitements de compétition, i.e. absence de compétition, compétition intra-spécifique et compétition interspecifique avec quatre espèces végétales communément associées à *A. thaliana* dans les communautés végétales en France, à savoir le pâturin annuel *Poa annua*, le mouron des oiseaux *Stellaria media*, le trèfle des champs *Trifolium arvense* et la véronique des champs *Veronica arvensis*. Ils ont notamment mis en évidence des interactions 'traitement de compétition x accessions' très significatives pour les neufs traits phénotypiques aériens mesurés, incluant la production totale de graines. La

population TOU-A semble donc adaptée pour réaliser des analyses phénotypiques en réponse aux interactions plante-plante.

Au niveau de la diversité génétique, les 80 accessions provenant de la population TOU-A1 (parmi un panel de 5707 accessions mondiales) ont été génotypés pour un jeu de 149 SNPs (Platt *et al.* 2010). Ces 80 accessions ont été associées à 57 groupes haplotypiques distincts, suggérant une diversité génétique importante au sein de cette population (Anastasio *et al.* 2011). Dans un second temps, 48 accessions provenant toujours de la population TOU-A1 (parmi un panel de 1307 accessions mondiales) ont été génotypées pour 250k SNPs (Horton *et al.* 2012). Encore une fois, une forte diversité génétique a été observée parmi ces 48 accessions. Malgré tout, le faible nombre d'accessions génotypées pour 250k SNPs demeurait un frein pour réaliser des analyses de GWA mapping avec suffisamment de puissance statistique (Baron *et al.* 2015).

Fort de ces résultats obtenus tant au niveau de la diversité phénotypique que de la diversité génétique, deux étapes supplémentaires et complémentaires ont été mises en place :

- dans le but d'identifier le caractère adaptatif de la réponse de la population TOU-A non seulement à la compétition interspécifique mais aussi au réchauffement climatique, l'évolution phénotypique de cette population a été étudiée entre 2002 et 2010. Pour cela, une expérience de résurrection a été mise en place *in situ* (dans l'habitat natif de la population TOU-A) par Fabrice Roux avant le début de ma thèse. Plus précisément, cette expérience a consisté à phénotyper 29 traits sur 195 accessions (populations TOU-A1 et TOU-A6) placées dans six micro-habitats différents. Ces six micro-habitats correspondent à la combinaison de trois types de sol natifs avec la présence/absence de *P. annua*, une espèce fréquemment observée dans la communauté végétale de la population TOU-A (Fabrice Roux, communication personnelle).
- afin de cartographier plus finement les bases génétiques sous-jacentes à la variation naturelle des interactions plante-plante, le séquençage génomique des 195 accessions provenant des populations TOU-A1 et TOU-A6 a été effectué avec la technologie Illumina®. Ces données génomiques nous ont aussi permis d'effectuer un scan génomique afin d'identifier des traces de sélection temporelle sur moins de 8 générations.

Dans cette première partie de mon travail de thèse, nous avons cherché à répondre à plusieurs questions :

i. Observe-t-on une évolution phénotypique en moins de 8 générations? Si oui, peut-on établir un lien entre les stratégies phénotypiques sélectionnées et des pressions de sélection potentielles telles que la compétition interspécifique ou bien encore le réchauffement climatique ?
ii. Quelle est l'architecture génétique sous-jacente aux variations phénotypiques? Cette architecture dépend-elle d'une hétérogénéité abiotique (i.e. sol) et biotique (i.e. présence d'un compétiteur) sur une très courte échelle spatiale ?
iii. Observe-t-on un lien entre évolution phénotypique et traces de sélection génomique?

Dans le cadre de ma thèse, ce chapitre m'a permis de (i) caractériser cette population au niveau génomique et (ii) d'estimer son potentiel pour identifier les bases génétiques adaptatives dans le cadre des interactions plante-plante.

NB : dans ce chapitre, mon travail a consisté (i) à effectuer les analyses bio-informatiques des données de séquençage Illumina (en collaboration avec Sébastien Carrère, IR au sein de la plate-forme de bioinformatique du LIPM), (ii) à effectuer les analyses GWA mapping en collaboration avec Léa Frachon (ancienne doctorante au sein de l'équipe), (iii) à caractériser l'architecture génétique sous-jacente aux 29 trait phénotypiques mesurés dans chacun des 6 micro-habitats et (iv) caractériser le niveau de pléiotropie des QTLs identifiés ainsi que leur sélection sur une période de 8 générations. Les 29 traits phénotypiques ont été mesurés soit par Fabrice Roux au cours de l'expérience (n = 5), soit par Léa Frachon après récolte des plantes à la fin de leur cycle de vie (n = 24). Le scan génomique de différenciation génétique au niveau temporel a été effectué par Miguel Navascuès et Renaud Vitalis (laboratoire CBGP, INRA Montpellier) ainsi que par Laurène Gay (laboratoire AGAP, INRA Montpellier).

B) Manuscrit: Intermediate degrees of synergistic pleiotropy drive adaptive evolution in ecological time.

Léa Frachon,[1¶] Cyril Libourel,[1¶] Romain Villoutreix,[2] Sébastien Carrère,[1] Cédric Glorieux,[2] Carine Huard-Chauveau,[1] Miguel Navascués,[3,4] Laurène Gay,[5] Renaud Vitalis,[3,4] Etienne Baron,[2] Laurent Amsellem,[2] Olivier Bouchez,[6,7] Marie Vidal,[6,8] Valérie Le Corre,[9] Dominique Roby,[1] Joy Bergelson,[10] Fabrice Roux[1,2]*

Affiliations :

[1] LIPM, Université de Toulouse, INRA, CNRS, Castanet-Tolosan, France

[2] Laboratoire Evolution, Ecologie et Paléontologie, UMR CNRS 8198, Université de Lille, Villeneuve d'Ascq Cedex, France

[3] INRA, UMR CBGP, F-34988 Montferrier-sur-Lez, France

[4] Institut de Biologie Computationnelle, F- 34095 Montpellier, France

[5] UMR AGAP, INRA, Montpellier, France

[6] INRA, GeT-PlaGe, Genotoul, Castanet-Tolosan, France

[7] GenPhySE, Université de Toulouse, INRA, INPT, INP-ENVT, Castanet Tolosan, France

[8] INRA, UAR1209, Castanet-Tolosan, France

[9] INRA, UMR1347 Agroécologie, France,

[10] Department of Ecology and Evolution, University of Chicago, Chicago, IL USA

[¶] These authors contributed equally to this work.

* To whom correspondence should be addressed. E-mail: fabrice.roux@toulouse.inra.fr

Manuscrit

“Intermediate degrees of synergistic pleiotropy drive adaptive evolution in ecological time”

Léa Frachon, Cyril Libourel, Romain Villoutreix, Sébastien Carrère, Cédric Glorieux, Carine Huard-Chauveau, Miguel Navascués, Laurène Gay, Renaud Vitalis, Etienne Baron, Laurent Amsellem, Olivier Bouchez, Marie Vidal, Valérie Le Corre, Dominique Roby, Joy Bergelson, Fabrice Roux

Rapid phenotypic evolution of quantitative traits can occur in natural populations on a timescale of decades or even years[1], but little is known about its underlying genetic architecture[2]. Theoretical investigations have revealed that genes with intermediate pleiotropy will, under certain conditions, drive adaptive evolution[3-4] but these predictions have rarely been tested, especially under ecologically realistic conditions. Here, we performed a resurrection experiment to compare the evolution of multiple traits across six *in situ* micro-habitats within a natural population of the plant *Arabidopsis thaliana*. We then used Genome Wide Association mapping to identify the SNPs associated with evolved and unevolved traits in each of these sites. Finally, a genome-wide analysis of temporal genetic differentiation allowed us to test for selection acting on these SNPs. Phenotypic evolution was consistent across all micro-habitats but GWAS revealed largely distinct genetic bases among sites. Adaptive evolutionary change was largely driven by rare QTLs with intermediate degrees of pleiotropy under strong selection; this pleiotropy was synergistic with the per-trait effect size of a SNP increasing with the degree of pleiotropy. In addition to these rare pleiotropic QTLs, weak selection was detected for frequent small micro-habitat-specific QTLs that shape single traits. In this French population, *A. thaliana* likely responded to both local warming and increased competition, in part mediated by central regulators of flowering time and circadian rhythm such as FLOWERING LOCUS C and TWIN SISTER OF FT. This genetic architecture, which includes both synergistic pleiotropic QTLs and distinct QTLs within particular micro-habitats, enables rapid phenotypic evolution while still maintaining genetic variation in wild populations.

Contemporary and rapid phenotypic evolution has been observed in many natural populations of plant and animal species[1,5], especially during invasion[6] and in response to both global climate change[7] and toxic pollution[8]. Although a handful of studies have identified the genetic architecture of contemporary adaptive evolution of qualitative traits (such as industrial melanism)[9] or single quantitative traits (such as herbicide detoxification in weeds or heavy-metal tolerance)[10,11], the genetic architecture of a suite of quantitative traits experiencing contemporary adaptive evolution remains largely unexplored.

Theoretical studies predict that the number and effect sizes of QTLs underlying multi-trait adaptive evolution depends, in part, on the magnitude of pleiotropy[3,4,12]. Based on Fisher's geometric model, in which every mutation potentially affects all traits, the rate of adaptation of a QTL should decrease with its degree of pleiotropy[4]. This results from the increased probability of antagonistic effects of a mutation when more traits are impacted. However; in contrast to the assumptions of the geometric model, laboratory studies have found an L-shaped distribution of the degree of pleiotropy such that most mutations affect only a small subset of traits[3,12]; this restricted pleiotropy should diminish the 'cost of complexity'. Of additional importance is the relationship between the degree of pleiotropy and the per-trait effect size of a mutation (termed pleiotropic scaling)[3,12]. Most theoretical models assume that the per-trait effect size of a mutation decreases (invariant total effect model) or remains constant (Euclidean superposition model) with the degree of pleiotropy[4]. However, laboratory studies have found synergistic pleiotropy in which the per-trait effect size of a mutation increases with the number of traits affected by that mutation[3]. The combination of restricted and synergistic pleiotropy leads to the prediction that polymorphisms with intermediate degrees of pleiotropy, while rare, should have the highest rate of adaptive evolution[3,4]. This prediction is yet to be tested empirically.

In its more general sense, pleiotropy refers to the shared impact of SNPs. This can include the effect of a SNP on (i) alternative phenotypic traits in one environment, (ii) a single phenotypic trait among environments, or (iii) alternative traits in multiple environments. Because wild populations evolve in complex abiotic and biotic environments, an exploration of the role of pleiotropy therefore requires consideration of the role of spatial environmental heterogeneity. In particular, when the same SNPs are favored in distinct micro-habitats, then the suite of selective effects may combine to drive rapid adaptive evolution whereas competing demands on a SNP across micro-sites could inhibit adaptive evolution.

In this study, we aimed to describe the genetic architecture underlying rapid phenotypic evolution of multiple quantitative traits of the annual plant *A. thaliana in situ*. More specifically, we aimed to test whether intermediate degrees of synergistic pleiotropy drive contemporary evolution of *A. thaliana* within a local population evolving in a spatially abiotic and biotic heterogeneous environment.

RESULTS AND DISCUSSION

Our study focused on the local population TOU-A (East of France; **Supplementary Fig. 1**) that experienced an increase in mean annual temperature of more than 1°C over the last 30 years (**Supplementary Fig. 2**). The site occupancy by *A. thaliana* additionally increased between 2002 and 2007 and remained stable thereafter (**Supplementary Fig. 1**). Seeds of 80 and 115 individual plants (hereafter named accessions) were collected in 2002 and 2010, respectively. Previous studies conducted on accessions collected in 2002 showed that this population has an estimated outcrossing rate of 6%[13] and is highly diverse at both genetic (based on genotyping at 149 SNPs) and phenotypic levels[13-16]. In addition, the TOU-A population

presents fine-scale spatial variation for a broad range of soil characteristics and is located between two permanent meadows dominated by grasses (**Supplementary Figs. 1 and 3**).

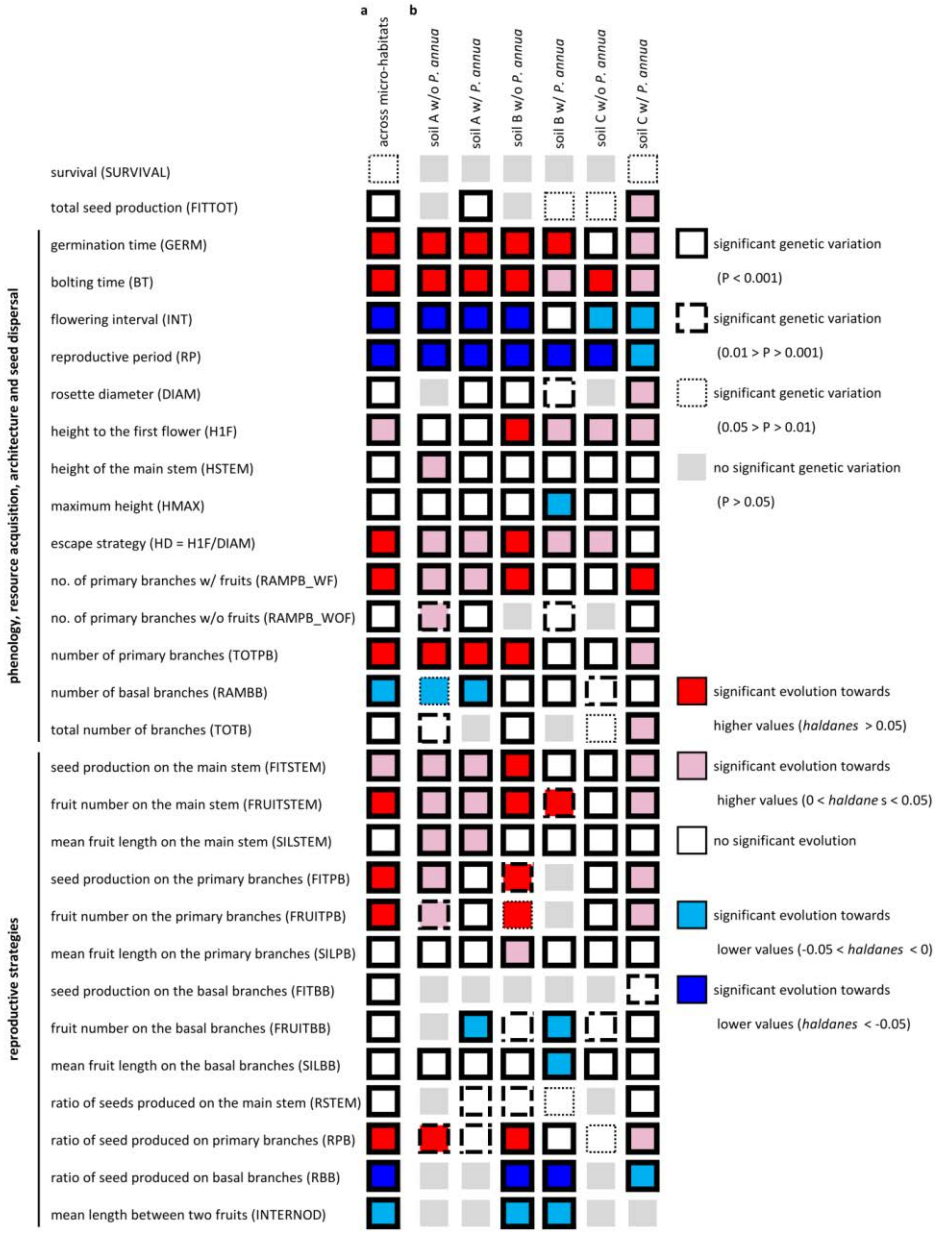

Figure 1 | Genetic variation among accessions and phenotypic evolution between 2002 and 2010. (**a**) Across the six micro-habitats. Genetic variation was detected for the 29 measured phenotypic traits. (**b**) Within each 'soil x competition' micro-habitat. The letters A, B and C stand for the three types of soil. 'w/o *P. annua*' and 'w/*P. annua*' correspond to the absence and presence of *P. annua*, respectively. The number of genetically variable traits varied between 21 (soil A in absence of *P. annua*) and 28 (soil C in presence of *P. annua*). The percentage of evolved genetically variable traits varied between 22.7% (soil C in absence of *P. annua*) and 76.2% (soil A in absence of *P. annua*). Each genetically variable trait (white and colored squares) in a given *in situ* experimental condition was defined as an eco-phenotype (n = 144). The rates of evolution are expressed in *haldanes* (a metric that scales the magnitude of change by incorporating trait standard deviations).

A resurrection experiment revealed rapid phenotypic evolution.

To identify phenotypic traits exhibiting evolutionary change within eight years, we established a resurrection experiment in which the 195 accessions collected in 2002 and 2010 were grown under common environmental conditions. This design enabled us to differentiate plastic from genetic responses[17]. After homogenizing for maternal effects, the 195 accessions were grown *in situ* in six representative micro-habitats, consisting of three contrasting soil types crossed with the presence or absence of the bluegrass *Poa annua*, a species frequently associated with *A. thaliana*[16] (**Supplementary Fig. 1**). A total of 5,850 plants were scored for 29 traits related to phenology, resource acquisition, shoot architecture, seed dispersal, fecundity, reproductive strategy and survival[18]. Interestingly, although no evolutionary change was observed for average total seed production across the six micro-habitats, we detected significant genetic evolution for 16 out of the 28 remaining traits (**Fig. 1a, Supplementary Table 1**). For example, we found a significant mean delay of 6.1 days for bolting time and a significant mean increase of ~7% in the number of fruits produced on the main stem (**Fig. 2a**). These results demonstrate that constant seed numbers can be maintained through evolution of flexible life-history and individual reproductive traits. A comparison of our results with the rates of evolution in other plant species[19] suggests a moderate rate of mean phenotypic evolution in the TOU-A population (**Fig. 2a**).

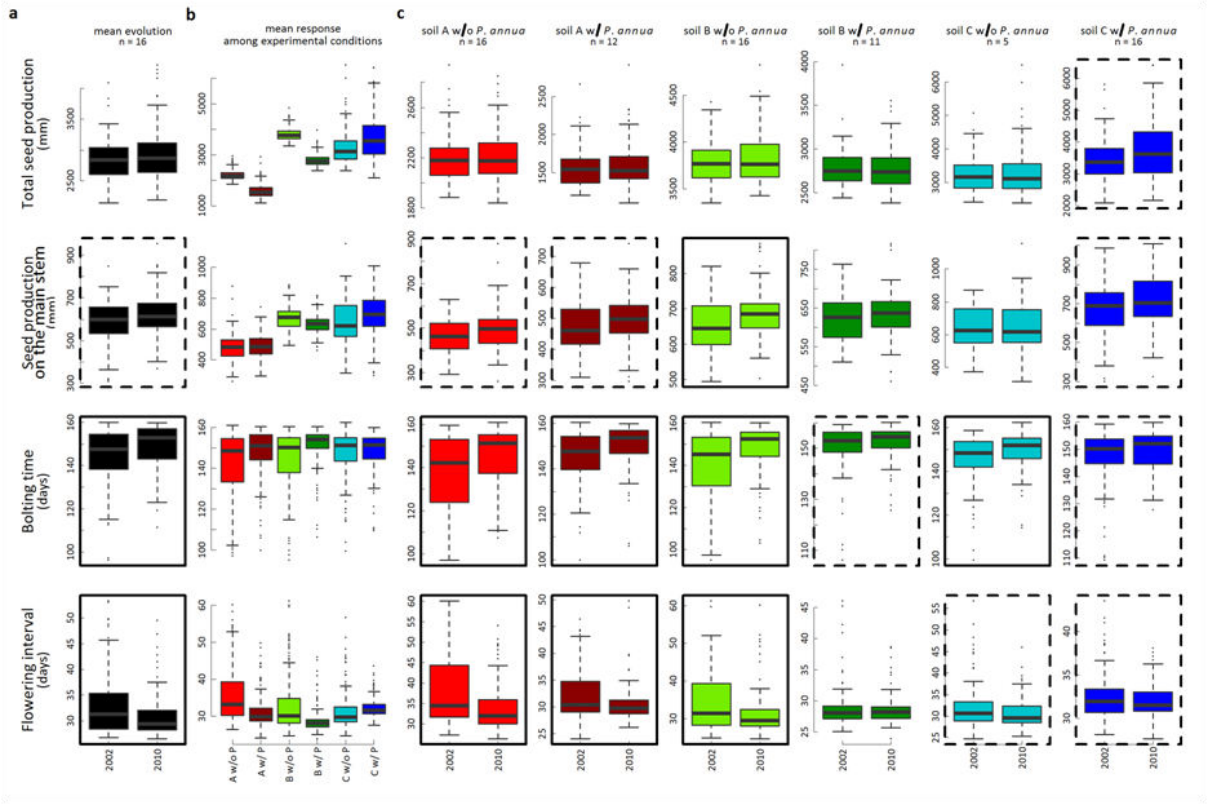

Figure 2 | Phenotypic changes in the TOU-A population over 8 generations. (**a**) Mean phenotypic evolution across the six micro-habitats. The total number of seeds produced can be maintained through evolution of phenological (bolting time and flowering interval) and individual reproductive (seed production on the main stem) traits. (**b**) Comparison among the six *in situ* 'soil x competition' micro-habitats. Average values of the phenotypes differed substantially among the six micro-habitats. (**c**) Evolution within each *in situ* micro-habitat. 'n' indicates the number of evolved phenotypic traits (**Fig. 1**). The identity of genetically variable traits that evolved between 2002 and 2008 depended on the micro-habitat. Each box plot is based on the genotypic values (BLUPs) of the TOU-A accessions (year 2002: n = 80, year 2010: n = 115). (**b**) and (**c**) The letters A, B and C stand for the three types of soil. 'w/o *P. annua*' and 'w/*P. annua*' correspond to the absence and presence of *P. annua*, respectively. (**a**) and (**c**): solid and dashed boxes indicate significant evolution with absolute *haldanes* > 0.05 and with absolute *haldanes* < 0.05, respectively (**Fig. 1**).

To confirm that the mean phenotypic change we observed was not the result of immigration from other phenotypically diverse populations[20], we sequenced the genomes of the 195 accessions collected in 2002 and 2010 (~25x coverage). We detected 1,902,592 Single Nucleotide Polymorphisms, only 5.6 times less than observed in a panel of 1135 worldwide accessions[21]. In addition, the TOU-A population appears strongly genetically isolated from

other local populations sampled within 1km (**Fig. 3a**), confirming the negligible role of immigration in the observed phenotypic change.

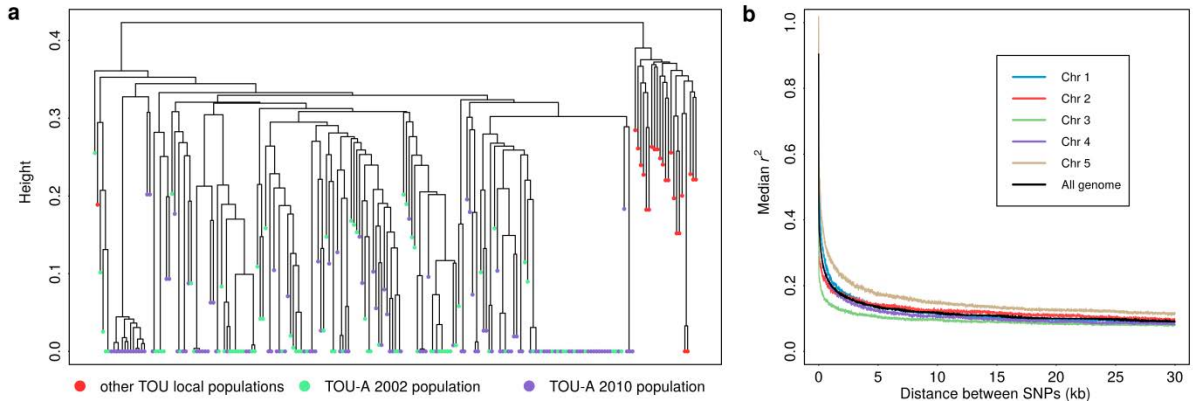

Figure 3 | Genomic patterns of the TOU-A population. (**a**) Hierarchical clustering analysis of the 195 TOU-A accessions and 24 accessions from 10 populations located within 1 km of the TOU-A population. (**b**) Decay of linkage disequilibrium (r^2) with physical distance over the five chromosomes of *A. thaliana*.

Similar phenotypic evolution associated with strong genotype-by-environment interactions.

We dissected the phenotypic evolution within each micro-habitat to test whether local abiotic and biotic growing conditions affect the genotype-phenotype relationships in the TOU-A population. Across the 29 traits measured in the six micro-habitats, 144 of these 174 eco-phenotypes displayed significant genetic variance (**Fig. 1b**), with broad-sense heritability estimates ranging from 0.20 to 0.87 (mean $H^2 = 0.57$, median $H^2 = 0.60$; **Supplementary Table 2**). Average values of the phenotypes differed substantially among the six micro-habitats (**Fig. 2b, Supplementary Table 1**). The proportions (ranging from 22.7% to 76.2%) and identities of genetically variable traits that evolved in our eight-year timespan also depended on the

micro-habitat (**Figs. 1b and 2c**). These results highlight the need to consider fine-scale environmental conditions to obtain an accurate picture of the diversity of micro-evolutionary phenotypic processes occurring within a population.

Although each trait that evolved was consistent in its direction in all micro-habitats (**Fig. 1b**), we observed highly significant changes in the ranking of accessions among micro-habitats that resulted from genotype-by-environment interactions for most traits (**Supplementary Table 1**). For example, increased allocation of reproduction to the main stem was consistently observed but different accessions most strongly manifested this allocation pattern among micro-habitats (**Supplementary Fig. 4**). These results are in accordance with previous studies revealing genotype-by-environment interactions for plant fitness-related traits at the scale of a few meters[22,23]. However, the existence of genotype-by-environment interactions does not clarify the extent of pleiotropy governing phenotypes in alternative micro-habitats: phenotypic evolution toward the same optimum may be driven by loci harboring alleles differing in the magnitude of allelic effects across micro-habitats and/or by distinct genetic bases in different micro-habitats[24].

Pleiotropy is restricted and synergistic

To characterize the genetics underlying these environmentally dependent genotype-phenotype relationships, we used GWA mapping to determine the genetic architecture, the magnitude of pleiotropy and the extent of pleiotropic scaling. The TOU-A population is well-suited for GWA mapping because it is phenotypically diverse and linkage disequilibrium (LD) decays to $r^2 = 0.5$ within 18 base pairs on average (**Fig. 3b**). In agreement with limited LD, we observed an L-shaped distribution of the size of LD blocks, with a median size of 780bp (mean

size = 5.5kb) (**Supplementary Fig. 5**). To verify our ability to finely map genomic regions associated with phenotypic variation, we first tested for the presence of significant associations of known functional polymorphisms. We successfully identified three known functional genes conferring either qualitative or quantitative resistance against bacterial pathogens when the 195 TOU-A accessions were infected under controlled conditions. In two of the three cases, the most highly associated SNP (hereafter named top SNP) was located within the gene (*RPS2* and *RKS1*)[15,25] and in the third case it was located 15 bp away (*RPM1*)[26] (**Supplementary Fig. 6**).

To further assess the efficacy of GWAS mapping in the TOU-A population, we followed the methodology used in Brachi *et al.* (2010)[27] to calculate enrichments for *a priori* candidate genes for bolting time in the six *in situ* micro-habitats (**Fig. 1**). Because bolting time is a quantitative trait for which the genetic network has been extensively studied, it is well suited for calculating enrichments for *a priori* candidate genes. Similar to previous results for a field trial utilizing 197 worldwide accessions[27], the enrichment ratio quickly dropped with the number of top SNPs in five out of the six micro-habitats, demonstrating that candidate genes were overrepresented among top-ranking SNPs (**Fig. 4a, Supplementary Fig. 7**).

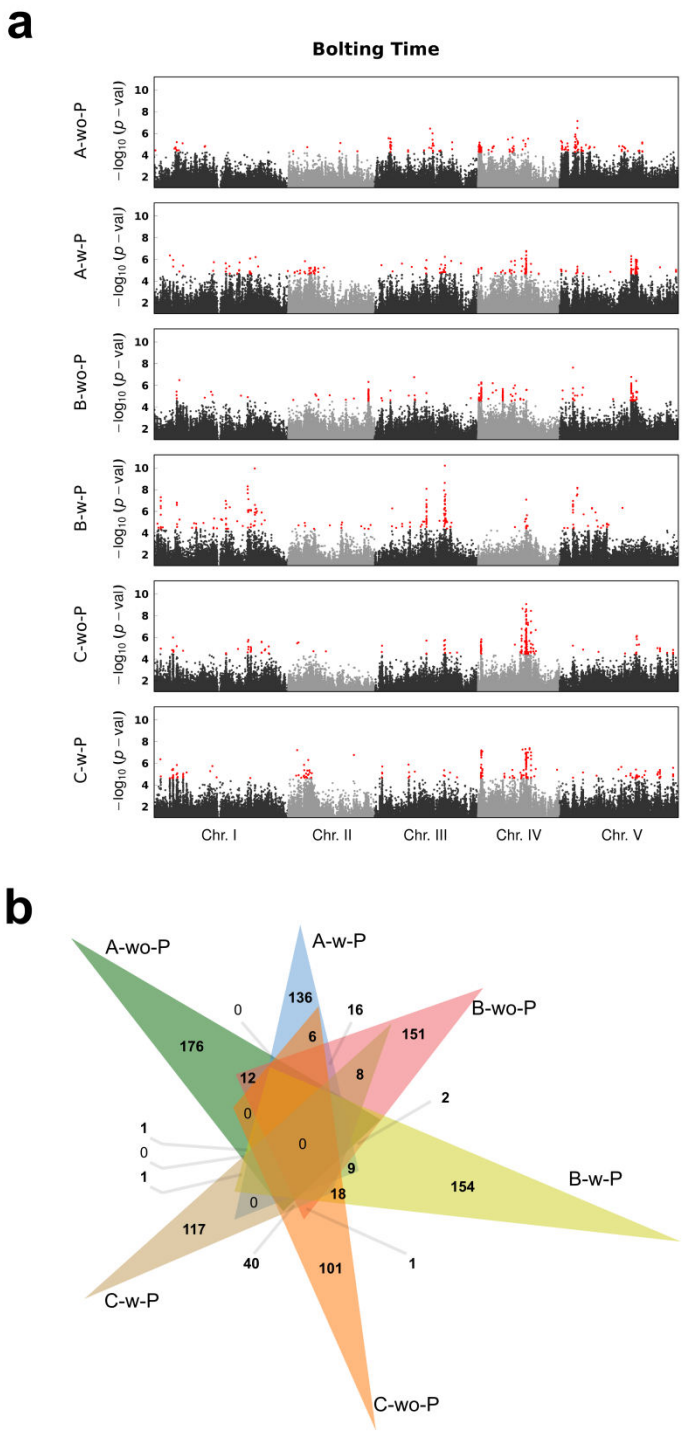

Figure 4 | Identification of genomic regions associated with bolting time variation in the TOU-A population. (**a**) Manhattan plots of mapping results for each of the six *in situ* 'soil x competition' treatments. The *x*-axis indicates the physical position along the chromosome. The *y*-axis indicates the -$\log_{10}$ *p*-values using the EMMAX method. MARF > 7%. For each experimental condition, the 200 top SNPs are highlighted in red. (**b**) Venn diagram partitioning the bolting time SNPs detected among the lists of 200 top SNPs for each *in situ* 'soil x competition' treatment. Genetic bases underlying bolting time are largely distinct across micro-habitats.

Here, we illustrate the impacts of genetic architecture, magnitude of pleiotropy and pleiotropic scaling when considering the 200 top SNPs (0.01% of the total number of SNPs) for each of the 144 eco-phenotypes. Although we observed significant enrichment for up to the 500 SNPs, focus on the 200 top SNPs is conservative in defining pleiotropy and increases the fraction of true positives. Our choice of threshold does not matter: our biological conclusions are robust to successive cutoffs of top SNPs within the range of 50-500 SNPs, and to three successive cutoffs in terms of the significance of SNPs ($-\log_{10} p$-value > 6, $-\log_{10} p$-value > 5, $-\log_{10} p$-value > 4; chosen based on van Rooijen *et al.* 2015, Thoen *et al.* 2016, Kooke *et al.* 2017)[28-30].

We first compared the genetic architecture among micro-habitats for GWA results from each of the 144 heritable eco-phenotypes (**Supplementary Fig. 8**). The number of genes located within 2kb of the 200 top SNPs ranged from 45 (fruit number on basal branches in soil B with *P. annua*) to 141 (maximum height scored in soil B without *P. annua*) (mean = 105 genes, median = 108 genes; **Supplementary Fig. 9**). For a given phenotypic trait, the numbers of associated genes sometimes varied widely across micro-habitats, even when broad-sense heritabilities were similar (**Fig. 4a, Supplementary Fig. 9, Supplementary Table 2**).

The extent of pleiotropy for each top SNP was determined by calculating an effective number of eco-phenotypes, N_{eff}, sharing a given top SNP according to Pavlicev *et al.* (2009)[31]. This statistic corrects for correlations among eco-phenotypes to produce a measure of pleiotropy that is not inflated. In agreement with previous laboratory observations on yeast, nematode and mouse[3], we found that N_{eff} follows an L-shaped distribution (**Fig. 5a**). More than 78% of top SNPs impacted a single trait in a single micro-habitat, indicating that genetic bases are largely distinct across micro-habitats (**Supplementary Fig. 10 and 11**), as illustrated for bolting time (**Fig. 4b**). This pattern of restricted pleiotropy in our study is more consistent with

the notion of modular pleiotropy (with genes being organized into structured networks) than universal pleiotropy in Fisher's geometric model (i.e. each gene affects every trait)[3,4].

We found that the total effect size of a top SNP, calculated by either the Manhattan distance (T_M) or the Euclidean distance (T_E), increased with N_{eff} faster than linearly ($T_M = c^* {N_{eff}}^d$, $d = 1.226 \pm 0.003$; $T_E = a^* Neff^b$, $b = 0.724 \pm 0.0035$; **Fig. 5b, Supplementary Fig. 11 and 12, Supplementary Tables 3 and 4**). This contrasts with most theoretical models, which typically assume that the per-trait effect size of a mutation decreases ($d = 0.5$ or $b = 0$, invariant total effect model) or remains constant ($d = 1$ or $b = 0.5$, Euclidean superposition model) with the degree of pleiotropy[4]. While previously observed in controlled laboratory conditions[3], our study reveals that such a pattern of synergistic pleiotropy can also extend to phenotypes scored in ecological realistic conditions.

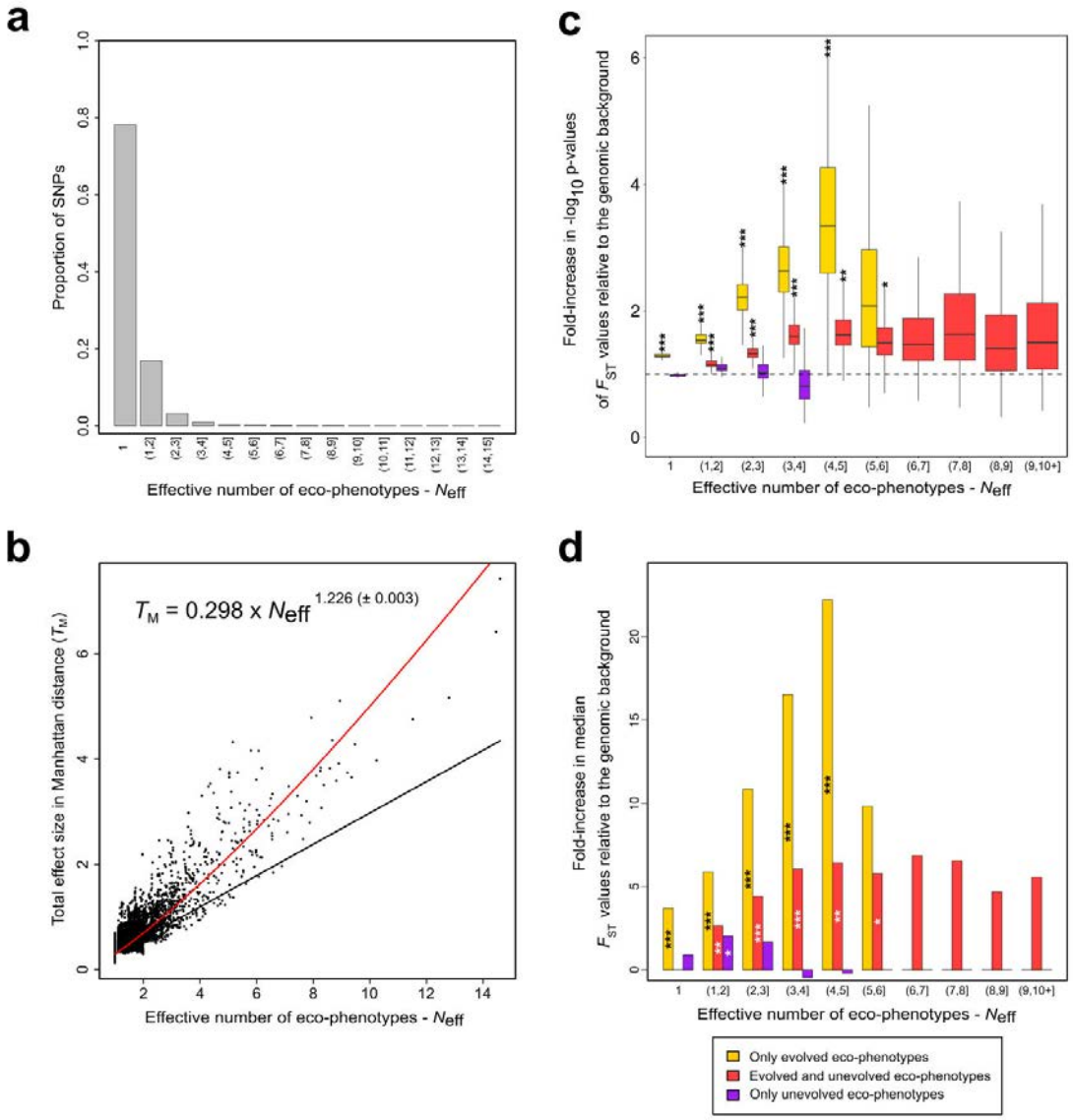

Figure 5 | Genetic architecture underlying *in situ* phenotypic evolution in the TOU-A population when considering a threshold of 200 top SNPs. (**a**) Frequency distribution of the effective number of eco-phenotypes affected by a SNP (N_{eff}, accounting for the correlations between eco-phenotypes)[31] among the 21,268 unique top SNPs. (**b**) Regression of total effect size T_M (total effect size by the Manhattan distance) on N_{eff}. The formula corresponds to the pleiotropic scaling relationship $\boldsymbol{T_M = c^* N_{eff}^{\ d}}$. A scaling component d exceeding 1 indicates that the mean per-trait effect size of a given top SNP increased with N_{eff}[3,4]. Solid red line: fitted relationship between T_M and N_{eff}, solid black line: linear dependence ($d = 1$). (**c**) Fold-increase in median $-\log_{10}$ (p-values) of neutrality tests based on temporal differentiation for SNPs that hit only evolved eco-phenotypes, only unevolved eco-phenotypes or both types of eco-phenotypes, according to different classes of effective number of eco-phenotypes. The dashed line corresponds to a fold-increase of 1, i.e. no increase in median significance of neutrality tests based on temporal differentiation. (**d**) Fold-increase in median F_{ST} values for SNPs that hit only evolved eco-phenotypes, only unevolved eco-phenotypes or both types of eco-phenotypes, according to different classes of N_{eff} (median F_{ST} across the genome = 0.00293). Significance against a null distribution obtained by bootstrapping: *0.05 > P > 0.01, **0.01 > P > 0.001, ***P < 0.001, absence of symbols: non-significant.

Intermediate degrees of synergistic pleiotropy drive adaptive evolution.

According to theoretical predictions[3,4], the combination of an L-shape distribution of N_{eff} and synergistic pleiotropy should lead polymorphisms with intermediate degrees of pleiotropy, while rare, to experience the highest rates of adaptive evolution. A genome-wide scan for selection based on temporal differentiation (F_{ST}) (**Supplementary Fig. 13**) revealed a signature of selection for top SNPs associated with evolved eco-phenotypes, but not for top SNPs associated with unevolved eco-phenotypes; top SNPs jointly associated with evolved and unevolved eco-phenotypes revealed an intermediate signature of selection (**Fig. 5c, Supplementary Fig. 11**). Because this temporal differentiation is tested against changes in the genomic background, this result rejects the hypothesis of selectively neutral evolution for evolved eco-phenotypes. When focusing attention on top SNPs associated with evolved eco-phenotypes, we found that single-trait micro-habitat-specific SNPs were under weak selection while SNPs exhibiting an intermediate degree of pleiotropy revealed the largest fold-increase of median temporal F_{ST} values (**Fig. 5d, Supplementary Fig. 11**). This pattern is strengthened when considering only the top SNPs for evolved phenotypes that have a polarity of effects in line with the direction of phenotypic evolution (~75.4% of the total number of top SNPs associated with evolved eco-phenotypes; **Supplementary Fig. 14**). Altogether, these results confirm that the evolved multi-trait combinations identified *in situ* are under selection.

As previously highlighted for the patterns of restricted pleiotropy and synergistic pleiotropy, the relationships between degree of pleiotropy and signatures of selection were robust to different number of top SNPs and thresholds of significance (within the range considered; **Supplementary Fig. 11**).

Identity of candidate genes under directional selection.

The most pleiotropic genes underlying adaptive evolution in the TOU-A population, were determined by retrieving all genes associated with 11 or more evolved eco-phenotypes. Among the 14 candidate genes (**Supplementary Table 5**), was the floral integrator *TWIN SISTER OF FT* (*TSF*), which was associated with bolting time (three microhabitats), flowering interval (one micro-habitat), the length of reproductive period (three micro-habitats), the number of primary branches (one micro-habitat) and the escape strategy (three micro-habitats). Interestingly, based on a panel of 948 worldwide accessions of *A. thaliana*, *TSF* has been found to be significantly associated with climate variation (i.e. number of consecutive cold days)[32], suggesting that *TSF* may play a major role in the adaptation of *A. thaliana* to climate at different geographical scales.

We additionally tested for biological processes that were enriched in the extreme tail of our genome-wide temporal differentiation scan (**Supplementary Table 6**). In total, 24 biological processes were enriched, 15 of which were supported by genes associated with phenotypic traits measured in this study (**Supplementary Table 6**). Enrichment for vernalization response was supported by *VERNALIZATION2* (*VRN2*) associated with six eco-phenotypes including two proxies of fitness (i.e. survival and seed production, **Supplementary Table 6**). We also detected many related, enriched functions such as stamen development, pollen maturation and callose deposition (**Supplementary Table 6**), which are consistent with the simultaneous evolution of fecundity traits observed in this study (**Fig. 1**). For instance, the candidate gene *POWDERY MILDEW RESISTANT 4* is traditionally regarded as a defense response to wounding and pathogens due to its role in reinforcing the cell wall, although it is also essential for pollen viability and cell division[33]. In this study, *POWDERY MILDEW*

RESISTANT 4 was associated with two fecundity traits: mean fruit length on primary branches (in soil A without *P. annua*) and the number of fruits on the main stem (in soil C with *P. annua*; **Supplementary Table 6**). The simultaneous evolution of fecundity traits suggests an adaptive strategy of short-lived semelparous species like *A. thaliana* in crowded environments, where plants tend to escape competition[16,34]. In agreement with this hypothesis, we observed an evolution of the escape strategy trait in five out of six micro-habitats (**Fig. 1**).

The remaining nine enriched biological processes were supported by genes that were not associated with any measured phenotype. This is not surprising in that we missed the entire seed and seedling stage, and did not capture the entire suite of biotic and abiotic factors that can impact selection over time. Among these candidate genes was the MADS-box transcription factor *FLOWERING LOCUS C* (*FLC*) that, in agreement with the recent local warming experienced by the TOU-A population, supported the strong enrichment detected for vernalization response, response to temperature stimulus and regulation of circadian rhythm (**Supplementary Table 6**). *FLC* is a well-known pleiotropic gene[35] that affects many traits that we did not measure (such as vernalization response, water use efficiency and regulation of seed dormancy by maternal temperature)[36-39], suggesting that one or more of these traits may have undergone contemporary and rapid phenotypic evolution in the TOU-A population.

It is interesting to note that we identified two central regulators of flowering time and circadian clock in our set of candidate pleiotropic genes, *i.e. FLC* and *TSF*. In two *Brassica rapa* populations that evolved rapidly following drought in Southern California[40], rapid evolution was in part mediated by a homologue of *SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1* (*SOC1*), a target of *FLC*-mediated transcriptional repression[41], suggesting that

central regulators of flowering time and circadian clock play a major role in the response to global warming.

CONCLUSION

Our ecological genomic comparison of plants separated by eight generations revealed rapid multi-trait adaptive evolution that was similar among six micro-habitats, but largely mediated by different genes. The strong genotype-by-environment interactions highlight the importance of considering fine-scale ecological variation. By limiting the erosion of standing genetic variation, this micro-habitat dependent genetic architecture should allow populations like TOU-A to continue to respond to future environmental changes.

In addition, the combination of GWAS and an *in situ* resurrection experiment validated the prediction that polymorphisms with intermediate degrees of pleiotropy, while rare, should have the highest rate of adaptive evolution. This result reinforces the importance of simultaneous evolution of multiple traits in shaping the genomic adaptive trajectory of natural populations. On-going resurrection projects in plants[42] and long-term population surveys of wild animals[43] represent an exciting opportunity to test whether restricted pleiotropy combined with synergistic pleiotropy also underlies contemporary and rapid adaptive evolution in other plant and animal species.

ACKNOWLEDGEMENTS

This work was funded by the Région Midi-Pyrénées (CLIMARES project), the INRA Santé des Plantes et Environnement department (RESURRECTION project), the INRA-ACCAF metaprogram (SELFADAPT project), the LABEX TULIP (ANR-10-LABX-41, ANR-11-IDEX-0002-02) and the National Institute of Health.

AUTHOR CONTRIBUTIONS

F.R. supervised the project. F.R. conceived of and designed the experiments. E.B., L.A., Ro.V and F.R. conducted the *in situ* experiment. L.F., C.G., C.H.C. and F.R. measured the phenotypic traits. L.F. and F.R. analyzed the phenotypic traits. O.B. and M.V. generated the sequencing data. S.B. and C.L. performed the bioinformatics analyses. L.F., C.L. and F.R. performed the GWA mapping. L.F., C.L., D.R. and F.R. performed and analyzed the enrichment tests. M.N., L.G. and Re.V. developed a methodology in selfing species to perform a genome-wide scan for selection based on temporal differentiation. V.L.C and J.B guided the analysis of phenotypic and genomic data. F.R. and J.B. wrote the manuscript, with contributions from L.F., C.L, Ro.V., M.N., L.G., Re.V. and D.R. All authors contributed to the revisions.

REFERENCES

1. Franks, S.J., Weber, J.J. & Aitken, S.N. Evolutionary and plastic responses to climate changes in terrestrial plant populations. *Evol. Appl.* 7, 123-139 (2014).
2. Bay, A.B., Rose, N., Barrett, R., Bernatchez, L., Ghalambor, C.K., Lasky, J.R., Brem, R.B., Palumbi, S.R. & Ralph, P. Predicting responses to contemporary environmental change using evolutionary response architectures. *Am. Nat.* **189** (2017).
3. Wang, Z., Liao, B.-Y. & Zhang, J. Genomic patterns of pleiotropy and the evolution of complexity. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 18034-18039 (2010).
4. Wagner, G.P. & Zhang, J. The pleiotropic structure of the genotype – phenotype map: the evolvability of complex organisms. *Nat. Rev. Genet.* **12,** 204-213 (2011).
5. DeLong, J.P., Forbes, V.E., Galic, N., Gibert, J.P., Laport, R.G., Phillips, J.S. & Vavra J.M. How fast is fast? Eco-evolutionary dynamics and rates of change in populations and phenotypes. *Ecol. Evol.* **6,** 573-581 (2016).
6. Buswell, J.M., Moles, A.T. & Hartley, S. Is rapid evolution common in introduced plant species. *J. Ecol.* **99,** 214-224 (2011).
7. Franks, S.J., Sim, S. & Weis, A.E. Rapid evolution of flowering time by an annual plant in response to a climate fluctuation. *Proc. Natl. Acad. Sci. U.S.A.* **104**: 1278-1282 (2007).
8. Reid, N.M., Proestou, D.A., Clark, B.W., Warren, W.C., Colbourne, J.K., Shaw, J.R., Karchner, S.I., Hahn, M.E., Nacci, D., Oleksiak, M.F., Crawford, D.L. & Whitehead, A. The genomic landscape of rapid repeated evolutionary adaptation to toxic pollution in wild fish. *Science* **354,** 1305-1308.
9. van't Hof, A.E., Edmonds, N., Dalikova, M., Marec, F. & Saccheri, I.J. Industrial melanism in British peppered moths has a singular and recent mutational origin. *Science* **332,** 958-960.
10. Hanikenne, M., Talke, I.N., Haydon, M.J., Lanz, C., Nolte, A., Motte, P., Kroymann, J., Weigel, D. & Krämer, U. Evolution of metal hyperaccumulation required *cis*-regulatory changes and triplication of *HMA4*. *Nature* **453,** 391-395.
11. Délye, C., Jasieniuk, M. & Le Corre V. Deciphering the evolution of herbicide resistance in weeds. *Trends Genet.* **29,** 649-658.
12. Wagner, G.P. *et al.* Pleiotropic scaling of gene effects and the 'cost of complexity'. *Nature* **452,** 470-472 (2008).

13. Platt, A. *et al.* The scale of population structure in *Arabidopsis thaliana*. *PLoS Genet.* **6**, e1000843 (2010).
14. Brachi, B. *et al.* Investigation of the geographical scale of adaptive phenological variation and its underlying genetics in *Arabidopsis thaliana*. *Mol. Ecol.* **22**, 4222-4240 (2013).
15. Huard-Chauveau, C. *et al.* An atypical kinase under balancing selection confers broad-spectrum disease resistance in Arabidopsis. *PLoS Genet.* **9**, e1003766 (2013).
16. Baron, E., Richirt, J., Villoutreix, R., Amsellem, L. & Roux, F. The genetics of intra- and interspecific competitive response and effect in a local population of an annual plant species. *Funct. Ecol.* **29**, 1361-1370 (2015).
17. Franks, S.J. *et al.* The resurrection initiative: storing ancestral genotypes to capture evolution in action. *BioScience* **58**, 870-873 (2008).
18. Roux, F., Mary-Huard, T., Barillot, E., Wenes, E., Botran, L., Durand, S., Villoutreix, R., Martin-Magniette, M.-L., Camilleri, C. & Budar, F. Cytonuclear interactions affect adaptive phenotypic traits of the annual plant *Arabidopsis thaliana* in ecologically realistic conditions. *Proc. Natl. Acad. Sci. U.S.A.* **113**: 3687-3692 (2016).
19. Bone, E. & Farres, A. Trends and rates of microevolution in plants. *Genetica* **112-113**, 165-182 (2001).
20. Hansen, M.M., Olivieri, I., Waller, D.M., Nielsen, E.E. & The GeM working group. Monitoring adaptive genetic responses to environmental change. *Mol. Ecol.* **21**, 1311-1329 (2012).
21. The 1001 Genomes Consortium. 1,135 genomes reveal the global pattern of polymorphism in *Arabidopsis thaliana*. *Cell* **166**, 1-11 (2016).
22. Kalisz, S. Variable selection on the timing of germination in *Collinsia verna* (Scrophulariaceae). *Evolution* **40,** 479-491 (1986).
23. Stratton, D.A. Spatial scale of variation in fitness of *Erigeron annuus*. *Am. Nat.* **146,** 608-324 (1995).
24. Des Marais, D.L., Hernandez, K.M. & Juenger, T.E. Interaction and plasticity : exploring genomic responses of plant to the abiotic environment. *Annu. Rev. Ecol. Evol. Syst.* **44**, 5-29 (2013).
25. Bent, A.F., Kunkel, B.N., Dahlbeck, D., brown, K.L., Schmidt, R., Giraudat, J., Leung, J. & Staskawicz, B.J. *RPS2* of *Arabidopsis thaliana*: a leucine-rich repeat class of plant disease resistance genes. *Science* **265,** 1856-1860 (1994).

26. Grant, M.R., Godirad, L., Straube, E., Ashfield, T., Lewald, J., Sattler, A., Innes, R.W & Dangl J.L. Structure of the Arabidopsis *RPM1* gene enabling dual specificity disease resistance. *Science* **269,** 843-846 (1995).
27. Brachi, B. *et al.* Linkage and association mapping of *Arabidopsis thaliana* flowering time in nature. *PLoS Genet.* 6:e1000940 (2010).
28. Van Rooijen, R., Aarts, M.G.M. & Harbinson, J. Natural genetic variation for acclimation of photosynthetic light use efficiency to growth irradiance in Arabidopsis. *Plant Physiol.* **167,** 1412-1429 (2015).
29. Kooke, R., Kruijer, W., Bours, R., Becker, F., Kuhn, A., van de Geest, H., Buntjer, J., Doeswijk, T., Guerra, J., Bouwmeester, H., Vreugdenhil, D. & Keurentjes, J.B. Genome-wide association mapping and genomic prediction elucidate the genetic architecture of morphological traits in Arabidopsis. *Plant Physiol.* **170,** 2187-2203 (2016).
30. Thoen, M.P.M. *et al.* Genetic architecture of plant stress resistance: multi-trait genome-wide association mapping. *New Phytol.* **213,** 1346-1362 (2016).
31. Pavlicev, M., Cheverud, J.M. & Wagner, G.P. Measuring morphological integration using eigenvalues variance. *Evol. Biol.* **36,** 157-170 (2009).
32. Hancock, A.M. *et al.* Adaptation to climate across the *Arabidopsis thaliana* genome. Science **334**, 83-86 (2011).
33. Ellinger, D. & Voigt, C.A. Callose biosynthesis in Arabidopsis with a focus on pathogen response: what we have learned with the last decade. *Annals of Botany* **114**, 1349-1358 (2014).
34. Bonser, S.P. High reproduction efficiency as an adaptive strategy in competitive environments. *Funct. Ecol.* **27**, 876-885 (2013).
35. Deng, W. *et al. FLOWERING LOCUS C* (*FLC*) regulates development pathways throughout the life cycle of *Arabidopsis*. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 6680-6685 (2011).
36. McKay, J.K., Richards, H. & T. Mitchell-Olds. Genetics of drought adaptation in *Arabidopsis thaliana*: I. Pleiotropy contributes to genetic correlations among ecological traits. *Mol. Ecol.* **12,** 1137-1151 (2013).
37. Li, P., Filiault, D., Box, M.S., Kerdaffrec, E., van Oosterhout, C., Wilczek, A.M., Schmitt, J., McMullan, M., Bergelson, J., Nordborg, M. & Dean, C. Multiple *FLC* haplotypes

defined by independent *cis*-regulatory variation underpin life history diversity in *Arabidopsis thaliana*. *Genes Dev.* **28,** 1635-1640 (2014).

38. Blair, L., Auge, G. & Donohue, K. Effect of *FLOWERING LOCUS C* on seed germination depends on dormancy. *Funct. Plant Biol.* **44,** 493-506 (2017).
39. Auge, G., Blair, L.K., Neville, H. & Donohue, K. Maternal vernalization and vernalization-pathway genes influence seed germination. *New Phytol.* doi: 10.1111/nph.14520 (2017).
40. Franks, S.J., Kane, N.C., O'Hara, N.B., Tittes, S. & Rest, J.S. Rapid genome-wide evolution in *Brassica rapa* populations following drought revealed by sequencing of ancestral and descendant gene pools. *Mol. Ecol.* **25,** 3622-3631 (2016).
41. Salathia, N. *et al. FLOWERING LOCUS C* - dependent and – independent regulation of the circadian clock by the autonomous and vernalization pathways. *BMC Plant Biol.* **6,** 10 (2006).
42. Etterson, J.R. *et al.* Project Basline : an unprecedented resoruce to study plant evolution across space and time. *Am. J. Bot.* **103,** 164-173 (2016).
43. Kruuk, L.E.B., Garant, D. & Charmantier, A. The study of quantitative genetics in wild populations. Pages 1-15 *In Quantitative genetics in wild populations*. Edited by A. Charmantier, D. Garant and L.E.B. Kruuk. Oxford University Press, Oxford, U.K. pp 1-15 (2014).

METHODS

Plant material. The population TOU-A is located under a 350m electric fence separating two permanent meadows experiencing cycles of periodic grazing by cattle in the village of Toulon-sur-Arroux (France, Burgundy, N 46°38'57.302'', E 4°7'16.892''). Seeds from individual plants were collected in 2002 (TOU-A-2002, n = 80) and 2010 (TOU-A-2010, n = 115) according to a sampling scheme allowing us to take into account the density of *A. thaliana* plants along a 350m transect (**Supplementary Fig. 1**). Differences in maternal effects among the 195 accessions collected in 2002 and 2010 were reduced by growing one plant per family under controlled greenhouse conditions, for one generation (16-h photoperiod, 20°C).

Ecological characterization. Eighty-three soil samples collected along the 350m transect were characterized for 14 edaphic factors[14]: pH, maximal water holding capacity (WHC), total nitrogen content (N), organic carbon content (C), C/N ratio, soil organic matter content (SOM), concentrations of P_2O_5, K, Ca, Mg, Mn, Al, Na and Fe. Climate data was generated with the ClimateEU v4.63 software package[44].

Phenotypic characterization. An experiment of 5,850 plants was set up at the local site of the TOU-A population. The 195 accessions collected in 2002 and 2010 were grown in six representative 'soil x competition' micro-habitats. Each of these micro-habitats was organized in five blocks. Each of the five blocks corresponded to an independent randomization of 195 plants with one replicate per accession collected in 2002 and 2010. Seeds were sown in late September to mimic the main natural germination cohort observed in the TOU-A population (**Supplementary Fig. 1**). Each plant was scored for a total of 29 phenotypic traits chosen to characterize the life history of *A. thaliana* including the timing of offspring production or seed

dispersal, or because they are involved in the response to competition and/or are good estimators of life-time fitness and reproductive strategies[18].

Phenotypic analyses, natural variation, phenotypic evolution and evolutionary rates. We explored natural variation of all phenotypic traits using the following statistical mixed model:

$$Y_{ijklm} = \mu_{trait} + block_i\,(soil_j * comp_k) + soil_j + comp_k + soil_j * comp_k + year_l + soil_j * year_l + comp_k * year_l + soil_j * comp_k * year_l + accession_m\,(year_l)) + accession_m\,(year_l)) * soil_j + accession_m\,(year_l)) * comp_k + accession_m\,(year_l)) * soil_j * comp_k + \varepsilon_{ijklm} \quad (1)$$

In this model, '*Y*' is one of the 29 phenotypic traits, 'μ' is the overall phenotypic mean; 'block' accounts for differences between the five experimental blocks within each type of 'soil * absence/presence of *P. annua*' experimental combination; 'soil' corresponds to the effects of the three types of soil; 'comp' measure the effect of the presence of *P. annua*; 'year' corresponds to effect of the two sampling years 2002 and 2010; 'accession' measures the effect of accessions within year; interaction terms involving the 'accession' term account for genetic variation in reaction norms of accessions between the three types of soil and the absence or presence of *P. annua*; and 'ε' is the residual term.

All factors were treated as fixed effects, except 'accession' that was treated as a random effect. For fixed effects, terms were tested over their appropriate denominators for calculating *F*-values. Significance of the random effects was determined by likelihood ratio tests of model with and without these effects. When necessary, raw data were either log transformed or Box-Cox transformed to satisfy the normality and equal variance assumptions of linear regression. A correction for the number of tests was performed for each modeled effect to control the False Discovery Rate (FDR) at a nominal level of 5%.

Inference was performed using ReML estimation, using the PROC MIXED procedure in SAS 9.3 (SAS Institute Inc., Cary, North Carolina, USA) for all traits with the exception of SURVIVAL, which was analyzed using the PROC GLIMMIX procedure in SAS 9.3.

For all traits, Best Linear Unbiased Predictions (BLUPs) were obtained for each accession in each of the six experimental conditions, using the PROC MIXED procedure in SAS 9.3 (SAS Institute Inc., Cary, North Carolina, USA):

$$Y_{imc} = \mu_{trait} + block_i + accession_m + \varepsilon_{im} \quad (2)$$

For each trait, significant genetic variation among the accessions was detected by testing the significance of the 'accession' term in equation (2). A correction for the number of tests was performed for the modeled 'accession' effect (across the 29 traits within each of the six experimental conditions) to control the FDR at a nominal level of 5%. Because *A. thaliana* is a highly selfing species[13], BLUPs correspond to the genotypic values of accessions.

In each of the six experimental conditions, rates of evolutionary change based on genotypic values of accessions were calculated in *haldanes* (h_g) for all eco-phenotypes with significant genetic variation among the 195 accessions collected in 2002 and 2010. *haldanes* is a metric that scales the magnitude of change by incorporating trait standard deviations[45,46]. h_g values were calculated between 2002 and 2010, as:

$$h_g = \frac{(x_2/s_p) - (x_1/s_p)}{g} \quad (3)$$

where 'x' corresponds to the mean genotypic value at year 1 (TOU-A population collected in 2002) and year 2 (TOU-A population collected in 2010), 's_p' is the standard deviation of the genotypic values of the trait pooled across the two years, and 'g' is the number of generations. Because only one germination cohort was observed every year between 2002 and 2010 (i.e. fall germination cohort), only one generation per year was considered in the calculation of *haldanes* values. For a given trait, 95% confidence intervals were estimated based on the distribution of 1000 *haldanes* values obtained by bootstrapping 1000 random samplings with replacement of genetic values within each year. A *haldanes* value was considered significantly different from zero if its 95% confidence intervals did not overlap zero.

Sequencing and polymorphism detection. DNA-seq experiments were performed on an Illumina HiSeq2500 using a paired-end read length of 2x100 pb with the Illumina TruSeq SBS v3 Reagent Kits. Raw reads of each of the 195 accessions were mapped onto the TAIR10 *A. thaliana* reference genome Col-0 with a maximum of 5 mismatches on at least 80 nucleotides. A semi-stringent SNPCalling across the genome was then performed for each accession with SAMtools mpileup (v0.01019)[47] and VarScan (v2.3)[48] with the parameters corresponding to a theoretical sequencing coverage of 30X and the search for homozygous sites.

Patterns of linkage disequilibrium and geographic structure. Considering only SNPs with a Minor Allele Relative Frequency (MARF) > 0.07, the LD extent within 30kb-windows on each chromosome were estimated using *VCFtools*[49]. LD blocks across the genome were identified in the PLINK environment using the following parameters --blocks no-pheno-req --maf 0.07 --blocks-max-kb 200, leading to the identification of 19,607 blocks with at least two SNPs (mean number of SNPs per block = 47.6, median number of SNPs per block = 12, mean block length = 5.5kb, median block length = 0.78kb). To position the TOU-A population within the French geographic structure, we retrieved the positions of the 214,051 SNPs genotyped on

24 accessions within 10 populations located within 1km of the TOU-A population[50] across the genomes of the TOU-A population. Clustering genotype analysis was performed using the packages gdsfmt and SNPRelate in the *R* environment[51], using the snpgdspLD pruning command with the following parameters *ld.threshold=0.8 slide.max.bp=500 maf=0.07*, leaving us with 90,883 SNPs.

Genome-Wide Association mapping and MARF threshold. GWA mapping was run using a mixed-model approach implemented in the software EMMAX (Efficient Mixed-Model Association eXpedited)[52]. This model includes a genetic kinship matrix as a covariate to control for population structure.

Because of bias due to rare alleles[27,52,53], we estimated a MARF threshold above which the *p*-value distribution is not dependent on the MARF. We plotted the 99% quantile of the *p*-value distribution of all 144 eco-phenotypes (i.e. 'micro-habitat x trait' combinations) displaying significant genetic variance (**Fig. 1**) along 50 MARF values (with an increment of 0.01 from 0.01 to 0.5). A locally-weighted polynomial regression indicated that *p*-value distributions were dependent on MARF value. From visual inspection, we considered a threshold of MARF value > 0.07, which resulted in a total number of 981,617 SNPs for the following analyses (**Supplementary Fig. 15**).

Enrichment for *a priori* candidate genes. To determine the threshold number of top SNPs (i.e. SNPs with the highest associations) above which additional top SNPs would behave like the rest of the genome, we calculated enrichments for *a priori* candidate genes for natural genetic variation of bolting time observed in the six *in situ* experimental conditions (**Fig. 1**). Based on an algorithm described in Brachi *et al.*(2010)[27] and a list of 328 candidate genes for bolting time[14], enrichment was calculated for progressively fewer selective sets of top SNPs

within a 20Kb window of an *a priori* candidate gene. For each set of top SNPs, a null distribution of enrichment was computed to determine a 95% confidence interval[27].

Degree of pleiotropy and pleiotropic scaling. Each trait displaying significant genetic variance in a given *in situ* micro-habitat was considered an "eco-phenotype". The degree of pleiotropy of a given top SNP was defined as the number of eco-phenotypes that shared this top SNP. To account for the correlations between eco-phenotypes that can overestimate the degree of pleiotropy, we followed Wagner *et al.* (2008)[12] by estimating for each top SNP an effective number of eco-phenotypes as $N_{\text{eff}} = N - \text{var}(\lambda)$ where $\text{var}(\lambda)$ is the variance of the eigenvalues of the error-corrected matrix.

The allelic effects were calculated using the mixed model implemented in the software EMMAX after fitting the pairwise genetic kinship effect[52]. Because different units were used to measure the 29 traits scored in this study, we calculated a standardized allelic effect for each eco-phenotype affected by a top SNP according to Wagner *et al.* (2008)[12]. The standardized effect on eco-phenotype *i*, denoted by A_i, is half the difference in genotypic means between the two homozygous genotypes. The total size of the phenotypic effects of a top SNP was then calculated by the Manhattan distance[54] $T_{\text{M}} = \sum_{i=1}^{n} |A_i|$, where n is the degree of pleiotropy and A_i is the standardized allelic effect[3,4,12]. The pleiotropic scaling relationship between the total effect size and the effective number of eco-phenotypes was calculated as $T_{\text{M}} = c^* {N_{\text{eff}}}^d$.

The pleiotropic scaling relationship between the total effect size and the effective number of eco-phenotypes was also calculated as $T_{\text{E}} = a^* Neff^b$, with T_{E} corresponding to the Euclidean distance and calculated as $T_{\text{E}} = \sqrt{\sum_{\text{i=1}}^{\text{n}} A_{\text{i}}^2}$, where n is the degree of pleiotropy and A_i is the standardized allelic effect.

The degree of pleiotropy and the pleiotropic scaling relationship were calculated for (i) five threshold number of top SNPs (i.e. 50 SNPs, 100 SNPs, 200 SNPs, 300 SNPs and 500 SNPs) and (ii) three thresholds of significance ($-\log_{10} p$-value > 6, $-\log_{10} p$-value > 5, $-\log_{10} p$-value > 4). To avoid pseudo-replication due to the presence of several top SNPs in a given LD block (n = 19,607 blocks with at least two SNPs), the pleiotropic scaling was also calculated for each threshold number of top SNPs and each threshold of significance, (i) by considering the mean value of T_M (or T_E) and N_{eff} per LD block containing top SNPs and (ii) by randomly sampling one top SNP per LD block (this step was repeated 1,000 times).

Genome-wide scan for selection based on temporal differentiation. In the following, we outline a procedure inspired by Goldringer & Bataillon (2004)[55] to test for the homogeneity of differentiation across SNP markers between two temporal samples. If all SNP markers are selectively neutral, they should provide estimates of temporal differentiation drawn from the same distribution, which depends on the strength of genetic drift in the population (and therefore on its effective size). In contrast, if some marker loci are targeted by selection (or if they are in linkage disequilibrium with selected variants), then some heterogeneity in locus-specific measures of temporal differentiation should be observed. This is due to selection that will tend to drive measures of differentiation to values greater (or smaller) than expected under drift alone. The rationale of our approach is therefore to identify those SNPs that show outstanding differentiation, compared to neutral expectation.

We measure temporal differentiation between sample pairs using F_{ST}. Although the F_C statistic[56] was used in Goldringer & Bataillon (2004)[55], estimators of F_{ST} have better statistical properties in terms of bias and variance, and multilocus estimates have been precisely defined and thoroughly evaluated[57].

Using a multilocus estimate of F_{ST} from the pair of temporal samples, we infer the

effective size of the population. Because the 195 *A. thaliana* accessions are considered highly homozygous across the genome, heterozygous sites were discarded (see above) and the data therefore consist of haploid genotypes. We considered a single haploid population of constant size N_e, which has been sampled at generation 0, and τ generations later. Generations do not overlap. New mutations arise at a rate μ, and follow the infinite allele model (IAM). Following Skoglund *et al.* (2014)[58], the pairwise parameter F_{ST} between the two samples can be read:

$$F_{ST} = \frac{1 - e^{-\theta T/2}}{1 + \theta - e^{-\theta T/2}}$$

where $T \equiv \tau / N_e$ and $\theta \equiv 2N_e\mu$. In the low mutation limit (i.e., as $\mu \to 0$):

$$F_{ST} \approx \frac{T}{T+2} = \frac{\tau}{\tau + 2N_e}$$

This suggests that a simple moment-based estimator of effective population size can be derived as:

$$\widehat{N}_e = \frac{\tau(1 - \hat{\bar{F}}_{ST})}{2\hat{\bar{F}}_{ST}}$$

where $\hat{\bar{F}}_{ST}$ is a multilocus estimate of the parameter F_{ST}. In what follows, we use the estimator of Weir & Cockerham (1984)[57]; preliminary analyses showed that these estimates of effective size have lower bias and variance than averaged estimates based on single-locus estimates of F_C.

In this study, the pairwise differentiation between the 195 *A. thaliana* accessions samples collected in 2002 and 2010 based on the full set of 1,902,592 SNP markers was: $\hat{\bar{F}}_{ST}$ = 0.0215, which gives an estimate of $\hat{N}_e$ = 182 (measured as a number of gene copies).

For each SNP, we tested the null hypothesis that the locus-specific differentiation measured at this focal marker was only due to genetic drift. For this purpose, we computed the

expected distribution of F_{ST} for each SNP, conditional upon the estimated effective size (using the same estimated value for all markers: $\hat{N}_e = 182$), and the allele frequencies at the focal SNP in the initial sample (i.e. 80 accessions collected in 2002). We simulated individual gene frequency trajectories, as follows:

Suppose that we observe k_0 copies of the reference allele, out of n_0 sampled genes, in the 2002 sample. We assume that these observed counts are drawn from a binomial distribution $B(n_0,\pi_0)$ where π_0 is the (unknown) allele frequency of the reference allele in the population. Assuming a Beta(1,1) prior distribution for π_0 (uniform distribution), and using the Bayes inversion formula, the posterior distribution of π_0 is a Beta($k_0 + 1, n_0 - k_0 + 1$). For each marker and for each simulation, we therefore draw the initial allele frequency $\tilde{\pi}_0$ from a Beta($k_0 + 1, n_0 - k_0 + 1$). We then draw "pseudo-observed" allele counts using a random draw from $B(n_0, \tilde{\pi}_0)$. This procedure allows accounting for the sampling variance in initial allele frequencies, instead of fixing $\tilde{\pi}_0$ to the observed frequency in the sample, as previously done in Goldringer & Bataillon (2004)[55].

Then, we simulated eight generations of drift, using successive binomial draws with parameters $\hat{N}_e = 182$ and the allele frequency in the previous generation. In the last generation, a sample of genes is taken as a binomial draw with parameters n_τ (the sample size in 2010), and $\tilde{\pi}_\tau$ (the simulated allele frequency in the last generation).

Last, we computed locus-specific estimates of temporal F_{ST} from the simulated allele counts at the initial and last generation. The whole procedure was repeated at least 10,000 times for each marker (additional simulations were performed for some markers to obtain non-null *p*-values).

Finally, we assigned a *p*-value to each SNP marker, computed as the proportion of

simulations giving a locus-specific estimate of F_{ST} larger than or equal to the observed value at the focal SNP. We checked that the distribution of *p*-values was fairly uniform (data not shown).

Note that all SNP markers with a MARF ≤ 0.07 (computed as the overall frequency across the two temporal samples) were discarded from the analysis. There were 981,617 remaining loci (**Supplementary Fig. 7**). To avoid any potential bias, all the distributions of F_{ST} were obtained using only simulated markers with a MARF > 0.07.

Enrichment analysis of top SNPs for signals of selection. Based on the effective number of eco-phenotypes affected by a SNP, we tested whether top SNPs related to evolved eco-phenotypes rejected the hypothesis of selectively neutral evolution more often than top SNPs related to unevolved eco-phenotypes for any given degree of pleiotropy. For each set of top SNPs (i.e. top SNPs that hit only evolved eco-phenotypes, top SNPs that hit only unevolved eco-phenotypes and top SNPs that hit both types of eco-phenotypes), we first computed a fold-increase in median significance of F_{ST} values using the following ratio: $ratio_{significance}$ = median of $-\log_{10}$(*p*-values) of F_{ST} values of *n* top SNPs / median of $-\log_{10}$(*p*-values) of F_{ST} values of *n* SNPs randomly sampled across the genome, where *n* = number of top SNPs. This step was repeated 1,000 times, generating a distribution of fold-increase in median significance of F_{ST} values of top SNPs. We assigned a *p*-value by computing the proportion of $ratio_{significance}$ smaller or equal to 1. The random sampling was done according to a scheme that results in sets of SNPs that resemble the original set with respect to linkage disequilibrium[32].

We then tested whether the strength of selection differed among the degrees of pleiotropy by computing a fold-increase in median F_{ST} values for each set of top SNPs, using the following ratio: $ratio_{values}$ = median of F_{ST} values of *n* top SNPs / median of F_{ST} values of all SNPs. This step was repeated 1,000 times, by randomly sampling the same number *n* of SNPs across the genome. This procedure generated a null distribution of fold-increase in median F_{ST} values. We assigned a *p*-value by comparing $ratio_{values}$ calculated for the set of top SNPs to the quantiles at 95%, 99% and 99.9% of the null distribution.

The enrichment analysis of top SNPs for signals of selection was calculated for (i) five threshold number of top SNPs (i.e. 50 SNPs, 100 SNPs, 200 SNPs, 300 SNPs and 500 SNPs) and (ii) three thresholds of significance ($-\log_{10}$ *p*-value > 6, $-\log_{10}$ *p*-value > 5, $-\log_{10}$ *p*-value > 4).

Identity of candidate genes under directional selection and enrichment in biological processes.

To identify pleiotropic candidate genes associated with the 76 evolved eco-phenotypes, we first selected the 50 SNPs the most associated with each evolved eco-phenotype, leading to a total of 3800 SNPs. We then retrieved all the annotated genes located within a 2kb window on each side of those top SNPs, leading to a final list of 4855 unique candidate genes. We finally focused on genes associated with 11 or more evolved eco-phenotypes.

To determine which biological processes were important for adaptation of the TOU-A population over eight generations, we tested whether SNPs in the 0.1% upper tail of the F_{ST} value distribution were over-represented in each of 736 Gene Ontology Biological Processes from the GOslim set[59]. 10,000 permutations were run to assess significance using the same methodology as described in Hancock *et al.* (2011)[32]. For each significantly enriched biological process, we retrieved the identity of all the genes containing SNPs in the 0.1% upper tail of the F_{ST} values distribution.

Data availability. The raw sequencing data used for this study will be available at the NCBI Sequence Read Archive (http://ncbi.nlm.nih.gov/sra) through the Study accession SRP077483. The phenotypic data that support the findings of this study are available from the authors on a reasonable request. The genomic SNP data files will be archived through the Dryad digital repository upon acceptance for publication.

Code availability. Custom scripts and phenotypic and genomic files used in this study have been archived in a local depository (https://lipm-browsers.toulouse.inra.fr/pub/Frachon2017-NEE/) that can be accessed by the reviewers with the login '**reviewersNEE'** and the password '**FaupKinmyad4**'. All the scripts and data sets will be made available available in the Dryad database upon acceptance of the manuscript. The code for performing genome-wide scan for selection based on temporal differentiation will be made available on the Zenodo database upon acceptance of the manuscript (Vitalis R, Gay L and Navascues M (2016) TempoDiff: a computer program to detect selection from temporal genetic differentiation. INRA. http://dx.doi.org/10.5281/zenodo.375600).

44. Hamann, A., Wang**,** T., Spittlehouse, D.L. & Murdock, T.Q. A comprehensive, high-resolution database of historical and projected climate surfaces for western North America. *B. Am. Meteorol. Soc.* **94**, 1307 (2013).
45. Hendry, A.P. & Kinnison, M.T. The pace of modern life: measuring rates of contemporary microevolution. *Evolution* **53**, 1637-1653 (1999).
46. Gingerich, P.D. Rates of evolution on the time scale of the evolutionary process. *Genetica* **112-113**, 127-144 (2001).
47. Li, H. *et al.* The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078-2079 (2009).
48. Koboldt, D.C. *et al.* VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res.* **22**, 568-576 (2012).
49. Danecek, P. *et al.* The variant call format and VCFtools. *Bioinformatics* **27**, 2156-2158 (2011).
50. Horton, M.W. *et al.* Genome-wide patterns of genetic variation in worldwide *Arabidopsis thaliana* accessions for the RegMap panel. *Nat. Genet.* **44**, 212-216 (2012).
51. Zheng, X. *et al.* A High-performance computing toolset for relatedness and Principal Component Analysis of SNP data. *Bioinformatics* **28,** 326-3328 (2012).
52. Kang, H.M. *et al.* Variance component model to account for sample structure in genome-wide association studies. *Nat. Genet.* **42**, 348-354 (2010).

53. Atwell, S. *et al.* Genome-wide association study of 107 phenotypes in a common set of *Arabidopsis thaliana* inbred lines. *Nature* **465**, 627-631 (2010).
54. Hermisson J. & McGregor A.P. Pleiotropic scaling and QTL data. *Nature* **456**, E3-E4 (2008).
55. Goldringer, I. & Bataillon, T. On the distribution of temporal variations in allele frequency consequences for the estimation of effective population size and the detection of loci undergoing selection. *Genetics* **168**, 563-568 (2004).
56. Waples, R.S. A generalized approach for estimating effective population size from temporal changes in allele frequency. *Genetics* **121**, 379-391 (1989).
57. Weir, B.S. & Cockerham, C.C. Estimating *F*-statistics for the analysis of population structure. *Evolution* **38**, 1358-1370 (1984).
58. Skoglund, P., Sjödin, P., Skoglund, T., Lascoux, M. & Jakobsson, M. Investigating population history using temporal genetic differentiation. *Mol. Biol. Evol.* **31**, 2516-2527 (2014).
59. The Gene Ontology Consortium. The Gene Ontology project in 2008. *Nucleic Acids Res.* **36**, D440-D444 (2008).

Supplementary Information

Intermediate degrees of synergistic pleiotropy drive adaptive evolution in ecological time

Léa Frachon, Cyril Libourel, RomainVilloutreix, Sébastien Carrère, Cédric Glorieux, Carine Huard-Chauveau, Miguel Navascués, Laurène Gay, Renaud Vitalis, Etienne Baron, Laurent Amsellem, Olivier Bouchez, Marie Vidal, Valérie Le Corre, Dominique Roby, Joy Bergelson, Fabrice Roux*

*To whom correspondence should be addressed. E-mail: fabrice.roux@inra.fr

This file includes:

Supplementary Information: text

Supplementary Figures 1-15

Supplementary Tables 1-6

SUPPLEMENTARY INFORMATION

Plant material

In this study, we focused on the population TOU-A located under a 350m electric fence separating two permanent meadows experiencing cycles of periodic grazing by cattle (**Supplementary Fig. 1A**) in the village of Toulon-sur-Arroux (Burgundy, East of France, N 46°38'57.302'', E 4°7'16.892''). Seeds from individual plants were collected in 2002 (TOU-A1), 2007 (TOU-A5) and 2010 (TOU-A6) according to a sampling scheme allowing us to take into account the density of *A. thaliana* plants along the transect: (1) from the starting point of the transect (**Supplementary Fig. 1A**), walk along the transect until a plant is found and collect seeds from this plant, (2) if this plant is at the beginning of a patch, then collect seeds from plants located every 50 cm along this patch, (3) else, walk along the transect until a new plant is found and collect seeds from this plant. According to this sampling scheme, seeds of 80, 115 and 115 individual plants were collected in 2002, 2007 and 2010, respectively (**Supplementary Fig. 1**). Seeds collected from those 310 individual plants constitute seed families, hereafter named accessions. Given the outcrossing rate of ~6% observed in the TOU-A population[1], the 310 accessions were considered as relatively homozygous across the genome.

Seeds from the 80 accessions collected in 2002 were grown individually in a controlled greenhouse at The University of Chicago (USA) and seeds for each TOU-A1 accession collected. The analysis of these 80 accessions genotyped at 149 SNPs gave an estimate of selfing rate of ~94%[1].

Differences in the maternal effects between the 310 accessions were reduced by growing one plant of each family for one generation under controlled greenhouse conditions (16-h photoperiod, 20°C) in early 2011 at the University of Lille 1. For this purpose, we planted seeds produced at The University of Chicago for accessions from the TOU-A1 population, and seeds collected in the field for accessions from the TOU-A5 and TOU-A6 populations. For the purpose of this study, we only used seeds from the 80 accessions collected in 2002 and from the 115 accessions collected in 2010.

Ecological characterization

Climate characterization

Data for the mean annual temperature, the mean warmest month temperature, the mean coldest month temperature, the sum of degree-days above 5°C, the sum of degree-days below 0°C and the mean annual precipitation were retrieved from 1970 to 2013. Climate data was generated with the ClimateEU v4.63 softwarepackage, available at http://tinyurl.com/ClimateEU, based on methodology described in Hamann *et al.* (2013)[2].

Soil characterization

A sample of the 5-cm upper soil layer was collected at 83 positions scattered along the transect in 2010 (**Supplementary Fig. 1**). These samples were air-dried in the greenhouse (20-22°C), and then stored at room temperature. As described in Brachi *et al.*(2013)[3], each soil sample was characterized for 14 edaphic factors: pH, maximal water holding capacity (WHC), total nitrogen content (N), organic carbon content (C), C/N ratio, soil organic matter content (SOM), concentrations of P_2O_5, K, Ca, Mg, Mn, Al, Na and Fe. Iron concentration (Fe) was excluded from further analyses due to a lack of variation among the 83 samples. In order to reduce multicollinearity, the set of remaining 13 edaphic variables was pruned based on the pairwise Spearman correlations of the variables, so that no two variables had a Spearman *rho* greater than 0.8. In cases where variables were strongly inter-correlated, we selected the one with the most obvious link to the ecology of *A. thaliana*. The final set of 10 edaphic variables considered in this study was N, C/N ratio, pH, WHC, P_2O_5, K, Mg, Mn, Na and Al.

To visualize the edaphic space of the TOU-A population, we conducted a principal component analysis (PCA) based on the 83 values of the 10 edaphic traits (*R* package ade4)[4].

Phenotypic characterization

Experimental design

An experiment of 5850 plants was set up at the local site of the TOU-A population. The experimental design and the experimental conditions are illustrated on **Supplementary Fig. 1**. Based on the edaphic space (**Supplementary Fig. 3**), we defined three contrasting edaphic areas under the electric fence, hereafter named soil types A, B and C. In late August 2012, a 12.3-m² (4.4m * 2.8) plot was delimited by an electric fence for protection against cattle in each

soil type. In each plot, one subplot of 2.88-m² (4.8m * 0.6m, experimental condition without the presence of *P. annua*, see below) and one subplot of 3.36-m² (4.8m * 0.7m, experimental condition with the presence of *P. annua*, see below) were arranged at 80-cm spacing. In late August 2012, each subplot was manually weeded and tilled for the 10-cm upper soil layer. The 24th of September 2012, subplots were surrounded by green plastic covers for weed control. To mimic the main natural germination cohort observed in the TOU-A population in late September 2012 (**Supplementary Fig. 1**), seeds were sown on the 24th of September 2012 for the experimental conditions 'soil A without *P. annua*', ''soil A with *P. annua*' and 'soil B without *P. annua*', and on the 25th of September 2012 for the experimental conditions 'soil B with *P. annua*', ''soil C without *P. annua*' and 'soil C with *P. annua*'. Each of the six *in situ* experimental conditions was organized in five blocks, each one being represented by 3 arrays of 66 individual wells (Ø4 cm, vol. ~38 cm^3) (TEKU, JP 3050/66). Across the five blocks, the 15 arrays were stuck some on the others and organized according to a grid of 15 columns and one line. To buffer against possible border effects in the experimental conditions with *P. annua*, the 15 arrays were surrounded by one row of wells sown with both *P. annua* and *A. thaliana* (accession TOU-A6-69 collected in 2010). All the wells were first filled with 3 cm of the respective native soil, then with an additional 1cm of the respective native soil that was oven dried for two days at 65°C. The oven dried native soil prevented germination from the seed bank, whereas the 3-cm native soil allowed the colonization of the oven dried native soil by native microbiota.

In each of the six *in situ* experimental conditions, each of the five blocks corresponded to an independent randomization of 195 plants with one replicate per accession collected in 2002 and 2010. In each block, the remaining three wells were left empty. Five seeds of *A. thaliana* were sown in each well. For the three *in situ* experimental conditions with *P. annua*, a mean number of five seeds of *P. annua* were additionally sown in each well. Seeds for *P. annua* were ordered to the company Herbiseeds (http://www.herbiseed.com/home.aspx). After sowing, arrays were directly transported *in situ* and slightly buried in their dedicated soil types. Arrays were covered for 10 days with an agricultural fleece that allowed the seeds to be exposed to rain and sunlight while preventing them from disturbance by rain drops.

Germination date was monitored daily for 10 days (see below). Seeds germinated in more than 97.74 % of the wells. Wells were thinned to one seedling of *A. thaliana* and/or one seedling of *P. annua* between 18 and 22 days after sowing. During the course of the experiment

(late September 2012 – late June 2013), plants were protected from herbivory by slugs as described in Brachi *et al.* (2010)[5].

Measured traits

Each plant was scored for a total of 29 phenotypic traits related to phenology (n = 4), resource acquisition (n = 1), architecture and seed dispersal (n = 9), fecundity (n = 14) and survival (n = 1). These traits were chosen to characterize the life history of *A. thaliana* including the timing of offspring production or seed dispersal[3,6-8], or because they are involved in the response to competition[9,10], and/or are good estimators of life-time fitness and reproductive strategies[7,11-14]. Most of these traits have been fully described in Roux *et al.* (2016)[14]:

- *Phenology*: Germination time (GERM) was measured as the number of days between sowing and the emergence of the first seedling (opening of both cotyledons). Bolting time (BT), flowering interval (INT) and the reproductive period (RP) were scored as the interval between germination date and bolting date (inflorescence distinguishable from the leaves at a size < 5 mm), between bolting date and flowering date (appearance of the first open flower) and between flowering date and date of maturation of the last fruit, respectively.
- *Resource acquisition*: At the start of flowering, the maximum diameter of the rosette measured to the nearest millimeter was used as a proxy for plant size (DIAM).
- *Architecture and seed dispersal*: After maturation of the last fruit, the above-ground portion was harvested and stored at room temperature. Plants were later phenotyped for the following architectural and seed dispersal related traits: height from soil to the first fruit on the main stem (H1F, in mm), height of the main stem (HSTEM, in mm), maximum height (HMAX, in mm), number of primary branches on the main stem with fruits (RAMPB_WF) or without fruits (RAMPB_WOF), total number of primary branches (TOTPB), total number of basal branches (RAMBB) and total number of branches (TOTB = TOTPB + RAMBB). We also evaluated a response strategy to competition (ratio HD = H1F / DIAM)[9].
- *Fecundity*: Because the number of seeds in a fruit is highly correlated with fruit length[11], total seed production was approximated by total fruit length (FITTOT, in mm). Seed production is a good proxy for fecundity in a highly selfing annual species like *A. thaliana*. FITTOT was obtained by adding the fruit length produced on the main stem (FITSTEM, in mm), the primary branches on the main stem (FITPB, mm) and the basal branches (FITBB, in mm). These estimates of fruit length were obtained by counting the number of fertilized fruits produced on each type of branches (FRUITSTEM, FRUITPB and FRUITBB) and multiplying these counts by an estimate of their corresponding fruit (or silique) length (SILSTEM, SILPB and SILBB, in mm), estimated as the average of three haphazardly selected representative fruits. We also calculated three ratios corresponding to the percentage of seeds produced by one branch type as a function of

the total amount of seed produced: RSTEM = FITSTEM / FITTOT, RPB = FITPB / FITTOT and RBB = FITBB / FITTOT. Finally, we estimated the average length between two fruits on the main stem (INTERNOD = (HSTEM – H1F) / (FRUITSTEM - 1); in mm).

- *Survival*: All plants that germinated but did not survive were counted as dead (SURVIVAL = 0). Harvested plants were counted as alive (SURVIVAL = 1).

Genomic characterization

DNA extraction, libraries preparation and genome sequencing

Genomic DNA for the 195 accessions collected in 2002 and 2010 was extracted as described in Brachi *et al.* (2013)[3]. DNAseq was performed at the GeT-PlaGe core facility (INRA Toulouse). DNA-seq libraries were prepared according to Illumina's protocol using the Illumina TruSeq Nano LT Kit. Briefly, DNA was fragmented by sonication on a covaris M220, size selection was performed using CLEANNA CleanPCR beads and adaptators were ligated for sequencing. Library quality was assessed using an Advanced Analytical Fragment Analyser and libraries were quantified by QPCR using the Kapa Library Quantification Kit. DNA-seq experiments were performed on an Illumina HiSeq2500 using a paired-end read length of 2x100 pb with the Illumina TruSeq SBS v3 Reagent Kits. Each PCR product with tag-sequence was first quantified using PicoGreen® dsDNA Quantitation Reagent. Then a mix was made depending on these quantities in order to obtain an equimolar pool.

Mapping and SNP calling

Raw reads of each of the 195 accessions were mapped onto the *A. thaliana* reference genome Col-0 (genome size: 119Mb, TAIR10, https://www.arabidopsis.org/portals/genAnnotation/gene_structural_annotation/annotation_data.jsp) using glint software (1.0.rc8; Faraut & Courcelle, unpublished software) with the following parameters: a maximum of 5 mismatches on at least 80 nucleotides and keep alignments with the best score (*glint mappe --no-lc-filtering --best-score --mmis 5 --lmin 80 --step 2*). The mapped reads were filtered for proper pairs with SAMtools (v0.01.19)[15] (*samtools view -f 0x02*). The mean and the median coverage to a unique position in the reference genome was ~25.5x and ~24.5x, respectively.

A stringent SNPCalling across the genome was then performed for each accession with SAMtools mpileup (v0.01019)[15] and VarScan (v2.3)[16] with the parameters corresponding to a

theoretical sequencing coverage of 30X and the search for homozygous sites (*samtools mpileup -B ; VarScan mpileup2snp --min-coverage 5 --min-reads2 4 --min-avg-qual 30 --min-var-freq 0.97 --p-value 0.01*). Due to the relatively high selfing rate observed in *A. thaliana* and the generation(s) of selfing performed in greenhouse conditions (see the subsection 'Plant material'), the frequency of heterozygous sites should be low; those sites were not considered in this study in order to avoid paralogs. All polymorphic sites were then identified among the 195 accessions. Finally, a SNP calling based on all accessions was performed on all polymorphic sites to differentiate null values from the reference value. Sites with more than 50% missing values were discarded from the set of polymorphic sites.

Testing whether the mean Linkage Disequilibrium extent in the TOU-A population is short enough for fine mapping of genomic regions associated with natural phenotypic variation

The presence of significant associations at loci known to be involved in well described phenotypes provides a proof-of-concept for the power of conducting GWAS in a given mapping population. To estimate the power of fine mapping in the TOU-A population, we focused (i) on the *R* genes *RPM1* and *RPS2* responsible for the hypersensitive cell death response (HR) against the engineered bacterial strain of *Pseudomonas syringae* DC3000 expressing either *AvrRpm1* (DC3000::AvrRpm1) or *AvrRpt2* (DC3000::AvrRpt2), respectively, and (ii) on the atypical kinase *RKS1* conferring quantitative broad-spectrum resistance against the vascular bacterial pathogen *Xanthomonas campestris* pv. *campestris* (reviewed in Roux & Bergelson (2016)[17]). The 195 accessions collected in 2002 and 2010 were grown, inoculated and phenotyped for (i) qualitative resistance against DC3000::AvrRpm1 (leaf collapse scored at 6hpi) and DC3000::AvrRpt2 (leaf collapse scored at 1dpi) as described in Vailleau *et al.* (2002)[18], and (ii) quantitative resistance against the strain *Xcc568* (disease index scored using a scale from 0 to 4 at 10dpi) as described in Huard-Chauveau *et al.* (2013)[19]. Given the broad-sense heritability values close to one observed for qualitative resistance[20], four leaves of a single plant were inoculated for each accession. For quantitative resistance against *Xcc568*, a randomized complete block design was set up with two blocks, each being an independent randomization of one replicate per accession. In the latter case, the following general linear model was used to analyze disease index (GLM procedure in SAS9.1, SAS Institute Inc., Cary, North Carolina, USA):

$$\text{disease index}_{ij} = \mu + \text{block}_i + \text{accession}_j + \varepsilon_{ij}$$

where 'μ' is the overall mean; 'block' accounts for differences among the two experimental blocks; 'accession' corresponds to the 195 natural accessions; and 'ε' is the residual term. Normality of the residuals was not improved by transformation of the data. Least-square mean (LSmean) was obtained for each natural accession

GWA mapping was run using a mixed-model approach implemented in the software EMMAX (Efficient Mixed-Model Association eXpedited)[21]. This model includes a genetic kinship matrix as a covariate to control for population structure.GWA mapping was based on (i) raw means for qualitative resistance against DC3000::AvrRpm1 and DC3000::AvrRpt2, and (ii) LSmeans for quantitative resistance against *Xcc568*.

References to Supplementary Information

1. Platt, A. *et al.* The scale of population structure in *Arabidopsis thaliana*. *PLoS Genet.* **6**, e1000843 (2010).
2. Hamann, A., Wang**,** T., Spittlehouse, D.L. & Murdock, T.Q. A comprehensive, high-resolution database of historical and projected climate surfaces for western North America. B. Am. Meteorol. Soc. **94**, 1307 (2013).
3. Brachi, B. *et al.* Investigation of the geographical scale of adaptive phenological variation and its underlying genetics in *Arabidopsis thaliana*. *Mol. Ecol.* **22**, 4222-4240 (2013).
4. Chessel, C., Dufour, A.B. & Thioulouse, J. The ade4 package – I – One-table methods. *R News* **4**, 5 (2004).
5. Brachi, B. *et al.* Linkage and association mapping of *Arabidopsis thaliana* flowering time in nature. *PLoS Genet.* 6:e1000940 (2010).
6. Weinig, C. *et al.* Novel loci control variation in reproductive timing in *Arabidopsis thaliana* in natural environments. *Genetics* **162**, 1875-1884 (2002).
7. Reboud, C. *et al.* Natural variation among accessions of *Arabidopsis thaliana*: beyond the flowering date, what morphological traits are relevant to study adaptation? In Plant adaptation: molecular biology and ecology. Edited by Q. C. Cronk, J. Whitton and I. Taylor. NRC Research Press, Ottawa, Canada. pp 135-142 (2004).
8. Wender, N.J., Polisetty, C.R. & Donohue, K. Density-dependent processes influencing the evolutionary dynamics of dispersal: a functional analysis of seed dispersal in *Arabidopsis thaliana* (Brassicaceae). *Am. J. Bot.* **92**, 960-971 (2005).
9. Baron, E., Richirt, J., Villoutreix, R., Amsellem, L. & Roux, F. The genetics of intra- and interspecific competitive response and effect in a local population of an annual plant species. *Funct. Ecol.* **29**, 1361-1370 (2015).

10. Brachi, B., Aimé, C., Glorieux, C., Cuguen, J. & Roux, F. Adaptive value of phenological traits in stressful environments: predictions based on seed production and Laboratory Natural Selection. *PLoS One* **7**, e32069 (2012).
11. Roux, F., Gasquez, J. & Reboud, X. The dominance of the herbicide resistant cost in several *Arabidopsis thaliana* mutant lines. *Genetics* **166**, 449-460 (2004).
12. Roux, F., Giancola, S., Durand, S. & Reboud, X. Building of an experimental cline with *Arabidopsis thaliana* to estimate herbicide fitness cost. *Genetics* **173**,1023-1031 (2006).
13. Bac-Molenaar, J.A. *et al.* Genome-wide association mapping of fertility reduction upon heat stress reveals developmental stage-specific QTLs in *Arabidopsis thaliana*. *The Plant Cell* **27**, 1857-1874 (2015).
14. Roux, F. *et al.* Cytonuclear interactions affect adaptive phenotypic traits of the annual plant *Arabidopsis thaliana* in ecologically realistic conditions. *Proc. Natl. Acad. Sci. U.S.A.* **113**: 3687-3692 (2016).
15. Li, H. *et al.* The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078-2079 (2009).
16. Koboldt, D.C. *et al.* VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res.* **22**, 568-576 (2012).
17. Roux, F. & Bergelson, J. The genetics underlying natural variation in the biotic interactions of *Arabidopsis thaliana*: the challenges of linking evolutionary genetics and community ecology. *Curr. Top. Dev. Biol.* **119**, 111-156 (2016).
18. Vailleau, F. *et al.* A R2R3-MYB gene, *AtMYB30*, acts as a positive regulator of the hypersensitive cell death program in plants in response to pathogen attack. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 10179-10184 (2002).
19. Huard-Chauveau, C. *et al.* An atypical kinase under balancing selection confers broad-spectrum disease resistance in Arabidopsis. *PLoS Genet.* **9**, e1003766 (2013).
20. Atwell, S. *et al.* Genome-wide association study of 107 phenotypes in a common set of *Arabidopsis thaliana* inbred lines. *Nature* **465**, 627-631 (2010).
21. Kang, H.M. *et al.* Variance component model to account for sample structure in genome-wide association studies. *Nat. Genet.* **42**, 348-354 (2010).

Figure S1 | General picture of the TOU-A population. (**a**) Photograph showing the habitat type. The population is located under a 350m electric fence separating two permanent meadows. (**b**) Position of plants for which seeds have been collected in 2002, 2007 and 2010. (**c**) Position of soil samples collected in 2010. The letters A, B and C indicate the three edaphic areas (i.e. soil types) in which the *in situ* experiment has been performed (see **Supplementary Fig. 3**). (**d**) Tillage of the 10-cm upper soil layer in late August 2012 and protection from cattle by electric fences. (**e**) Soil cover with green plastic for weed control in late September 2012. (**f**) Observed natural germination flushes in late September 2012.

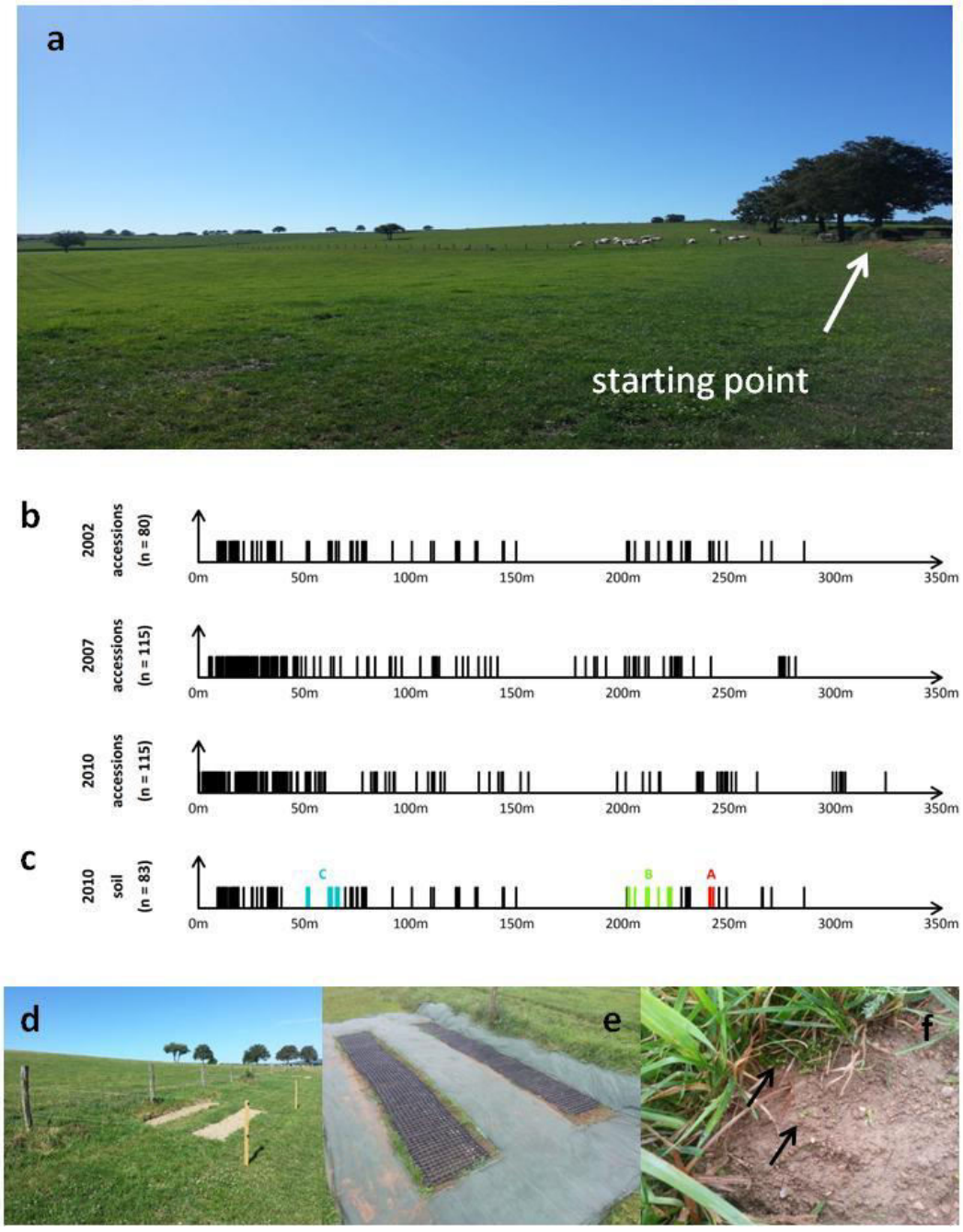

Figure S2 | Climate change since 1970 in the locality of the TOU-A population. Blue dots indicate the three sampling years (2002, 2007 and 2010). The green dot indicates the year of the *in situ* experiment. Red lines correspond to the mean of the last five consecutive years. A significant change over time was detected for the mean annual temperature (Spearman's *rho* = 0.63, $P = 5.5 \times 10^{-6}$), the mean warmest month temperature (Spearman's *rho* = 0.35, $P = 0.019$) and the sum of degree-days above 5°C (Spearman's *rho* = 0.69, $P = 7.1 \times 10^{-7}$), but not for the mean coldest month temperature (Spearman's *rho* = 0.-0.026, $P = 0.865$), the mean annual precipitation (Spearman's *rho* = 0.025, $P = 0.869$) and the sum of degree-days below 0°C (Spearman's *rho* = -0.090, $P = 0.560$).

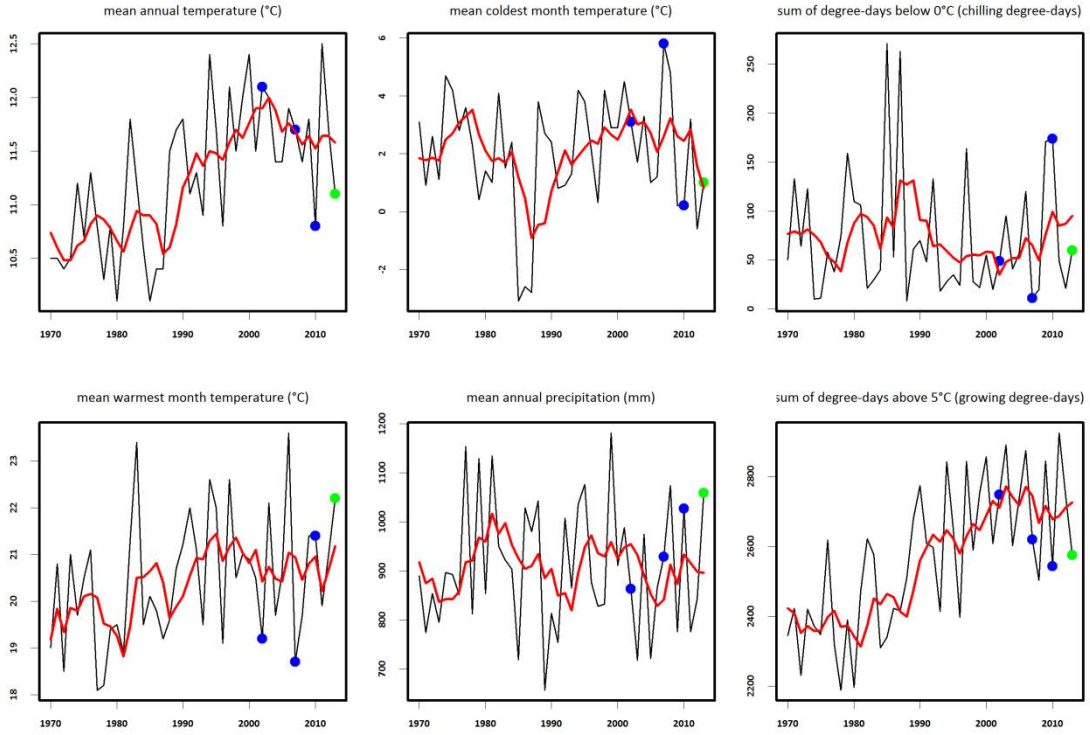

Figure S3 | Edaphic variation in the TOU-A population. (**a**) Factor loading plot resulting from principal components analysis. Factor 1 and factor 2 explained 31.54% and 24.22% of total soil variance. Maximum water holding capacity (WHC), content of total nitrogen (N), organic carbon / total nitrogen ratio (C.N), concentrations of P_2O_5, K, Mg, Mn, Al and Na. (**b**) Distribution of eigenvalues against the ranked component number. (**c**) Position of the 83 soil samples in the 'Factor1 – Factor 2' edaphic space. Red, green and blue dots correspond to the soil samples located in three soil areas 'soil A', 'soil B' and 'soil C', respectively.

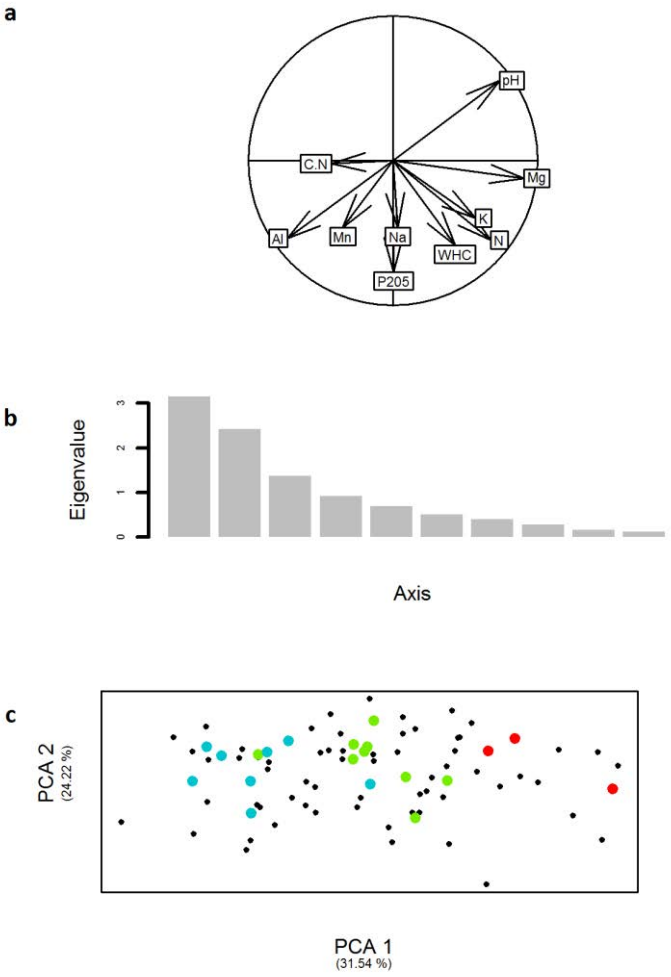

Figure S4 | Illustration of the genotype-by-environment interactions across the six *in situ* 'soil x competition' micro-habitats. (**a**) Genetic variation for reaction norms of bolting time. (**b**) Genetic variation for reaction norms of seed production on the main stem. Solid red lines: reaction norms of the 80 accessions collected in 2002, dashed blue lines: reaction norms of the 115 accessions collected in 2010.

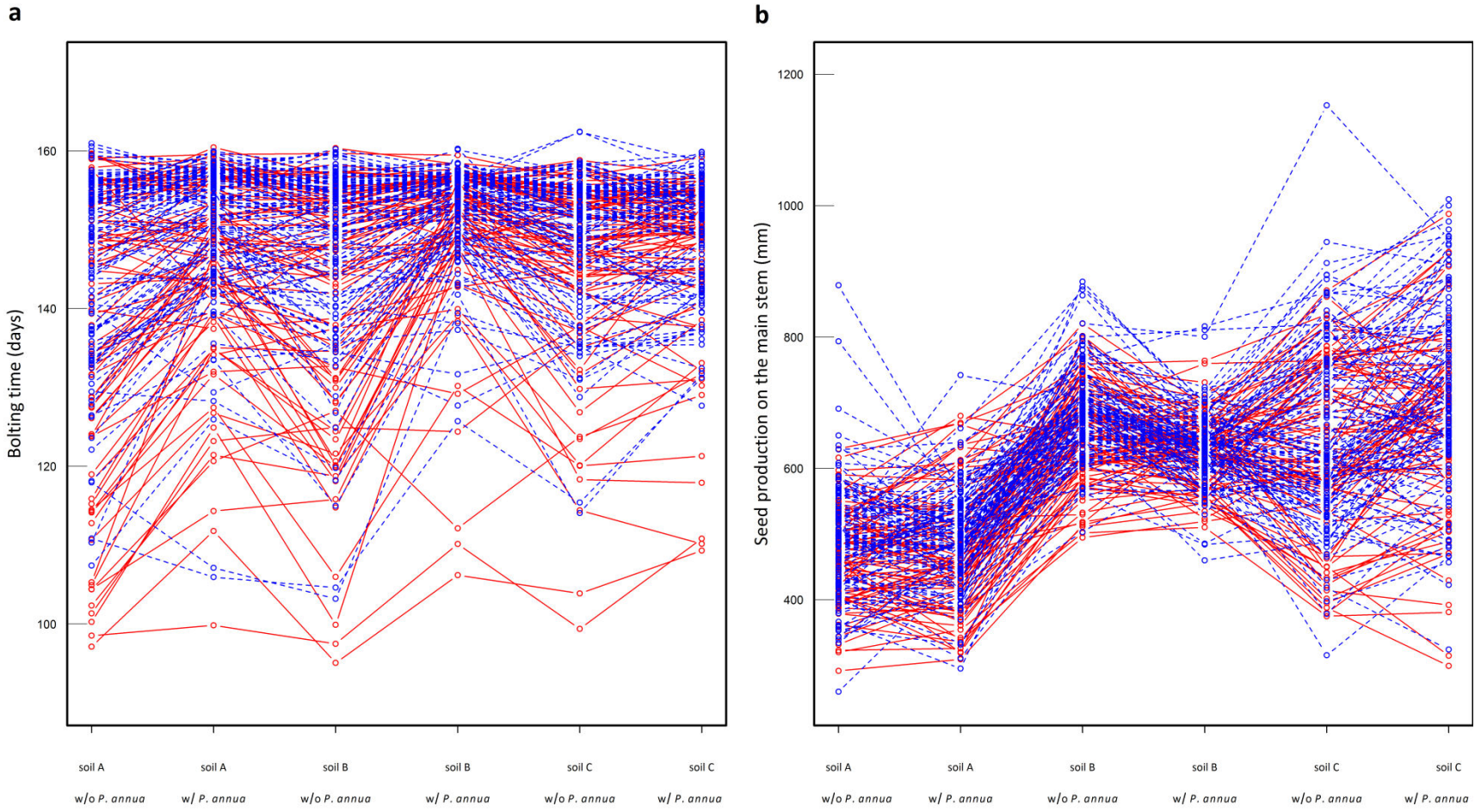

Figure S5 | Distribution of the size of LD blocks in the TOU-A population.

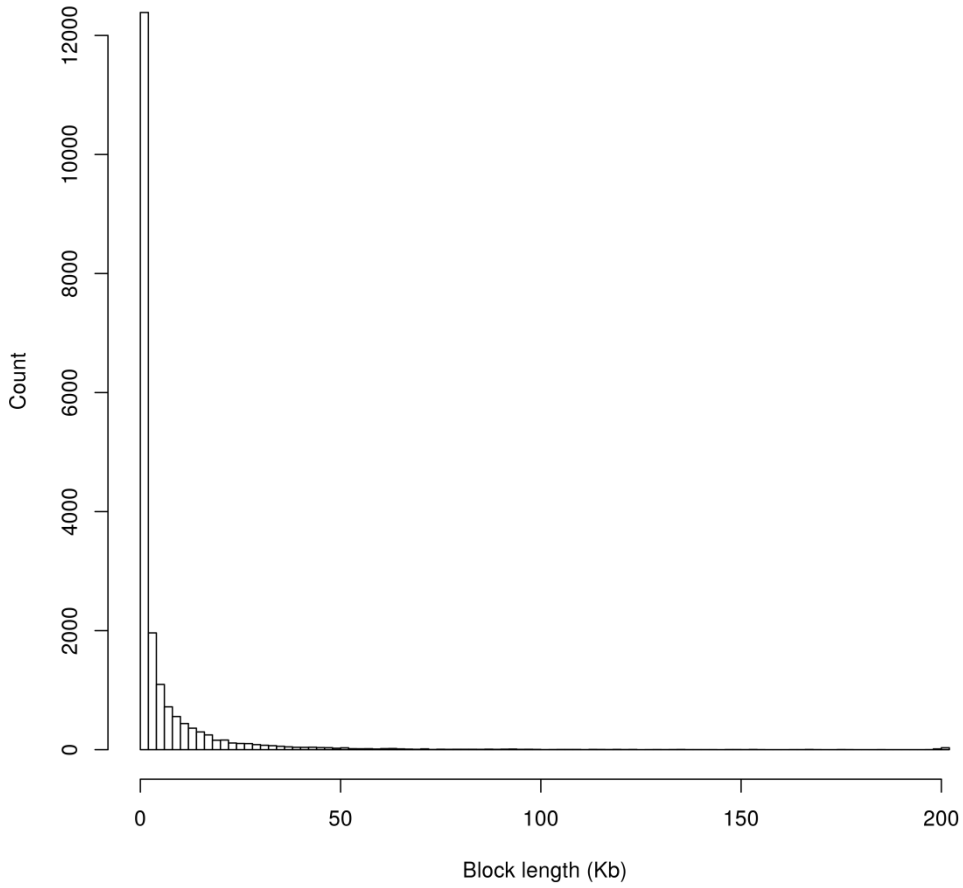

Figure S6 | GWA analysis of hypersensitive response to the bacterial elicitors *AvrRpm1* (a) and *AvrRpt2* (b) and quantitative resistance to *Xanthomonas campestris* pv. *campestris* strain *Xcc568* (c). The top SNPs are located 15bp from *RESISTANCE TO PSEUSOMONAS SYRINGAE PV MACULICOLA* (*RPM1*), within *RESISTANT TO PSEUDOMONAS SYRINGAE 2* (*RPS2*) and within *RESISTANCE RELATED KINASE 1* (*RKS1*). The *x*-axis indicates the physical position along the chromosome. The *y*-axis indicates the -log^{10} *p*-values of phenotype-SNP associations using the EMMAX method. MARF > 7%.

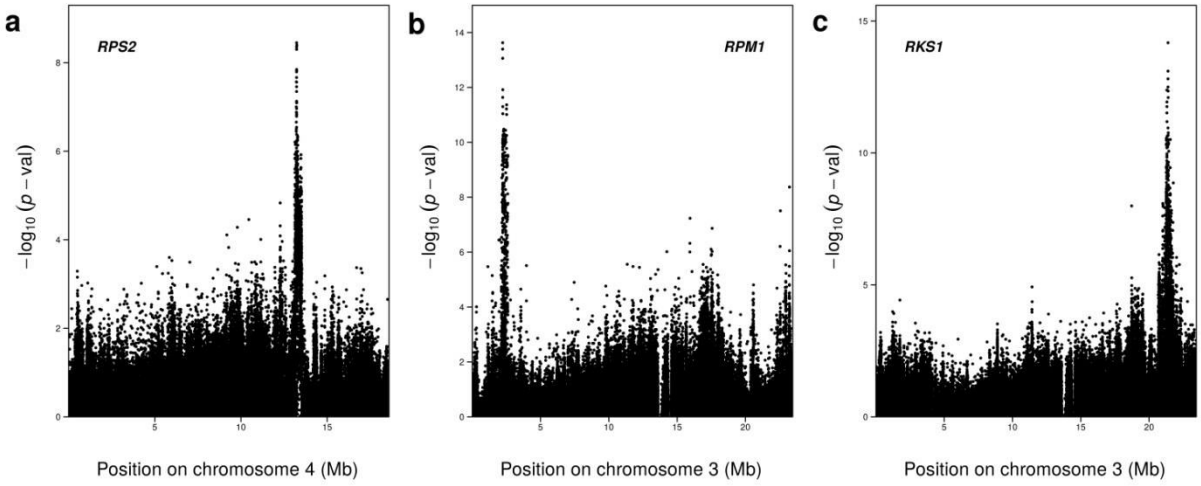

Figure S7 | Enrichment ratios in flowering time candidate genes for the six *in situ* 'soil x competition' micro-habitats (i.e. three soils A, B and C x absence or presence of *P. annua*), as a function of the number of top SNPs chosen in the GWA mapping results for bolting time using the EMMAX method. The corresponding 95% confidence intervals from the null distributions are represented by the green area.

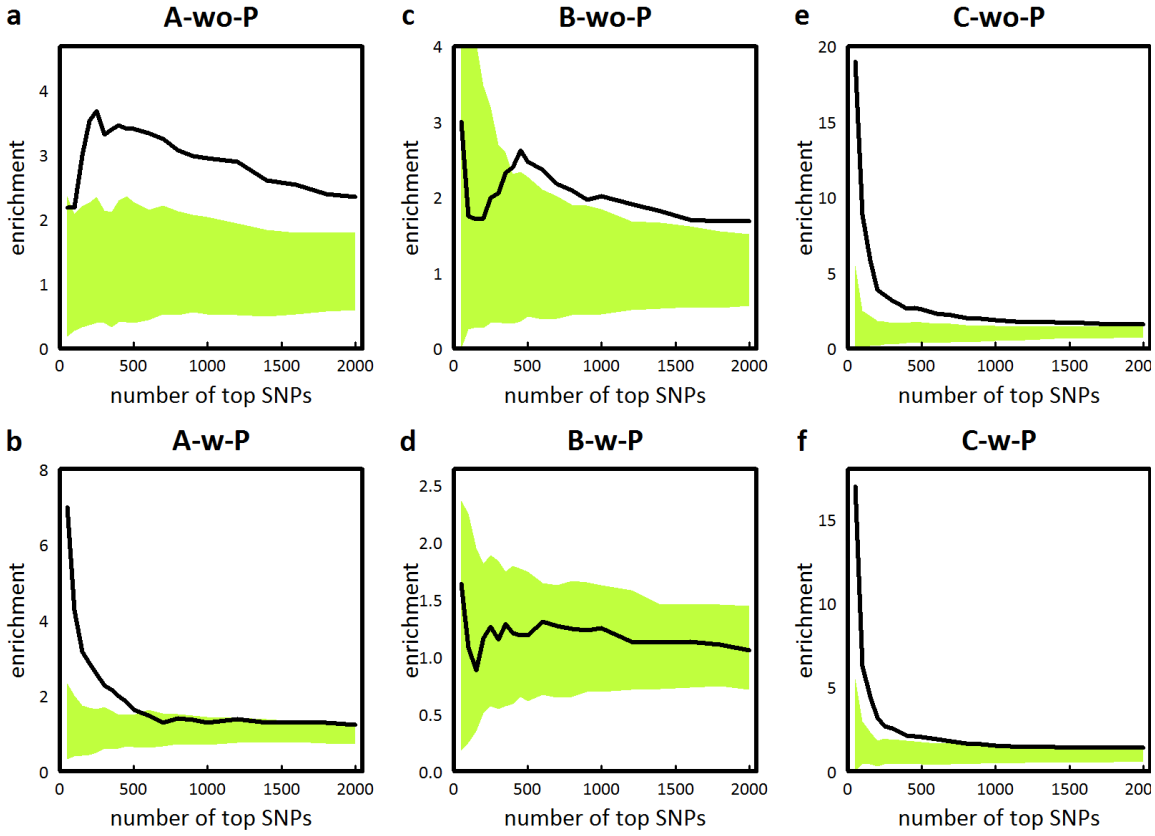

Figure S8 | Identification of genomic regions associated with the 144 heritable eco-phenotypes in the TOU-A population. The *x*-axis indicates the physical position along the chromosome. The *y*-axis indicates the -log^{10} *p*-values using the EMMAX method. MARF > 7%. On each Manhattan plot, the 100 and 200 top SNPs are highlighted in blue and red, respectively.

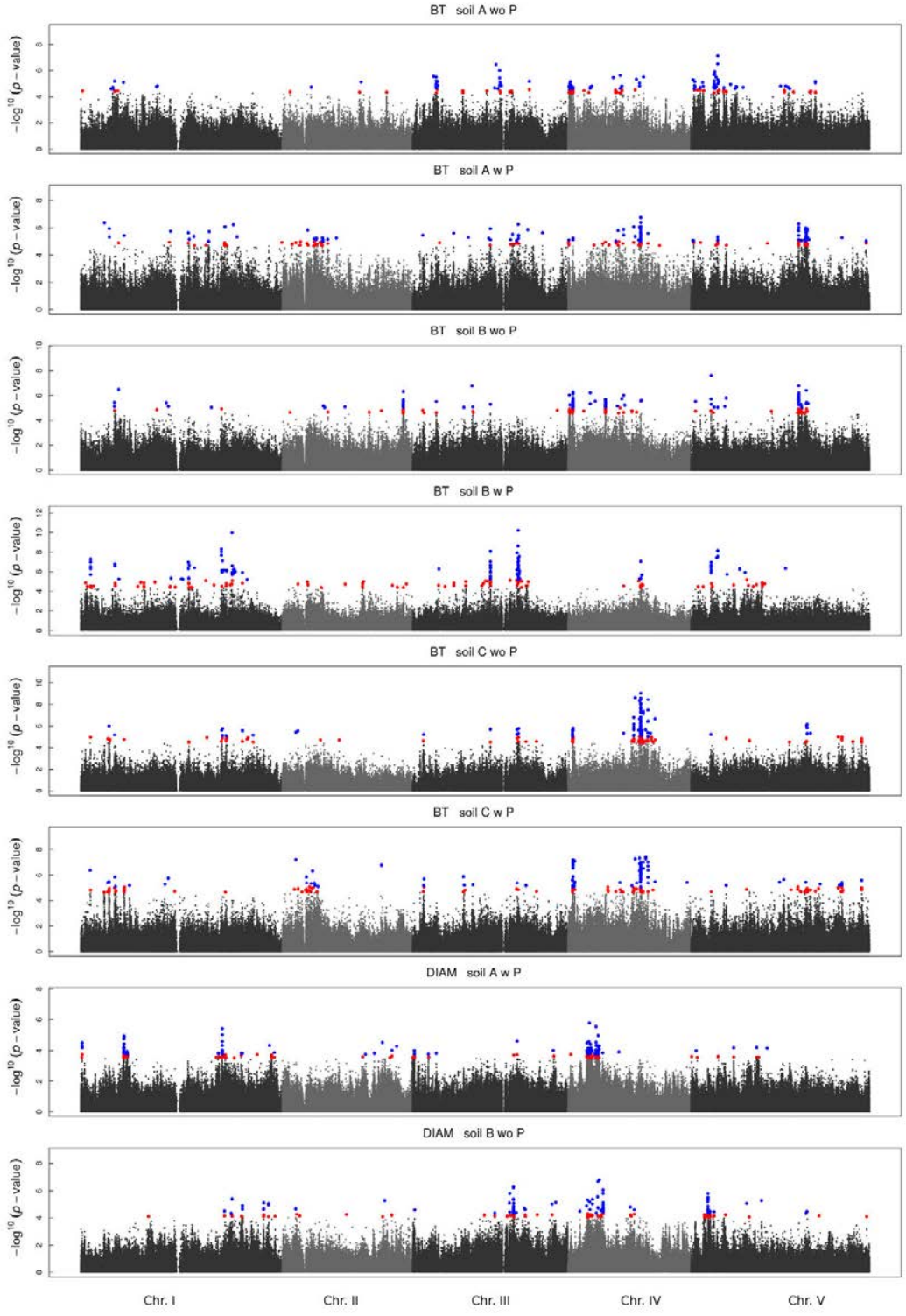

Figure S8 (continued)

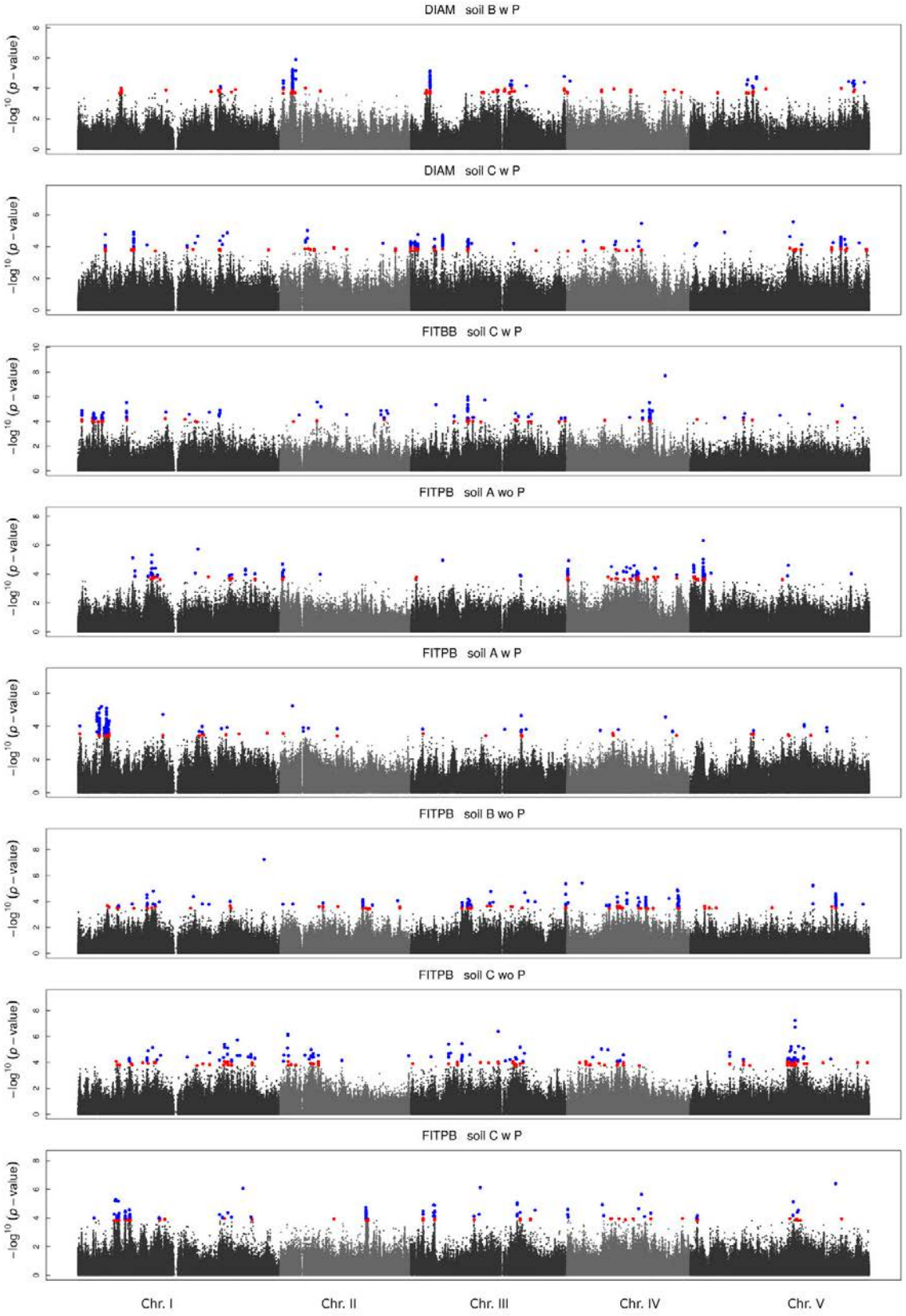

Figure S8 (continued)

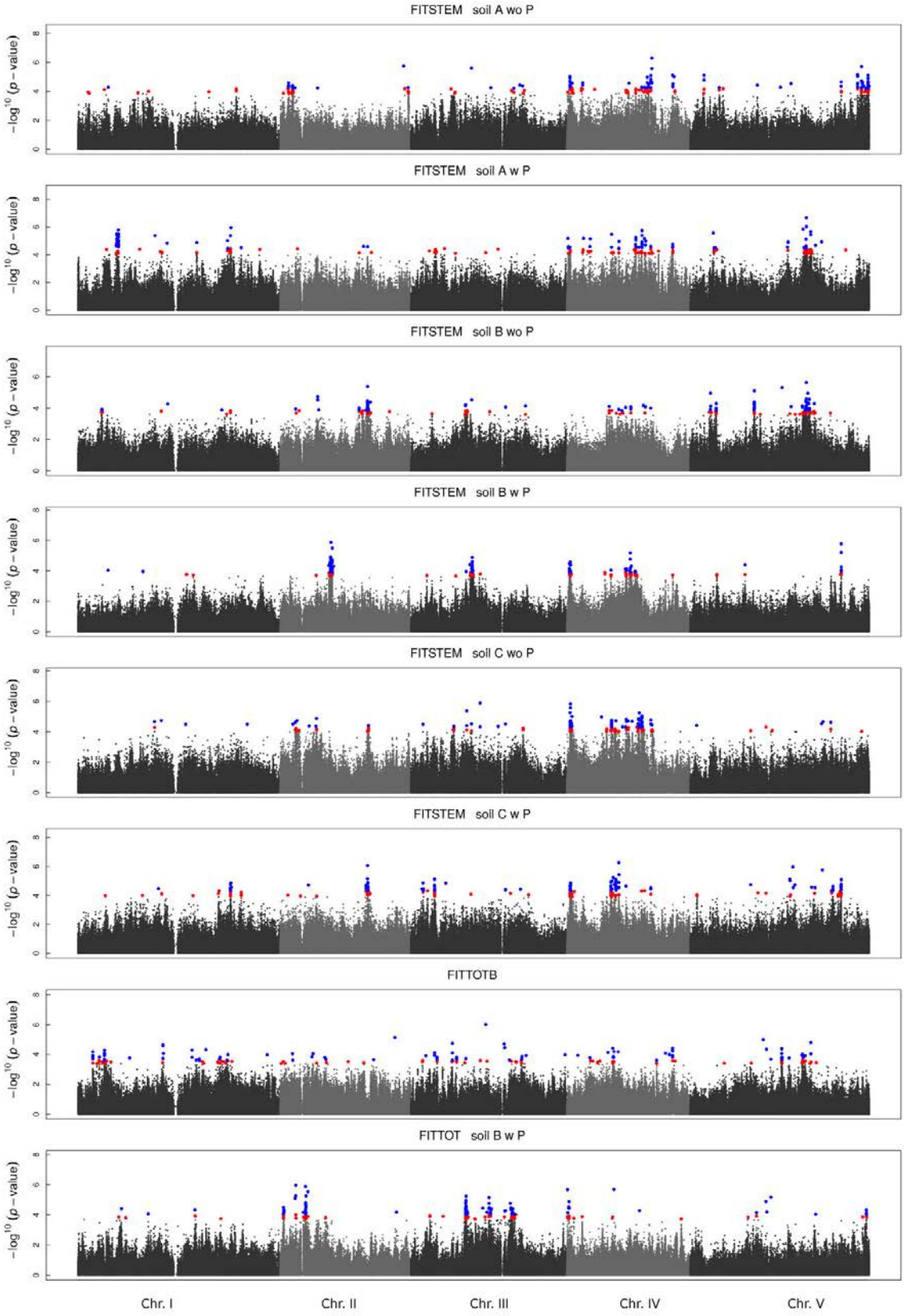

Figure S8 (continued)

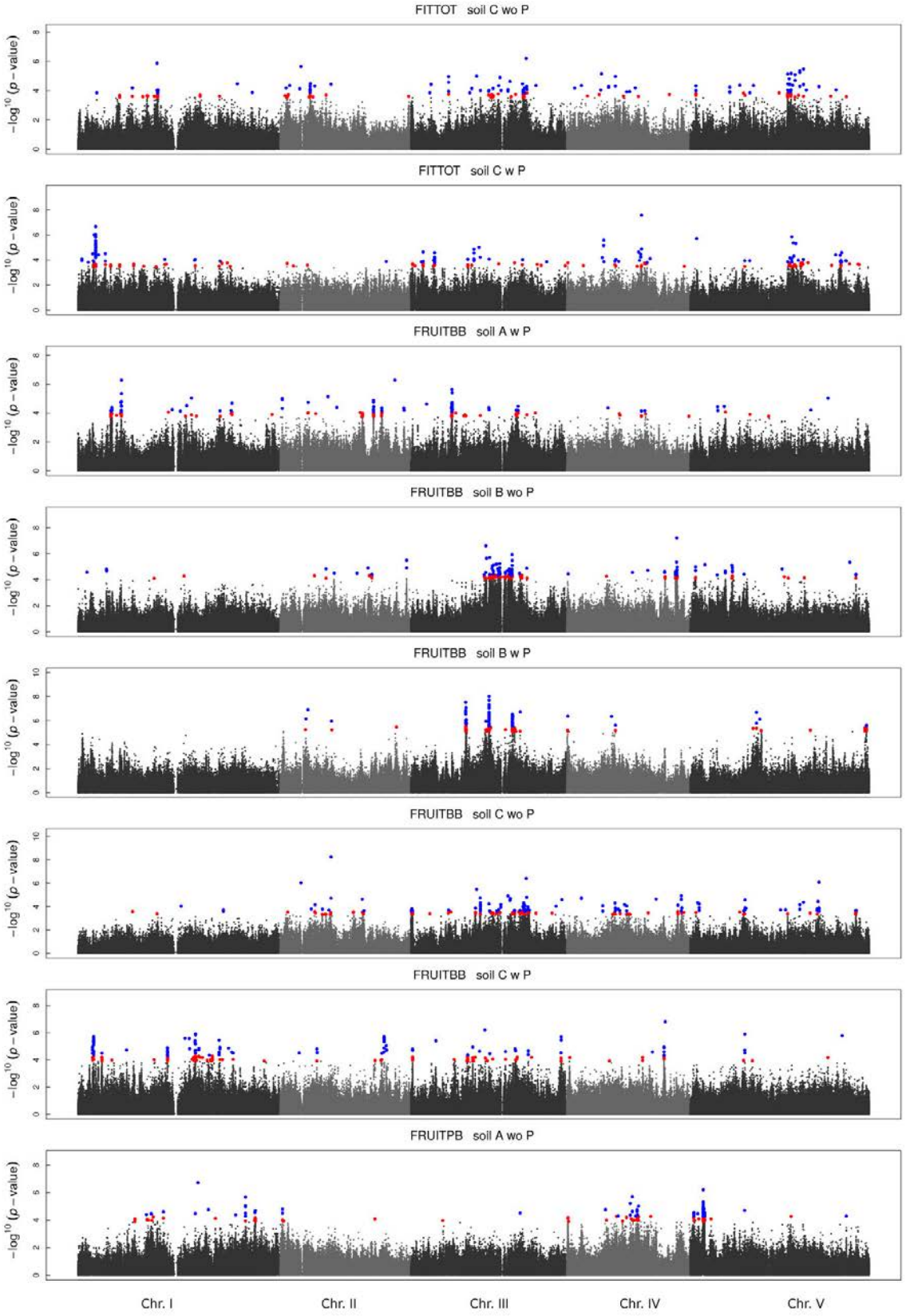

Figure S8 (continued)

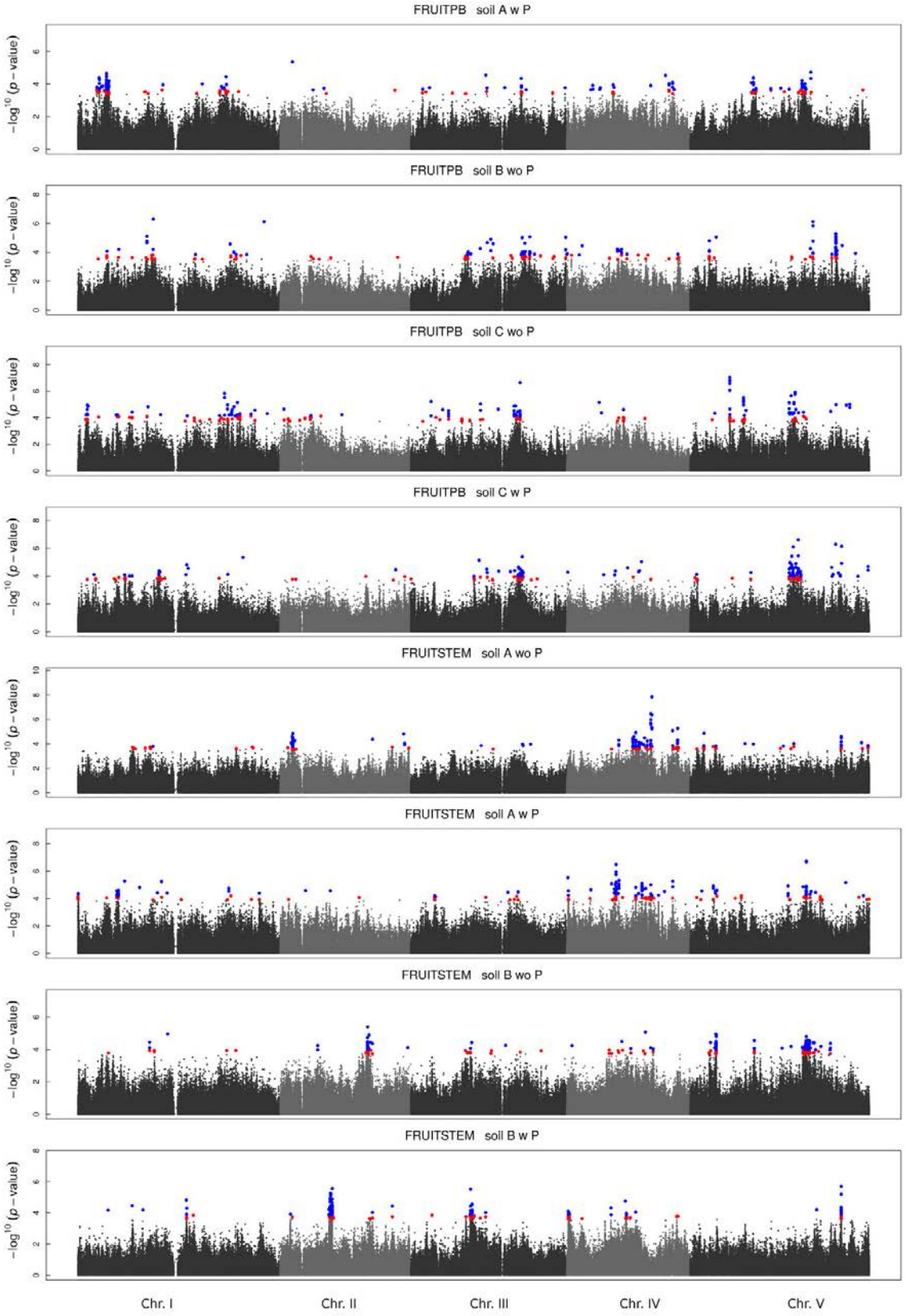

Figure S8 (continued)

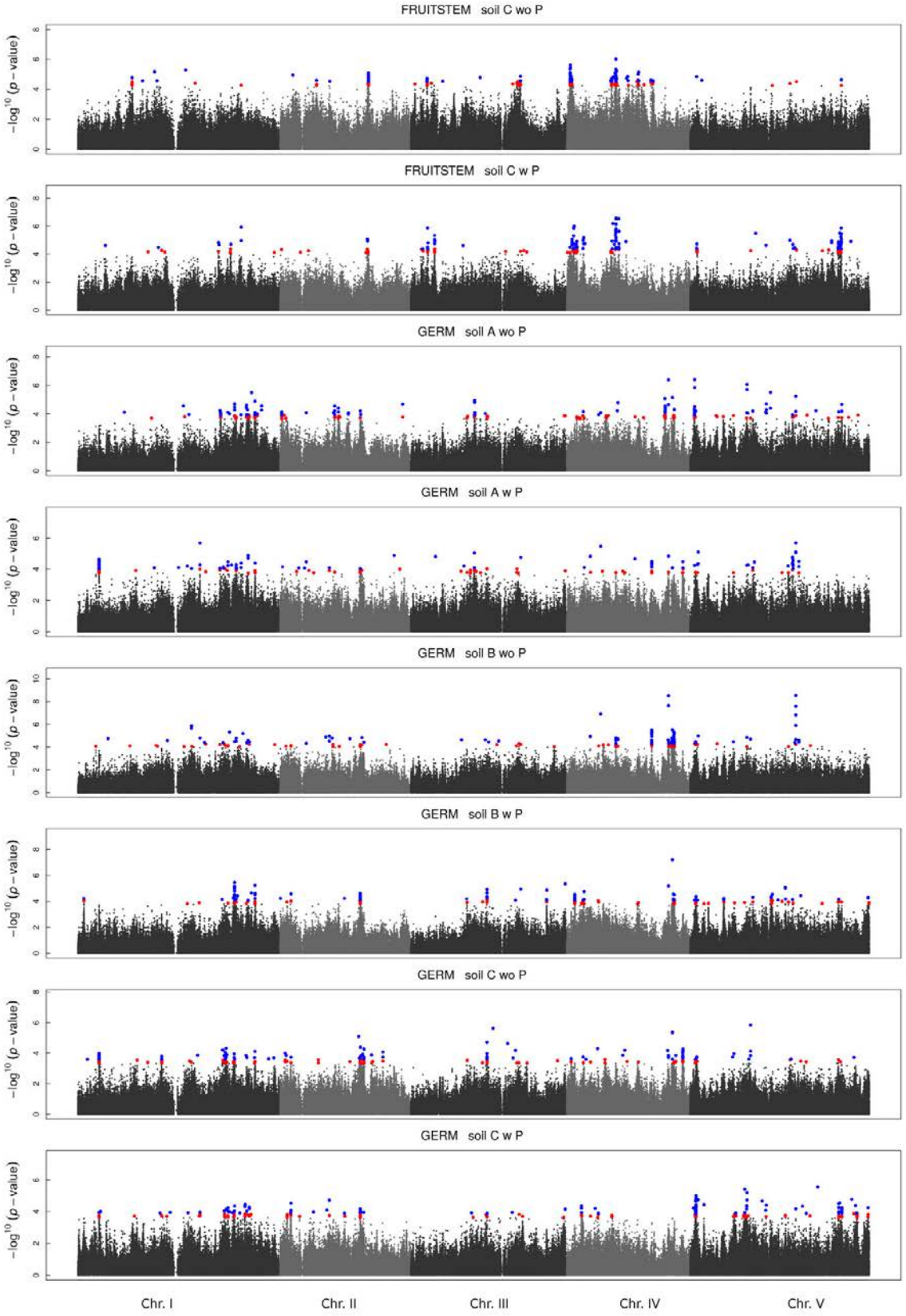

Figure S8 (continued)

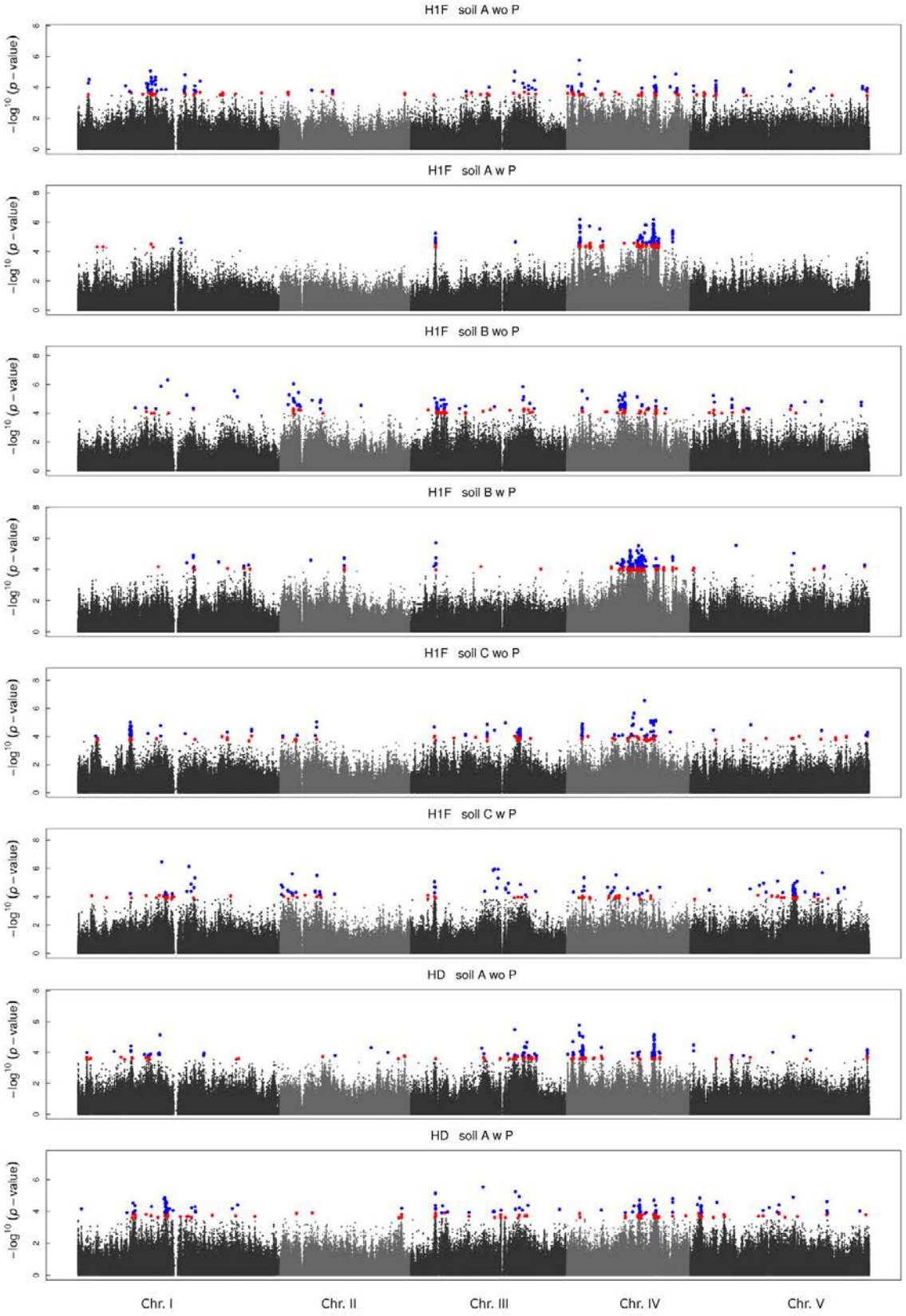

Figure S8 (continued)

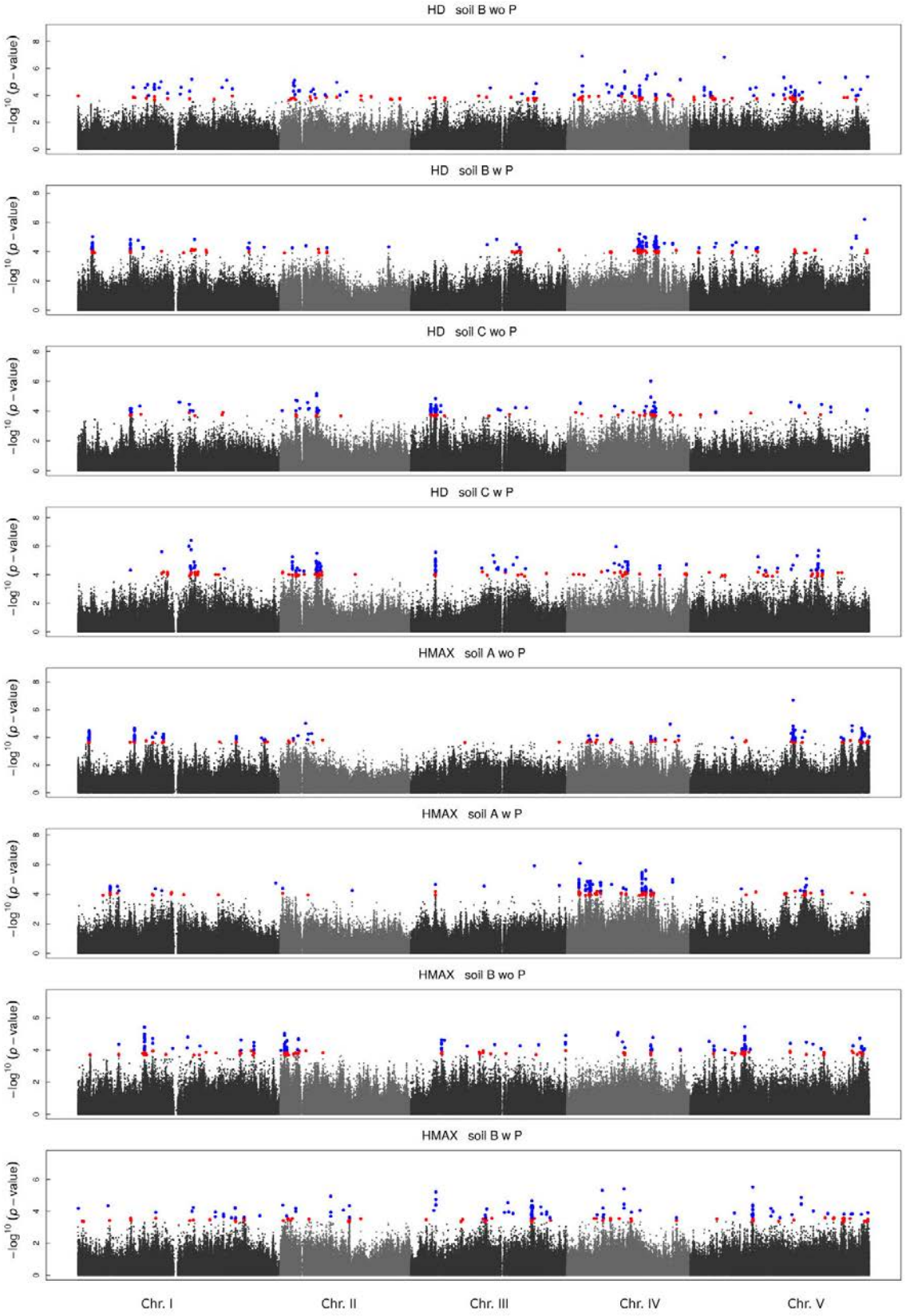

Figure S8 (continued)

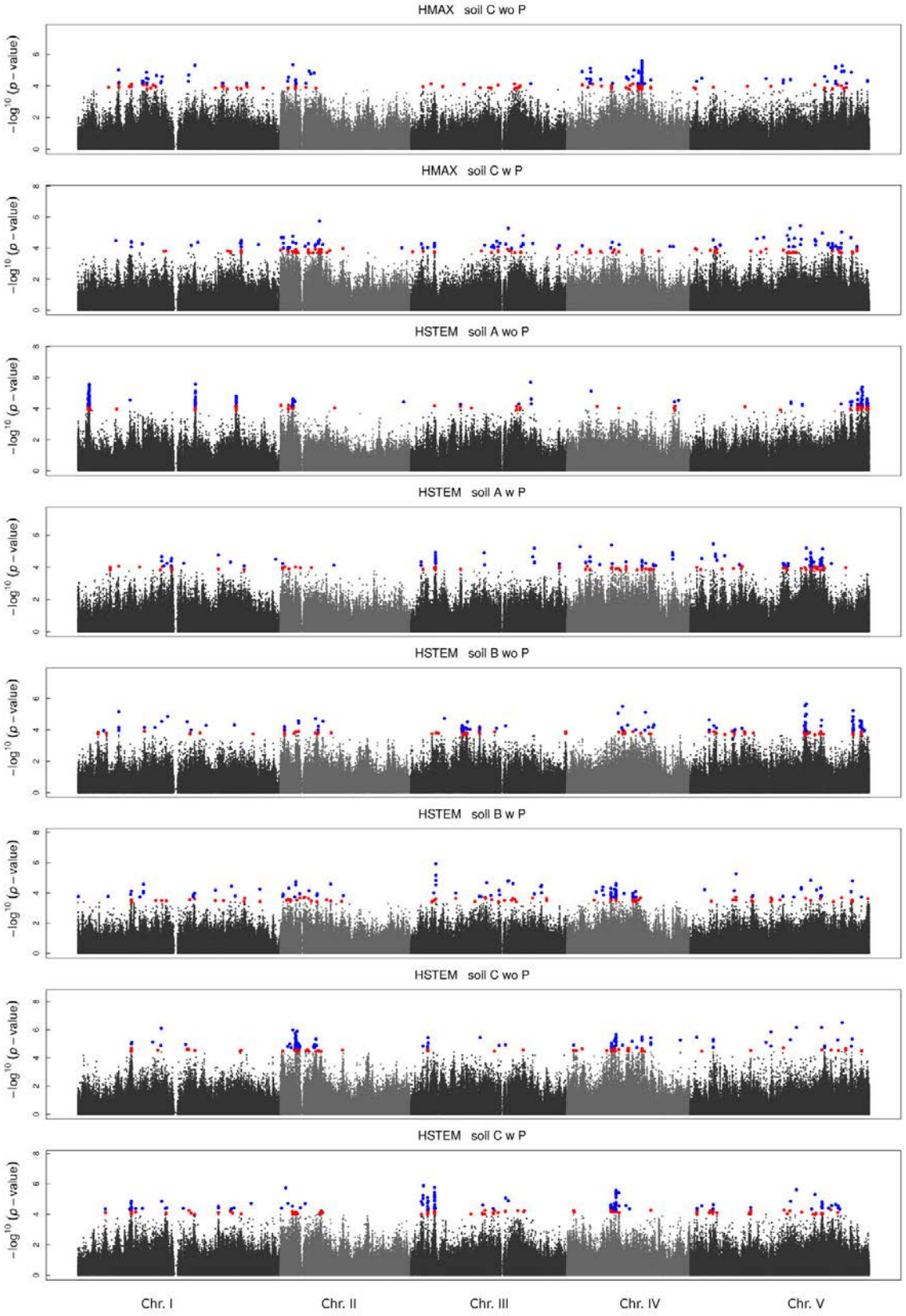

Figure S8 (continued)

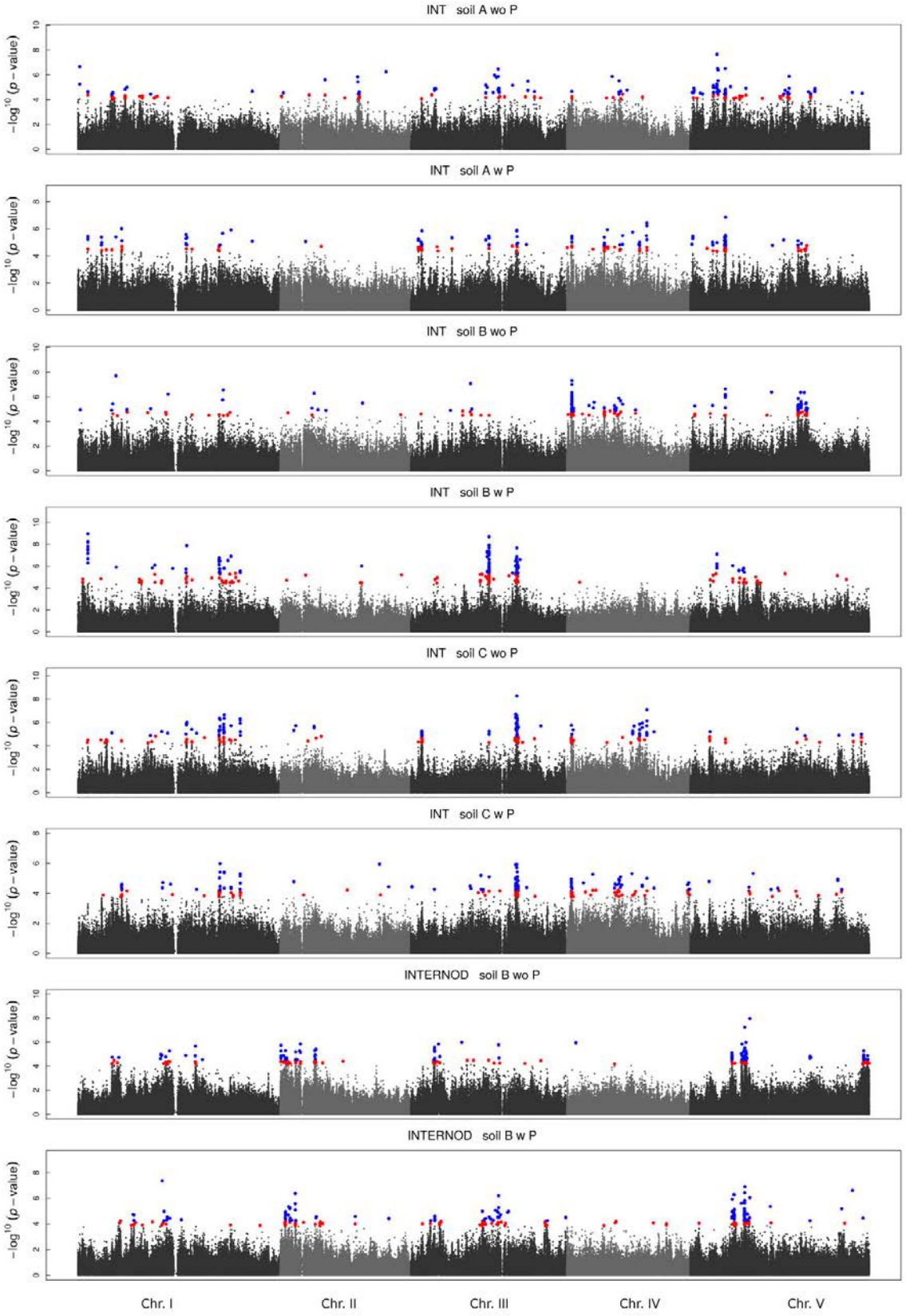

Figure S8 (continued)

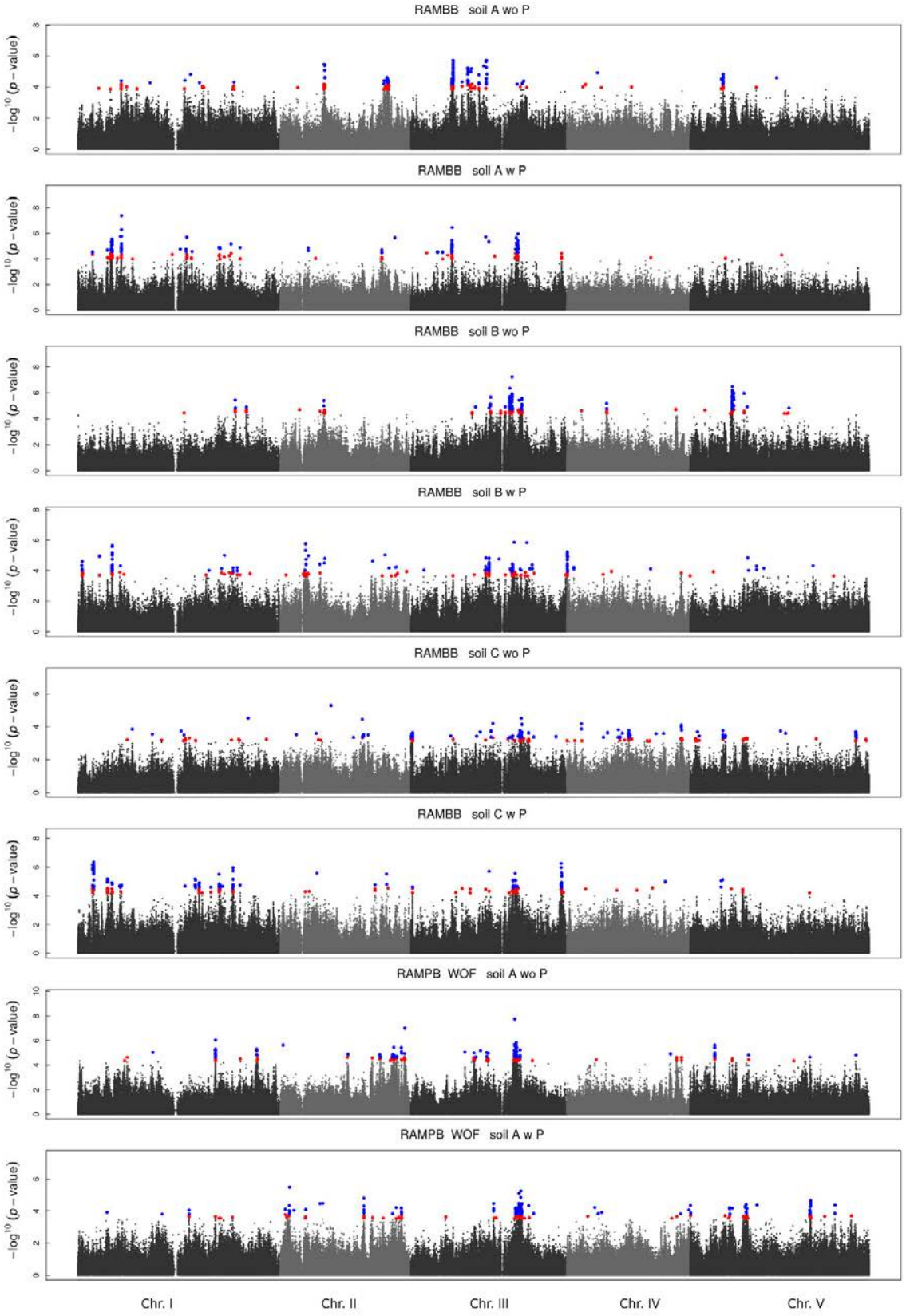

Figure S8 (continued)

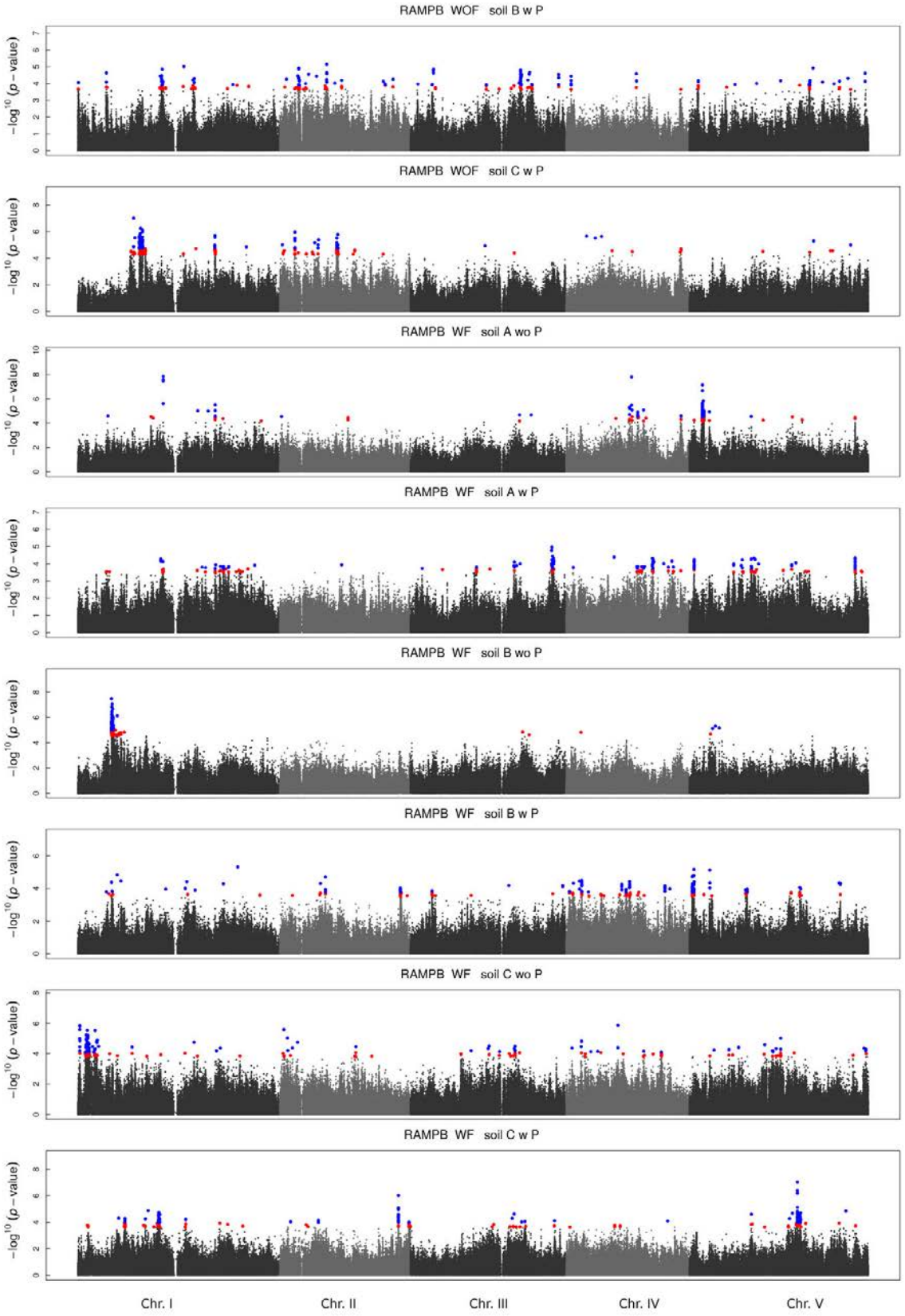

Figure S8 (continued)

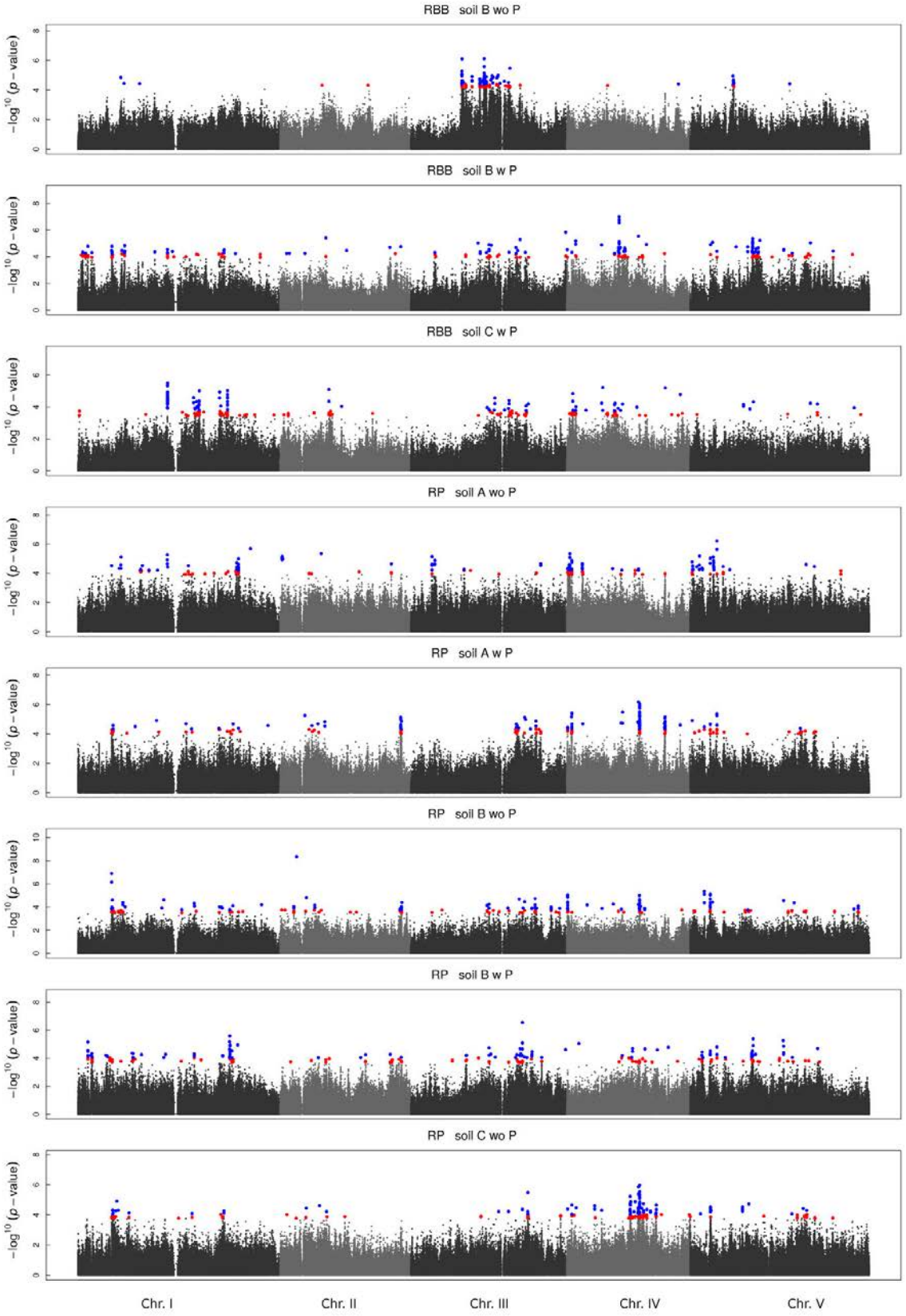

Figure S8 (continued)

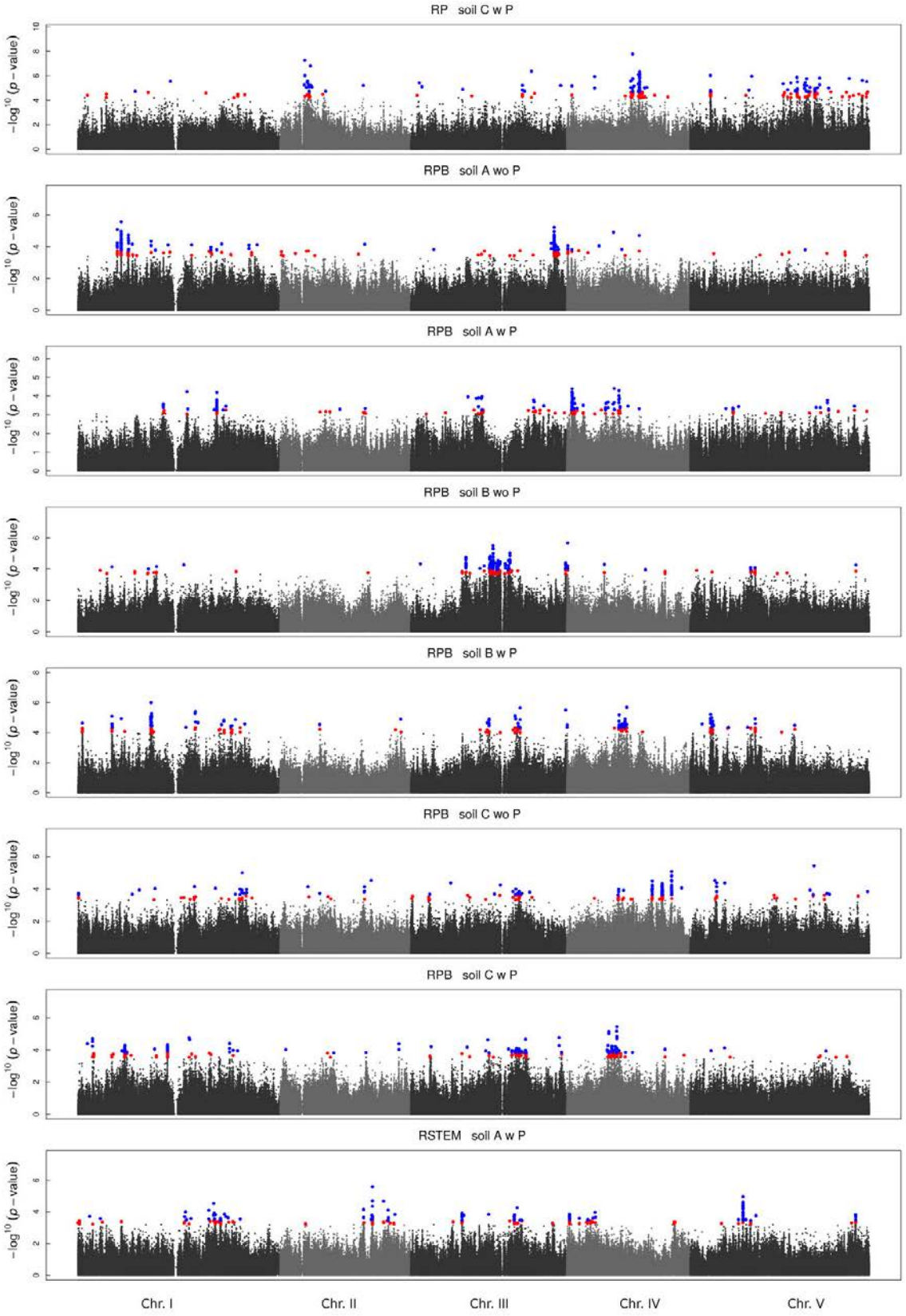

Figure S8 (continued)

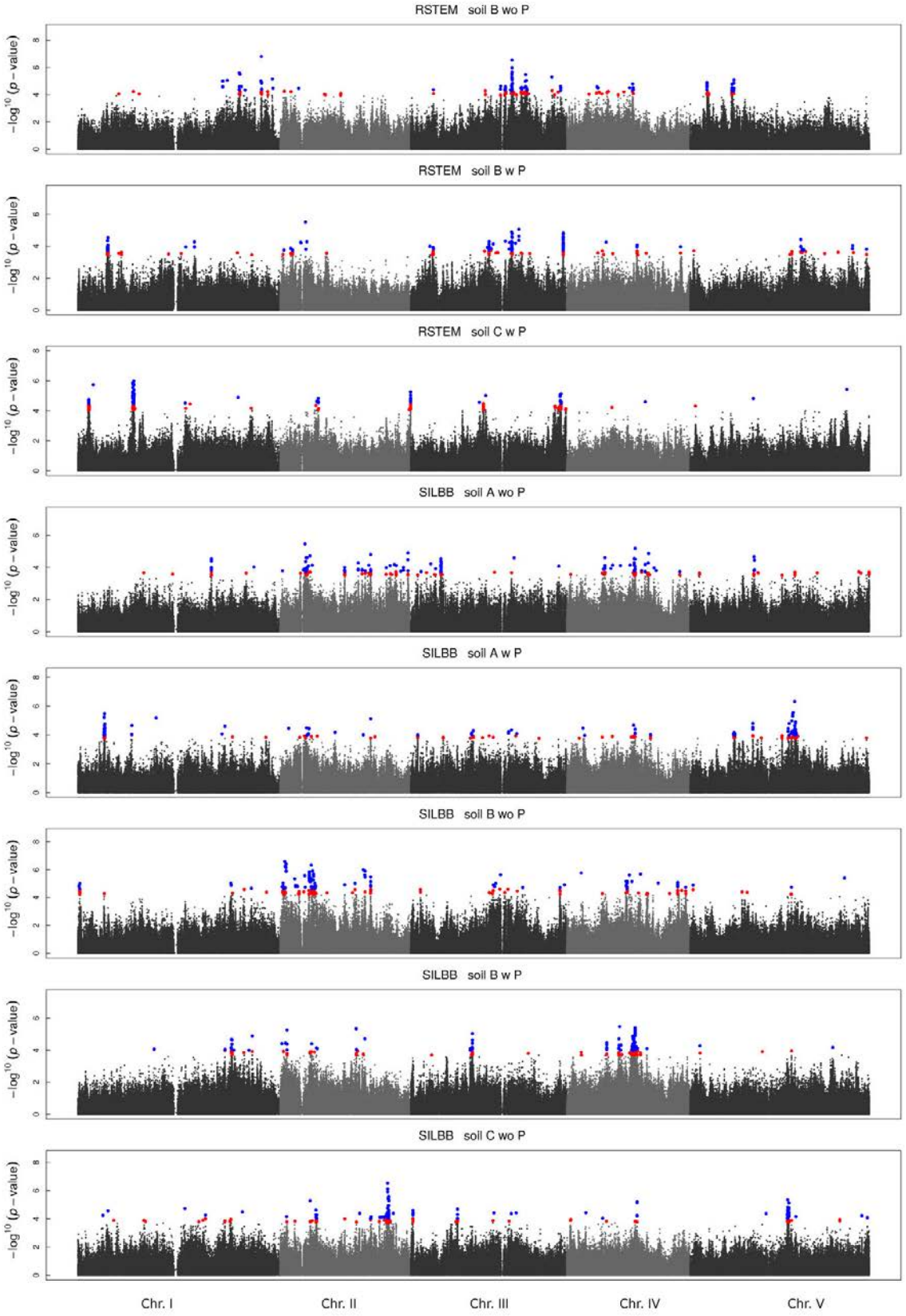

Figure S8 (continued)

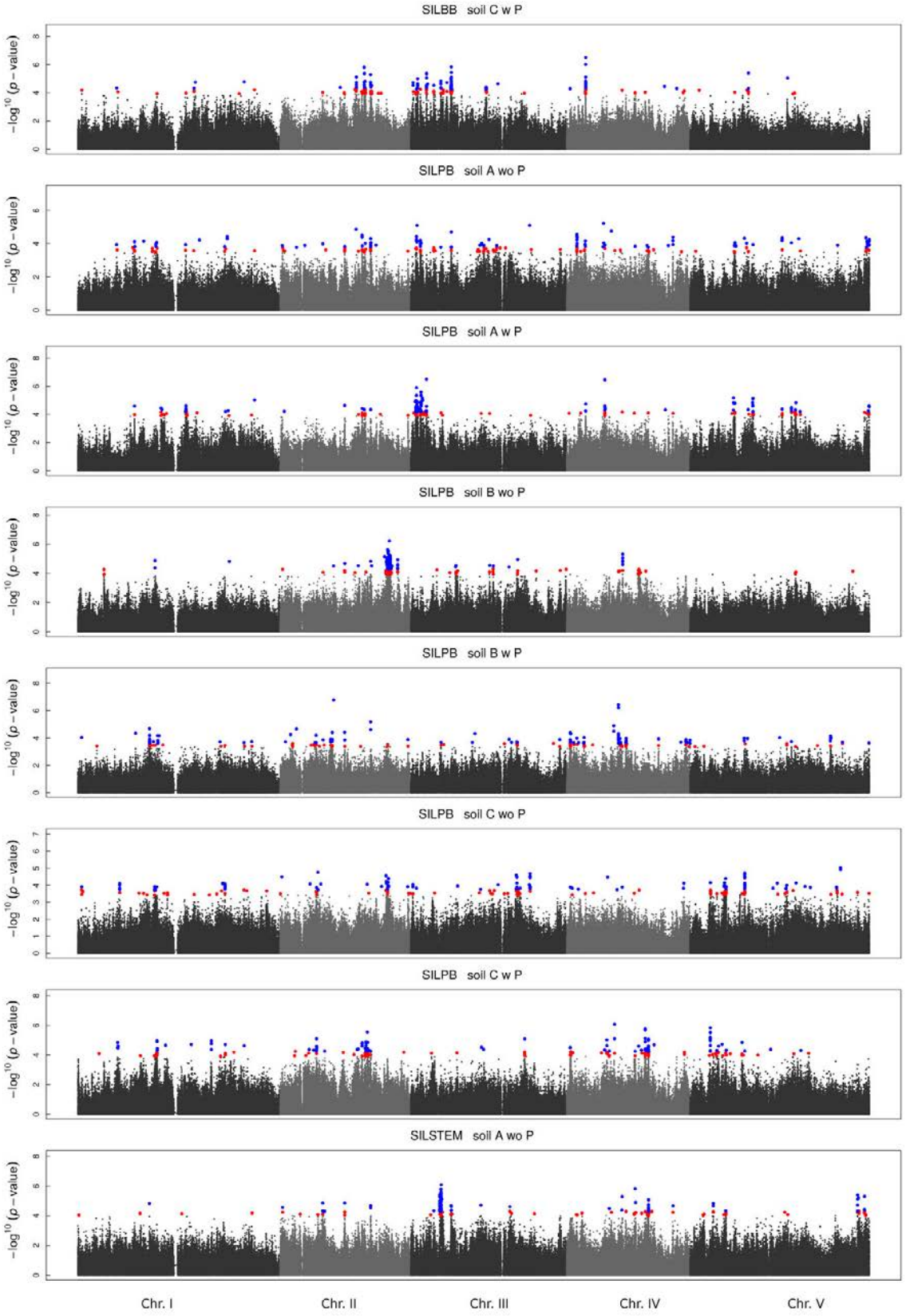

Figure S8 (continued)

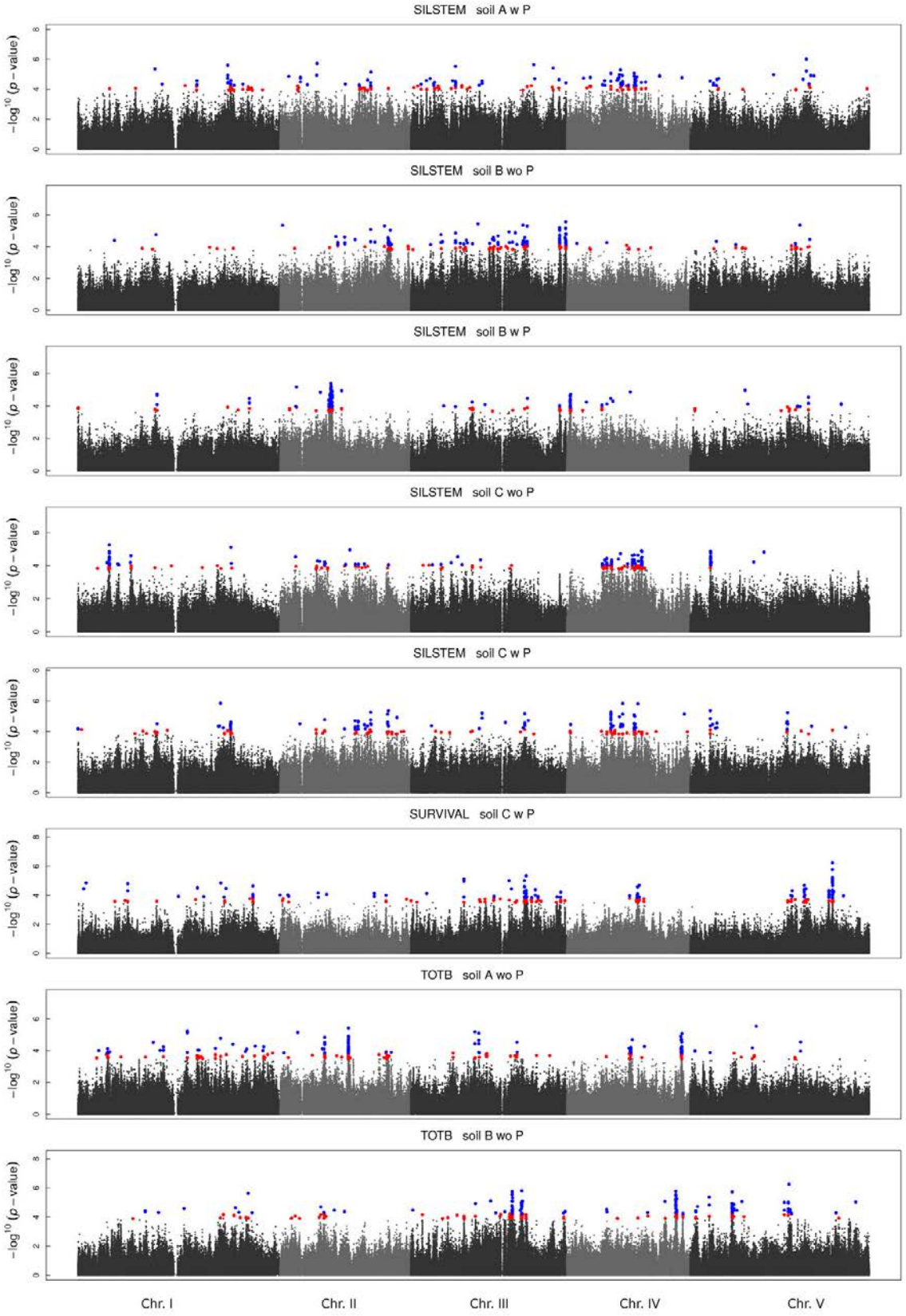

Figure S8 (continued)

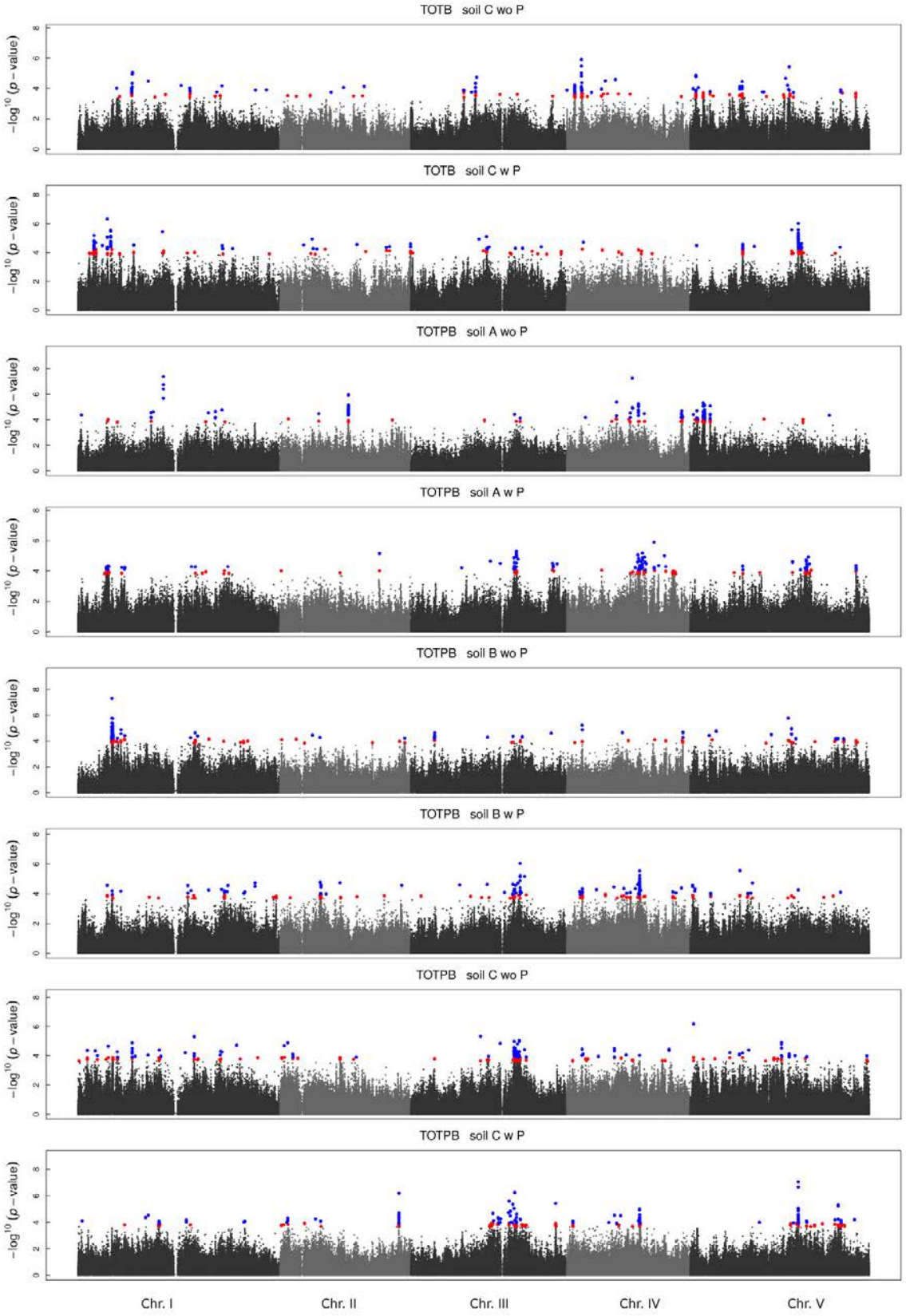

Figure S9 | Number of genes represented in the top 200 SNPs for each of the 144 heritable eco-phenotypes. Genes have been retrieved in a 1kb window size on each side of each top SNP.

Figure S10 | Non proportional Venn diagram presenting the partitioning of top SNPs associated with the 144 heritable eco-phenotypes between the six *in situ* 'soil x competition' micro-habitats. (i.e. three soils A, B and C x absence or presence of *P. annua*). Numbers in brackets indicate the number of eco-phenotypes for each *in situ* 'soil x competition' micro-habitat.

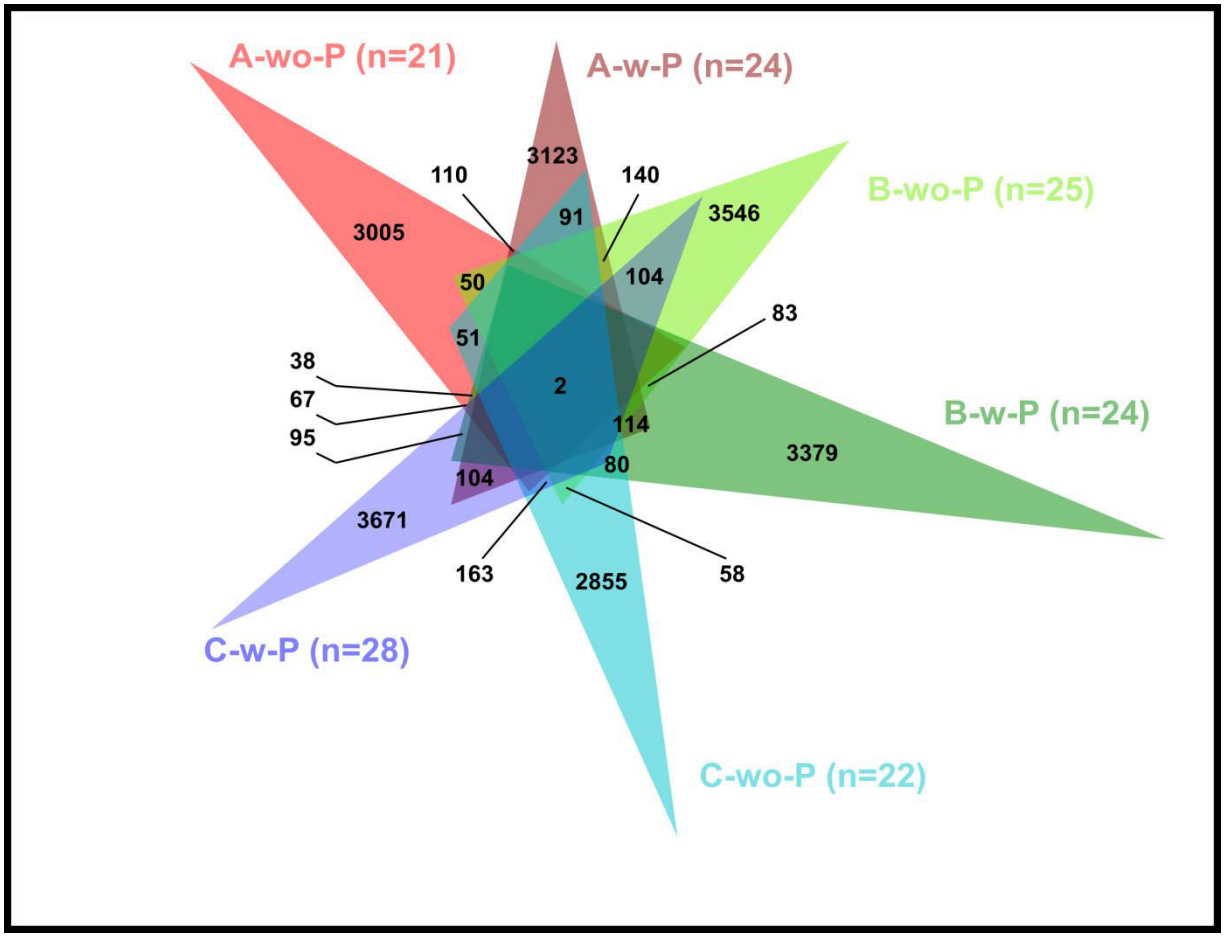

Figure S11A | Degree of pleiotropy and pleiotropic scaling in the TOU-A population when considering a threshold of 50, 100, 300 and 500 top SNPs. (**Top panels**) Frequency distribution of the effective number of eco-phenotypes affected by a SNP (N_{eff}, accounting for the correlations between eco-phenotypes) among the 21,268 unique top SNPs. (**Bottom panels**) Regression of total effect size T_M (total effect size by the Manhattan distance) on N_{eff}. The formula corresponds to the pleiotropic scaling relationship $\boldsymbol{T_M = c^* N_{eff}^{\ d}}$. A scaling component d exceeding 1 indicates that the mean per-trait effect size of a given top SNP increased with N_{eff}[8]. Solid red line: fitted relationship between T_M and N_{eff}, solid black line: linear dependence ($d = 1$).

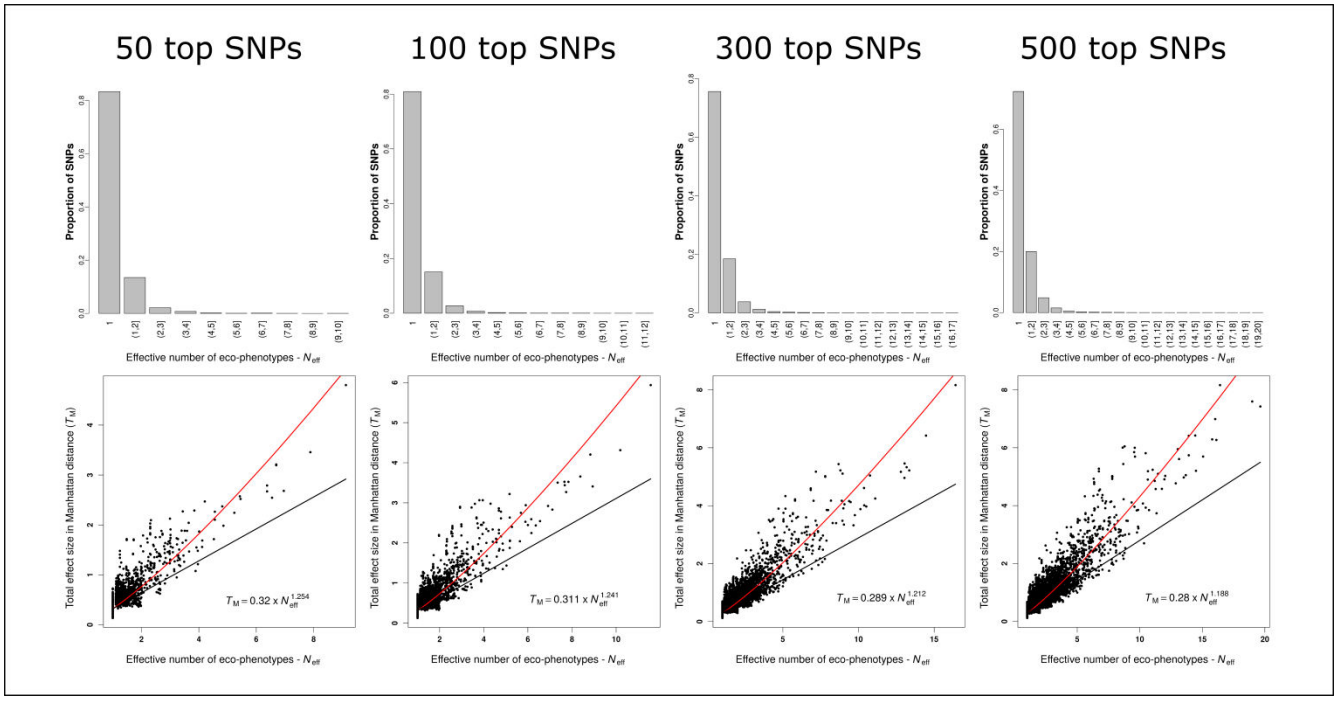

Figure S11B | Significance and strength of selection in the TOU-A population when considering a threshold of 50, 100, 300 and 500 top SNPs. (**Top panels**) Fold-increase in median $-\log_{10}$ (p-values) of neutrality tests based on temporal differentiation for SNPs that hit only evolved eco-phenotypes, only unevolved eco-phenotypes or both types of eco-phenotypes, according to different classes of effective number of eco-phenotypes. The dashed line corresponds to a fold-increase of 1, i.e. no increase in median significance of neutrality tests based on temporal differentiation. (**Bottom panels**) Fold-increase in median F_{ST} values for SNPs that hit only evolved eco-phenotypes, only unevolved eco-phenotypes or both types of eco-phenotypes, according to different classes of N_{eff} (median F_{ST} across the genome = 0.00293). Significance against a null distribution obtained by bootstrapping: *$0.05 > P > 0.01$, **$0.01 > P > 0.001$, ***$P < 0.001$, ns: non-significant.

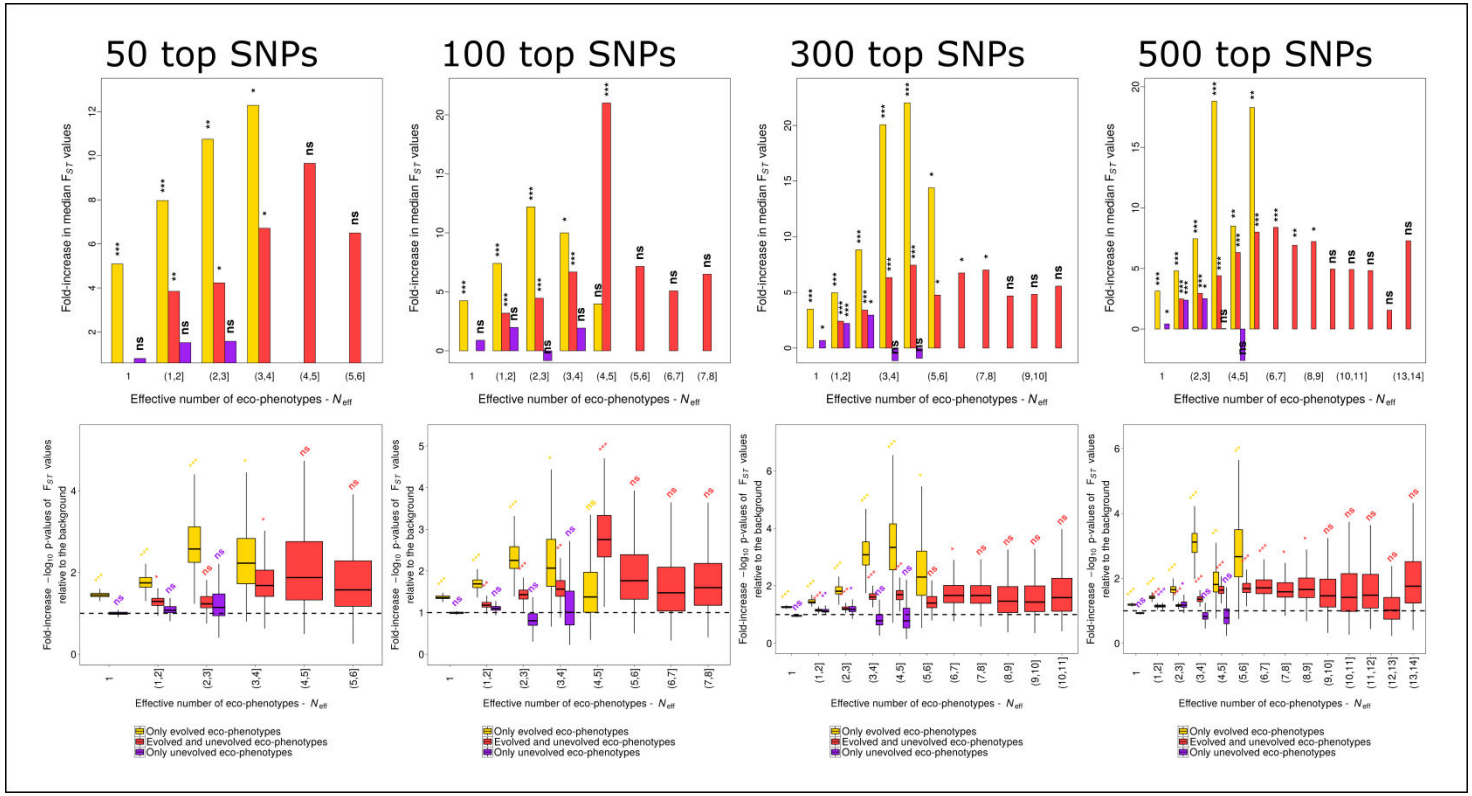

Figure S11C | Degree of pleiotropy and pleiotropic scaling in the TOU-A population when considering SNPs with a $-\log_{10}$ *p*-value above 6, 5 and 4. (**Top panels**) Frequency distribution of the effective number of eco-phenotypes affected by a SNP (N_{eff}, accounting for the correlations between eco-phenotypes) among the 21,268 unique top SNPs. (**Bottom panels**) Regression of total effect size T_M (total effect size by the Manhattan distance) on N_{eff}. The formula corresponds to the pleiotropic scaling relationship $\boldsymbol{T_M = c^* N_{eff}^{\ d}}$. A scaling component d exceeding 1 indicates that the mean per-trait effect size of a given top SNP increased with N_{eff}[8]. Solid red line: fitted relationship between T_M and N_{eff}, solid black line: linear dependence ($d = 1$).

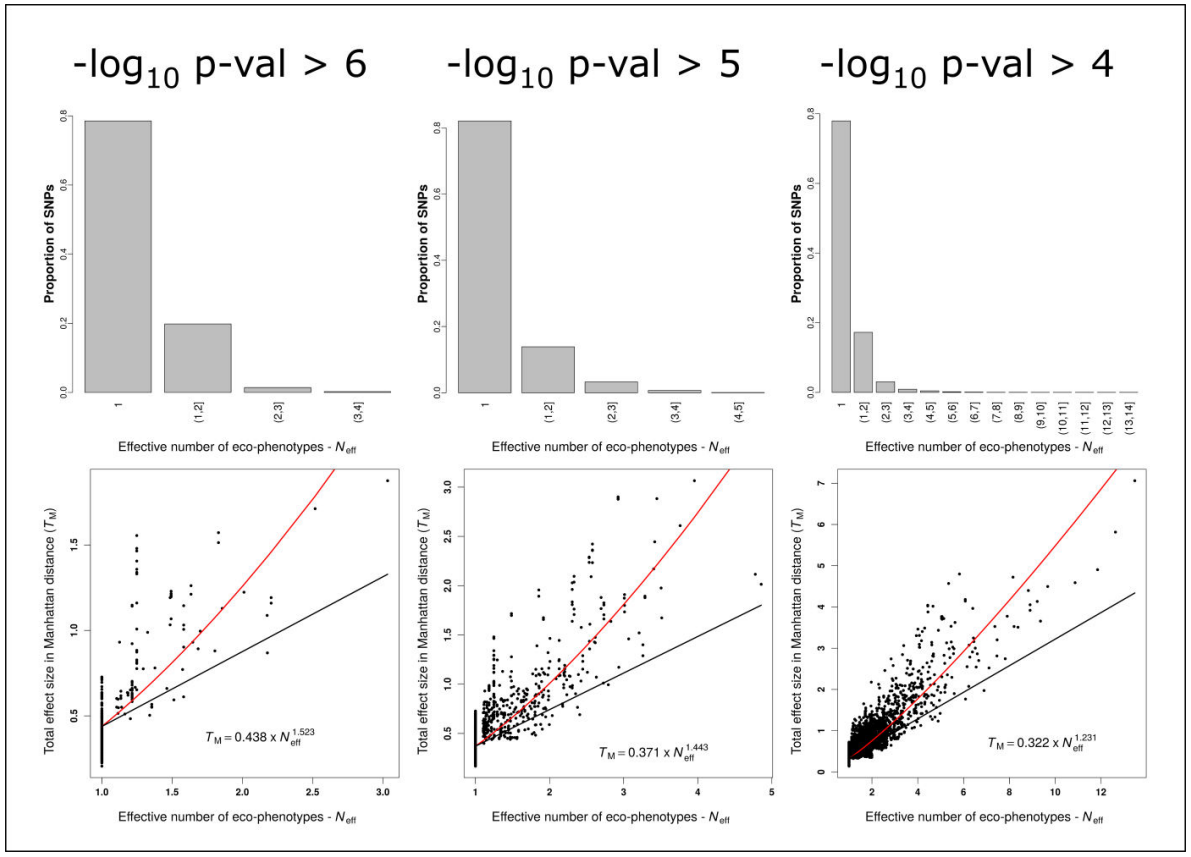

Figure S11D | Significance and strength of selection in the TOU-A population when SNPs with a $-\log_{10}$ *p*-value above 6, 5 and 4. (**Top panels**) Fold-increase in median $-\log_{10}$ (*p*-values) of neutrality tests based on temporal differentiation for SNPs that hit only evolved eco-phenotypes, only unevolved eco-phenotypes or both types of eco-phenotypes, according to different classes of effective number of eco-phenotypes. The dashed line corresponds to a fold-increase of 1, i.e. no increase in median significance of neutrality tests based on temporal differentiation. (**Bottom panels**) Fold-increase in median F_{ST} values for SNPs that hit only evolved eco-phenotypes, only unevolved eco-phenotypes or both types of eco-phenotypes, according to different classes of N_{eff} (median F_{ST} across the genome = 0.00293). Significance against a null distribution obtained by bootstrapping: *0.05 > *P* > 0.01, **0.01 > *P* > 0.001, ***$P < 0.001$, ns: non-significant.

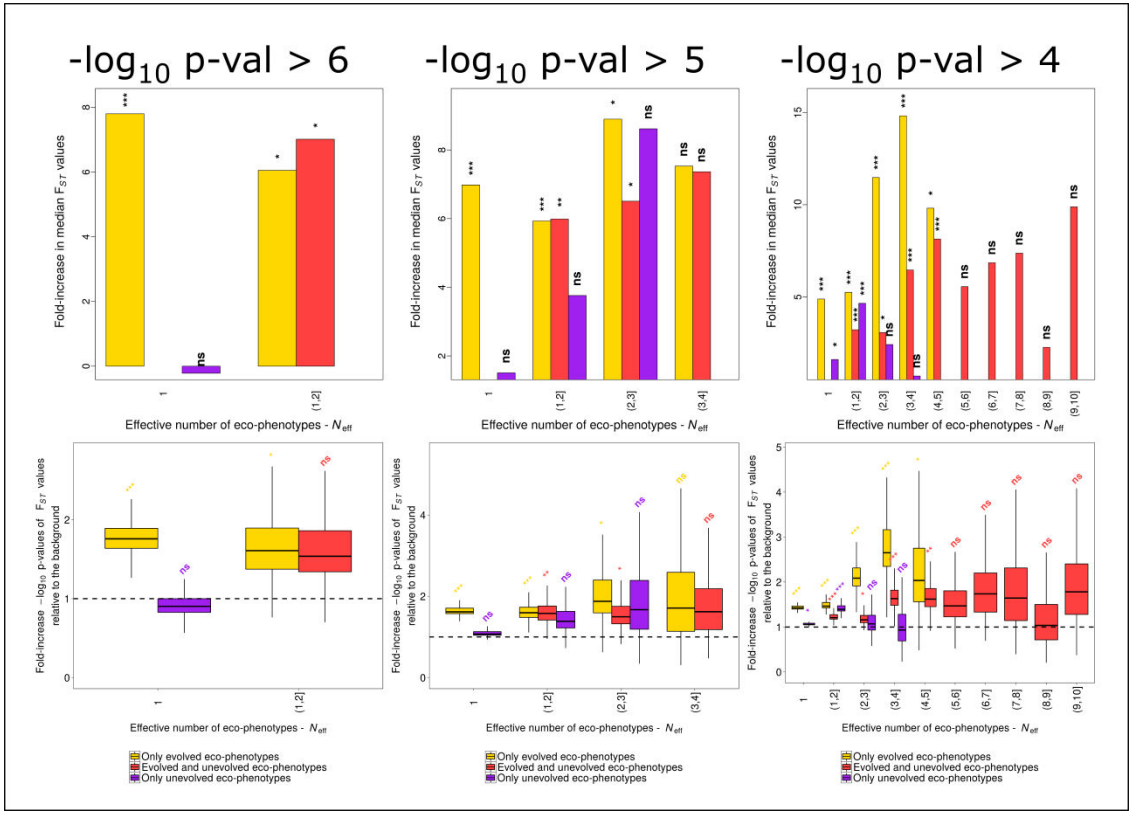

Figure S12 | Scaling relationships between total phenotypic effect size of the 200 top SNPs and the number of eco-phenotypes (*N*, left panels) or the effective number of eco-phenotypes (N_{eff}, right panels). The pleiotropic scaling relationship was calculated as (i) $T_M = c^* {N_{eff}}^d$, with T_M corresponding to the Manhattan distance (bottom panels) and (ii) $T_E = a^* Neff^b$, with T_E corresponding to the Euclidean distance (top panels).

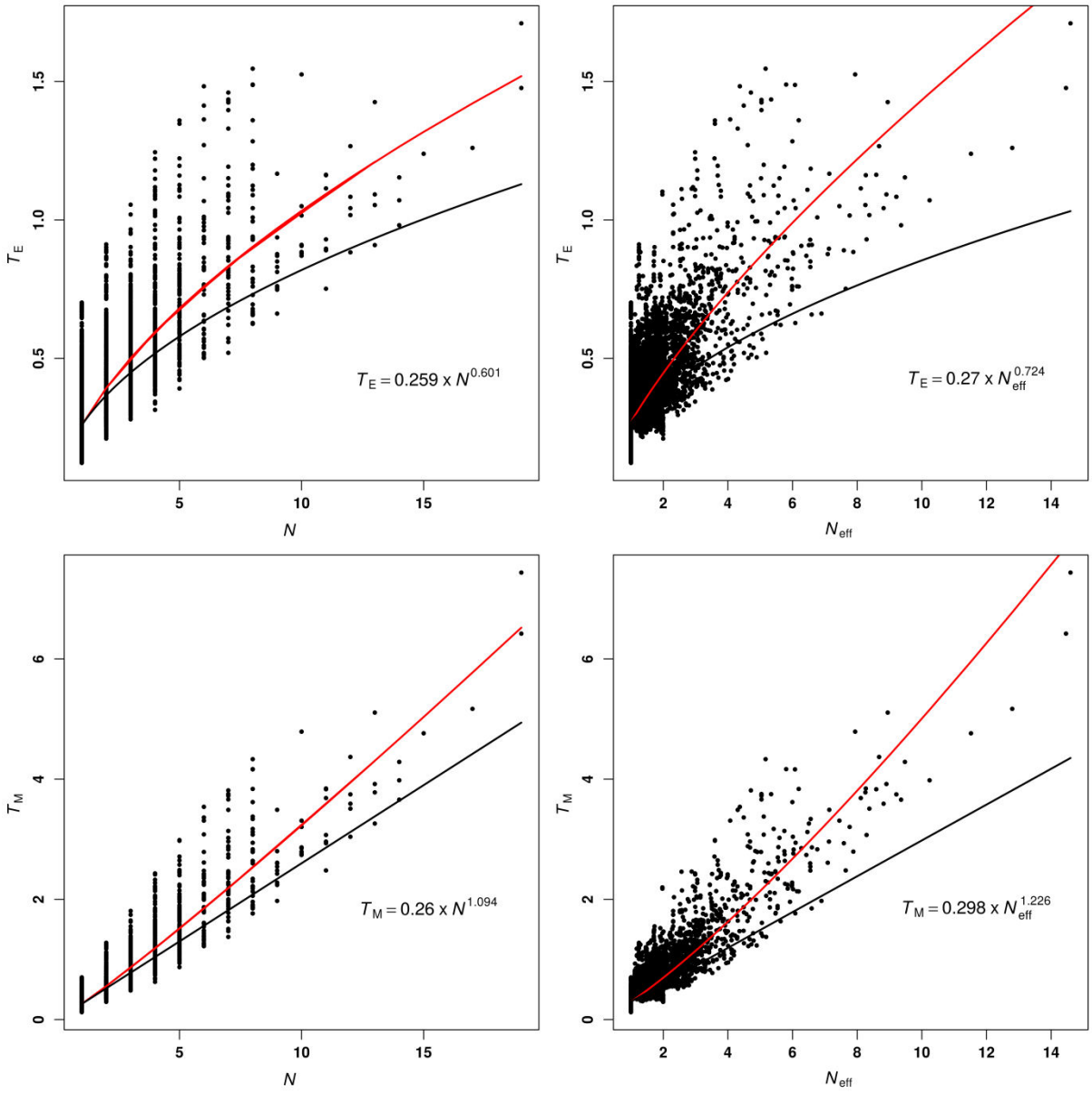

Figure S13 | Genome-wide scan for selection based on temporal differentiation. (**a**) Manhattan plot of F_{ST} at each SNP marker (dots) along the *A. thaliana* genome. The blue dashed line corresponds to the 0.1% upper tail of the F_{ST} value distribution (n = 982). Median F_{ST} across the genome = 0.00293. (**b**) -$\log_{10}$ (*p*-value) of the simulation-based test of the null hypothesis that the locus-specific differentiation measured at each SNP is only due to genetic drift. Only SNP markers with MARF > 7% are considered.

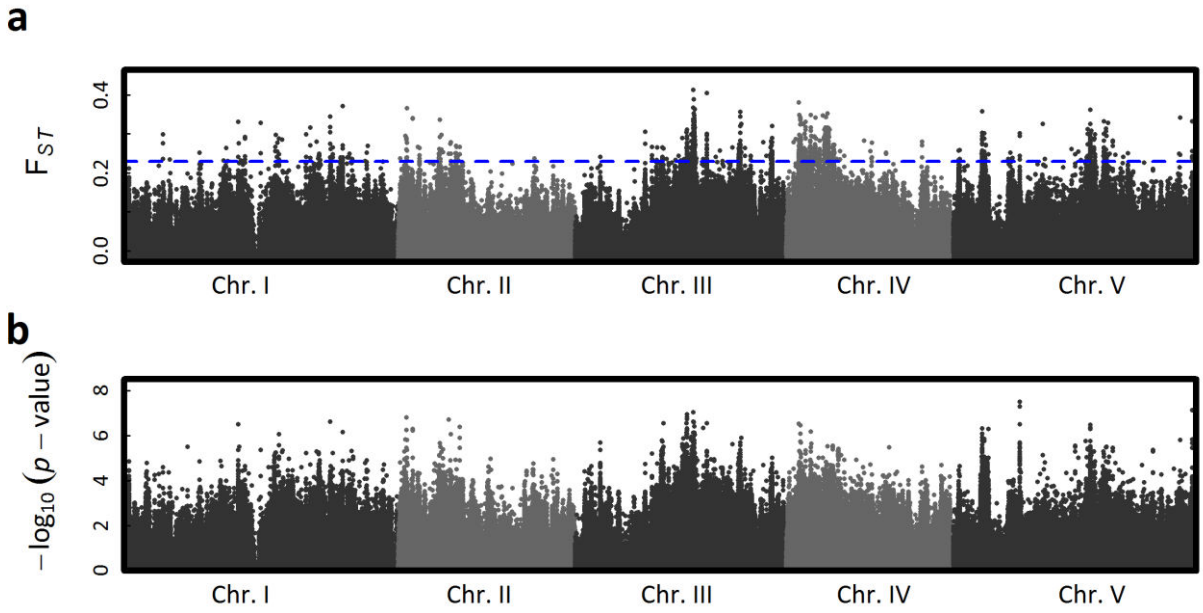

Figure S14 | Polarity of effects. (**a**) Proportion of top SNPs associated with evolved eco-phenotypes with a polarity of effects in line with the direction of phenotypic evolution, according to different classes of N_{eff}. (**b**) Effect of polarity effects on the fold-increase in median F_{ST} values for SNPs that hit only evolved eco-phenotypes, according to different classes of N_{eff} (median F_{ST} across the genome = 0.00293). Significance against a null distribution obtained by bootstrapping: *0.05 > P > 0.01, **0.01 > P > 0.001, ***P < 0.001, absence of symbols: non-significant. Due to the small number of SNPs with an effective number of eco-phenotypes above 4, those SNPs were grouped for testing the significance of fold-increase in median F_{ST} values.

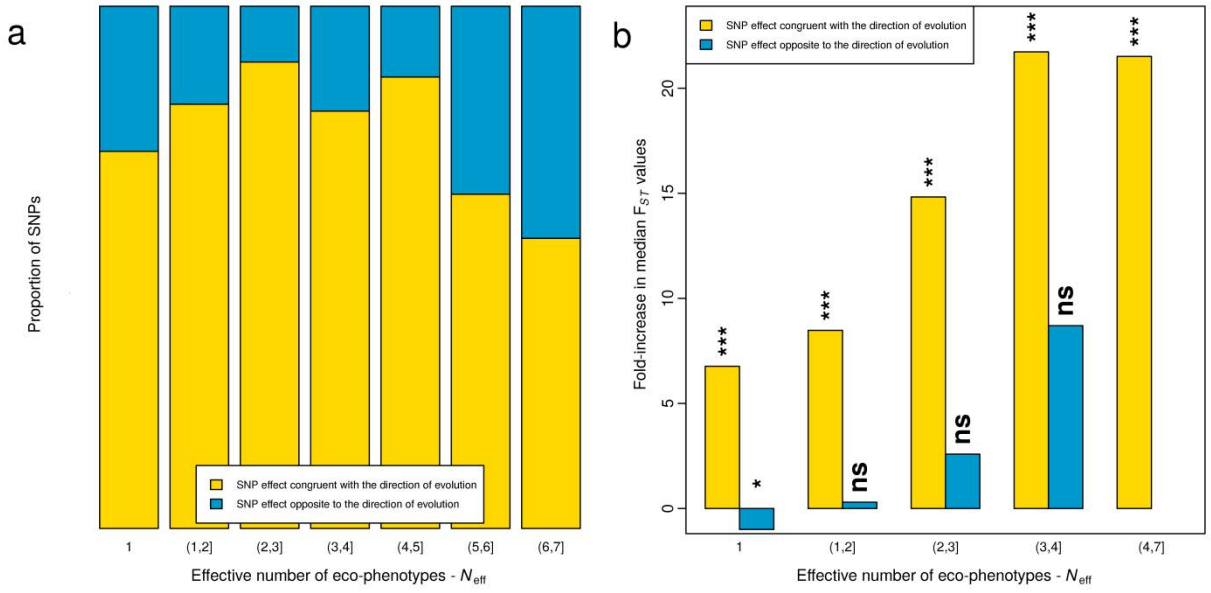

Figure S15 | The distribution dependence of *p*-value distribution on minor allele relative frequency (MARF) for EMMAX across the 144 eco-phenotypes (see Fig. 1). For a given MARF value, each point corresponds to the quantile at 99% of the *p*-value distribution of one of the 144 heritable eco-phenotypes. A locally-weighted polynomial regression is illustrated by a red solid line. A MARF threshold above 7% is depicted by a dashed blue line.

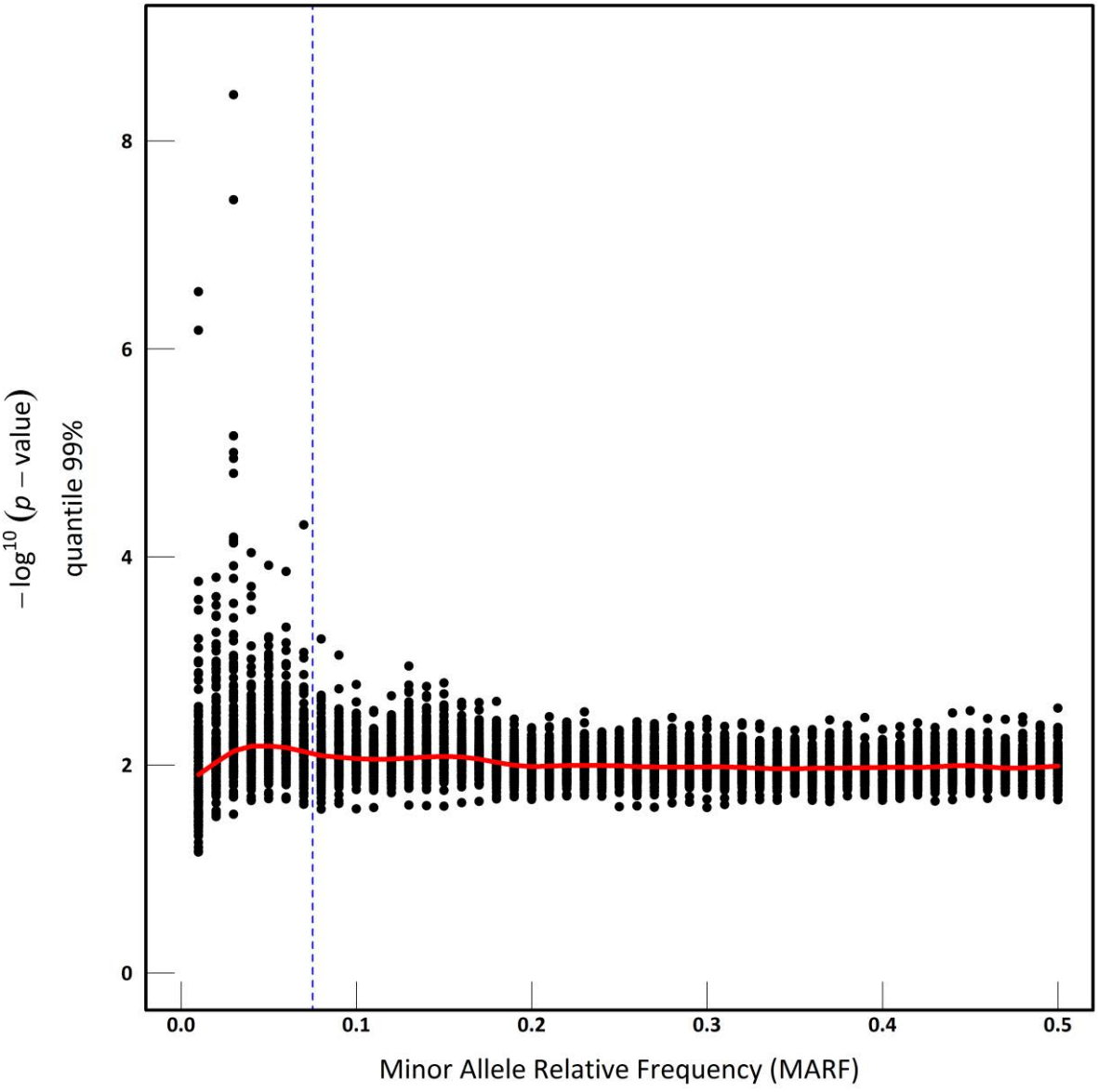

Table S1 | Phenotypic variation of 195 accessions sampled in 2002 and 2010 and scored across six *in situ* 'soil x competition' micro-habitats.

	Model terms§																					
	block (soil*comp)		soil		comp		soil*comp		year		soil*year		comp*year		soil*comp* year		acc (year)		acc(year)* soil		acc(year)* comp		acc(year)* soil*comp	
Traits ǂ	*F*	*P*	*F*	*P*	*F*	*P*	*F*	*P*	*F*	*P*	*F*	*P*	*F*	*P*	*F*	*P*	*LRT*	*P*	*LRT*	*P*	*LRT*	*P*	*LRT*	*P*
GERM	5.40	***	53.44	***	104.47	***	64.64	***	8.61	**	10.13	***	0.51	ns	0.08	ns	299.1	***	0.0	ns	0.0	ns	7.9	*
BT	3.90	***	23.57	***	120.85	***	25.07	***	13.46	**	11.13	***	22.40	***	4.54	***ns***	280.7	***	5.0	*	13.5	**	5.8	***ns***
INT	1.65	*	29.20	***	66.50	***	42.51	***	13.22	**	7.93	**	16.85	***	2.39	ns	140.4	***	0.6	ns	16.1	**	4.1	***ns***
RP	8.24	***	132.02	***	45.37	***	20.95	***	19.18	***	3.65	***ns***	12.60	**	1.62	ns	287.4	***	17.3	***	2.7	ns	0.0	ns
DIAM	5.28	***	75.04	***	57.82	***	46.57	***	0.16	ns	5.34	*	0.11	ns	0.12	ns	40.1	***	2.8	ns	0.0	ns	0.2	ns
H1F	2.41	***	177.58	***	31.60	***	86.96	***	7.29	*	1.41	ns	0.64	ns	2.37	ns	125.5	***	10.7	**	0.8	ns	1.3	ns
HSTEM	4.01	***	342.68	***	14.99	***	55.05	***	0.71	ns	0.08	ns	1.51	ns	2.41	ns	178.0	***	49.9	***	0.2	ns	0.0	ns
HMAX	6.63	***	584.30	***	4.24	***ns***	84.33	***	0.39	ns	0.88	ns	0.17	ns	2.72	ns	162.5	***	43.4	***	0.2	ns	0.0	ns
HD	1.82	*	77.84	***	99.75	***	27.52	***	8.96	**	2.19	ns	2.90	ns	0.89	ns	175.8	***	0.0	ns	0.0	ns	8.9	*
RAMPB_WF	2.73	***	34.20	***	16.26	***	37.44	***	13.43	**	0.20	ns	0.19	ns	3.42	***ns***	47.8	***	6.7	*	7.1	***ns***	0.0	ns
RAMPB_WOF	1.43	ns	2.01	ns	3.89	ns	2.58	ns	0.41	ns	1.88	ns	4.45	***ns***	0.36	ns	53.8	***	1.3	ns	0.0	ns	0.0	ns
TOTPB	2.43	***	53.77	***	13.13	***	48.04	***	12.57	**	1.92	ns	3.18	ns	6.41	*	118.9	***	5.3	*	4.8	***ns***	0.0	ns
RAMBB	2.90	***	9.94	***	120.87	***	14.61	***	3.28	ns	3.00	ns	0.84	ns	0.12	ns	68.2	***	1.9	ns	1.3	ns	0.1	ns
TOTB	3.48	***	42.53	***	113.24	***	49.52	***	1.95	ns	0.35	ns	0.15	ns	1.57	ns	42.3	***	2.1	ns	3.3	ns	0.0	ns
FITTOT	5.07	***	201.80	***	28.37	***	35.21	***	0.29	ns	0.31	ns	0.22	ns	1.61	ns	20.5	***	13.1	**	0.0	ns	0.0	ns
FITSTEM	3.23	***	161.17	***	0.79	ns	14.78	***	7.09	*	0.14	ns	0.46	ns	4.16	***ns***	119.6	***	44.5	***	0.1	ns	0.0	ns
FRUITSTEM	3.01	***	83.92	***	0.49	ns	9.86	***	8.97	**	0.09	ns	0.00	ns	3.18	***ns***	85.6	***	47.9	***	0.1	ns	0.1	ns
SILSTEM	3.76	***	434.03	***	101.02	***	31.50	***	1.78	ns	2.39	ns	2.52	ns	1.96	ns	256.6	***	20.8	***	3.0	ns	0.0	ns
FITPB	3.24	***	197.91	***	0.19	ns	42.05	***	9.71	**	0.99	ns	0.78	ns	8.94	**	25.0	***	4.8	*	0.0	ns	8.1	*
FRUITPB	3.49	***	158.53	***	1.19	ns	45.88	***	10.64	**	0.72	ns	0.01	ns	6.29	*	22.7	***	12.7	**	0.0	ns	2.0	ns
SILPB	3.86	***	450.20	***	33.62	***	25.65	***	0.34	ns	1.11	ns	1.42	ns	4.17	***ns***	211.2	***	8.1	*	0.0	ns	0.0	ns
FITBB	1.82	*	20.38	***	36.04	***	2.59	ns	0.21	ns	2.51	ns	0.23	ns	0.17	ns	7.9	**	1.7	ns	0.0	ns	0.1	ns
FRUITBB	2.81	***	24.36	***	95.10	***	8.08	***	2.16	ns	2.18	ns	0.56	ns	0.14	ns	41.4	***	8.3	*	0.0	ns	0.7	ns
SILBB	2.48	***	148.46	***	12.90	***	16.34	***	0.03	ns	1.91	ns	0.14	ns	2.66	ns	149.6	***	6.0	*	0.0	ns	0.0	ns
RSTEM	3.12	***	76.90	***	34.61	***	19.77	***	0.97	ns	0.09	ns	1.87	ns	0.17	ns	25.9	***	5.7	*	0.3	ns	0.0	ns
RPB	1.61	*	67.83	***	42.64	***	5.22	**	7.69	*	0.56	ns	1.43	ns	4.39	***ns***	55.9	***	6.9	*	0.0	ns	0.2	ns
RBB	1.73	*	19.98	***	27.71	***	0.17	ns	15.53	***	1.14	ns	0.91	ns	0.51	ns	6.0	***ns***	3.3	ns	0.9	ns	0.2	ns
INTERNOD	1.21	ns	32.07	***	1.77	ns	1.10	ns	3.42	ns	0.03	ns	0.00	ns	1.79	ns	2.7	ns	2.0	ns	0.0	ns	0.0	ns
SURVIVAL	39.31	***	57.15	***	0.06	ns	47.30	***	Inf	***	Inf	***	Inf	***	Inf	**	1.0	ns	3.5	ns	0.0	ns	ne	ne

*0.05 > *P* > 0.01, **0.01 > *P* > 0.001, ***$P < 0.001$. ns: non-significant, *ns* : significant before a false discovery rate (FDR) correction at the nominal level of 5%, ne: not estimated. ǂ All traits were measured quantitatively with the exception of survival which is a binary trait. § Each trait was modeled separately using a mixed model. Model random terms were tested with likelihood ratio tests (LRT) of models with and without these effects. A correction for the number of tests was performed for each modeled effect (*i.e.* per column) to control the FDR at a nominal level of 5%.

Table S2 | Broad-sense heritability values (H^2) of the 174 eco-phenotypes scored across six *in situ* 'soil x competition' micro-habitats. *P*: bold values indicate significant broad-sense heritability estimates after a false discovery rate (FDR) correction at the nominal level of 5%.

Ecophenotype	H^2	*P*
BT_A_wo_P	0.868	**0.00E+00**
BT_A_w_P	0.847	**0.00E+00**
BT_B_wo_P	0.864	**0.00E+00**
BT_B_w_P	0.827	**0.00E+00**
BT_C_wo_P	0.864	**0.00E+00**
BT_C_w_P	0.843	**0.00E+00**
DIAM_A_wo_P	0.084	6.27E-01
DIAM_A_w_P	0.480	**4.50E-08**
DIAM_B_wo_P	0.329	**1.01E-03**
DIAM_B_w_P	0.303	**4.90E-03**
DIAM_C_wo_P	0.174	1.38E-01
DIAM_C_w_P	0.470	**5.40E-08**
FITBB_A_wo_P	0.261	2.64E-01
FITBB_A_w_P	0.005	1.00E+00
FITBB_B_wo_P	0.256	1.69E-01
FITBB_B_w_P	0.260	1.60E-01
FITBB_C_wo_P	0.323	8.22E-02
FITBB_C_w_P	0.451	**2.21E-03**
FITPB_A_wo_P	0.442	**3.14E-04**
FITPB_A_w_P	0.398	**3.84E-04**
FITPB_B_wo_P	0.341	**2.06E-03**
FITPB_B_w_P	0.174	1.93E-01
FITPB_C_wo_P	0.490	**1.05E-06**
FITPB_C_w_P	0.602	**1.51E-12**
FITSTEM_A_wo_P	0.626	**2.07E-11**
FITSTEM_A_w_P	0.644	**3.58E-16**
FITSTEM_B_wo_P	0.495	**3.65E-08**
FITSTEM_B_w_P	0.419	**2.80E-05**
FITSTEM_C_wo_P	0.709	**0.00E+00**
FITSTEM_C_w_P	0.716	**0.00E+00**
FITTOT_A_wo_P	0.230	7.29E-02
FITTOT_A_w_P	0.399	**6.65E-04**
FITTOT_B_wo_P	0.170	1.86E-01
FITTOT_B_w_P	0.202	**2.89E-02**
FITTOT_C_wo_P	0.418	**2.63E-02**
FITTOT_C_w_P	0.566	**8.90E-07**
FRUITBB_A_wo_P	0.256	5.33E-02
FRUITBB_A_w_P	0.342	**6.65E-04**
FRUITBB_B_wo_P	0.334	**2.20E-03**
FRUITBB_B_w_P	0.400	**5.20E-05**
FRUITBB_C_wo_P	0.303	**6.05E-03**
FRUITBB_C_w_P	0.635	**0.00E+00**
FRUITPB_A_wo_P	0.326	**8.95E-03**
FRUITPB_A_w_P	0.401	**4.18E-05**
FRUITPB_B_wo_P	0.252	**2.21E-02**
FRUITPB_B_w_P	0.147	2.45E-01
FRUITPB_C_wo_P	0.447	**2.94E-06**
FRUITPB_C_w_P	0.591	**7.08E-15**

Ecophenotype	H^2	*P*
FRUITSTEM_A_wo_P	0.515	**1.53E-07**
FRUITSTEM_A_w_P	0.601	**2.61E-14**
FRUITSTEM_B_wo_P	0.370	**3.27E-04**
FRUITSTEM_B_w_P	0.323	**3.51E-03**
FRUITSTEM_C_wo_P	0.676	**0.00E+00**
FRUITSTEM_C_w_P	0.749	**0.00E+00**
GERM_A_wo_P	0.827	**0.00E+00**
GERM_A_w_P	0.796	**0.00E+00**
GERM_B_wo_P	0.781	**0.00E+00**
GERM_B_w_P	0.773	**0.00E+00**
GERM_C_wo_P	0.659	**0.00E+00**
GERM_C_w_P	0.738	**0.00E+00**
H1F_A_wo_P	0.536	**6.72E-09**
H1F_A_w_P	0.567	**4.27E-12**
H1F_B_wo_P	0.705	**0.00E+00**
H1F_B_w_P	0.541	**2.66E-09**
H1F_C_wo_P	0.574	**6.09E-12**
H1F_C_w_P	0.695	**0.00E+00**
HD_A_wo_P	0.518	**1.10E-07**
HD_A_w_P	0.673	**0.00E+00**
HD_B_wo_P	0.728	**0.00E+00**
HD_B_w_P	0.666	**0.00E+00**
HD_C_wo_P	0.529	**1.80E-09**
HD_C_w_P	0.714	**0.00E+00**
HMAX_A_wo_P	0.607	**2.19E-12**
HMAX_A_w_P	0.627	**0.00E+00**
HMAX_B_wo_P	0.625	**0.00E+00**
HMAX_B_w_P	0.574	**1.30E-11**
HMAX_C_wo_P	0.725	**0.00E+00**
HMAX_C_w_P	0.720	**0.00E+00**
HSTEM_A_wo_P	0.615	**2.30E-12**
HSTEM_A_w_P	0.640	**0.00E+00**
HSTEM_B_wo_P	0.620	**0.00E+00**
HSTEM_B_w_P	0.614	**1.25E-14**
HSTEM_C_wo_P	0.748	**0.00E+00**
HSTEM_C_w_P	0.761	**0.00E+00**
INT_A_wo_P	0.755	**0.00E+00**
INT_A_w_P	0.641	**0.00E+00**
INT_B_wo_P	0.771	**0.00E+00**
INT_B_w_P	0.623	**0.00E+00**
INT_C_wo_P	0.720	**0.00E+00**
INT_C_w_P	0.513	**4.52E-10**
INTERNOD_A_wo_P	0.026	1.00E+00
INTERNOD_A_w_P	0.103	5.66E-01
INTERNOD_B_wo_P	0.595	**1.73E-14**
INTERNOD_B_w_P	0.472	**9.37E-07**
INTERNOD_C_wo_P	0.105	4.71E-01
INTERNOD_C_w_P	0.160	1.00E+00

Table S2 (continued)

Ecophenotype	H^2	*P*
RAMBB_A_wo_P	0.316	**1.19E-02**
RAMBB_A_w_P	0.343	**6.65E-04**
RAMBB_B_wo_P	0.464	**9.28E-07**
RAMBB_B_w_P	0.388	**1.73E-04**
RAMBB_C_wo_P	0.344	**1.25E-03**
RAMBB_C_w_P	0.669	**0.00E+00**
RAMPB_WOF_A_wo_P	0.314	**8.95E-03**
RAMPB_WOF_A_w_P	0.555	**1.17E-11**
RAMPB_WOF_B_wo_P	0.182	1.54E-01
RAMPB_WOF_B_w_P	0.305	**5.83E-03**
RAMPB_WOF_C_wo_P	0.152	2.14E-01
RAMPB_WOF_C_w_P	0.405	**1.95E-05**
RAMPB_WF_A_wo_P	0.493	**8.70E-07**
RAMPB_WF_A_w_P	0.438	**2.94E-06**
RAMPB_WF_B_wo_P	0.484	**1.18E-07**
RAMPB_WF_B_w_P	0.398	**2.14E-04**
RAMPB_WF_C_wo_P	0.349	**9.08E-04**
RAMPB_WF_C_w_P	0.497	**1.09E-08**
RBB_A_wo_P	0.398	1.95E-01
RBB_A_w_P	0.440	1.05E-01
RBB_B_wo_P	0.520	**3.27E-04**
RBB_B_w_P	0.607	**1.90E-06**
RBB_C_wo_P	0.299	1.60E-01
RBB_C_w_P	0.664	**3.17E-05**
RP_A_wo_P	0.801	**0.00E+00**
RP_A_w_P	0.827	**0.00E+00**
RP_B_wo_P	0.798	**0.00E+00**
RP_B_w_P	0.742	**0.00E+00**
RP_C_wo_P	0.842	**0.00E+00**
RP_C_w_P	0.817	**0.00E+00**
RPB_A_wo_P	0.403	**2.81E-03**
RPB_A_w_P	0.345	**2.23E-03**
RPB_B_wo_P	0.555	**4.52E-10**
RPB_B_w_P	0.430	**1.65E-04**
RPB_C_wo_P	0.255	**1.56E-02**
RPB_C_w_P	0.505	**1.09E-08**
RSTEM_A_wo_P	0.253	7.42E-02
RSTEM_A_w_P	0.369	**1.43E-03**
RSTEM_B_wo_P	0.355	**5.97E-03**
RSTEM_B_w_P	0.293	**1.61E-02**
RSTEM_C_wo_P	0.074	1.00E+00
RSTEM_C_w_P	0.405	**3.47E-04**
SILBB_A_wo_P	0.677	**2.55E-06**
SILBB_A_w_P	0.681	**3.64E-04**
SILBB_B_wo_P	0.803	**0.00E+00**
SILBB_B_w_P	0.718	**2.63E-07**
SILBB_C_wo_P	0.713	**1.40E-11**
SILBB_C_w_P	0.729	**2.45E-10**

Ecophenotype	H^2	*P*
SILPB_A_wo_P	0.638	**3.52E-12**
SILPB_A_w_P	0.604	**9.79E-12**
SILPB_B_wo_P	0.779	**0.00E+00**
SILPB_B_w_P	0.702	**0.00E+00**
SILPB_C_wo_P	0.635	**8.41E-14**
SILPB_C_w_P	0.725	**0.00E+00**
SILSTEM_A_wo_P	0.774	**0.00E+00**
SILSTEM_A_w_P	0.781	**0.00E+00**
SILSTEM_B_wo_P	0.797	**0.00E+00**
SILSTEM_B_w_P	0.797	**0.00E+00**
SILSTEM_C_wo_P	0.743	**0.00E+00**
SILSTEM_C_w_P	0.755	**0.00E+00**
SURVIVAL_A_wo_P	0.022	8.72E-01
SURVIVAL_A_w_P	0.000	1.00E+00
SURVIVAL_B_wo_P	0.135	2.02E-01
SURVIVAL_B_w_P	0.113	2.90E-01
SURVIVAL_C_wo_P	0.000	1.00E+00
SURVIVAL_C_w_P	0.227	**2.29E-02**
TOTB_A_wo_P	0.359	**2.53E-03**
TOTB_A_w_P	0.216	6.31E-02
TOTB_B_wo_P	0.372	**4.52E-04**
TOTB_B_w_P	0.175	1.75E-01
TOTB_C_wo_P	0.254	**2.65E-02**
TOTB_C_w_P	0.542	**6.87E-11**
TOTPB_A_wo_P	0.565	**7.37E-10**
TOTPB_A_w_P	0.498	**3.19E-08**
TOTPB_B_wo_P	0.668	**0.00E+00**
TOTPB_B_w_P	0.580	**1.30E-11**
TOTPB_C_wo_P	0.525	**6.60E-10**
TOTPB_C_w_P	0.618	**4.60E-16**

Table S3 | Manhattan distance: scaling relationships between total phenotypic effect size of SNPs with the highest association and the effective number of eco-phenotypes (N_{eff}). The pleiotropic scaling relationship between the total effect size and the effective number of eco-phenotypes was calculated as $T_M = c^* {N_{eff}}^d$, with T_M corresponding to the Manhattan distance and calculated as $T_M = \sum_{i=1}^{n} |A_i|$, where n is the degree of pleiotropy and A_i is the standardized allelic effect. To avoid pseudo-replication due to the presence of several top SNPs in a given LD block, the pleiotropic scaling was also calculated for each threshold number of top SNPs and each threshold of significance, (i) by considering the mean value of the total effect size and N_{eff} per LD block containing top SNPs ('Mean per block' column) and (ii) by randomly sampling one top SNP per LD block (this step was repeated 1,000 times) ('Random' column).

T_M	Threshold	Total SNPs	Unique SNPs	% pleiotropic SNPs	All unique SNPs		Mean per block		Random	
					c	d	c	d	c	d
number of top SNPs	50 SNPs	7200	5728	16.69	0.317	1.255	0.317	1.3	0.32 (0.315 - 0.324)	1.268 (1.224 - 1.323)
	100 SNPs	14400	11100	19.05	0.308	1.242	0.309	1.275	0.311 (0.307 - 0.316)	1.253 (1.214 - 1.294)
	200 SNPs	28800	21268	21.86	0.294	1.228	0.295	1.255	0.296 (0.292 - 0.3)	1.243 (1.208 - 1.274)
	300 SNPs	43200	30854	24.4	0.286	1.215	0.289	1.223	0.289 (0.285 - 0.293)	1.217 (1.187 - 1.249)
	500 SNPs	72000	48851	27.64	0.277	1.19	0.283	1.178	0.282 (0.278 - 0.287)	1.181 (1.152 - 1.204)
$-\log_{10} p$-value	> 6	538	424	21.46	0.433	1.545	0.423	1.799	0.425 (0.416 - 0.438)	1.736 (1.503 - 1.92)
	> 5	3165	2457	17.91	0.366	1.446	0.361	1.51	0.362 (0.35 - 0.372)	1.49 (1.382 - 1.637)
	> 4	22822	16720	22.06	0.319	1.232	0.318	1.267	0.32 (0.314 - 0.326)	1.241 (1.197 - 1.293)

Table S4 | Euclidean distance: scaling relationships between total phenotypic effect size of SNPs with the highest association and the effective number of eco-phenotypes (N_{eff}). The pleiotropic scaling relationship between the total effect size and the effective number of eco-phenotypes was calculated as $T_E = a^* Neff^b$, with T_E corresponding to the Euclidean distance and calculated as $T_E = \sqrt{\sum_{i=1}^{n} A_i^2}$, where n is the degree of pleiotropy and A_i is the standardized allelic effect. To avoid pseudo-replication due to the presence of several top SNPs in a given LD block, the pleiotropic scaling was also calculated for each threshold number of top SNPs and each threshold of significance, (i) by considering the mean value of the total effect size and N_{eff} per LD block containing top SNPs ('Mean per block' column) and (ii) by randomly sampling one top SNP per LD block (this step was repeated 1,000 times) ('Random' column).

T_E	Threshold	Total SNPs	Unique SNPs	% pleiotropic SNPs	All unique SNPs		Mean per block		Random	
					a	b	a	b	a	b
number of top SNPs	50	7200	5728	16.69	0.292	0.739	0.295	0.766	0.296 (0.294 - 0.298)	0.757 (0.718 - 0.803)
	100	14400	11100	19.05	0.28	0.743	0.283	0.771	0.284 (0.282 - 0.286)	0.764 (0.729 - 0.797)
	200	28800	21268	21.86	0.267	0.726	0.268	0.751	0.269 (0.267 - 0.271)	0.747 (0.722 - 0.772)
	300	43200	30854	24.4	0.26	0.712	0.26	0.728	0.261 (0.259 - 0.263)	0.73 (0.705 - 0.755)
	500	72000	48851	27.64	0.25	0.692	0.25	0.697	0.252 (0.25 - 0.253)	0.705 (0.682 - 0.725)
$-\log_{10} p$-value	> 6	538	424	21.46	0.398	0.919	0.393	1.158	0.395 (0.39 - 0.401)	1.104 (0.884 - 1.285)
	> 5	3165	2457	17.91	0.34	0.838	0.335	0.917	0.335 (0.331 - 0.339)	0.906 (0.835 - 0.99)
	> 4	22822	16720	22.06	0.287	0.727	0.288	0.743	0.29 (0.288 - 0.292)	0.74 (0.706 - 0.776)

Table S5 | List of candidate genes associated with 11 or more evolved eco-phenotypes.

Atg number	no eco-phenotypes	Locus name	Molecular function
AT4G01820	17	ABCB3	member of MDR subfamily
AT4G01830	11	PGP5	P-glycoprotein 5 (PGP5)
AT4G14660	12	NRPE7	Non-catalytic subunit specific to DNA-directed RNA polymerase V
AT4G18350	12	NCED2	Encodes 9-*cis* -epoxycarotenoid dioxygenase, a key enzyme in the biosynthesis of abscisic acid.
AT4G19960	24	AtKUP/HAK/KT9	Encodes a potassium ion transmembrane transporter.
AT4G20325	12		unknown
AT4G20330	11		Transcription initiation factor TFIIE, beta subunit
AT4G20340	13		Transcription factor TFIIE, alpha subunit
AT4G20350	18		oxidoreductases
AT4G20362	15	SORF6	Potential natural antisense gene, locus overlaps with AT4G20360
AT4G20370	11	TSF	Encodes a floral inducer that is a homolog of FT.
AT4G24520	12	ATR1	Encodes a cyp450 reductase likely to be involved in phenylpropanoid metabolism.
AT5G12430	14	TPR16	Encodes one of the 36 carboxylate clamp (CC)-tetratricopeptide repeat (TPR) proteins
AT5G43430	13	ETFBETA	Encodes the electron transfer flavoprotein ETF beta, a putative subunit of the mitochondrial electron transfer flavoprotein complex

Table S6 | Enrichment of biological process in the 0.1% tail of the F_{ST} values.

Biological process	Enrichment	*P* value	Atg number	Locus name	Molecular function	Associated eco-phenotypes[1]
vernalization response	22	**	AT5G10140	*FLC*	MADS-box protein	
			AT4G16845	*VRN2*	nuclear-localized zinc finger protein	H1F_B_w_P, RSTEM_B_wo_P, SURVIVAL_C_w_P, DIAM_B_wo_P, H1F_C_wo_P, SILBB_B_w_P, FITTOT_C_wo_P
regulation of circadian rhythm	21	**	AT5G10140	*FLC*	MADS-box protein	
response to temperature stimulus	21	**	AT5G10140	*FLC*	MADS-box protein	
negative regulation of flower development	21	*	AT5G10140	*FLC*	MADS-box protein	
regulation of cell shape	17	*	AT3G59100	*GLUCAN SYNTHASE-LIKE 11*	protein similar to callose synthase	
			AT4G03550	*POWDERY MILDEW RESISTANT 4*	callose synthase	FRUITSTEM_C_w_P, SILPB_A_wo_P
beta-D-glucan biosynthetic process	17	*	AT3G59100	*GLUCAN SYNTHASE-LIKE 11*	protein similar to callose synthase	
			AT4G03550	*POWDERY MILDEW RESISTANT 4*	callose synthase	FRUITSTEM_C_w_P, SILPB_A_wo_P
pollen tube development	15	*	AT4G05450	*MFDX1*	mitochondrial ferredoxin 1	FRUITSTEM_C_w_P, SILPB_A_wo_P
electron transport chain	15	*	AT4G05450	*MFDX1*	mitochondrial ferredoxin 1	FRUITSTEM_C_w_P, SILPB_A_wo_P
polar nucleus fusion	14	*	AT4G05440	*EMBRYO SAC DEVELOPMENT ARREST 35*	unknown	FRUITSTEM_C_w_P, SILPB_A_wo_P
			AT5G42020	*BIP2*	luminal binding protein	DIAM_C_w_P, TOT_B_C_w_P, RAMPB_WF_C_w_P
stamen development	14	*	AT4G03190	*AFB1*	F box protein belonging to the TIR1 subfamily	
			AT5G41700	*UBIQUITIN CONJUGATING ENZYME 8*	one of the polypeptides that constitute the ubiquitin-conjugating enzyme E2	FITTOT_C_w_P, FRUITPB_C_w_P, RSTEM_B_w_P, SILPB_C_w_P, INT_B_wo_P, SILSTEM_B_wo_P, RAMPB_WF_C_w_P
defense response by callose deposition in cell wall	14	*	AT4G03550	*POWDERY MILDEW RESISTANT 4*	callose synthase	FRUITSTEM_C_w_P, SILPB_A_wo_P
salicylic acid mediated signaling pathway	14	*	AT4G03550	*POWDERY MILDEW RESISTANT 4*	callose synthase	FRUITSTEM_C_w_P, SILPB_A_wo_P
defense response signaling pathway, resistance gene-dependent	14	*	AT4G03550	*POWDERY MILDEW RESISTANT 4*	callose synthase	FRUITSTEM_C_w_P, SILPB_A_wo_P
cell cycle arrest	13	*	AT4G05440	*EMBRYO SAC DEVELOPMENT ARREST 35*	unknown	FRUITSTEM_C_w_P, SILPB_A_wo_P
calcium-mediated signaling	12	*	AT4G03560	*TPC1*	depolarization-activated Ca(2+) channel	
trehalose biosynthetic process	12	*	AT5G10100	*TPPI*	haloacid dehalogenase-like hydrolase (HAD) superfamily protein	FITPB_A_wo_P, FRUITPB_A_wo_P
calcium ion transmembrane transport	12	*	AT4G03560	*TPC1*	depolarization-activated Ca(2+) channel	
calcium ion transport	12	*	AT4G03560	*TPC1*	depolarization-activated Ca(2+) channel	
regulation of salicylic acid mediated signaling pathw	10	*	AT4G03440		Ankyrin repeat family protein	FRUITSTEM_C_w_P
			AT4G03460		Ankyrin repeat family protein	GERM_A_wo_P, SILPB_B_w_P, SILPB_A_wo_P, SILSTEM_B_wo_P
			AT4G03470		Ankyrin repeat family protein	SILPB_B_w_P, SILPB_A_wo_P, SILSTEM_B_wo_P
			AT4G03500		Ankyrin repeat family protein	FRUITSTEM_C_w_P
cellular response to salicylic acid stimulus	10	*	AT4G03440		Ankyrin repeat family protein	FRUITSTEM_C_w_P
			AT4G03450		Ankyrin repeat family protein	FRUITSTEM_C_w_P, GERM_A_wo_P
			AT4G03460		Ankyrin repeat family protein	GERM_A_wo_P, SILPB_B_w_P, SILPB_A_wo_P, SILSTEM_B_wo_P
			AT4G03470		Ankyrin repeat family protein	SILPB_B_w_P, SILPB_A_wo_P, SILSTEM_B_wo_P
			AT4G03500		Ankyrin repeat family protein	FRUITSTEM_C_w_P
photosynthetic electron transport chain	10	*	AT4G03280	*PGR1*	Encodes the Rieske FeS center of cytochrome b6f complex	FRUITSTEM_C_w_P
developmental growth	7	*	AT4G03190	*AFB1*	F box protein belonging to the TIR1 subfamily	
pollen maturation	7	*	AT4G03190	*AFB1*	F box protein belonging to the TIR1 subfamily	
regulation of auxin mediated signaling pathway	5	*	AT3G59060	*PIF5*	novel Myc-related bHLH transcription factor	

*0.05 > *P* > 0.01, **0.01 > *P* > 0.001. The significance of enrichment was tested against a null distribution using 10,000 permutations.

[1] The letters A, B and C stand for the three types of soil. 'wo_P' and 'w_P' correspond to the absence and presence of *P. annua*, respectively.

C) Conclusions

Dans ce chapitre, les deux principaux objectifs étaient (i) de tester l'adaptation d'une population locale d'*A. thaliana* à des facteurs abiotiques (i.e. climat) et biotique (i.e. interactions plante-plante) sur une courte échelle de temps, et (ii) de développer une population de GWA mapping adaptée à l'étude des interactions plante-plante.

Pour le premier objectif, nous avons pu démontrer que la population TOU-A présentait une variation génétique importante pour la majorité des éco-phénotypes (144/174 éco-phénotypes) et notamment en réponse à la présence de *P. annua*. En phénotypant des lignées échantillonnées en 2002 et 2010, nous avons pu mettre en évidence la sélection simultanée d'un ensemble de traits dont l'évolution s'est faite vers un nouvel optimum phénotypique similaire entre les six micro-habitats. Parmi ces traits, nous avons notamment observé l'évolution d'un trait relatif à la réponse à la compétition (HD = H1F/DIAM) vers de plus grandes valeurs, avec une sélection de plantes ayant une tige plus grande au moment de la floraison mais sans modification du diamètre de la rosette. Ce résultat suggère une évolution de la population vers une stratégie d'évitement de la compétition. Attention, ceci n'est qu'une hypothèse ! En effet, bien que la compétition interspécifique semble très fréquente au sein de la communauté TOU-A, nous ne savons pas si elle s'est intensifiée entre 2002 et 2010. Il aurait été intéressant d'étudier l'évolution de la composition de la communauté végétale TOU-A et l'intensité de la compétition sur cette même période. Malheureusement, ce type de données n'a pas été collecté.

Pour atteindre le deuxième objectif, il fallait que la population TOU-A présente une diversité génomique importante et un LD court compatible avec des analyses de GWA mapping. Ces deux pré-requis ont été validés. En effet, en collaboration avec la plateforme bioinformatique du laboratoire, l'analyse des données de séquençage des 195 accessions de cette population nous a permis de mettre en évidence une diversité génomique très importante à un niveau local. En effet, sur la base de ces 195 accessions, on retrouve environ 1/6ème de la diversité génomique observée dans un panel de 1135 accessions échantillonnées au niveau mondial (The 1001 Genomes Consortium, 2016). Les fortes interactions 'accession x micro-habitat' identifiées dans notre étude pour la majorité des traits (et notamment la production de graines) pourraient expliquer le maintien d'une telle diversité. Ainsi, une forte hétérogénéité spatiale sur de courtes distances physiques aussi bien au niveau abiotique (i.e. édaphique) qu'au

niveau biotique (i.e. interactions plante-plante) favoriserait le maintien de génotypes spécialistes au cours des générations.

La forte diversité génétique observée au sein de la population TOU-A s'accompagne d'un déséquilibre de liaison très court (~18bp). Ce résultat a été une réelle surprise. En effet, le LD observé sur un jeu d'accessions mondiales est plutôt de l'ordre de 5-10kb (Kim *et al.* 2007). Comment alors expliquer un LD si court pour la population TOU-A ? Une hypothèse reposerait sur un taux d'allogamie particulièrement élevée dans cette population (de l'ordre de 7%, Platt *et al.* 2010), ce qui augmenterait le taux de recombinaison efficace. Par ailleurs, il a été observé la présence de quelques plantes 'femelles' (absence d'étamines sur la majorité des fleurs) au sein de cette population, ce qui pourrait faciliter les croisements entre plantes et donc aussi augmenter le taux de recombinaison efficace.

Par une approche de résurrection couplée à des analyses de GWA mapping et de différentiation génétique temporelle, nous avons donc pu montrer que l'utilisation des 195 accessions provenant de la population TOU-A était adaptée pour l'identification des bases génétiques associées aux interactions plante-plante. Cependant, dans cette étude, nous n'avons considéré l'interaction d'*A. thaliana* qu'avec une seule autre espèce végétale. Or, dans la majorité des cas, les plantes interagissent de manière simultanée et/ou séquentielle avec plusieurs espèces tout au long de leur cycle de vie. C'est aussi le cas pour *A. thaliana* au sein de la communauté végétale TOU-A (Figure 1.2). Il est donc nécessaire de considérer cette complexité dans l'identification des bases génétiques associées aux interactions plante-plante. C'est dans ce contexte d'interactions complexes que s'inscrit mon second chapitre de thèse.

Figure 1.2. Illustration de la complexité des interactions plante-plante que peuvent rencontrer des plantes d'*A. thaliana* au sein de la population TOU-A.

Chapitre 2

Identification des bases génétiques sous-jacentes aux interactions mono- et pluri-spécifiques

A) Introduction

Dans le chapitre précédent, nous avons pu montrer que la population locale TOU-A était adaptée à l'identification des bases génétiques associées à la variation naturelle des interactions plante-plante. En effet, elle présente de nombreux avantages (forte diversité génomique et LD court) pour cartographier finement des régions génomiques associées à la variation naturelle phénotypique. Nous avons aussi mis en évidence que la présence d'un seul individu d'une autre espèce végétale pouvait induire des réponses phénotypiques très différentes entre les accessions composant cette population. Néanmoins, l'interaction avec un seul individu mais surtout avec une seule espèce végétale semble biologiquement et écologiquement peu réaliste, tant les plantes interagissent avec une multitude d'individus d'espèces différentes (notamment dans les communautés végétales naturelles). Il apparait donc important de replacer l'étude des bases génétiques associées aux interactions plante-plante dans un contexte écologiquement plus réaliste. En particulier, il reste à déterminer si l'architecture génétique sous-jacente à la réponse à la compétition plurispécifique peut être prédite à partir des architectures génétiques obtenues en conditions monospécifiques.

Dans ce chapitre, nous avons tout d'abord cherché à estimer l'ampleur de la variation génétique de la réponse à la compétition pour un ensemble de 96 accessions provenant de la population TOU-A, sur lesquelles quatre traits phénotypiques ont été mesurés. Ces accessions ont été soumises à 12 traitements de compétition monospécifiques et plurispécifiques. Ces traitements sont basés sur la combinaison simultanée d'une, deux ou trois espèces fréquemment associées à *A. thaliana* dans des communautés végétales en France (*Poa annua*, *Stellaria media* et *Veronica arvensis*, Frachon *et al.* 2019) et présentes au sein de la communauté végétale associée à la population TOU-A. Sur la base des données individuelles de séquençage des 96 accessions, nous avons comparé l'architecture génétique de la réponse à la compétition pour l'ensemble des quatre traits mesurés dans les 12 traitements de compétition. Enfin, nous avons identifié les gènes à proximité des SNP les plus associés à la réponse à la compétition pour l'ensemble des 48 combinaisons 'trait phénotypique x traitement de compétition'. Cette liste de gènes nous a permis de tester si les processus biologiques les plus significativement associés aux interactions plante-plante différaient entre interactions monospécifiques et interactions plurispécifiques.

Lors de ce chapitre, nous avons donc cherché à répondre à trois questions principales :

i. Quelle est l'étendue de la variation génétique au sein de la population locale d'*A. thaliana* dans différents contextes d'interactions plante-plante mono- et plurispécifiques?
ii. Peut-on prédire l'architecture génétique en conditions plurispécifiques à partir des QTL identifiés en conditions monospécifiques (hypothèse d'additivité)? Ou observe-t-on l'émergence de nouveaux QTL?
iii. Quels sont les principaux processus biologiques associés à la réponse à la compétition dans différents contextes d'interactions plante-plante mono- et plurispécifiques?

NB : dans ce chapitre, mon travail a consisté (i) à effectuer les analyses statistiques des données de phénotypage, (ii) à effectuer les analyses de GWA mapping, (iii) à caractériser l'architecture génétique de l'ensemble des traits, (iv) à identifier les gènes sous-jacents et (v) à identifier les processus biologiques surreprésentés dans les différentes conditions de compétition. L'ensemble des mesures phénotypiques ont été faites par Etienne Baron (ancien doctorant de Fabrice Roux) et Juliana Lenglet (stagiaire de master 2 d'Etienne Baron).

B) Manuscrit: The genomic architecture of competitive response of *Arabidopsis thaliana* is highly flexible between monospecific and plurispecific neighborhoods.

Short running title: The genetics of plurispecific interactions

Cyril Libourel[1], Etienne Baron[1,2], Juliana Lenglet[2], Laurent Amsellem[2], Dominique Roby[1,3*], Fabrice Roux[1,2,3*]

Affiliations :

[1] LIPM, Université de Toulouse, INRA, CNRS, Castanet-Tolosan, France

[2] Laboratoire Evolution, Ecologie et Paléontologie, UMR CNRS 8198, Université de Lille, Villeneuve d'Ascq Cedex, France

[3] These authors contributed equally to this work.

* To whom correspondence should be addressed. E-mail: fabrice.roux@inra.fr & dominique.roby@inra.fr

Manuscrit

"The genomic architecture of competitive response of *Arabidopsis thaliana* is highly flexible between monospecific and plurispecific neighborhoods"

Cyril Libourel, Etienne Baron, Juliana Lenglet, Laurent Amsellem, Dominique Roby, Fabrice Roux

Abstract

Although plants simultaneously interact with multiple neighboring species throughout their life cycle, there is still very limited information about the genetics of the competitive response in the context of plurispecific interactions. Using a local mapping population of *Arabidopsis thaliana*, we set up a Genome Wide Association study to estimate the extent of genetic variation of the competitive response in presence of 12 plant species assemblages, and to compare the genetic architecture of the competitive response between monospecific and plurispecific neighborhoods. Based on four phenotypic traits, we detected strong crossing reaction norms not only among the three monospecific neighborhoods, but also among the different plant assemblages. Accordingly, the genetic architecture of the competitive response was highly dependent on the identity and the relative abundance of the neighboring species. In addition, enriched biological processes underlying the competitive response largely differ between monospecific and plurispecific neighborhoods. In particular, receptor-like kinases and transporters were significantly enriched in plurispecific neighborhoods. Our results suggest that plants can integrate and respond to different species assemblages depending on the identity and number of each neighboring species, through a large range of genes associated mainly with perception and signaling processes leading to developmental and stress responses.

Key words: *Arabidopsis thaliana*, genetic variation, GWAS, local population, plant-plant interactions, plurispecific interactions.

Introduction

Because plant-plant interactions are recognized as a major factor mediating plant community structure, diversity and dynamics (Tilman 1985; Goldberg & Barton 1992; Chesson 2000; Martorell & Freckelton 2014), deciphering the genetic and molecular bases underlying plant-plant interactions appears fundamental to predict the evolutionary dynamics of plant communities in ecological time (Pierik *et al.*, 2013; Frachon *et al.*, 2017). This is especially relevant in the context of current anthropogenic modifications of plant assemblages, which may in part result from the intertwined effect of increased plant biomass and reduced plant diversity under climate warming (Baldwin *et al.*, 2014) or from native species having different geographical range shifts under climate change (Bachelet *et al.*, 2001; Gilman *et al.*, 2010; Singer *et al.*, 2013). Because the average potential to reduce crop yield is significantly higher for weeds than any crop pests (Oerke *et al.*, 2004; Neve *et al.*, 2009), identifying and characterizing the function of genes underlying plant-plant interactions appears also fundamental to accelerate breeding programs aimed at increasing crop competitiveness (Worthington & Reberg-Horton 2013; Onishi *et al.*, 2018). In addition, in the context of complementarity in using resources, optimizing species assemblages in cropping species may be facilitated by the understanding of the genetics underlying overyielding (Litrico & Violle 2015; Pakeman *et al.*, 2015; Weiner *et al.*, 2017).

However, in comparison to other types of biotic interactions such as plant response to virus, bacteria, fungi, oomycetes and to a lesser extent herbivores, there is still very limited information about the genetics associated with natural variation of plant-plant interactions, i.e. when plants have been directly challenged by other plants (and not in artificial environments simulating plant-plant interactions) (Bergelson & Roux 2010; Bartoli & Roux 2017). For

example, a recent review listed only 47 Quantitative Trait Loci (QTL) mapping studies (including three Genome Wide Association studies, GWAS) that have been designed to study the genetic architecture underlying natural variation of plant-plant interactions (Subrahmaniam *et al.*, 2018). About two-thirds of these QTL mapping studies focused on asymmetric interactions (i.e. when one of the interacting partners benefits at the expense of the other), including allelopathy underlying weed suppressive ability and response to parasitic plants (Subrahmaniam *et al.*, 2018). Surprisingly, despite the importance of competition in driving plant community assemblages, only six QTL mapping studies (including two GWAS) focused on competitive interactions in a heterospecific context, i.e. when both interacting species suffer significant cost by investing in competing and therefore compromising on the benefit (Dudley 2015). In agreement with other types of biotic interactions (Roux & Bergelson 2016; Bartoli & Roux 2017), plant-plant interactions are mainly driven by a complex genetic architecture, ranging from the identification of few medium-effect QTLs to the identification of up to tens of small-effect QTLs (Subrahmaniam *et al.*, 2018).

While informative, most of these QTL mapping studies are based on monospecific heterospecific interactions (i.e. one single pair of interacting species; Subrahmaniam *et al.*, 2018). However, throughout their life cycle, focal plants often interact simultaneously with several neighboring species, either in crop fields or in natural communities (Wilson *et al.*, 2012). This highlights the need to study the genetic architecture underlying plant-plant interactions by considering the response of a focal species to plurispecific interactions. In particular, whether the genetic architecture underlying the response of a focal species in a plurispecific neighborhood corresponds to the sum of QTLs that are specific to a neighbor species and/or to the emergence of new QTLs remains on open question (Subrahmaniam *et al.*, 2018).

To address this question, we set up a GWAS to compare the genetic architecture of competitive response of the model plant *Arabidopsis thaliana* between monospecific and plurispecific neighborhoods. Although *A. thaliana* has long been considered as not being often challenged by other species in natural plant communities, several studies recently challenged this view (i) by revealing extensive genetic diversity associated with the response to interspecific competition (Brachi *et al.*, 2012; Bartheleimer *et al.*, 2015), in particular at the within-population scale (Baron *et al.*, 2015; Frachon *et al.*, 2017), and (ii) by finding an adaptation of a genetically polymorphic local population likely to increased interspecific competition in less than eight generations (Frachon *et al.*, 2017). In addition, in the context of monospecific interactions, the genetic architecture of competitive response of *A. thaliana* was found to be highly dependent on the identity of the neigboring species (Baron *et al.*, 2015), making *A. thaliana* an attractive model to test the absence or presence of new QTLs in a plurispecific neighborhood in comparison to related monospecific neighborhoods.

In this study, we first estimated the extent of genetic variation of competitive response in a set of 96 local *A. thaliana* French accessions, which were submitted to monospecific and plurispecific competition treatments based on all one-way, two-way and three-way combinations of three species frequently associated with *A. thaliana* in natural plant communities in France, *i.e. Poa annua*, *Stellaria media* and *Veronica arvensis*. Based on the whole-genome sequence of the 96 accessions, we then run GWA mapping to compare the genetic architecture of competitive response of *A. thaliana* between monospecific and plurispecific neighborhoods. Finally, we examined whether biological processes overrepresented among SNPs involved in competitive response were different between monospecific and plurispecific neighborhoods and discussed the function of candidate genes.

Materials and Methods

Plant material

A set of 96 whole-genome sequenced accessions of *A. thaliana* collected in the TOU-A local population (Toulon-sur-Arroux, Burgundy, France, 46°38'53.80"N - 4° 7'22.65"E) were used for the purpose of this study. As previously described in Frachon *et al.* (2017), the TOU-A population is highly polymorphic at both the phenotypic and genomic levels. Importantly, the very short Linkage Disequilibrium (LD) observed in this population ($r^2_{0.5}$~18bp) allows the fine mapping of genomic regions associated with natural variation of phenotypic traits down to the gene level (Brachi *et al.*, 2013; Huard-Chauveau *et al.*, 2013; Frachon *et al.*, 2017).

Maternal effects of the 96 accessions were reduced by growing one plant of each family for one generation under controlled greenhouse conditions (16-h photoperiod, 20°C) in early 2011 at the University of Lille. Given an estimated selfing rate of ~94% in this population (Platt *et al.*, 2010), the 96 accessions were considered as mostly homozygous along the genome.

In this study, we used three neighboring species commonly associated with *A. thaliana* in natural plant communities in France and detected in the TOU-A plant community (F. Roux, personal observation). These species are the meadow grass *P. annua* (Poaceae) with a low spreading growth form, the chickweed *S. media* (Caryophyllaceae) and the speedwell *V. arvensis* (Scrophulariaceae) both with a crawling growth form. Seeds for these three species have been obtained from the Herbiseeds company (http://www.herbiseed.com/home.aspx).

Phenotypic characterization

Experimental design

An experiment of 4,608 focal plants of *A. thaliana* and 12,672 neighbor plants was set up at the University of Lille 1 (North, France) in March 2013 using a split-plot design arranged as a randomized complete block design (RCBD) with 12 treatments nested within four blocks. These 12 competition treatments correspond to (Figure 1):

- one control treatment where *A. thaliana* was grown alone (i.e. absence of interaction; hereafter named treatment A).

- 10 interspecific interaction treatments corresponding to the full combination of the three neighboring species *P. annua* (P), *S. media* (S) and *V. arvensis* (V): PPP, SSS, VVV, PPS, PPV, PSS, PVV, SSV, SVV and PSV.

- one intraspecific interaction treatment (hereafter named treatment AAA). This treatment was included in the experiment to test whether the differences observed between the treatment where *A. thaliana* was grown alone (i.e. treatment A) and the treatments of interspecific interactions were not only due to the presence of a neighbor plant, but were rather dependent on either the identity of the neighboring species or the combination of neighboring species. With respect to the barochorous mode of seed dispersal in *A. thaliana* (Weinig *et al.*, 2006; Wender *et al.*, 2005), the intraspecific interaction treatment corresponds to intra-genotypic interaction.

Each 'block x competition treatment' combination was represented by 96 pots (7 cm x 7 cm x 7 cm, vol. ~250 cm^3; TEKU MQC) filled with damp standard culture soil (Huminsubstrat N3, Neuhaus) and disposed in staggered rows, each pot corresponding to one

of the 96 TOU-A accessions. The 17th of January 2013 (day 0), a minimum of five *A. thaliana* seeds were sown in the central position of each pot. For all the treatments (with the exception of treatment A), seeds for neighboring plants were evenly spaced, two cm away from the *A. thaliana* central position (Figure 1). Germination date of *A. thaliana* focal seedlings was daily monitored until six days after sowing. At this time, five accessions had a poor germination rate (between 0% and 6.25%) and were therefore discarded from further analyses. Plants that germinated after 6 days (i.e. 1.16%) were also discarded from further analyses.

A. thaliana focal seedlings and neighboring seedlings were thinned to one per pot 18 to 20 days after seed sowing. Plants were grown at 20 °C under natural light supplemented by artificial light to provide a 16-hr photoperiod and were top watered without supplemental nutrients. The experience lasted 87 days, from sowing to harvesting of the last plants.

Measured phenotypic traits

Three raw phenotypic traits were measured on each focal plant of *A. thaliana* at the time of their flowering, which was measured as the number of days between germination and flowering dates. The first trait corresponds to the height from the soil to the first flower on the main stem (H1F expressed in mm). H1F is related to seed dispersal (Wender *et al.*, 2005) and shade avoidance (Dorn *et al.*, 2000) in *A. thaliana*. The two other traits were used as proxies of resources accumulation. The maximum diameter of the rosette was measured at the nearest millimeter (DIAM; Weinig *et al.*, 2006). This trait is a proxy of the growth of the rosette of the focal plant from germination to flowering. The above-ground dry biomass (BIOMASS, expressed in grams, with a precision down to the tenth of a milligram) was estimated by drying the aboveground portion for 48H at 60°C.

We additionally quantified the strategy adopted by *A. thaliana* in response to neighboring plants by calculating the ratio HD as the height of the first flower on the rosette diameter (i.e. H1F/DIAM)**.** Values of HD below and above 1 would correspond to an aggressive and escape strategy, respectively (Baron *et al.*, 2015).

Plants that had not flowered 87 days after sowing (i.e. 1.15%) were assigned a flowering date value of 87. H1F and HD were therefore not available for these plants.

Statistical analysis

Exploring natural variation of plant-plant interactions at different levels of complexities

The following mixed model (PROC MIXED procedure, REML method, SAS 9.3, SAS Institute Inc) was used to explore the genetic variation of response among the 96 TOU-A accessions:

$$Y_{ijkl} = \mu_{trait} + block_i + treatment_j + block_i \times treatment_j + germ_l\,(treatment_j) + FLO_m\,(treatment_j) + FLO_m^2\,(treatment_j) + accession_k + treatment_j \times accession_k + \varepsilon_{ijkl} \quad (1)$$

where '*Y*' is one of the phenotypic traits scored on focal *A. thaliana* plants, 'μ' is the overall phenotypic mean; 'block' accounts for differences in micro-environment among the four experimental blocks; 'treatment' corresponds to effect of the 12 treatments (A, AAA, PPP, PPS, PPV, PSS, PSV, PVV, SSS, SSV, SVV and VVV); 'accession' measures the effect of the 91 accessions; the interaction term 'treatment x accession' accounts for genetic variation in reaction norms across the 12 treatments; the term 'germ(treatment)' is a covariate accounting for natural variation for the germination date between the 91 accessions; and 'ε' is the residual term. Because phenotypic traits were measured at flowering time, we controlled the putative

linear and non-linear effects of this phenological stage by including the linear term 'FLO' as well as the quadratic term 'FLO*FLO' in model (1).

All factors were treated as fixed effects; with the exception of the term 'accession' that was treated as a random effect. For the calculation of *F*-values, terms were tested over their appropriate denominators. Given the split-plot design used in this study, the variance associated with 'block x treatment' was used as the error term for testing the 'block' and 'treatment' effects. Model random terms were tested with likelihood ratio tests of models with and without these effects.

Heritability

Based on variance components estimated by REML (PROC VARCOMP procedure in SAS 9.3, SAS Institute Inc.), the broad-sense heritability of each phenotypic trait (H^2_{trait}) was estimated within each treatment using the following model:

$$Y_{ik} = \mu_{trait} + block_i + accession_k + \varepsilon_{ik} \qquad (2)$$

$$\text{as } H^2_{trait} = V_F / (V_F + (V_R/n))$$

where V_F is the estimated between-accession variance component, V_R is the residual variance and 'n' is the number of replicates per accession.

Significance of H^2_{trait} was assessed by testing the significance of the term '$accession_k$' by fitting model (2) using the PROC MIXED procedure in SAS 9.3 (REML method).

Estimation of genotypic values

- For each treatment, Best Linear Unbiased Predictors (BLUPs) were obtained for each accession using the following model (PROC MIXED procedure, REML method, SAS 9.3, SAS Institute Inc).

$$Y_{iklm} = \mu_{trait} + block_i + accession_k + germ_l + FLO_m + FLO_m^2 + \varepsilon_{iklm} \quad (3)$$

Because *A. thaliana* is a highly selfing species (Platt *et al.*, 2010), BLUPs correspond to genotypic values of accessions.

Genetic correlations

For each phenotypic trait, we estimated the strength of 'accession x treatment' interactions by estimating across-environment genetic correlations for each pairwise treatment combination. Genetic correlations were estimated by calculating the Pearson correlation coefficient based on BLUPs by using the *cor.test* function implemented in the *R* environment. Significant crossing reaction norms were detected by testing whether 95% confidence intervals of Pearson's *r* were not overlapping with the value of 1.

To test whether the four phenotypic traits were not redundant, we estimated for each treatment the genetic correlation for each pair of traits, by calculating the Pearson correlation coefficient based on BLUPs as described above.

Genome wide association mapping

The effects of population structure on phenotype-genotype associations has been demonstrated to be limited in the TOU-A population (Brachi *et al.*, 2013; Baron *et al.*, 2015). Nevertheless, GWA mapping was run using a mixed-model approach implemented in the software EFFICIENT MIXED-MODEL ASSOCIATION EXPEDITED (EMMAX, H.M. Kang

et al., 2010). This model includes a genetic kinship matrix as a covariate to control for the effect of the demographic history of the TOU-A population. This kinship matrix was estimated on the whole set of 1,692,194 SNPs detected among the 91 accessions.

In this study, we discarded SNPs with more than 16 missing values across the 91 accessions. In addition, because rare alleles may lead to an inflation of low *p*-values (Atwell *et al.*, 2010; Brachi *et al.*, 2010; H.M. Kang *et al.*, 2010), we only considered SNPs with a minor allele relative frequency (MARF) > 10%, leaving us with 630,234 SNPs.

Genetic architecture of plant-plant interactions

For each trait, we compared the genetic architecture among the 12 treatments by focusing on the 200 most associated SNPs (i.e. top SNPs) of each treatment. This number of top SNPs represents ~ 0.03% of the total number of SNPs and has been previously demonstrated to be appropriate and conservative to describe in the TOU-A population the genetic architecture of a set of 29 complex quantitative traits related to phenology, development and fecundity (Frachon *et al.*, 2017). The degree of flexibility of the genetic architecture among the 12 treatments was estimated by calculating the degree of environmental pleiotropy of a given top SNP, which is defined as the number of treatments that shared this top SNP (Wang *et al.*, 2010). Specific comparisons within a subset of treatments were illustrated by Venn diagrams using the *jvenn* online plug-in (Bardou *et al.*, 2014).

Identification of candidate genes associated with response to monospecific and plurispecific interactions

To identify candidate genes associated with plant-plant interactions, we selected the 20 top SNPs for each 'trait x treatment' combination. Following Frachon *et al.* (2017), we then retrieved all the annotated genes located within or overlapping with a 2kb region around each

top SNP, using the TAIR 10 database (https://www.arabidopsis.org/), leaving us with 369 unique candidate genes.

To identify the biological processes involved in response to the different neighborhoods, we first merged the list of candidate genes across the four phenotypic traits according to the four following categories: Control = A; Intraspecific = AAA; Monospecific = PPP, SSS and VVV; Plurispecific = PPS, PSS, PPV, PVV, SSV, SVV and PSV. Based on the MAPMAN classification (Provart & Zhu 2003), the four resulting lists of unique candidate genes were then submitted to the classification superviewer tool (http://bar.utoronto.ca/ntools/cgi-bin/ntools_classification_superviewer.cgi) to identify the biological processes significantly over-represented ($P < 0.01$). Only candidate genes from significantly enriched biological processes were further considered for a functional interpretation of the molecular mechanisms involved in plant-plant interactions in our study.

Results

Extent of natural genetic variation of response to monospecific and plurispecific interactions

Highly significant genetic variation was found across the 12 treatments for the four phenotypic traits scored on focal *A. thaliana* plants (Table 1). Similar results were observed without considering the treatments where *A. thaliana* was grown alone or with clones (Tables S1 and S2). Broad-sense heritability values estimated for the four phenotypic traits within each treatment were all highly significant (Supporting information Table S3) and ranged from 0.65 to 0.96 (mean = 0.82, median = 0.81; Supporting information Table S4), suggesting that a large fraction of the phenotypic variation observed within each treatment was driven by genetic differences among the local *A. thaliana* accessions (Figure 1). For each treatment, genetic

correlations between the four phenotypic traits were significantly different from the unity (absolute values of Pearson's *r*: min = -0.72, max = 0.86, mean = -0.02, median = -0.21), suggesting that the traits scored in this study partly behave independently (Supporting information Figure S1).

Table 1: Natural variation of four phenotypic traits scored on *A. thaliana* plants in 12 treatments. Bold *P*-values indicate significant effects after FDR correction. Model random terms were tested with likelihood ratio tests (LRT) of models with and without these effects. Random effects are in italics. H1F: height from the soil to the first flower on the main stem, DIAM: maximum diameter of the rosette, BIOMASS: aboveground dry biomass, HD = H1F / DIAM.

	Traits							
	H1F		**DIAM**		**BIOMASS**		**HD**	
Model terms	*F or LRT*	*P*	*F or LRT*	*P*	*F or LRT*	*P*	*F or LRT*	*P*
block	3.51	**3.27E-02**	0.54	6.65E-01	3.69	**2.86E-02**	0.70	6.01E-01
treatment	24.37	**2.55E-32**	3.34	**2.10E-04**	31.35	**2.55E-32**	1.46	1.64E-01
accession	969.70	**9.70E-212**	821.20	**1.24E-179**	556.70	**3.07E-122**	1094.40	**1.53E-238**
*treatment*accession*	14.10	**2.55E-04**	39.40	**6.04E-10**	174.90	**3.53E-39**	33.20	**1.37E-08**
germ(treatment)	1.61	1.01E-01	0.79	6.65E-01	2.42	**5.59E-03**	0.90	6.01E-01
flo(treatment)	53.12	**2.55E-32**	9.56	**4.66E-18**	52.68	**2.55E-32**	8.70	**4.02E-16**
flo*flo(treatment)	39.04	**2.55E-32**	7.59	**1.06E-13**	18.26	**2.55E-32**	6.97	**2.54E-12**

Importantly, as evidenced by highly significant 'Treatment x Accession' interactions, strong genetic variation of reaction norms was found for the four phenotypic traits, with or without considering the treatments where *A. thaliana* was grown alone or with clones (Table 1, Supporting information Tables S1 and S2). Across the four phenotypic traits, across-environment genetic correlations ranged from 0.19 to 0.86 (mean = 0.57, median = 0.57; Figure 2), indicating that the rank of accessions largely differed among the 12 treatments. This pattern of crossing reaction norms is well illustrated for the height from the soil to the first flower (Figure 1).

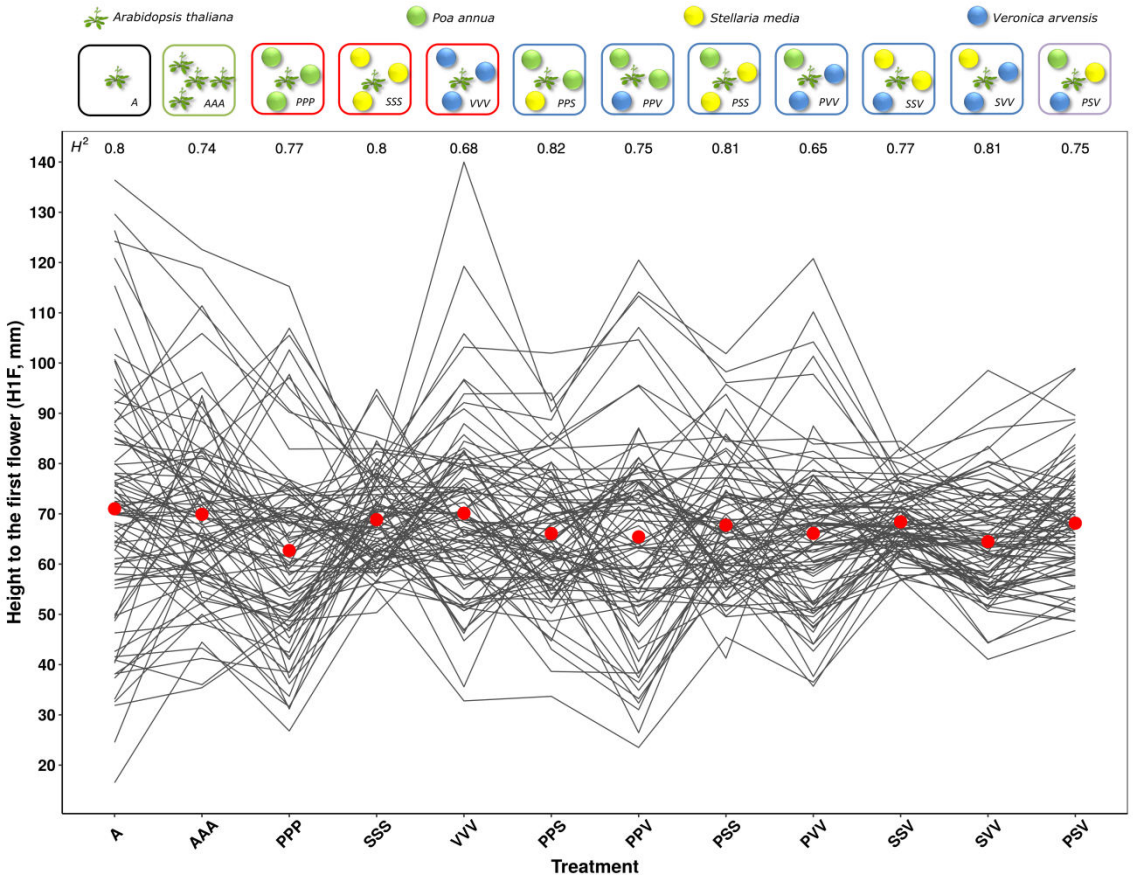

Figure 1. Natural genetic variation of reactions norms of 89 TOU-A accessions across 12 plant-plant interaction treatments. (**a**) Diagram illustrating the 12 treatments. (**b**) Height from the soil to the first flower on *A. thaliana* (H1F). Each line links the genotypic values of one of 89 TOU-A accessions. The two remaining accessions A1-69 and A1-117 are not represented due to missing BLUP values in the PPV and SSS treatments, respectively. For a given treatment, the mean H1F genotypic value among the accessions is represented by a red dot.

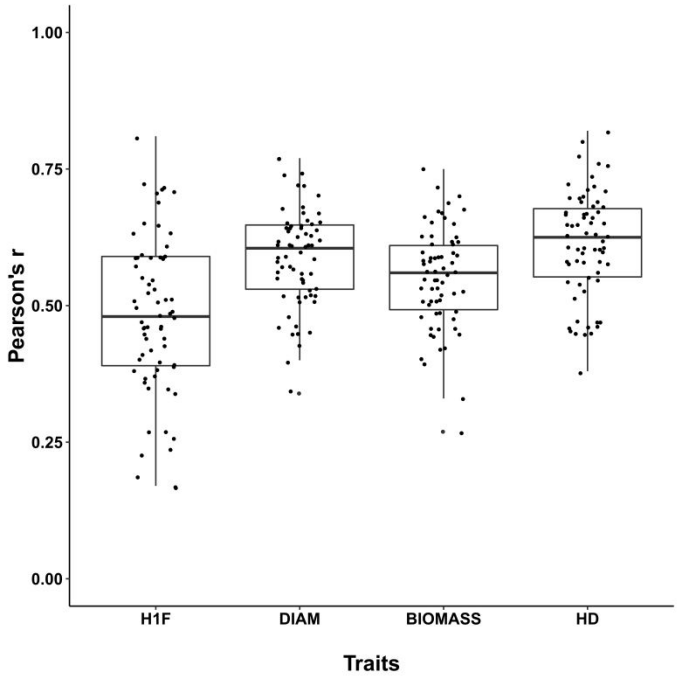

Figure 2. Pairwise genetic correlation coefficients of Pearson among the 12 treatments for each phenotypic trait. The dots correspond to the 66 pairwise treatment combinations.

Genetic architecture revealed by GWA mapping

To compare the genetic architecture underlying the competitive response of *A. thaliana* between monospecific and plurispecific neighborhoods, we used GWA mapping based on 630,234 SNPs (i.e. 1 SNP every 189bp). Following the methodology of a previous GWAS performed on the local TOU-A population (Frachon *et al.*, 2017), we described the genetic architecture by extracting for each treatment the 200 top SNPs associated with each of the four phenotypic traits (Supporting information Figure S2), leading to a final set of 4091 unique SNPs.

Across the 12 treatments, the degree of environmental pleiotropy of a given top SNP followed an L-shaped distribution (Figure 3). For the traits H1F, DIAM and HD, more than 75.6% of top SNPs were specific to a single treatment, indicating that the genetic bases are

largely distinct among the 12 treatments (Figure 3). The genetic architecture was less flexible among the 12 treatments for the trait BIOMASS with (i) the detection of less than 52% of top SNPs being specific to a single treatment and (ii) the identification of top SNPs shared among up to eight treatments (Figure 3). This latter observation likely resulted from the detection of a common association peak at the beginning of chromosome 4 between the 12 treatments (Supporting information Figure S2).

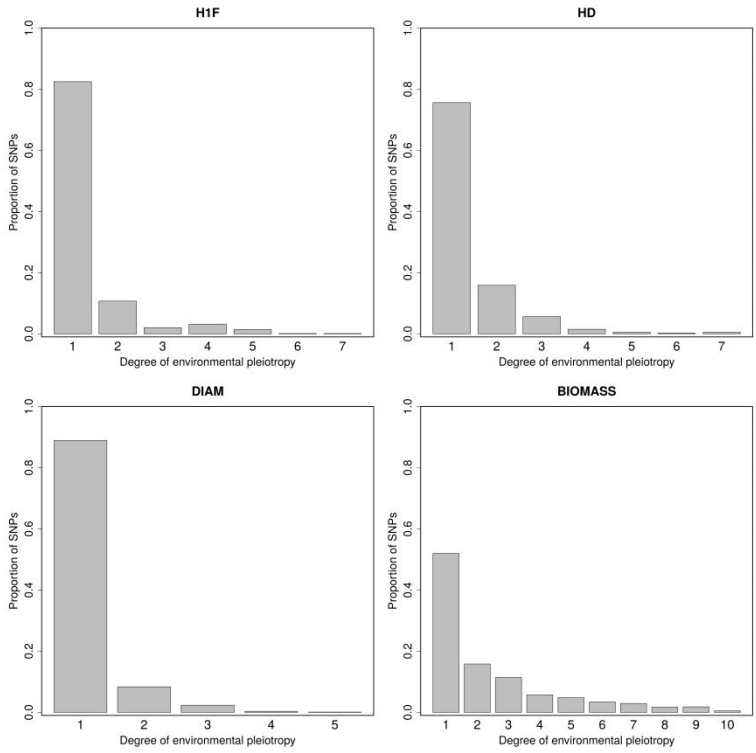

Figure 3. Degree of flexible genetic architecture of *A. thaliana* among the 12 treatments when considering a threshold of 200 top SNPs. For each phenotypic trait, bar plots represent the frequency distribution of the number of treatments that share a top SNP. H1F: height from the soil to the first flower on the main stem, DIAM: maximum diameter of the rosette, BIOMASS: aboveground dry biomass, HD = H1F / DIAM.

Therefore, in this study, the genetic architecture largely depends on both the composition and assemblage of the neighborhood of *A. thaliana*. Firstly, for the traits H1F, DIAM and HD, the genetic architecture of competitive response of *A. thaliana* to monospecific interactions was highly dependent on the identity of the neighboring species (Supporting

information Figures S3, S4 and S5). As illustrated for HD, on average 79% of top SNPs were specific to either the control treatment or one of the three monospecific interaction treatments (Figure 4). On the other hand, less than half of top SNPs (i.e. 44.5%) associated with BIOMASS were specific either the control treatment or one of the three monospecific interaction treatments (Supporting information Figure S6).

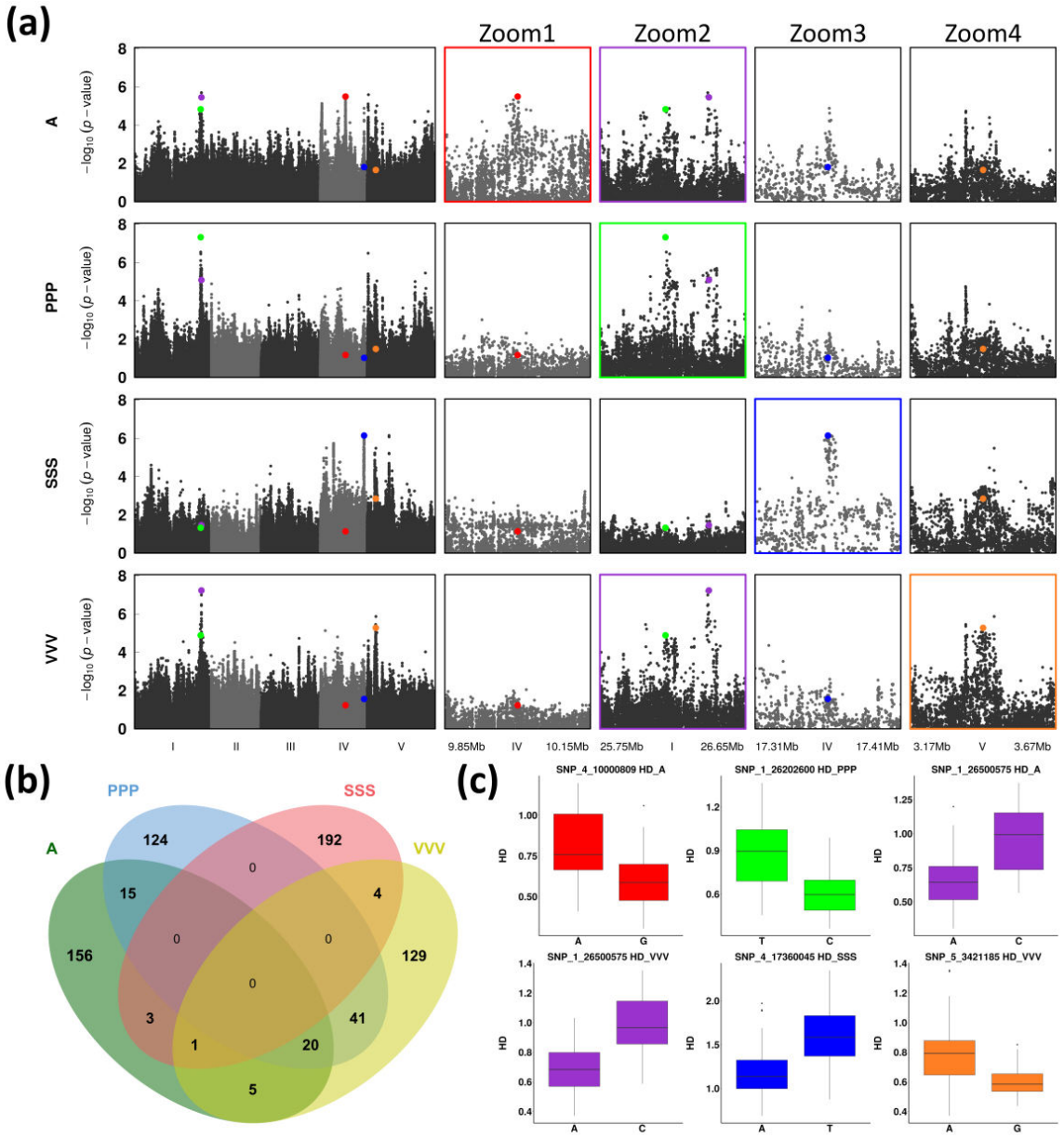

Figure 4. Identification of genomic regions associated with monospecific interactions for the ratio 'height of the first flower / the rosette diameter' (HD) in the TOU-A population. (**a**) Left panel: Manhattan plots of GWA mapping results for the A, PPP, SSS and VVV treatments. The *x*-axis indicates the physical position of the 630,234 SNPs along the five chromosomes. The *y*-axis indicates the -$\log_{10}$ *p*-values using the mixed model implemented in the software EMMAX using SNPs with MARF > 10% and missing data < 75%. Mid-panel and right panel: zooms on top SNPs illustrating the degree of specificity of genetic architecture among the treatments A (zoom1, red dot; zoom2, purple dot), PPP (zoom2, green dot), SSS (zoom3, blue dot) and VVV (zoom4, orange dot). (**b**) Venn diagram partitioning the HD SNPs detected among the lists of 200 top SNPs for the A, PPP, SSS and VVV treatments. **(c)** Box-plots illustrating the effects of the five top SNPs colored in panel (a) in their respective treatment.

Secondly, the genetic architecture of competitive response of *A. thaliana* was also highly dependent on the three-way combination of *P. annua*, *S. media* and *V. arvensis*. In particular, for the traits H1F, DIAM and HD, more than 93% of top SNPs associated with the response of *A. thaliana* to the simultaneous presence of *P. annua*, *S. media* and *V. arvensis* were not present in the sets of 200 top SNPs identified in the three monospecific interaction treatments (Supporting information Figures S3, S4 and S5). For example, a very neat peak of association for HD was identified at the beginning of chromosome 3 in the treatment PSV but not in the treatments PPP, SSS and VVV (Figure 5a).

Thirdly, the identity of top SNPs associated with the response of *A. thaliana* to the presence of a specific pair of neighboring species largely differed not only from the top SNPs identified in the corresponding monospecific interaction treatments (as previously observed for the treatment PSV), but also between the two assemblages based on this pair of neighboring species (Supporting information Figure S2). In the latter case, less than 16% of top SNPs were shared between treatments with the same composition (i.e. two neighboring species) but with different assemblages (i.e. two plants of species A + 1 plant of species B *vs* one plant of species A + two plants of species B) (Supporting information Figure S2). As an illustration, we detected a neat association peak for H1F at the end of chromosome 5 in the treatment PSS (i.e. one *P. annua* individual + two *S. media* individuals) but neither in the treatment PPS (i.e. two *P. annua* individuals + one *S. media* individual), nor in the treatments PPP and SSS (Figure 5b).

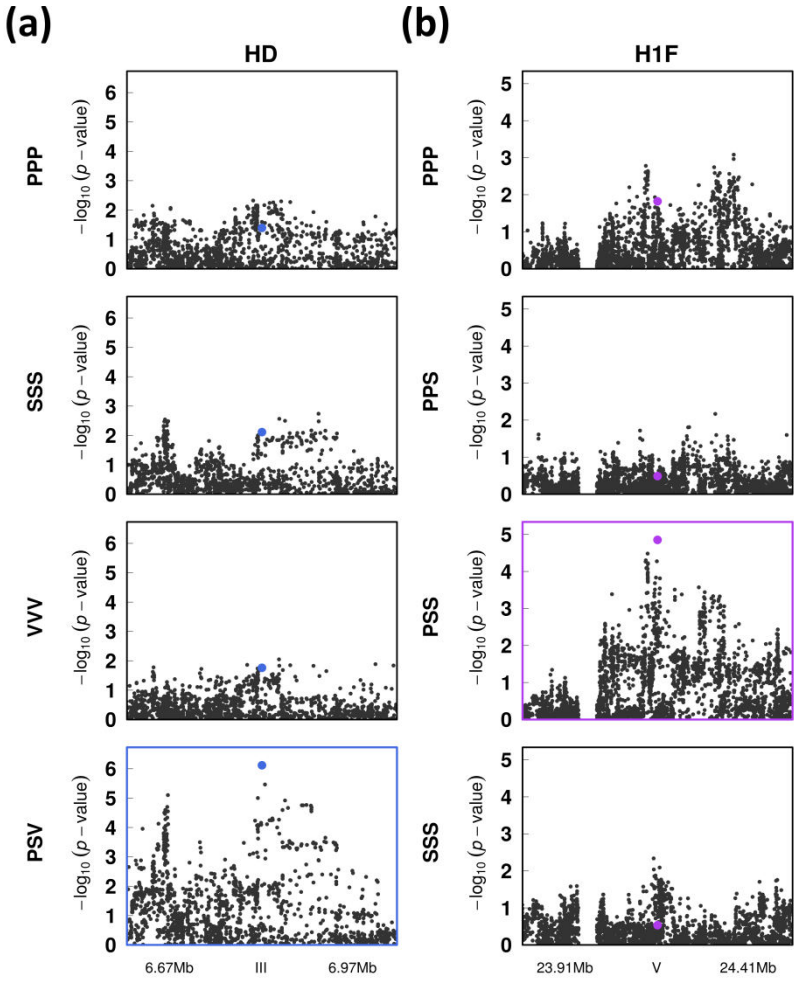

Figure 5. Comparison of the genetic architecture between monospecific and plurispecific interactions in the TOU-A population. (**a**) Zooms on an association peak identified for the ratio 'height of the first flower / the rosette diameter' (HD), which is specific to the plurispecific interaction treatment PSV. (**b**) Zooms illustrating an association peak identified for the height from the soil to the first flower (H1F), which depends on the assemblage between the neighboring species *P. annua* and *S. media*. The *x*-axis indicates the physical position of the SNPs along the considered genomic region. The *y*-axis indicates the -$\log_{10}$ *p*-values using the mixed model implemented in the software EMMAX using SNPs with MARF > 10% and missing data < 75%.

Identification of enriched biological processes and underlying candidate genes

We retrieved 369 unique genes located within or overlapping a 2kb region around the 20 top SNPs of each 'trait x treatment' (Dataset1). Considering this entire set of genes, only the 'transport' class was significantly over-represented in frequency compared to the overall class

frequency in the *Arabidopsis thaliana* MapMan annotation (Normed Frequency = 1.87, Supporting information Table S5), whereas the 'DNA' and 'not assigned' classes were significantly under-represented (Normed Frequency = 0.17 and 0.86, respectively, Supporting information Table S5).

Among the four interaction categories (i.e. Control, Intraspecific, Monospecific and Plurispecific), we observed strong differences in the number and identity of enriched biological processes (Supporting information Table S6). No biological process was found significantly enriched for the 'Control' and 'Intraspecific' interaction categories (Supporting information Table S6). In contrast, significantly enriched biological processes were detected in the context of interspecific interactions. For the 'Monospecific' interaction category, we detected a significant enrichment for the 'tetrapyrrole synthesis' class, which was represented by the genes *HEME OXYGENASE 3* (*HO3*) and *HEMB1* (Supporting information Table S7), both identified for HD in the PPP treatment. For the 'Plurispecific' interaction category, we identified three significantly enriched biological processes, i.e. signaling, transport and DNA with 18, 16 and 5 underlying candidate genes, respectively (Supporting information Table S6). In the 'signaling' class, 12 out of the 18 candidate genes encode receptor-like kinases (RLKs), i.e. one wall-associated receptor kinase (WAK), two proline-rich extensin-like receptor kinases (PERKs), two cystein-rich receptor-like kinases (CRKs) and seven leucine-rich repeat receptor-like protein kinases (LRR-RLKs) (Supporting information Table S6). The six remaining proteins were related to signaling (EPS15 HOMOLOGY DOMAIN 1 (EHD1), NO POLLEN GERMINATION 1 (NPG1) and IQ-DOMAIN 17 (IQD17)), a G-box family protein (G-BOX REGULATING FACTOR 6, GRF6), an exordium like protein (EXORDIUM LIKE 2, EXL2) and the RPM1-interacting protein 4 family protein (AT5G40645). In the 'transport' class, we identified several genes encoding diverse transporters (i) four ATP-BINDING CASSETTE

transport proteins (ABC transporters), which were P-GLYCOPROTEIN 18 (PGP18), ATP-BINDING CASSETTE A3 (ABCA3), ATP-BINDING CASSETTE F2 (ABCF2) and PLEIOTROPIC DRUG RESISTANCE 13 (PDR13) and (ii) two homologous phosphate transport proteins (PHOSPHATE 1 (PHO1) and its homolog PHO1;H1). We also identified four proteins related to the transport of mineral nutrients, i.e. one protein related to copper transport (COPPER TRANSPORTER 3, COPT3), two proteins related to magnesium transport (MAGNESIUM TRANSPORTER 4 (MRS2-3) and a magnesium transporter CorA-like family protein (*AT5G09710*)) and a nitrate transporter (NRT1 PTR FAMILY 5.13, NPF5.13). The six other genes are related to the transport of purine, calcium, sugars and peptides (Supporting information Table S7). In the 'DNA' class, the five candidate genes correspond to the DNA topoisomerase VI sub-unit A SPORULATION 11-1 (SPO11-1), a DNA glycosylase (AT3G50880), the RNA HELICASE-LIKE 8 (RH8), the DNA LIGASE IV (LIG4) and a histone superfamilly protein (AT5G59970).

Discussion

Because a focal plant rarely interacts with only one neighboring species either in crop fields or in more natural environments, the genetics of plant-plant interactions need to be studied in a community context. In this study, we compared the genetic architecture of the competitive response of *A. thaliana* between monospecific and plurispecific neighborhoods. To achieve this goal, we adopted a GWA mapping approach combined with the modern standards of ecological genomics. Indeed, the geographical scale at which selective agents act on a species should determine the mapping population used to identify genomic regions associated with ecologically relevant trait phenotypic variation (Bergelson & Roux 2010; Brachi *et al.*, 2013; Roux & Bergelson 2016). Because plants interact with neighbors over short distances, we focused on a highly genetically polymorphic local population of *A. thaliana* known to interact *in situ* with the three neighboring species considered in this study.

A flexible genetic architecture for competitive response between monospecific and plurispecific neighborhoods

Several studies reported extensive genetic variation of the competitive ability of *A. thaliana* in the context of pairwise heterospecific interactions, both at the worldwide and local scales (Bossdorf *et al.*, 2009; Brachi *et al.*, 2012; Bartheleimer *et al.*, 2015; Baron *et al.*, 2015; Frachon *et al.*, 2017). Based on four phenotypic traits related to resource accumulation and life-history trait such as seed dispersal (Reboud *et al.*, 2004; Wender *et al.*, 2005), we also found extensive local genetic variation of the competitive ability of *A. thaliana* in all the plurispecific neighborhoods tested in this study. More importantly, we detected strong crossing reaction norms not only among the three monospecific interaction treatments, but also among the

different plant assemblages surrounding the focal *A. thaliana* accessions. Altogether, these results suggest that the simultaneous interactions of *A. thaliana* with several plant partners can promote maintenance of the high genetic diversity observed in the TOU-A local population (i.e. only 5.6 times less than observed in a panel of 1,135 worldwide accessions) (Frachon *et al.*, 2017). This diversity may in turn confer a high potential for *A. thaliana* to respond to future modifications of the assemblages in the TOU-A plant community.

In agreement with the strong crossing reaction norms observed among the three monospecific interaction treatments, a GWA mapping approach reveals that the genetic architecture of *A. thaliana* competitive response in monospecific neighborhoods was highly dependent on the identity of the neighbor species. Similar results were observed (i) in *A. thaliana* challenged with four neighboring species (including the three species used in this study) in field conditions (Baron *et al.*, 2015) and (ii) in *Oryza sativa* challenged with the three weed species *Echinochloa oryzicola*, *Monochoria korsakowii* and *Schoenoplectus juncoides* in greenhouse conditions (Onishi *et al.*, 2018). Altogether, these results suggest that the effect of the identity of the neighboring species on the genetic architecture of competitive response is conserved among focal plant species and across diverse phenotyping environments. The genetic architecture of *A. thaliana* competitive response was also highly flexible between monospecific and plurispecific neighborhoods, suggesting that the genetic response of a particular accession of *A. thaliana* in a plurispecific neighborhood can be hardly predicted from the additivity of its genetic responses observed in the corresponding monospecific neighborhoods.

Three non-exclusive and interconnected hypotheses can be proposed to explain the emergence of new QTLs in plurispecific neighborhoods (Figure 6). They are based on (i) the putative generation of new signals or elimination/modification of pre-existing signals (e.g. light, aerial volatile organic compounds, root exudates and nutrient availability) in the context

of plurispecific interactions, and (ii) their perception by *A. thaliana*. Firstly, the amount of some signals produced by the neighboring species A is strongly reduced in a plurispecific neighborhood, thereby leading to the strong reduction of the perception/signaling events activated in the context of monospecific interactions. Secondly, the production of new or modified signals in a given neighboring species is triggered by the presence of another neighboring species. Thirdly, the simultaneous presence of different signals produced by the neighboring species A and B leads to the generation of a new signal. The two latter cases correspond to the production of new signals that emerged from higher-order interactions among the neighboring species (Levine *et al.*, 2017). Natural variation in genes involved in the perception, signaling and genetic program triggered by a new set of signals can explain the emergence of new QTLs in plurispecific neighborhoods (Figure 6).

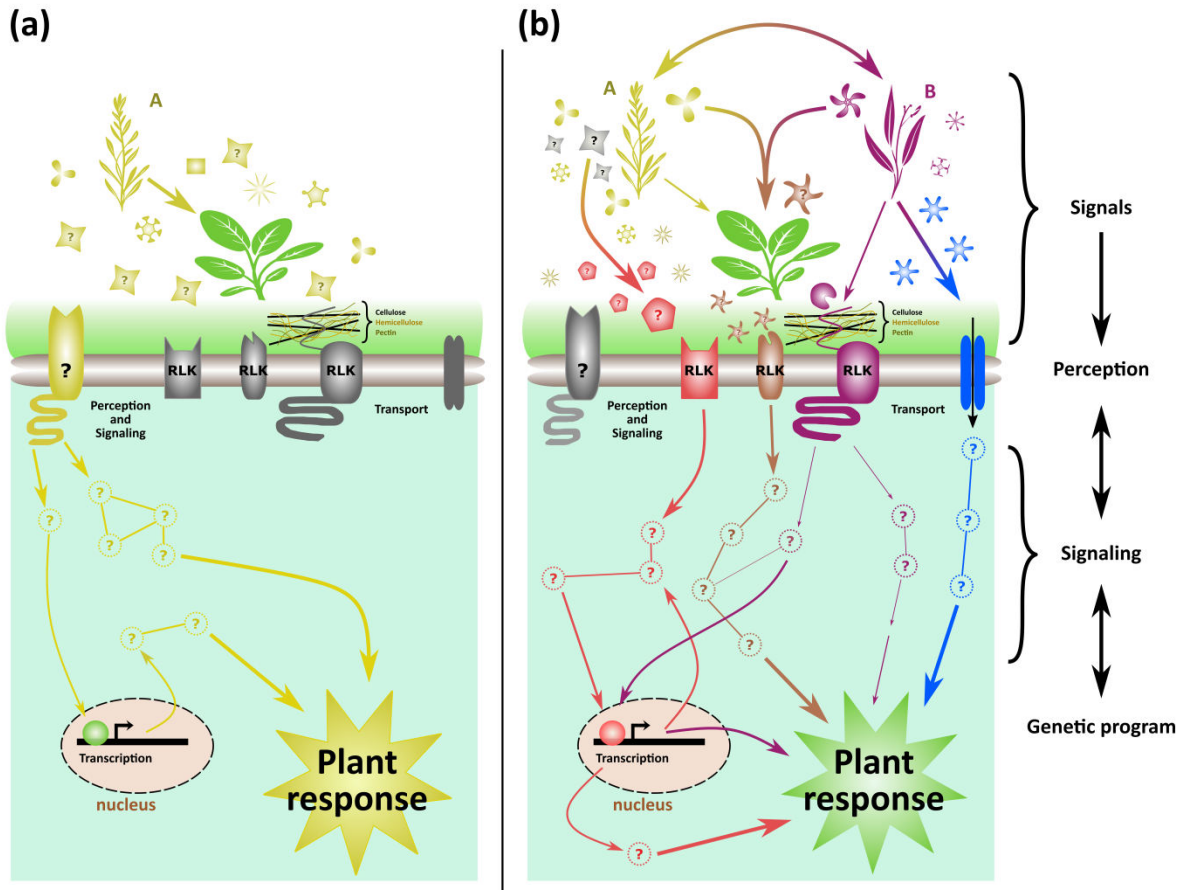

Figure 6. A schematic model to explain the differential identification of QTLs between mono- and plurispecific interactions. (a) Identification of QTLs in a monospecific neighborhood. In this diagram, the species A (yellow) produces different compounds (yellow symbols). For example, one of them is perceived by a cell wall receptor of the focal plant (green). The perception of this signal triggers the activation of different signaling pathways. These putative pathways lead to a specific plant response that could be variable among accessions due to genetic diversity of mechanisms underlying perception and/or signaling and/or gene expression events. **(b) Identification of different QTLs in a plurispecific neighborhood.** To explain the identification of different QTLs in response to a plurispecific neighborhood, we propose non-exclusive scenarios that are based on the putative generation of new signals or elimination/modification of pre-existing signals. (i) The amount of some signals produced by the species A could be strongly reduced in presence of species B (purple) leading to a strong reduction of the perception/signaling events occurring through the receptor (yellow) identified in (a). (ii) Another possibility could be that the species A, in response to species B, modifies a signal(s) present in the context of monospecific interactions into a new signal (red), which is perceived and transduced by a receptor like kinase (RLK) leading to a specific plant response. (iii) Likewise, the species B might produce a new signal (signal not produced in the context of monospecific interactions, in blue) that activates a transmembrane transporter leading to a specific plant response. (iv) A fourth case could be the simultaneous presence of different signals produced by the neighboring species A and B that together create a new signal (in brown), which is also perceived and transduced leading to a specific plant response. All these examples require (i) generation and /or modification and /or elimination of signals, and (ii) genetic variation in at least one of the three following mechanisms: perception, signaling and genetic programming. Vector plant patterns have been retrieved from the Vecteezy.com website.

Biological pathways and candidate genes associated with competitive response depend on the number of neighboring species

While significant enriched biological processes were identified in the context of interspecific interactions, no biological process was found significantly over-represented for the 'Control' and 'Intraspecific' categories. Technically, this observation might be explained by the fact that multiple functions are involved in such interactions, and/or most of the implicated genes correspond to unknown functions (26/92, Supporting information Table S6), leading to the absence of enrichment of any biological process.

Only the tetrapyrrole biosynthesis pathway represented by the genes *HO3* and *HEMB1* was found for the 'Monospecific' interaction category. In plants, tetrapyrroles play essential roles in photosynthesis, respiration, and signal transduction (Mochizuki *et al.*, 2010; Tanaka *et al.*, 2011). The tetrapyrrole biosynthesis pathway consists in two main branches, i.e. the chlorophyll and heme branches. In the chlorophyll biosynthetic pathway, two homologous transcription factors essential for phyA signaling, FAR-RED ELONGATED HYPOCOTYL 3 (FHY3) and FAR-RED IMPAIRED RESPONSE 1 (FAR1), activate *HEMB1* that in turn regulates chlorophyll biosynthesis and seedling growth (Tang *et al.*, 2012). *HO3* encodes a haem oxygenase protein. Haem oxygenases have recently emerged as players in plant cell protection to oxidative damage. In this context, *HO3* is involved in salinity tolerance in *A. thaliana* by controlling K+ retention (Bose *et al.*, 2013). Altogether, these findings suggest that photosynthesis, and more widely light perception and signaling, might be essential in the case of monospecific interactions.

In the 'Plurispecific' interaction category, the major over-represented biological process was related to signaling processes and mainly composed by RLKs (receptor like kinases).

Because RLKs can bind a large variety of ligands, they play essential roles in many plant processes such as plant immunity, development and growth (Tang *et al*., 2017). RLKs tend to be significantly over-represented among genes up-regulated under abiotic (UV-B, wounding and osmotic stress) and biotic (symbiotic or pathogenic interactions) stress conditions (Lehti-Shiu *et al.*, 2009; Tang *et al*., 2017). Our findings suggest that RLKs might also be key players in plant-plant interactions. Several of our candidate RLKs have been reported to be essential in plant development and morphogenesis. TMK1 (TRANSMEMBRANE KINASE 1) interacts with the auxin binding protein ABP1 and activates plasma membrane–associated ROPs (Rho-like guanosine triphosphatases (GTPase)), which control cytoskeleton modifications and the shape of leaf pavement cells in *A. thaliana* (Dai *et al.*, 2013; Xu *et al.,* 2014). IKU2 (HAIKU2) controls endosperm proliferation and seed size (X. Kang *et al.*, 2013). While RUL1 (REDUCED IN LATERAL GROWTH 1) is involved in secondary root growth (Agusti *et al.,* 2011), PERK13 (PROLINE EXTENSIN LIKE RLK) acts as a negative regulator of root hair growth (Y. Hwang *et al.*, 2016). Some other candidate RLKs detected in our study are related to abiotic stress responses: RPK1 is involved in a protein complex governing superoxide production and signaling at the cell surface and controlling senescence and cell death (Koo *et al*., 2017); CRK8 is transcriptionally regulated in response to light stress and ozone (Wrzaczek *et al.,* 2010) and EHD1 has been shown to confer salt tolerance when it is over-expressed in *A. thaliana* (Bar *et al.*, 2013). All these functions and others still unknown might participate to plant response to competition either *via* developmental or stress responses. Identification of the corresponding ligands together with mutant phenotyping in the context of plant-plant interactions should shed some light on the molecular mechanisms underlying these interactions.

The second major over-represented biological process in the ‘Plurispecific’ interaction category is related to transport functions. Interestingly, 4 out of the 16 transport related proteins

correspond to ABC transporter proteins. Although none of them have been functionally characterized, these transporters have been recently proposed as key players of plant adaptation to their abiotic or biotic environment (J.U. Hwang *et al*., 2016). Because of their diverse substrate specificities, they might constitute essential components of perception/signaling pathways activated during plurispecific plant-plant interactions Moreover, two proteins related to phosphate transport, PHO1 and its homolog PHO1;H1, were identified. The *pho1* mutant of *A. thaliana* exhibits inorganic phosphate (Pi) deficiency in the Pi export from roots to shoots, resulting in strong Pi deficiency in above-ground tissues (Hamburger *et al.*, 2002). *PHO1;H1* can complement the *pho1* mutant revealing some functional redundancy between these two proteins (Stefanovic *et al.*, 2007). Interestingly, *PHO1* has been identified in a genome-wide association study as a candidate gene underlying natural variation in root architecture and shown to be involved in lateral root plasticity response *via* its interplay with different signals (Rosas *et al*., 2013). This finding is particularly interesting in regard to the potential role of root architecture and Pi signaling in plant development in the context of interspecific interactions. Finally, the identification of candidate genes associated with copper (*COPT3*), magnesium (*MRS2-3* and *AT5G09710*) and nitrate (*NPF5.13*, Léran *et al.*, 2014) transport indicates that nutrient foraging might also be a major response strategy in the context of plurispecific plant-plant interactions (Pierik *et al.*, 2013).

Subrahmaniam *et al.* (2018) reported seven categories of functions previously identified in artificial environments simulating plant–plant interactions: photosynthesis, hormones, cell wall modification and degradation, defense against pathogens, ABC (ATP-binding cassette) transporters, histone modification and meristem identity/life history traits. Surprisingly, only 39 genes (out of our 369 candidate genes) are related to these functional categories (Supporting information Table S8), highlighting the added value of challenging focal plants directly with

neighbor plants to identify the molecular mechanisms underlying neighbor perception, signaling and the resulting cell reprogramming.

In conclusion, our results suggest that plants can integrate and respond to different species assemblages depending on the identity and number of each neighboring species, through a large range of genes associated mainly with perception and signaling processes leading to developmental and stress responses (Figure 6). Complementarily to our GWA study, transcriptomic and proteomic analyses of *A. thaliana* plants exposed to monospecific and plurispecific neighborhoods would help to identify genes and proteins that are differentially regulated under these conditions. To our knowledge, no such studies comparing global changes in protein and gene expression between monospecific and plurispecific neighborhoods have been reported so far (Subrahmaniam *et al.*, 2018). Another step to understand the mechanisms underlying natural variation of plant-plant competitive responses would be (i) to functionally validate the identified candidate genes. This would open the way to functional analyses to investigate (ii) the nature of the signals perceived by the plant, and (iii) decipher the signaling pathways resulting from signal perception, leading to the plant response.

Acknowledgments

Special thanks are given to Cédric Glorieux, Nathalie Faure and Angélique Bourceaux for their assistance during the common garden experiment. This work was funded by a PhD fellowship from the University of Paul Sabatier Toulouse and a PhD fellowship from the University of Lille 1 – Région Nord-Pas-de-Calais to EB. This study was also supported by the LABEX TULIP (ANR-10-LABX-41; ANR-11-IDEX-0002-02).

References

Agusti, J., Lichtenberger, R., Schwarz, M., Nehlin, L., & Greb, T. (2011). Characterization of transcriptome remodeling during cambium formation identifies *MOL1* and *RUL1* as opposing regulators of secondary growth. *PLoS Genetics*, 7, e1001312.

Atwell, S., Huang, Y. S., Vilhjálmsson, B. J., Willems, G., Horton, M., Li, Y., ... & Jiang, R. (2010). Genome-wide association study of 107 phenotypes in *Arabidopsis thaliana* inbred lines. *Nature*, 465, 627-631.

Bachelet, D., Neilson, R. P., Lenihan, J. M., & Drapek, R. J. (2001). Climate change effects on vegetation distribution and carbon budget in the United States. *Ecosystems*, 4, 164-185.

Baldwin, A. H., Jensen, K., & Schönfeldt, M. (2014). Warming increases plant biomass and reduces diversity across continents, latitudes, and species migration scenarios in experimental wetland communities. *Global change biology*, 20, 835-850.

Bar, M., Leibman, M., Schuster, S., Pitzhadza, H., & Avni, A. (2013). EHD1 functions in endosomal recycling and confers salt tolerance. *PloS ONE*, 8, e54533.

Bardou, P., Mariette, J., Escudié, F., Djemiel, C., & Klopp, C. (2014). jvenn: an interactive Venn diagram viewer. *BMC bioinformatics*, 15, 293.

Baron, E., Richirt, J., Villoutreix, R., Amsellem, L., & Roux, F. (2015). The genetics of intra- and interspecific competitive response and effect in a local population of an annual plant species. *Functional Ecology*, 29, 1361-1370.

Bartelheimer, M., Schmid, C., Storf, J., Hell, K., & Bauer, S. (2015). Interspecific competition in *Arabidopsis thaliana*: a knowledge gap is starting to close. In: Lüttge U, Beyschlag W, eds. *Progress in Botany*. Vol. 76. Berlin, Germany: Springer, 303–319.

Bartoli, C., & Roux, F. (2017). Genome-wide association studies in plant pathosystems: toward an ecological genomics approach. *Frontiers in plant science*, 8, 763.

Bergelson, J., & Roux, F. (2010). Towards identifying genes underlying ecologically relevant traits in Arabidopsis thaliana. *Nature Reviews Genetics*, 11, 867-879.

Bose, J., Xie, Y., Shen, W., & Shabala, S. (2013). Haem oxygenase modifies salinity tolerance in *Arabidopsis* by controlling K+ retention via regulation of the plasma membrane H+-ATPase and by altering SOS1 transcript levels in roots. *Journal of Experimental Botany*, 64, 471-481.

Bossdorf, O., Shuja, Z., & Banta, J. A. (2009). Genotype and maternal environment affect belowground interactions between *Arabidopsis thaliana* and its competitors. *Oikos*, 118, 1541-1551.

Brachi, B., Faure, N., Horton, M., Flahauw, E., Vazquez, A., Nordborg, M., ... & Roux, F. (2010). Linkage and association mapping of *Arabidopsis thaliana* flowering time in nature. *PLoS Genetics*, 6, e1000940.

Brachi, B., Aimé, C., Glorieux, C., Cuguen, J., & Roux, F. (2012). Adaptive value of phenological traits in stressful environments: predictions based on seed production and laboratory natural selection. *PLoS ONE*, 7, e32069.

Brachi, B., Villoutreix, R., Faure, N., Hautekèete, N., Piquot, Y., Pauwels, M., ... & Roux, F. (2013). Investigation of the geographical scale of adaptive phenological variation and its underlying genetics in *Arabidopsis thaliana*. *Molecular Ecology*, 22, 4222-4240.

Chesson, P. (2000). Mechanisms of maintenance of species diversity. *Annual review of Ecology and Systematics*, 31, 343-366.

Dai, N., Wang, W., Patterson, S. E., & Bleecker, A. B. (2013). The TMK subfamily of receptor-like kinases in *Arabidopsis* display an essential role in growth and a reduced sensitivity to auxin. *PLoS ONE*, 8, e60990.

Dorn, L. A., Pyle, E. H., & Schmitt, J. (2000). Plasticity to light cues and resources in *Arabidopsis thaliana*: testing for adaptive value and costs. *Evolution*, 54, 1982-199.

Dudley, S. A. (2015). Plant cooperation. *AoB Plants*, *7*.

Frachon, L., Libourel, C., Villoutreix, R., Carrère, S., Glorieux, C., Huard-Chauveau, C., ... & Amsellem, L. (2017). Intermediate degrees of synergistic pleiotropy drive adaptive evolution in ecological time. *Nature ecology & evolution*, 1, 1551–1561.

Gilman, S. E., Urban, M. C., Tewksbury, J., Gilchrist, G. W., & Holt, R. D. (2010). A framework for community interactions under climate change. *Trends in ecology & evolution*, 25, 325-331.

Goldberg, D. E., & Barton, A. M. (1992). Patterns and consequences of interspecific competition in natural communities: a review of field experiments with plants. *The American Naturalist* 139: 771-801.

Hamburger, D., Rezzonico, E., Petétot, J. M. C., Somerville, C., & Poirier, Y. (2002). Identification and characterization of the *Arabidopsis PHO1* gene involved in phosphate loading to the xylem. *The Plant Cell*, 14, 889-902.

Huard-Chauveau, C., Perchepied, L., Debieu, M., Rivas, S., Kroj, T., Kars, I., ... & Roby, D. (2013). An atypical kinase under balancing selection confers broad-spectrum disease resistance in *Arabidopsis*. *PLoS Genetics*, 9, e1003766.

Hwang, J. U., Song, W. Y., Hong, D., Ko, D., Yamaoka, Y., Jang, S., ... & Palmgren, M. (2016). Plant ABC transporters enable many unique aspects of a terrestrial plant's lifestyle. *Molecular plant*, 9, 338-355.

Hwang, Y., Lee, H., Lee, Y. S., & Cho, H. T. (2016). Cell wall-associated ROOT HAIR SPECIFIC 10, a proline-rich receptor-like kinase, is a negative modulator of *Arabidopsis* root hair growth. *Journal of experimental botany*, 67, 2007-2022.

Kang, H. M., Sul, J. H., Service, S. K., Zaitlen, N. A., Kong, S. Y., Freimer, N. B., ... & Eskin, E. (2010). Variance component model to account for sample structure in genome-wide association studies. *Nature genetics*, 42, 348-354.

Kang, X., Li, W., Zhou, Y., & Ni, M. (2013). A WRKY transcription factor recruits the SYG1-like protein SHB1 to activate gene expression and seed cavity enlargement. *PLoS Genetics*, 9, e1003347.

Koo, J. C., Lee, I. C., Dai, C., Lee, Y., Cho, H. K., Kim, Y., ... & Park, S. J. (2017). The protein trio RPK1–CaM4–RbohF mediates transient superoxide production to trigger age-dependent cell death in *Arabidopsis*. *Cell reports*, 21, 3373-3380.

Lehti-Shiu, M. D., Zou, C., Hanada, K., & Shiu, S. H. (2009). Evolutionary history and stress regulation of plant receptor-like kinase/pelle genes. *Plant Physiology*, 150, 12-26.

Léran, S., Varala, K., Boyer, J. C., Chiurazzi, M., Crawford, N., Daniel-Vedele, F., ... & Gassmann, W. (2014). A unified nomenclature of NITRATE TRANSPORTER 1/PEPTIDE TRANSPORTER family members in plants. *Trends in plant science*, 19, 5-9.

Levine, J. M., Bascompte, J., Adler, P. B., & Allesina, S. (2017). Beyond pairwise mechanisms of species coexistence in complex communities. *Nature*, 546, 56-64.

Litrico, I., & Violle, C. (2015). Diversity in plant breeding: a new conceptual framework. *Trends in plant science*, 20, 604-613.

Martorell, C., & Freckleton, R. P. (2014). Testing the roles of competition, facilitation and stochasticity on community structure in a species-rich assemblage. *Journal of Ecology*, 102, 74-85.

Mochizuki, N., Tanaka, R., Grimm, B., Masuda, T., Moulin, M., Smith, A. G., ... & Terry, M. J. (2010). The cell biology of tetrapyrroles: a life and death struggle. *Trends in plant science*, 15, 488-498.

Neve, P., Vila-Aiub, M., & Roux, F. (2009). Evolutionary-thinking in agricultural weed management. *New Phytologist*, 184, 783-793.

Oerke, E. C., & Dehne, H. W. (2004). Safeguarding production—losses in major crops and the role of crop protection. *Crop protection*, 23, 275-285.

Onishi, K., Ichikawa, N., Horiuchi, Y., Kohara, H., & Sano, Y. (2018). Genetic architecture underlying the evolutionary change of competitive ability in Asian cultivated and wild rice. *Journal of Plant Interactions*, 13, 442-449.

Pakeman, R. J., Karley, A. J., Newton, A. C., Morcillo, L., Brooker, R. W., & Schöb, C. (2015). A trait-based approach to crop–weed interactions. *European Journal of Agronomy*, 70, 22-32.

Pierik, R., Mommer, L., & Voesenek, L. A. (2013). Molecular mechanisms of plant competition: neighbour detection and response strategies. *Functional Ecology*, 27, 841-853.

Platt, A., Horton, M., Huang, Y. S., Li, Y., Anastasio, A. E., Mulyati, N. W., ... & Dunning, M. (2010). The scale of population structure in *Arabidopsis thaliana*. *PLoS Genetics*, 6, e1000843.

Provart, N., & Zhu, T. (2003). A browser-based functional classification SuperViewer for *Arabidopsis* genomics. *Currents in Computational Molecular Biology*, 2003, 271-272.

Reboud, X., Le Corre, V., Scarcelli, N., Roux, F., David, J. L., Bataillon, T., ... & McKhann, H. (2004). Natural variation among accessions of *Arabidopsis thaliana*: beyond the flowering date, what morphological traits are relevant to study adaptation. In: QCB Cronk, J. Whitton, RH Ree and IEP Taylor, eds. *Plant Adaptation: Molecular Genetics and Ecology*. Ottawa, Ontario, Canada: NRC Research Press, 135-142.

Rosas, U., Cibrian-Jaramillo, A., Ristova, D., Banta, J. A., Gifford, M. L., Fan, A. H., ... & Purugganan, M. D. (2013). Integration of responses within and across *Arabidopsis* natural accessions uncovers loci controlling root systems architecture. *Proceedings of the National Academy of Sciences*, 110, 15133-15138.

Roux, F., & Bergelson, J. (2016). The genetics underlying natural variation in the biotic interactions of *Arabidopsis thaliana*: The challenges of linking evolutionary genetics and community ecology. *Current topics in developmental biology*, 119, 111-156.

Singer, A., Travis, J. M., & Johst, K. (2013). Interspecific interactions affect species and community responses to climate shifts. *Oikos*, 122, 358-366.

Stefanovic, A., Ribot, C., Rouached, H., Wang, Y., Chong, J., Belbahri, L., ... & Poirier, Y. (2007). Members of the *PHO1* gene family show limited functional redundancy in phosphate transfer to the shoot, and are regulated by phosphate deficiency via distinct pathways. *The Plant Journal*, 50, 982-994.

Subrahmaniam, H. J., Libourel, C., Journet, E. P., Morel, J. B., Muños, S., Niebel, A., ... & Roux, F. (2018). The genetics underlying natural variation of plant–plant interactions, a beloved but forgotten member of the family of biotic interactions. *The Plant Journal*, 93, 747-770.

Tanaka, R., Kobayashi, K., & Masuda, T. (2011). Tetrapyrrole metabolism in *Arabidopsis thaliana*. *The Arabidopsis book/American Society of Plant Biologists* 9.

Tang, D., Wang, G., & Zhou, J. M. (2017). Receptor kinases in plant-pathogen interactions: more than pattern recognition. *The Plant Cell* 29: 618-637.

Tang, W., Wang, W., Chen, D., Ji, Q., Jing, Y., Wang, H., & Lin, R. (2012). Transposase-derived proteins FHY3/FAR1 interact with PHYTOCHROME-INTERACTING FACTOR1 to regulate chlorophyll biosynthesis by modulating HEMB1 during deetiolation in *Arabidopsis*. *The Plant Cell*, 24, 1984-2000.

Tilman, D. (1985). The resource-ratio hypothesis of plant succession. *The American Naturalist*, 125, 827-852.

Wang, Z., Liao, B. Y., & Zhang, J. (2010). Genomic patterns of pleiotropy and the evolution of complexity. *Proceedings of the National Academy of Sciences*, 107, 18034-18039.

Weinig, C., Johnston, J., German, Z. M., & Demink, L. M. (2006). Local and global costs of adaptive plasticity to density in *Arabidopsis thaliana*. *The American Naturalist*, 167, 826-836.

Weiner, J., Du, Y. L., Zhang, C., Qin, X. L., & Li, F. M. (2017). Evolutionary agroecology: individual fitness and population yield in wheat (*Triticum aestivum*). *Ecology*, 98, 2261-2266.

Wender, N. J., Polisetty, C. R., & Donohue, K. (2005). Density-dependent processes influencing the evolutionary dynamics of dispersal: a functional analysis of seed dispersal in *Arabidopsis thaliana* (Brassicaceae). *American Journal of Botany*, 92, 960-971.

Wilson, J. B., Peet, R. K., Dengler, J., & Pärtel, M. (2012). Plant species richness: the world records. *Journal of vegetation Science*, 23, 796-802.

Worthington, M., & Reberg-Horton, C. (2013). Breeding cereal crops for enhanced weed suppression: optimizing allelopathy and competitive ability. *Journal of Chemical Ecology*, 39, 213-231.

Wrzaczek, M., Brosché, M., Salojärvi, J., Kangasjärvi, S., Idänheimo, N., Mersmann, S., ... & Kangasjärvi, J. (2010). Transcriptional regulation of the CRK/DUF26 group of receptor-like protein kinases by ozone and plant hormones in *Arabidopsis*. *BMC Plant Biology*, 10, 95.

Xu, T., Dai, N., Chen, J., Nagawa, S., Cao, M., Li, H., ... & Wang, W. (2014). Cell surface ABP1-TMK auxin-sensing complex activates ROP GTPase signaling. *Science*, 343, 1025-1028.

Data Accessibility

Phenotypic data will be available in the Dryad database: doi:10.5061/dryad.XXX.

Author contributions

F.R., L.A. and D.R. supervised the project. E.B., L.A. and F.R. designed the experiments. E.B. and J.L. conducted the greenhouse experiment. J.L. and E.B measured the phenotypic traits. C.L. analyzed the phenotypic traits. C.L. performed the GWA mapping. C.L. performed and analyzed the enrichment tests. C.L., D.R. and F.R. wrote the manuscript. All authors contributed to the revisions.

Supplementary Information

The genomic architecture of competitive response of *Arabidopsis thaliana* is highly flexible between monospecific and plurispecific neighborhoods.

Cyril Libourel, Etienne Baron, Juliana Lenglet, Laurent Amsellem, Dominique Roby* and Fabrice Roux*

* Correspondence**:** Corresponding Author: fabrice.roux@inra.fr & dominique.roby@inra.fr

8 Supplementary Tables

6 Supplementary Figures

Figure S1. Pairwise genetic correlation coefficients of Pearson among the four phenotypic traits for each treatment. The dots for each pair of phenotypic traits correspond to the 12 treatments.

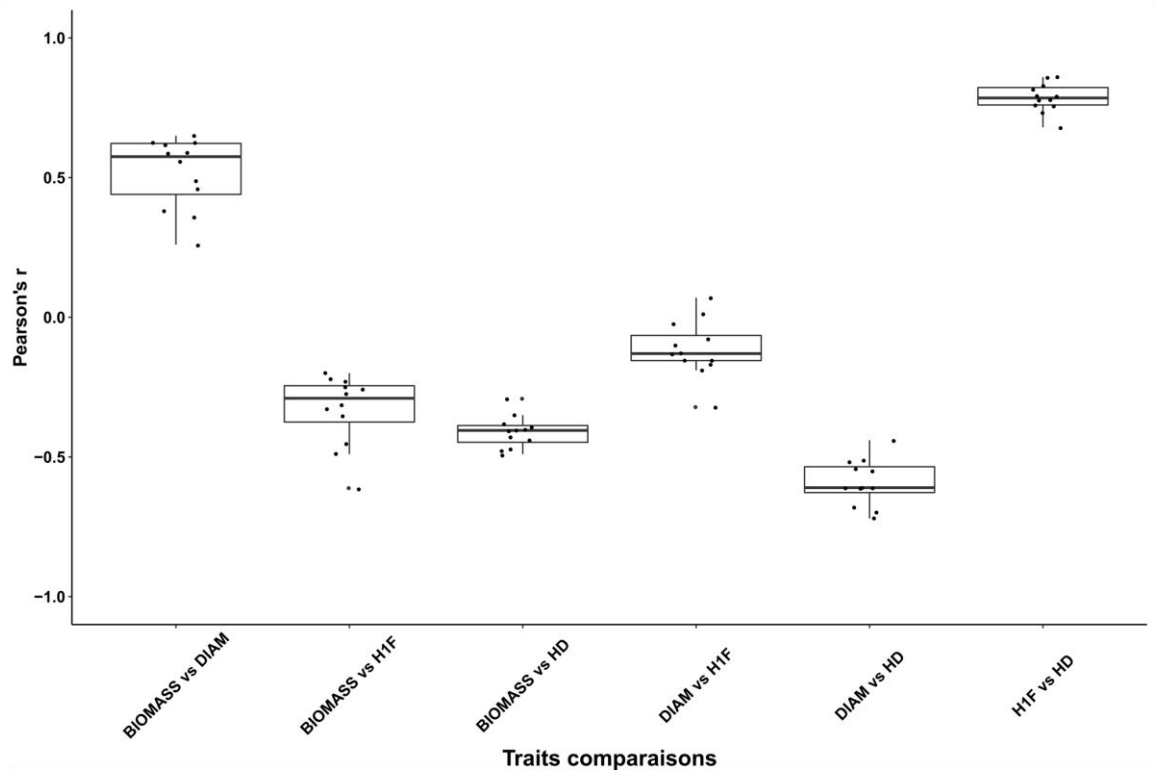

Figure S2. Identification of genomic regions associated with the four phenotypic traits scored on *A. thaliana* plants in 12 treatments. The *x*-axis indicates the physical position along the chromosome. The *y*-axis indicates the -log^{10} *p*-values using the EMMAX method. MARF > 10%. On each Manhattan plot, the 200 top SNPs are highlighted in red.

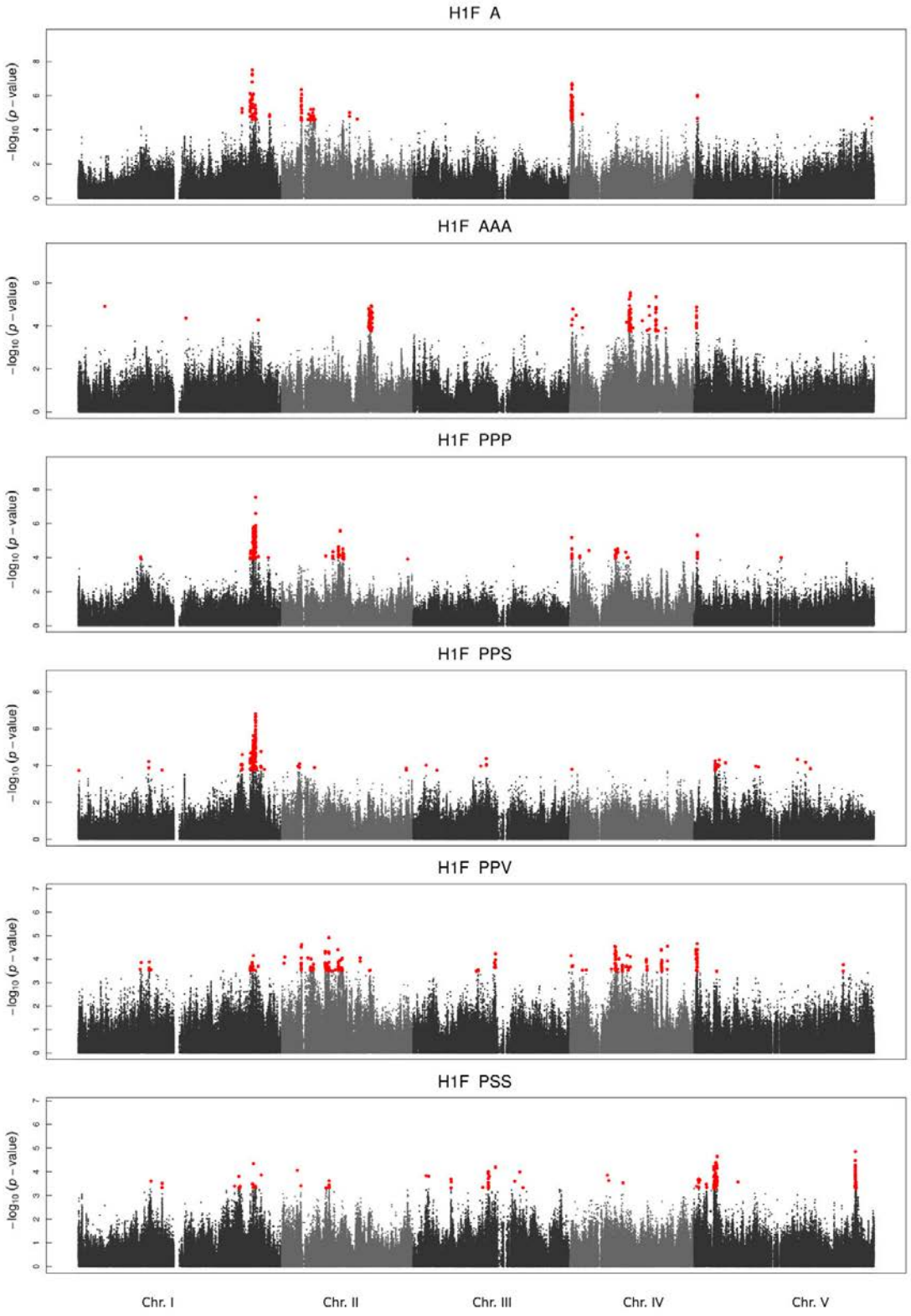

Figure S2 (continued)

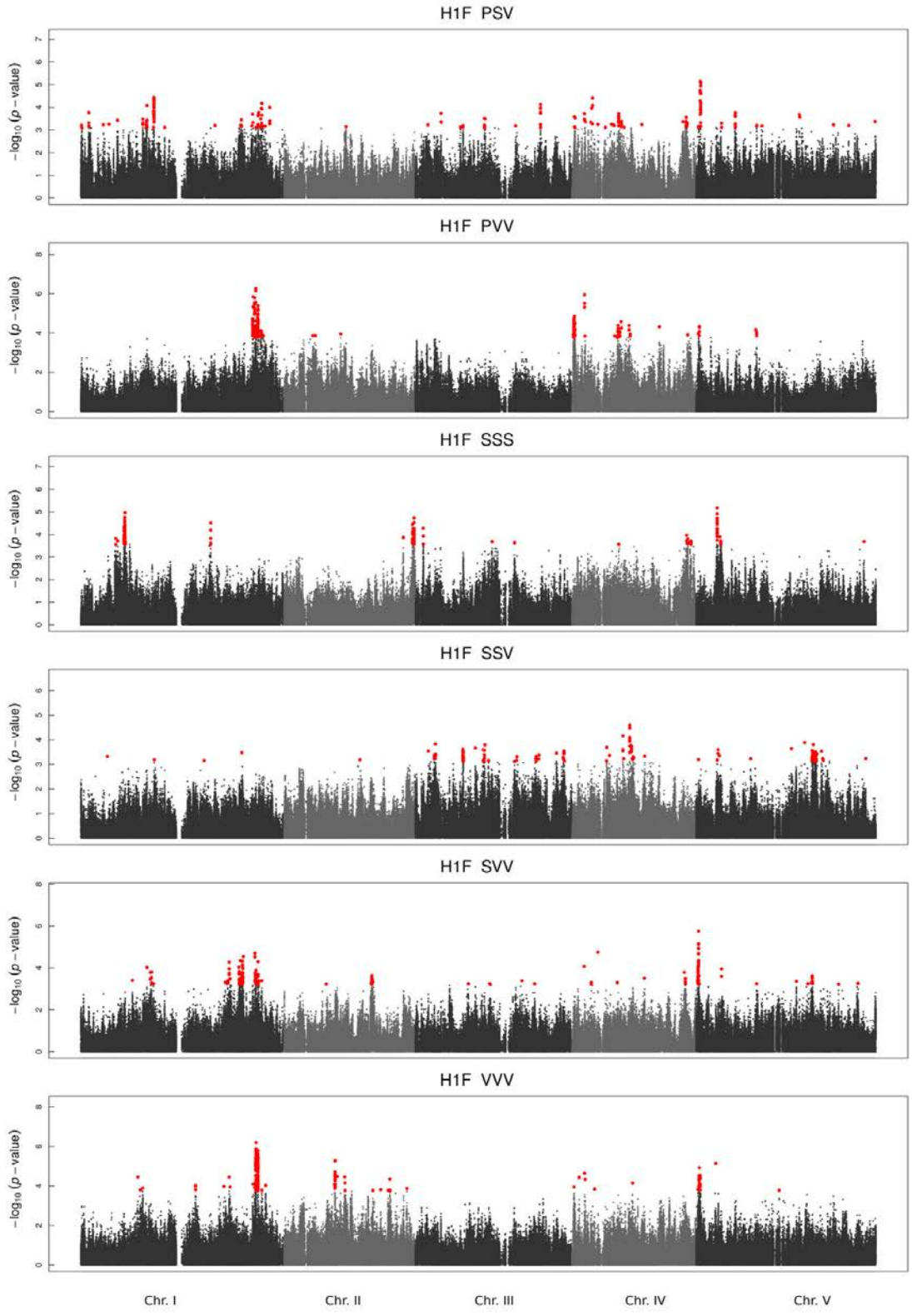

Figure S2 (continued)

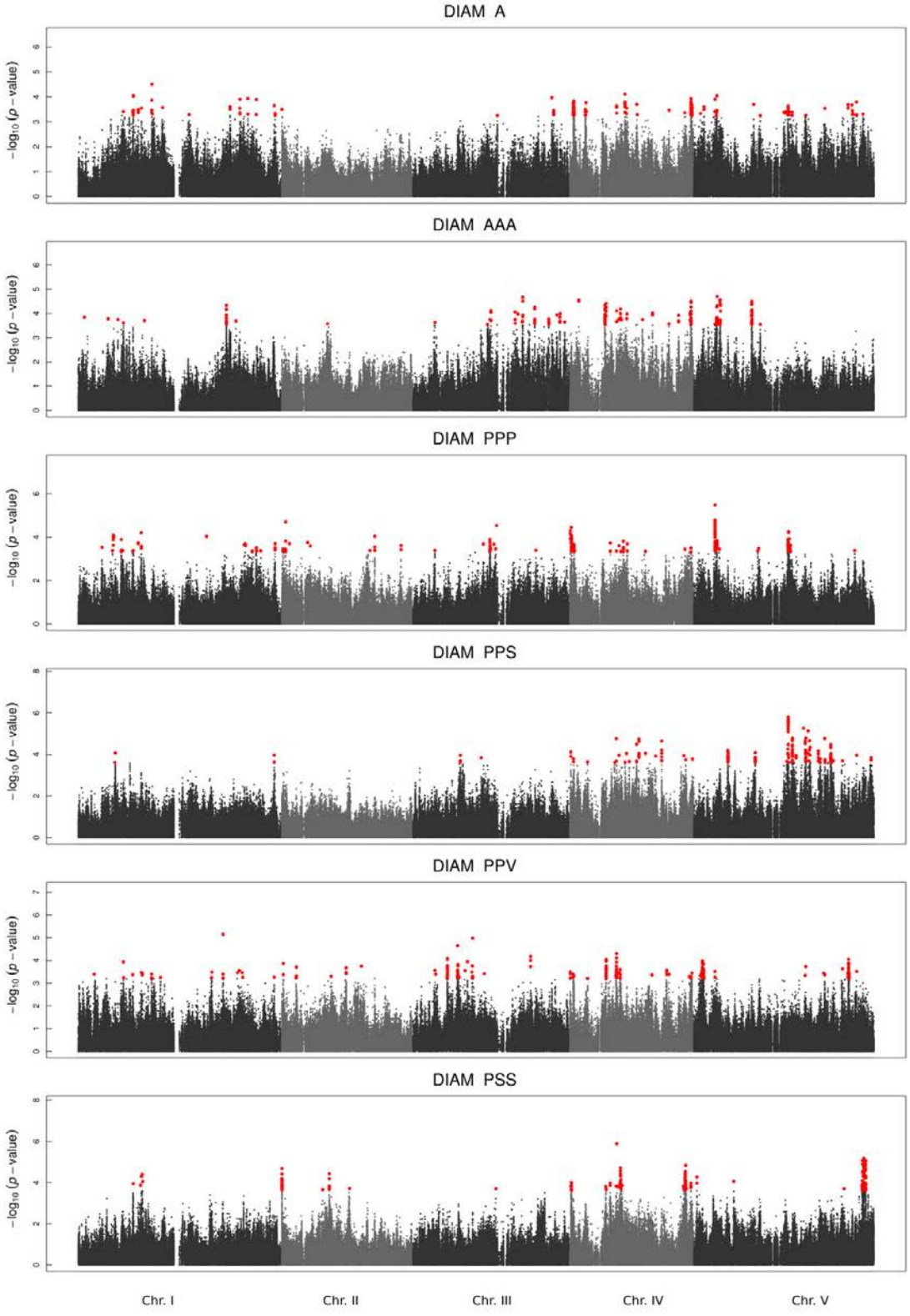

Figure S2 (continued)

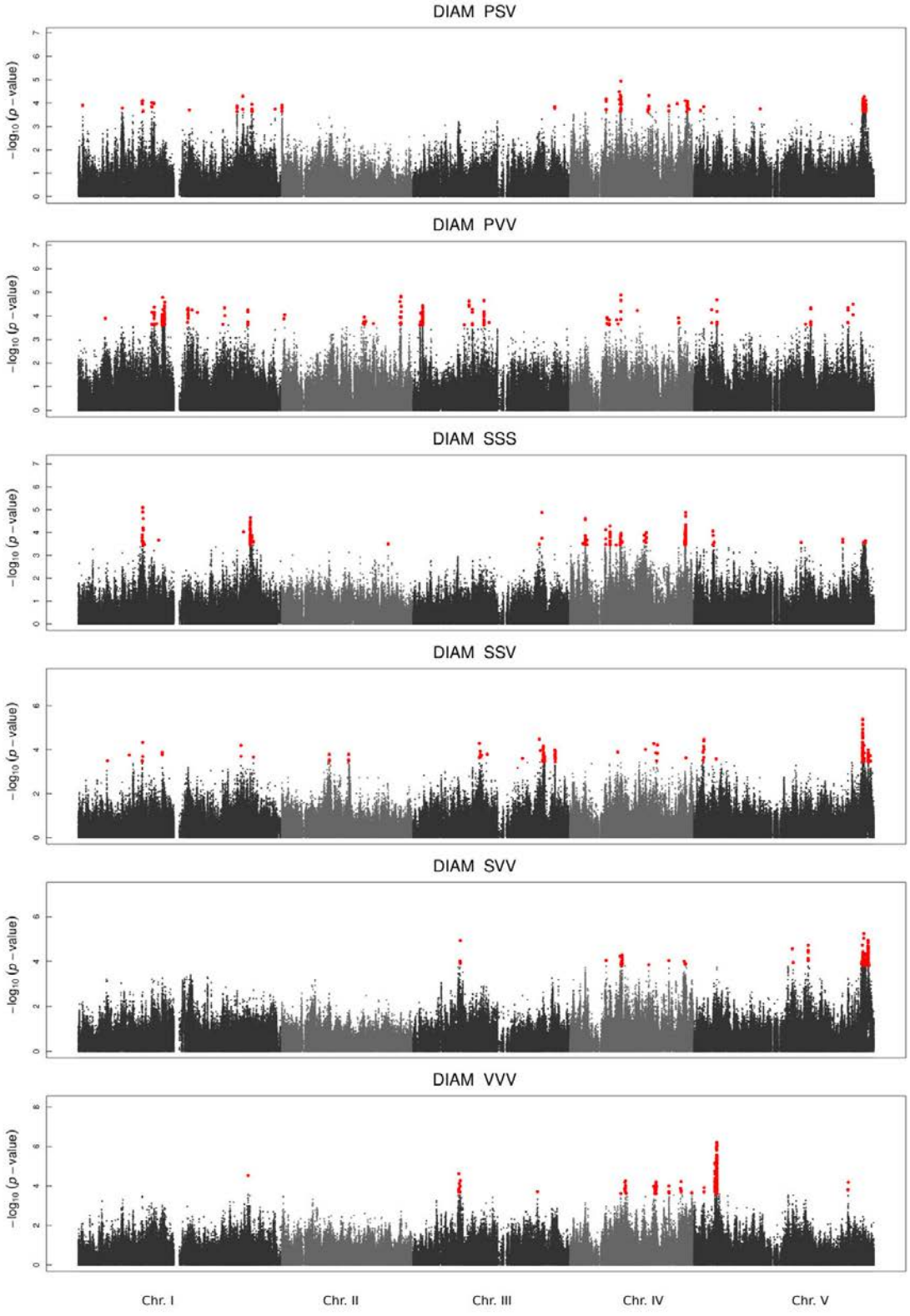

Figure S2 (continued)

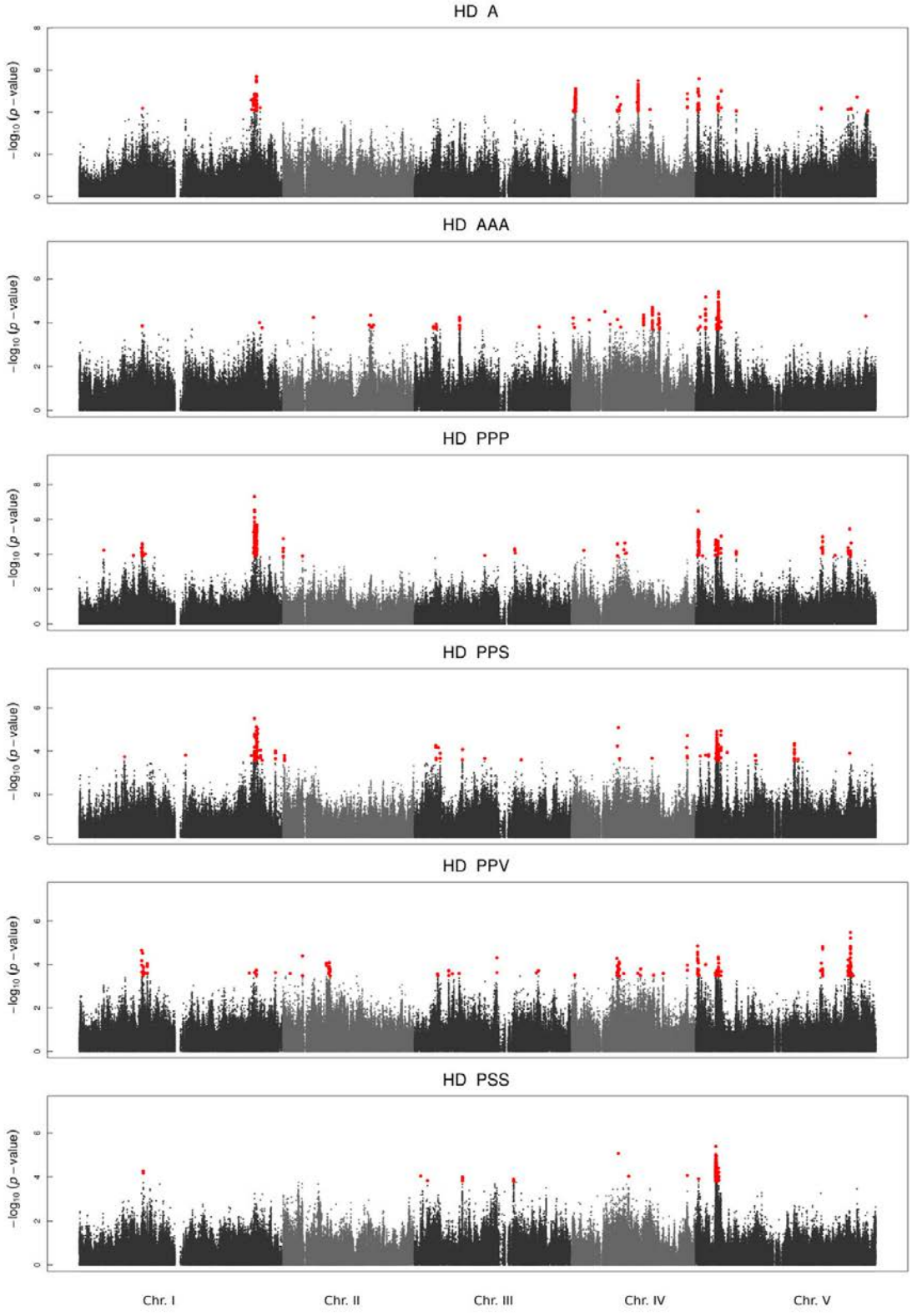

Figure S2 (continued)

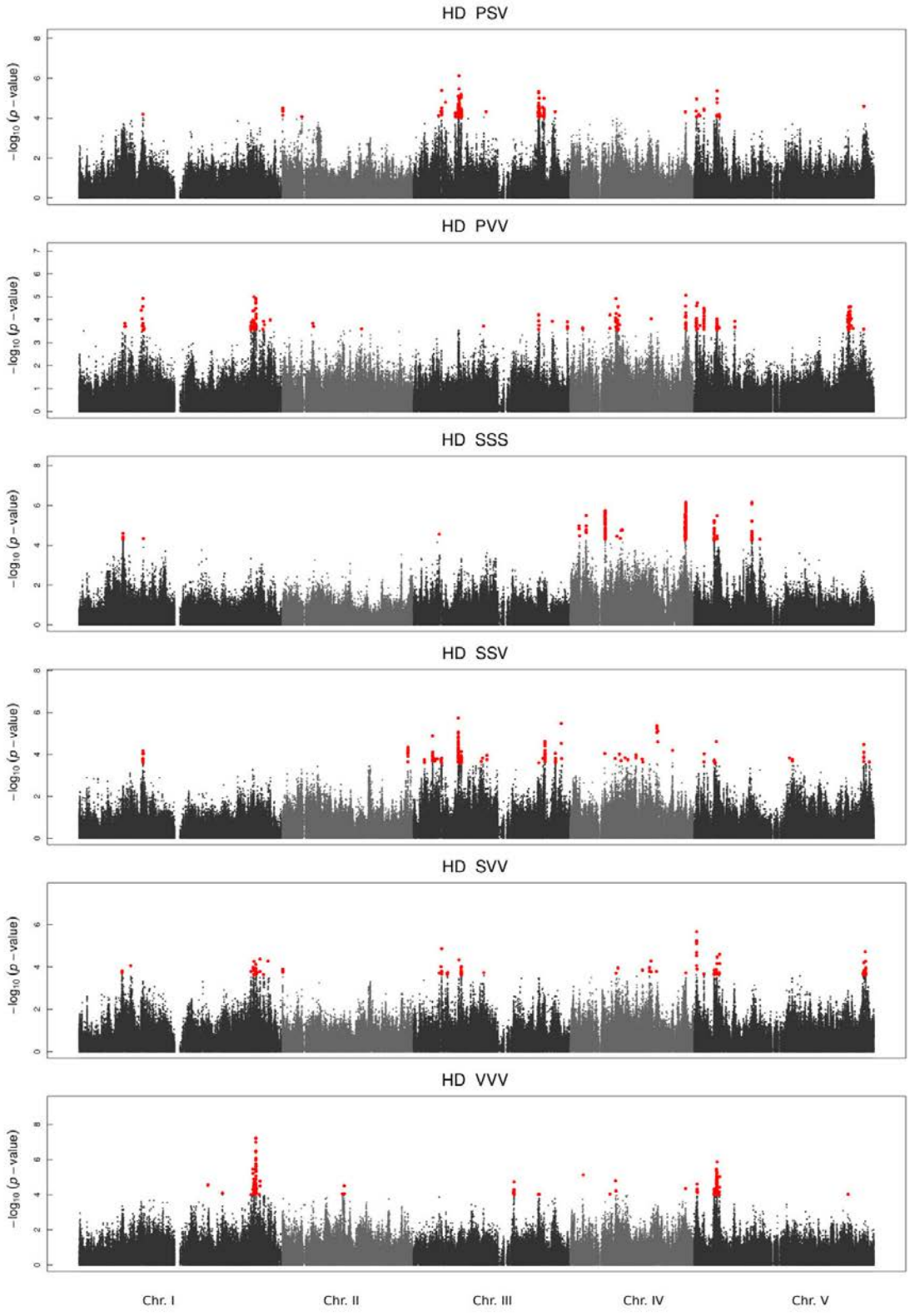

Figure S2 (continued)

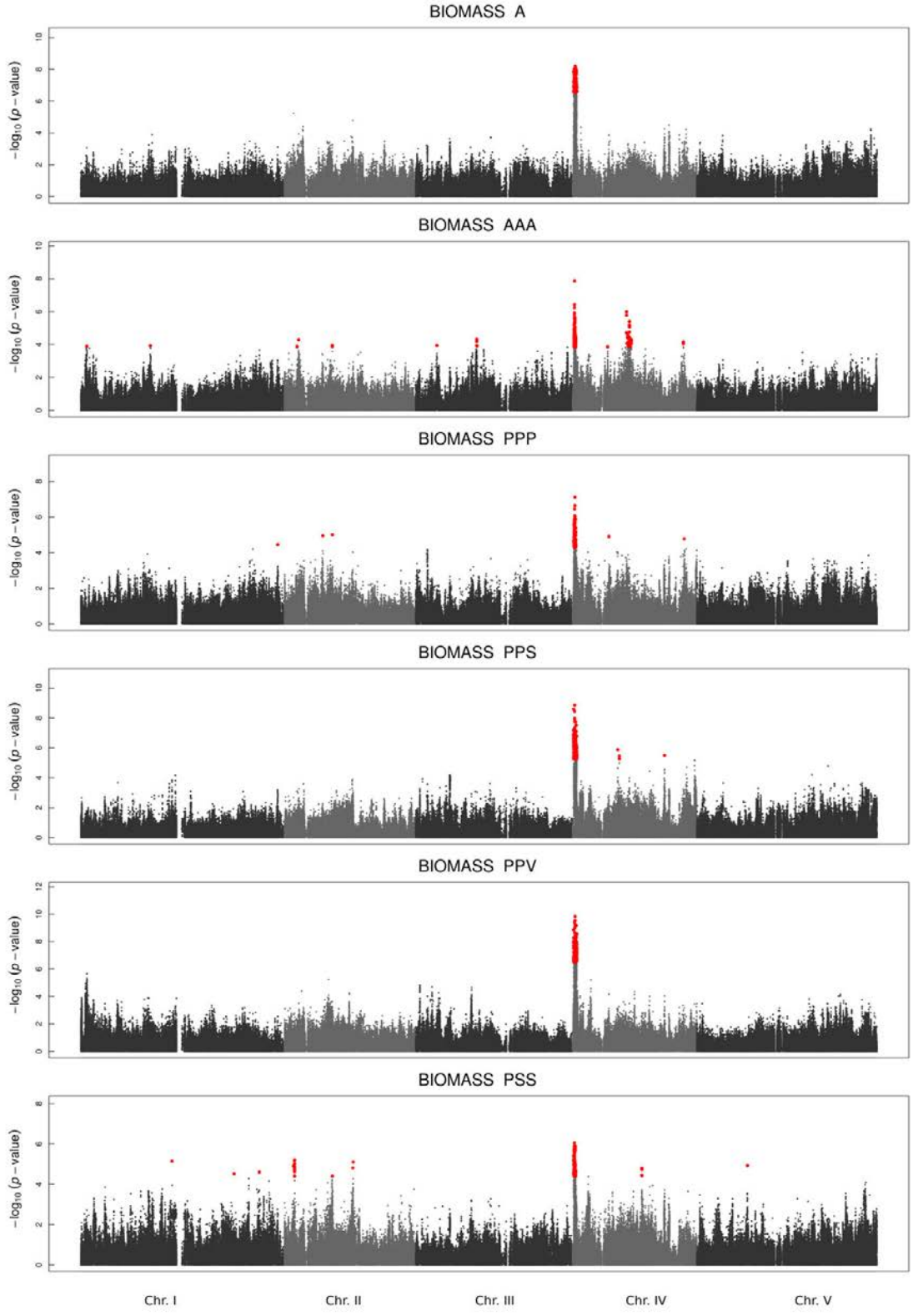

Figure S2 (continued)

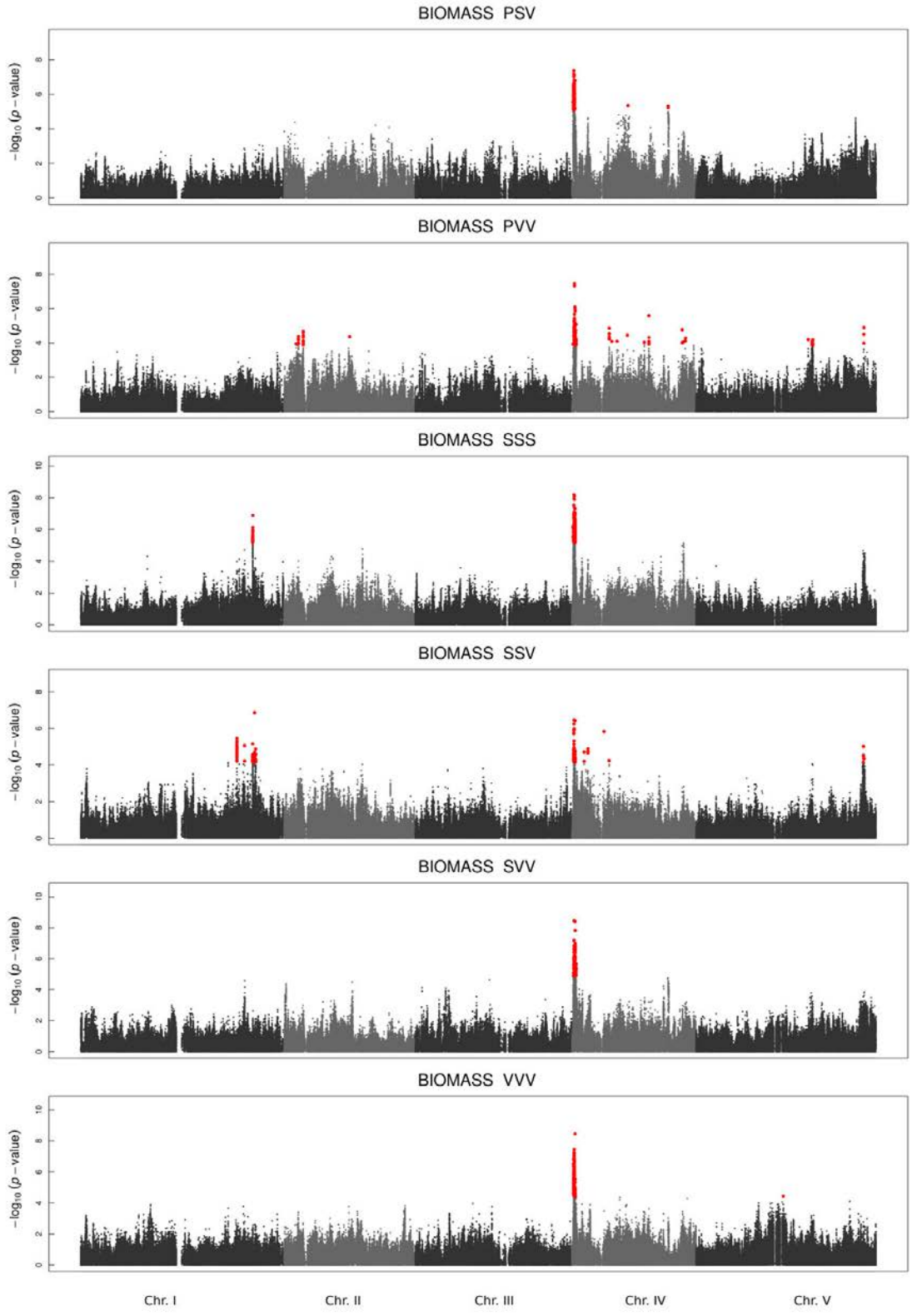

Figure S3. Non-proportional Venn diagram presenting the partitioning of H1F SNPs detected among the lists of 200 top SNPs for each treatment, according different subsets of treatments.

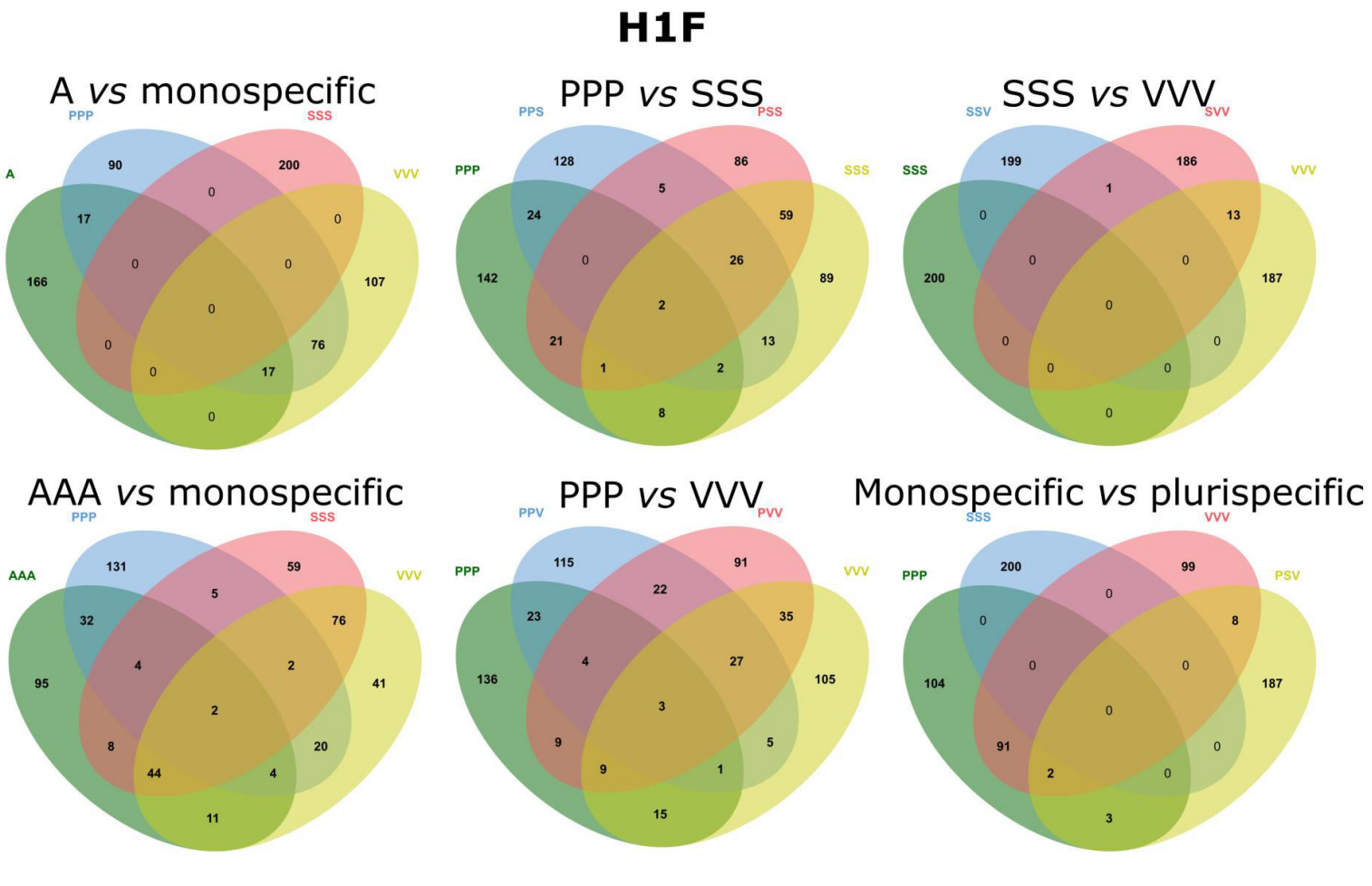

Figure S4. Non-proportional Venn diagram presenting the partitioning of DIAM SNPs detected among the lists of 200 top SNPs for each treatment, according different subsets of treatments.

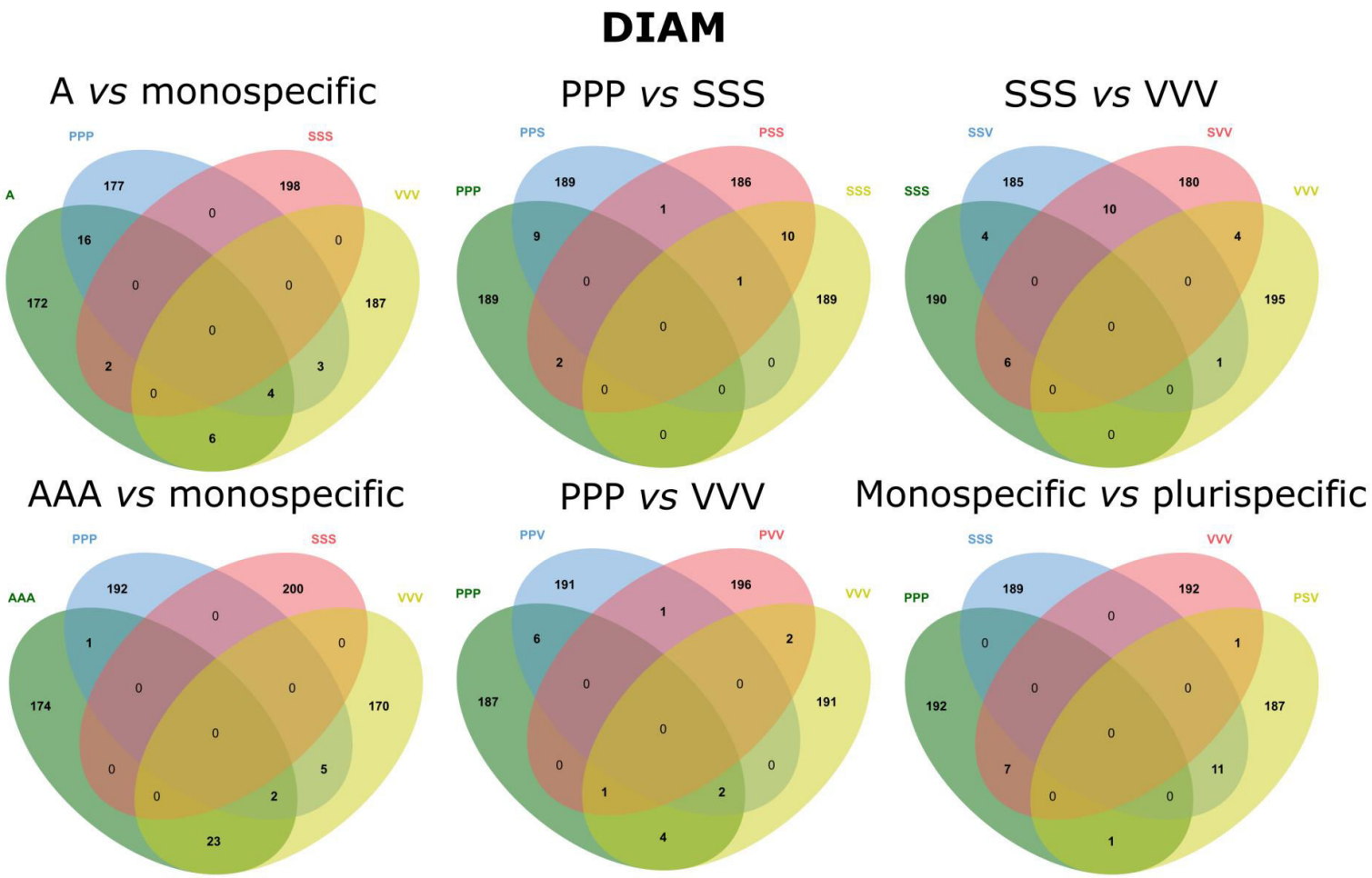

Figure S5. Non-proportional Venn diagram presenting the partitioning of HD SNPs detected among the lists of 200 top SNPs for each treatment, according different subsets of treatments.

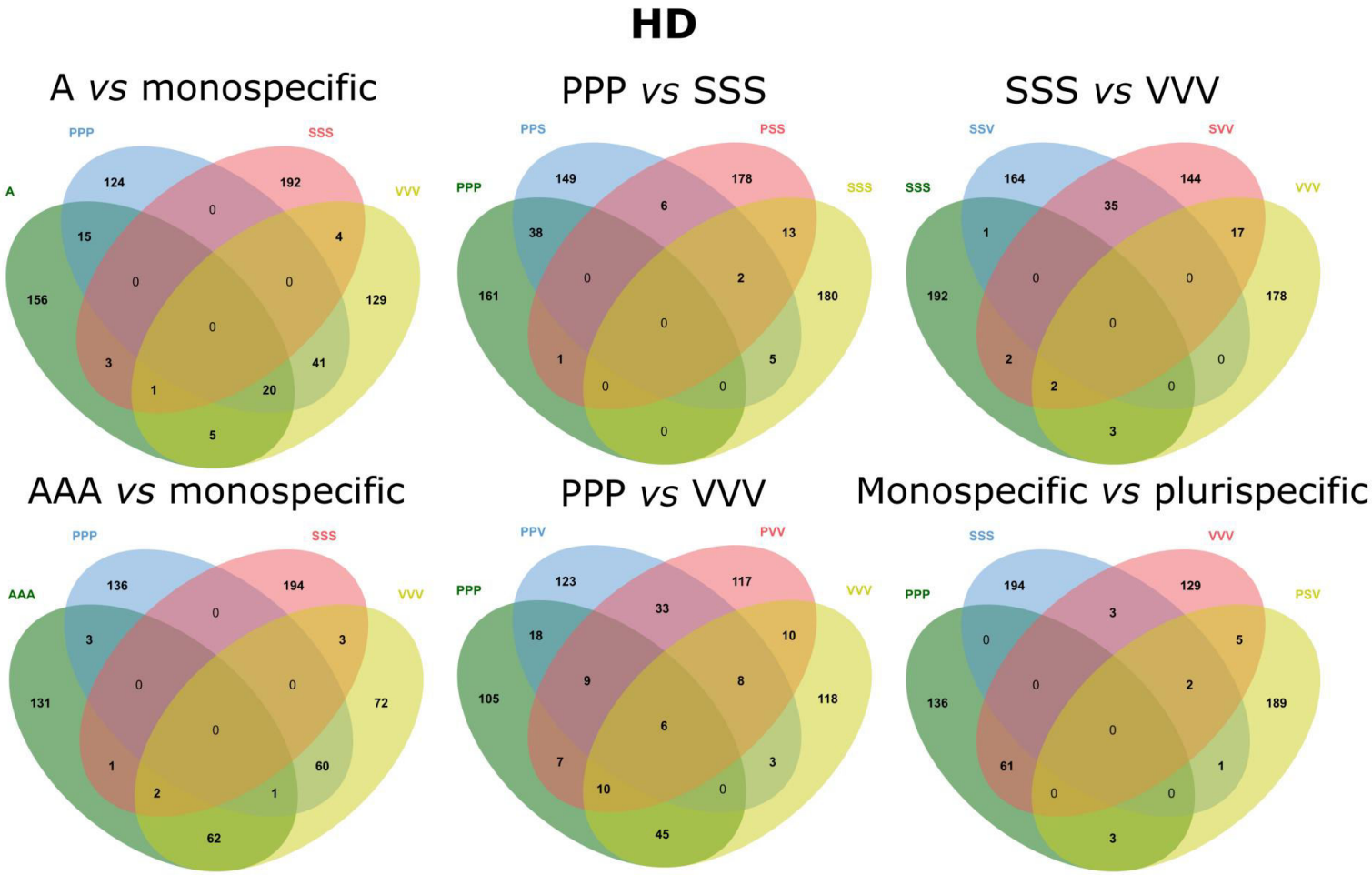

Figure S6. Non-proportional Venn diagram presenting the partitioning of BIOMASS SNPs detected among the lists of 200 top SNPs for each treatment, according different subsets of treatments.

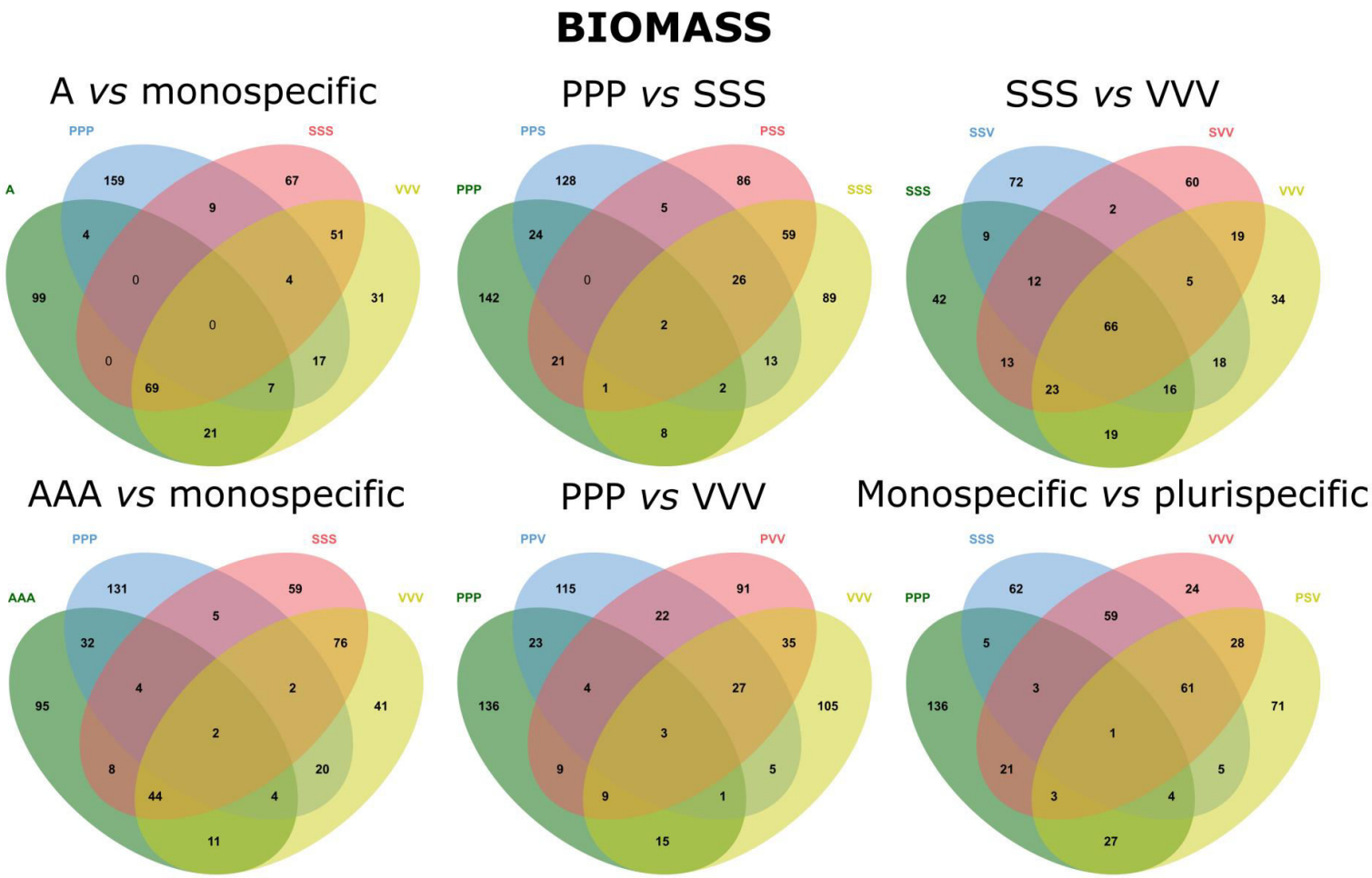

Table S1 Natural variation of four phenotypic traits scored on *A. thaliana* plants among all treatments, with the exception of the 'Control' treatment. Bold *P*-values indicate significant effects after FDR correction. Model random terms were tested with likelihood ratio tests (LRT) of models with and without these effects. Random effects are in italics. H1F: height from the soil to the first flower on the main stem, DIAM: maximum diameter of the rosette, BIOMASS: aboveground dry biomass, HD = H1F / DIAM.

	Traits							
	H1F		**DIAM**		**BIOMASS**		**HD**	
Model terms	*F or LRT*	*P*	*F or LRT*	*P*	*F or LRT*	*P*	*F or LRT*	*P*
block	2.97	**6.01E-02**	1.11	4.06E-01	3.36	**4.24E-02**	0.52	6.71E-01
treatment	22.58	**2.55E-32**	3.36	**3.44E-04**	22.89	**2.55E-32**	1.66	1.03E-01
accession	822.20	**1.13E-179**	718.50	**2.65E-157**	450.30	**4.35E-99**	984.50	**1.18E-214**
*treatment*accession*	19.10	**1.93E-05**	26.90	**3.52E-07**	171.50	**1.95E-38**	28.00	**2.12E-07**
germ(treatment)	1.31	2.50E-01	0.90	5.86E-01	2.04	**3.00E-02**	0.85	6.09E-01
flo(treatment)	49.74	**2.55E-32**	9.30	**3.42E-16**	43.33	**2.55E-32**	9.65	**6.94E-17**
flo*flo(treatment)	38.79	**2.55E-32**	8.25	**4.52E-14**	17.41	**2.55E-32**	7.72	**5.30E-13**

Table S2 Natural variation of four phenotypic traits scored on *A. thaliana* plants among all treatments, with the exception of the 'Control' and 'Intraspecific interaction' treatments. Bold *P*-values indicate significant effects after FDR correction. Model random terms were tested with likelihood ratio tests (LRT) of models with and without these effects. Random effects are in italics. H1F: height from the soil to the first flower on the main stem, DIAM: maximum diameter of the rosette, BIOMASS: aboveground dry biomass, HD = H1F / DIAM.

	Traits							
	H1F		**DIAM**		**BIOMASS**		**HD**	
Model terms	*F or LRT*	*P*	*F or LRT*	*P*	*F or LRT*	*P*	*F or LRT*	*P*
block	2.95	**6.39E-02**	1.32	3.23E-01	3.80	**2.89E-02**	0.44	7.28E-01
treatment	25.42	**2.80E-32**	3.38	**6.14E-04**	23.94	**2.80E-32**	1.84	6.87E-02
accession	762.00	**1.38E-166**	615.00	**8.54E-135**	390.00	**5.80E-86**	879.50	**7.87E-192**
*treatment*accession*	8.70	**4.69E-03**	25.60	**7.35E-07**	155.60	**5.82E-35**	24.70	**1.10E-06**
germ(treatment)	1.35	2.29E-01	0.89	5.62E-01	2.10	**2.89E-02**	0.92	5.51E-01
flo(treatment)	53.64	**2.80E-32**	9.56	**2.16E-15**	44.56	**2.80E-32**	10.73	**1.42E-17**
flo*flo(treatment)	40.04	**2.80E-32**	8.88	**3.66E-14**	17.48	**1.02E-30**	8.54	**1.55E-13**

Table S3 Natural variation of four phenotypic traits scored on *A. thaliana* plants within each treatment. Bold *P*-values indicate significant effects after FDR correction. Model random terms were tested with likelihood ratio tests (LRT) of models with and without these effects. Random effects are in italics. H1F: height from the soil to the first flower on the main stem, DIAM: maximum diameter of the rosette, BIOMASS: aboveground dry biomass, HD = H1F / DIAM.

		H1F		DIAM		BIOMASS		HD	
Treatment	**Model Term**	***F or LRT***	***P***	***F or LRT***	***P***	***F or LRT***	***P***	***F or LRT***	***P***
A	block	2.98	5.32E-02	0.66	6.49E-01	19.58	**9.02E-11**	4.17	**1.22E-02**
	accession	98.80	**9.57E-22**	50.80	**5.44E-12**	63.90	**1.21E-14**	78.00	**1.44E-17**
	germ	0.00	9.89E-01	1.51	2.93E-01	3.26	1.10E-01	0.04	8.46E-01
	flo	1.86	2.39E-01	179.53	**1.60E-29**	127.02	**1.42E-22**	13.71	**5.59E-04**
	flo*flo	0.72	4.78E-01	93.91	**6.62E-18**	32.65	**9.79E-08**	8.39	**7.98E-03**
AAA	block	0.45	7.63E-01	1.59	2.62E-01	1.33	3.49E-01	2.17	1.39E-01
	accession	70.70	**4.75E-16**	41.20	**5.30E-10**	14.80	**2.74E-04**	42.20	**3.47E-10**
	germ	5.04	**4.37E-02**	0.55	5.41E-01	0.17	7.31E-01	3.51	9.81E-02
	flo	0.85	4.45E-01	109.99	**2.39E-20**	100.12	**7.66E-19**	22.20	**1.09E-05**
	flo*flo	0.82	4.52E-01	87.52	**4.84E-17**	48.09	**1.45E-10**	16.42	**1.59E-04**
PPP	block	0.39	7.97E-01	2.32	1.15E-01	4.31	**1.04E-02**	1.15	4.21E-01
	accession	72.50	**2.00E-16**	56.80	**3.62E-13**	36.10	**6.70E-09**	64.50	**9.26E-15**
	germ	0.95	4.22E-01	1.64	2.72E-01	1.95	2.28E-01	2.61	1.59E-01
	flo	0.23	6.89E-01	115.96	**4.72E-21**	25.93	**1.93E-06**	14.27	**4.38E-04**
	flo*flo	0.05	8.46E-01	60.27	**1.36E-12**	0.26	6.81E-01	8.33	**8.23E-03**
PPS	block	3.35	**3.45E-02**	1.20	3.99E-01	7.26	**2.52E-04**	3.00	5.22E-02
	accession	41.40	**4.92E-10**	37.70	**3.09E-09**	16.40	**1.26E-04**	43.90	**1.57E-10**
	germ	0.63	5.10E-01	0.24	6.89E-01	1.17	3.67E-01	0.33	6.37E-01
	flo	2.12	2.08E-01	72.75	**1.25E-14**	68.30	**6.54E-14**	0.72	4.78E-01
	flo*flo	10.06	**3.51E-03**	91.57	**1.44E-17**	57.89	**3.18E-12**	0.19	7.25E-01

Table S3 (continued)

		H1F		DIAM		BIOMASS		HD	
Treatment	**Model Term**	***F or LRT***	***P***	***F or LRT***	***P***	***F or LRT***	***P***	***F or LRT***	***P***
PPV	block	1.08	4.45E-01	2.44	1.01E-01	7.65	**1.59E-04**	0.82	5.62E-01
	accession	66.30	**3.87E-15**	21.10	**1.18E-05**	61.20	**4.27E-14**	67.90	**1.88E-15**
	germ	0.75	4.75E-01	0.45	5.83E-01	4.30	6.30E-02	0.16	7.40E-01
	flo	5.01	**4.39E-02**	229.16	**1.20E-30**	96.16	**4.49E-18**	10.83	**2.40E-03**
	flo*flo	4.34	6.21E-02	147.03	**5.91E-25**	31.82	**1.48E-07**	6.13	**2.49E-02**
PSS	block	4.96	**4.76E-03**	1.06	4.53E-01	2.64	7.98E-02	6.80	**4.47E-04**
	accession	31.20	**7.66E-08**	52.70	**2.33E-12**	12.90	**6.99E-04**	44.10	**1.45E-10**
	germ	2.29	1.89E-01	14.94	**3.22E-04**	7.09	**1.49E-02**	0.29	6.61E-01
	flo	2.02	2.18E-01	47.03	**2.50E-10**	38.47	**8.07E-09**	0.02	8.94E-01
	flo*flo	8.98	**6.01E-03**	68.23	**6.99E-14**	27.81	**8.37E-07**	0.60	5.20E-01
PSV	block	10.25	**6.10E-06**	2.88	6.03E-02	13.62	**9.57E-08**	11.30	**1.62E-06**
	accession	31.30	**7.37E-08**	46.70	**4.05E-11**	9.00	**5.40E-03**	52.60	**2.39E-12**
	germ	0.07	8.22E-01	1.15	3.70E-01	2.16	2.04E-01	0.10	7.90E-01
	flo	2.71	1.51E-01	55.76	**7.20E-12**	61.28	**9.33E-13**	0.17	7.31E-01
	flo*flo	10.57	**2.71E-03**	63.54	**4.12E-13**	45.22	**4.92E-10**	2.37	1.82E-01
PVV	block	0.30	8.46E-01	8.30	**6.93E-05**	6.65	**5.32E-04**	1.81	2.06E-01
	accession	38.40	**2.19E-09**	48.90	**1.35E-11**	53.10	**2.00E-12**	27.00	**5.87E-07**
	germ	1.10	3.82E-01	0.06	8.30E-01	1.86	2.39E-01	0.73	4.78E-01
	flo	0.73	4.78E-01	117.67	**3.36E-21**	38.65	**7.55E-09**	19.31	**4.21E-05**
	flo*flo	0.43	5.91E-01	63.36	**4.42E-13**	6.77	**1.76E-02**	11.14	**2.05E-03**

Table S3 (continued)

		H1F		DIAM		BIOMASS		HD	
Treatment	**Model Terms**	***F or LRT***	***P***	***F or LRT***	***P***	***F or LRT***	***P***	***F or LRT***	***P***
SSS	block	2.38	1.08E-01	7.03	**3.41E-04**	9.03	**2.87E-05**	5.08	**4.13E-03**
	accession	16.10	**1.46E-04**	41.40	**4.92E-10**	18.70	**3.95E-05**	42.90	**2.51E-10**
	germ	0.11	7.84E-01	0.52	5.50E-01	2.20	2.00E-01	2.06	2.15E-01
	flo	5.40	**3.65E-02**	29.87	**3.42E-07**	8.80	**6.53E-03**	2.62	1.59E-01
	flo*flo	18.71	**5.78E-05**	55.21	**9.96E-12**	7.17	**1.45E-02**	5.07	**4.35E-02**
SSV	block	3.51	**2.82E-02**	0.70	6.28E-01	3.26	**3.83E-02**	4.19	**1.20E-02**
	accession	7.70	**1.04E-02**	28.20	**3.23E-07**	37.60	**3.20E-09**	15.30	**2.14E-04**
	germ	1.26	3.47E-01	0.11	7.84E-01	0.33	6.37E-01	2.48	1.71E-01
	flo	0.85	4.45E-01	45.10	**4.92E-10**	21.37	**1.62E-05**	4.40	6.04E-02
	flo*flo	7.71	**1.11E-02**	61.94	**7.34E-13**	16.22	**1.76E-04**	0.60	5.20E-01
SVV	block	6.81	**4.38E-04**	2.88	6.03E-02	13.37	**1.26E-07**	9.29	**2.02E-05**
	accession	33.40	**2.54E-08**	29.60	**1.63E-07**	18.30	**4.77E-05**	28.60	**2.67E-07**
	germ	0.00	9.89E-01	0.88	4.40E-01	3.40	1.03E-01	0.10	7.90E-01
	flo	1.48	2.98E-01	91.53	**1.44E-17**	77.54	**2.23E-15**	3.45	1.01E-01
	flo*flo	7.59	**1.18E-02**	96.48	**3.10E-18**	56.58	**5.18E-12**	0.36	6.26E-01
VVV	block	0.75	5.98E-01	4.32	**1.04E-02**	2.56	8.79E-02	0.51	7.31E-01
	accession	51.60	**3.80E-12**	36.30	**6.15E-09**	61.70	**3.43E-14**	53.00	**2.06E-12**
	germ	1.70	2.62E-01	0.05	8.46E-01	1.53	2.91E-01	2.75	1.48E-01
	flo	0.00	9.76E-01	195.71	**1.20E-30**	100.86	**7.82E-19**	36.83	**1.63E-08**
	flo*flo	0.25	6.86E-01	129.57	**8.50E-23**	46.31	**3.23E-10**	30.89	**2.12E-07**

Table S4 Broad-sense heritability values for the four phenotypic traits scored on *A. thaliana* plants within each treatment. Bold *P*-values indicate significant effects after FDR correction.

	Traits							
	H1F		DIAM		BIOMASS		HD	
Treatment	H^2	*P*	H^2	*P*	H^2	*P*	H^2	*P*
A	0.8	**9.57E-22**	0.94	**5.44E-12**	0.96	**1.21E-14**	0.86	**1.44E-17**
AAA	0.74	**4.75E-16**	0.89	**5.30E-10**	0.91	**2.74E-04**	0.82	**3.47E-10**
PPP	0.77	**2.00E-16**	0.95	**3.62E-13**	0.96	**6.70E-09**	0.89	**9.26E-15**
SSS	0.8	**1.46E-04**	0.67	**4.92E-10**	0.65	**3.95E-05**	0.8	**2.51E-10**
VVV	0.68	**3.80E-12**	0.94	**6.15E-09**	0.94	**3.43E-14**	0.85	**2.06E-12**
PPS	0.82	**4.92E-10**	0.71	**3.09E-09**	0.69	**1.26E-04**	0.84	**1.57E-10**
PPV	0.75	**3.87E-15**	0.93	**1.18E-05**	0.96	**4.27E-14**	0.89	**1.88E-15**
PSS	0.81	**7.66E-08**	0.75	**2.33E-12**	0.74	**6.99E-04**	0.82	**1.45E-10**
PVV	0.65	**2.19E-09**	0.94	**1.35E-11**	0.94	**2.00E-12**	0.83	**5.87E-07**
SSV	0.77	**1.04E-02**	0.73	**3.23E-07**	0.78	**3.20E-09**	0.78	**2.14E-04**
SVV	0.81	**2.54E-08**	0.76	**1.63E-07**	0.79	**4.77E-05**	0.83	**2.67E-07**
PSV	0.75	**7.37E-08**	0.77	**4.05E-11**	0.69	**5.40E-03**	0.81	**2.39E-12**

Table S5 Biological pathways (MapMan classification) represented for all unique genes identified. Bold lines indicate significant over-represented biological pathways ($P < 0.01$).

Normed Freq.	± bootstrap StdDev	p-value	Class
0.17	**0.078**	**7.96E-10**	**DNA**
1.87	**0.425**	**2.53E-03**	**transport**
0.86	**0.061**	**6.75E-03**	**not assigned**
1.25	0.166	0.021	RNA
1.4	0.313	0.029	signalling
1.14	0.142	0.029	protein
3.45	1.741	0.045	TCA / org transformation
1.69	0.665	0.054	lipid metabolism
4.55	3.366	0.062	C1-metabolism
1.2	0.216	0.063	misc
3.79	2.729	0.082	tetrapyrrole synthesis
0.81	0.237	0.094	stress
2.2	1.051	0.108	minor CHO metabolism
1.76	0.859	0.114	PS
0.87	0.31	0.132	cell
1.04	0.352	0.133	development
1.16	0.418	0.14	cell wall
1	0.415	0.163	hormone metabolism
1.02	0.435	0.177	secondary metabolism
0.98	0.363	0.178	micro RNA, natural antisense etc
1.8	1.331	0.204	major CHO metabolism
1.3	0.676	0.206	redox
3.5	2.423	0.217	N-metabolism
3.25	2.753	0.228	Biodegradation of Xenobiotics
0.7	0.458	0.236	amino acid metabolism
0.5	0.4	0.274	nucleotide metabolism
0.6	0.44	0.317	mitochondrial electron transport / ATP synthesis
1.15	0.907	0.367	glycolysis
1.12	0.622	0.368	Co-factor and vitamine metabolism
1.09	0.846	0.369	metal handling

Table S6 Biological pathways (MapMan classification) represented in the 'Control', 'Intraspecific', 'Monospecific' and 'Plurispecific' treatments. Bold lines indicate significant over-represented biological pathways ($P < 0.01$).

Type	Absolute values	± bootstrap	p-value	Class
Control	10	2.8	0.02	not assigned
Control	2	1.2	0.03	PS
Control	1	0.8	0.052	C1-metabolism
Control	1	0.8	0.062	tetrapyrrole synthesis
Control	3	1.5	0.071	development
Control	1	0.7	0.097	TCA / org transformation
Control	4	1.9	0.107	misc
Control	6	2.4	0.116	RNA
Control	8	2.5	0.131	protein
Control	1	0.6	0.144	minor CHO metabolism
Control	3	1.4	0.15	stress
Control	2	1.1	0.246	transport
Control	2	1.2	0.276	signalling
Control	1	0.9	0.33	lipid metabolism
Control	1	0.7	0.339	micro RNA, natural antisense etc
Control	1	0.8	0.358	cell wall
Intraspecific	2	1.2	0.092	DNA
Intraspecific	1	0.7	0.107	TCA / org transformation
Intraspecific	18	2.8	0.116	not assigned
Intraspecific	2	1.3	0.116	secondary metabolism
Intraspecific	1	1	0.132	major CHO metabolism
Intraspecific	6	2.1	0.139	RNA
Intraspecific	8	2.6	0.147	protein
Intraspecific	3	1.4	0.191	signalling
Intraspecific	1	0.7	0.214	misc
Intraspecific	2	1.1	0.219	development

Table S6 (continued)

Type	Absolute values	± bootstrap	p-value	Class
Intraspecific	2	1.1	0.26	transport
Intraspecific	1	0.8	0.268	amino acid metabolism
Intraspecific	1	0.7	0.343	lipid metabolism
Intraspecific	1	0.7	0.361	cell
Intraspecific	1	NaN	0.365	hormone metabolism
Intraspecific	1	0.7	0.366	cell wall
Monospecific	**2**	**1.2**	**8.33E-03**	**tetrapyrrole synthesis**
Monospecific	25	4.3	0.011	not assigned
Monospecific	2	1.1	0.021	TCA / org transformation
Monospecific	14	3	0.028	RNA
Monospecific	17	4.2	0.076	protein
Monospecific	2	1.1	0.099	PS
Monospecific	3	1.7	0.105	micro RNA, natural antisense etc
Monospecific	3	1.4	0.136	hormone metabolism
Monospecific	5	2.1	0.139	stress
Monospecific	5	2.2	0.159	signalling
Monospecific	4	2	0.19	misc
Monospecific	3	1.5	0.206	development
Monospecific	3	1.4	0.213	cell
Monospecific	2	1.1	0.226	lipid metabolism
Monospecific	3	1.8	0.228	transport
Monospecific	2	1	0.232	secondary metabolism
Monospecific	1	0.9	0.253	minor CHO metabolism
Monospecific	1	0.6	0.324	cell wall
Monospecific	1	0.6	0.334	redox

Table S6 (continued)

Type	Absolute values	± bootstrap	p-value	Class
Plurispecific	**5**	**2**	**1.24E-06**	**DNA**
Plurispecific	**16**	**3.7**	**3.26E-03**	**transport**
Plurispecific	**18**	**4**	**8.90E-03**	**signalling**
Plurispecific	78	7.5	0.014	not assigned
Plurispecific	31	5	0.021	RNA
Plurispecific	2	1.3	0.034	C1-metabolism
Plurispecific	40	6.2	0.06	protein
Plurispecific	15	3.5	0.076	misc
Plurispecific	2	1.2	0.1	TCA / org transformation
Plurispecific	1	0.7	0.103	micro RNA, natural antisense etc
Plurispecific	8	3.1	0.127	stress
Plurispecific	4	1.8	0.133	development
Plurispecific	3	1.6	0.139	redox
Plurispecific	6	2.2	0.161	cell
Plurispecific	1	0.9	0.164	N-metabolism
Plurispecific	5	2	0.165	cell wall
Plurispecific	1	0.9	0.174	Biodegradation of Xenobiotics
Plurispecific	2	1.1	0.175	minor CHO metabolism
Plurispecific	4	1.8	0.183	lipid metabolism
Plurispecific	3	1.5	0.189	hormone metabolism
Plurispecific	2	1.1	0.193	secondary metabolism
Plurispecific	1	0.7	0.256	tetrapyrrole synthesis
Plurispecific	2	1.3	0.259	PS
Plurispecific	2	1.1	0.273	amino acid metabolism
Plurispecific	1	0.9	0.333	glycolysis

Table S6 (continued)

Type	Absolute values	± bootstrap	p-value	Class
Plurispecific	1	0.6	0.336	Co-factor and vitamine metabolism
Plurispecific	1	0.7	0.339	metal handling
Plurispecific	1	0.6	0.349	nucleotide metabolism
Plurispecific	1	0.7	0.359	major CHO metabolism
Plurispecific	1	0.6	0.366	mitochondrial electron transport / ATP synthesis

C) Conclusion

Dans de ce chapitre, en phénotypant une population locale d'*A. thaliana* dans différents contextes d'interactions plante-plante mono- et pluri-spécifiques, j'ai cherché à :

i. déterminer l'étendue de la variation génétique au sein de cette population locale d'*A. thaliana*
ii. caractériser et comparer les architectures génétiques sous-jacentes.
iii. identifier les principaux processus biologiques associés.

Sur l'ensemble des quatre traits phénotypiques mesurés, nous avons détecté des normes de réactions fortement croisées, non seulement entre les trois traitements d'interaction monospécifiques, mais également entre les différents assemblages d'espèces entourant les accessions focales d'*A. thaliana*. En d'autres termes, malgré une réponse parfois similaire de l'ensemble de la population en réponse aux différents traitements d'interaction plante-plante, le rang des accessions diffère largement entre les 12 traitements d'interaction. Ce phénomène avait déjà été observé pour 48 accessions provenant de la population TOU-A placées dans différents contextes d'interactions monospécifiques sur un terrain expérimental (Baron *et al.* 2015). Cependant, il n'avait encore jamais été mis en évidence entre différents contextes d'interactions plurispécifiques, en particulier entre des assemblages composés des mêmes espèces mais dont les abondances relatives diffèrent. *A. thaliana* peut donc répondre de manière fine à différents niveaux de complexité végétale : (i) à l'identité de l'espèce à proximité, (ii) au nombre d'espèces (diversité), mais aussi et de façon plus surprenante (iii) à l'assemblage d'espèces (identité, nombre et abondances d'espèces). En effet, nous nous attendions à une bien plus grande additivité des QTL entre assemblages présentant les mêmes espèces mais pour lesquels seulement un individu était différent, comme par exemple entre les assemblages PPS et PSS.

Par une approche de GWA mapping, nous avons pu mettre en évidence que les QTL de réponse à la compétition d'*A. thaliana* étaient majoritairement très différents entre les 12 traitements d'interaction. De manière similaire à ce que nous avions observé entre les différents micro-habitats du chapitre 1, nous avons identifié une architecture génétique très flexible, car dépendante de l'assemblage d'espèces entourant *A. thaliana*. Cette architecture génétique complexe dépendante du traitement de compétition, permettrait là encore d'expliquer le

maintien d'une forte diversité génomique et phénotypique observée au sein de la population TOU-A. Pour tester cette hypothèse, il serait donc intéressant de déterminer s'il existe une relation entre la diversité génomique au sein d'une population d'*A. thaliana* et la diversité taxonomique et/ou fonctionnelle de la communauté végétale à laquelle elle appartient.

Nous avons également mis en lumière que les processus biologiques enrichis diffèrent largement entre les conditions de compétition mono- et plurispécifique. En particulier, les processus biologiques reliés à la perception/transduction de signaux ainsi qu'au transport sont considérablement enrichis dans les conditions d'interaction plurispécifique à la différence des conditions d'interaction monospécifique. De manière intéressante, les récepteurs kinases représentent une proportion importante des gènes associés à la perception/transduction de signaux.

La plupart des fonctions et processus identifiés restent assez éloignés de celles connues jusqu'à présent dans les interactions plante-plante (Subrahmaniam *et al.* 2018). Ces résultats nous suggèrent que l'étude des interactions plante-plante dans un contexte écologiquement réaliste permettrait d'identifier des mécanismes génétiques et moléculaires jusqu'à maintenant peu décrits dans les interactions plante-plante.

Bien qu'informative, cette étude ne repose que sur des approches corrélatives. Le nom des gènes candidats identifiés ne reste qu'hypothétique. Afin de comprendre les mécanismes sous-jacents à la variation naturelle de la réponse à la compétition chez *A. thaliana*, il faudrait donc valider fonctionnellement certains des gènes candidats identifiés. L'importance des récepteurs kinases dans de nombreux processus biologiques ainsi que leur rôle prépondérant dans la réponse aux interactions plurispécifiques en font des candidats intéressants pour initier une telle approche.

Chapitre 3

Identification d'un récepteur kinase contrôlant la réponse compétitive à *Poa annua*

A) Introduction

Dans le chapitre précédent, nous avons pu mettre en évidence pour l'ensemble des quatre traits phénotypiques mesurés, qu'*A. thaliana* pouvait élaborer une réponse différente à la présence de plantes dans son voisinage, et ce, en fonction des différents niveaux de complexité végétale de ce voisinage, notamment en fonction (i) de l'identité de l'espèce à proximité, (ii) du nombre d'espèces, mais aussi (iii) de l'assemblage des espèces présentes. Par une approche de GWA mapping, nous avons ainsi pu mettre en évidence des QTL de réponse à la compétition d'*A. thaliana*, majoritairement très différents entre les 12 traitements d'interaction.

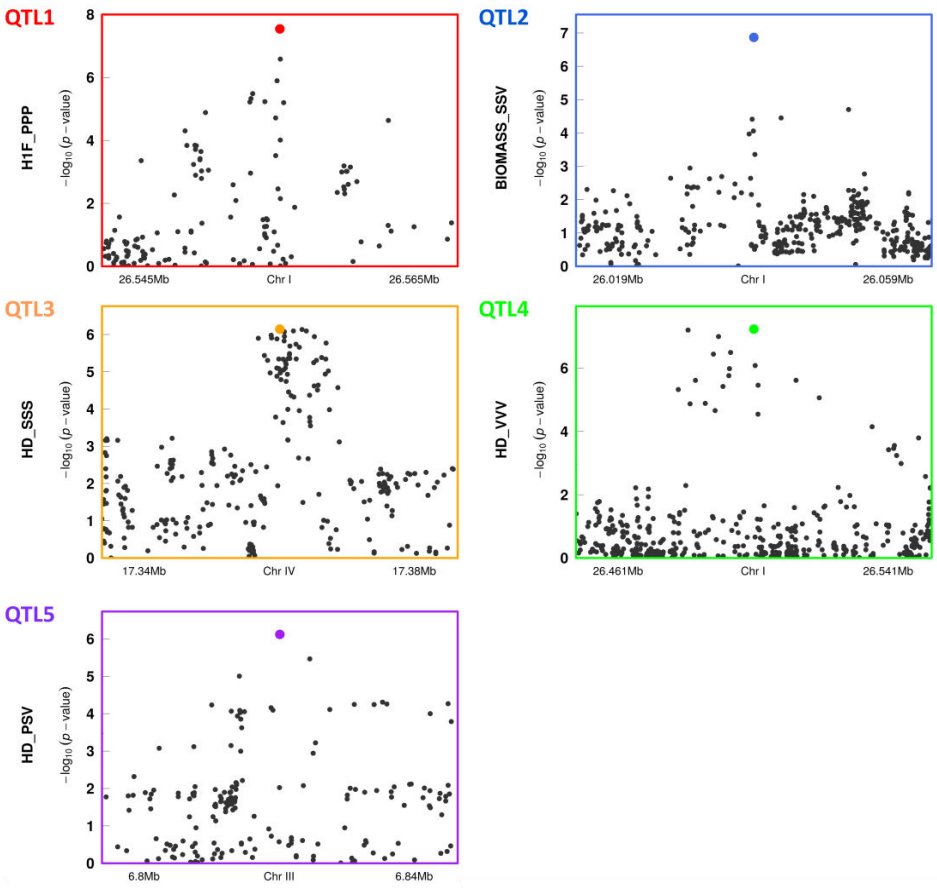

Figure 3.1. Identification de régions génomiques spécifiquement associées à une combinaison 'Trait x Traitement de compétition'. QTL1 identifié pour H1F en PPP ; QTL2 identifié pour BIOMASS en SSV ; QTL3 identifié pour HD en SS ; QTL4 identifié pour HD en VVV et QTL5 HD en PSV. L'axe des abscisses indique la position physique des SNP le long des cinq chromosomes. L'axe des ordonnées indique les -$\log_{10}$ *p*-valeurs obtenues à l'aide d'un modèle mixte (implémenté dans le logiciel EMMAX) en utilisant des SNP présentant une fréquence relative de l'allèle mineure (MARF) supérieure à 10% et moins de 75% de données manquantes. Pour chaque graphe, le point coloré correspond au SNP le plus significativement associé au trait phénotypique.

Parmi les QTL identifiés pour une combinaison 'Trait phénotypique x Traitement de compétition' donnée, cinq présentaient des niveaux de significativité très importants ($-\log_{10}(p\text{-val}) > 6$) et ont donc particulièrement retenu notre attention en vue de l'identification des gènes causaux sous-jacents à ces QTL (Figure 3.1.).

Pour choisir le QTL sur lequel concentrer mes efforts durant la thèse, plusieurs critères de choix ont été adoptés :

- La forme du pic d'association qui donne une approximation du nombre de gènes candidats sous-jacents au QTL: en effectuant un zoom sur ces cinq régions QTL, nous avons observé que les pics d'association présentaient des formes contrastées (Figure 3.1.). Alors que les QTL 1 et 2 présentent des pics d'association relativement fins (1-2kb), les pics d'association des QTL 4 et 5 restent assez mal définis avec des SNP ayant un niveau de significativité qui reste important sur plusieurs kilobases aux alentours du SNP le plus significativement associé. Entre ces cas extrêmes, le QTL3 correspond à un pic d'association bien défini mais sur une région génomique un peu plus large (~10kb) que celle observée pour les QTL 1 et 2. Il semble donc utile de nous focaliser sur un pic fin et présentant une association fortement significative.
- Le type d'interactions plante-plante considéré : les QTL 1, 3 et 4 ont été identifiés en conditions d'interaction monospécifique. Les QTL 2 et 5 ont été identifiés en conditions d'interactions entre 3 (SSV) ou 4 (PSV) espèces, respectivement. Etant donné qu'aucun gène de réponse à la compétition n'a été à ce jour identifié (à notre connaissance), il semblait préférable dans un premier temps de se focaliser sur un type d'interaction plante-plante simple n'incluant que deux espèces.
- L'identité des gènes candidats sous-jacents aux QTL: les fonctions moléculaires des gènes candidats sous-jacents aux QTL 3, 4 et 5 sont souvent inconnues ou bien ne présentent pas de liens évidents avec des fonctions potentiellement attendues dans le contexte des interactions plante-plante. D'un autre côté, les SNP les plus associés aux QTL 1 et 2 sont tous les deux situés dans la région promotrice (à moins de 500 pb) de deux gènes candidats déjà décrits dans la littérature et présentant des rôles potentiellement intéressants dans les interactions plante-plante : un régulateur négatif de la croissance des poils racinaires chez *A. thaliana*

(*Root Hair Specific 10*, *RHS10*, Hwang *et al.* 2016) pour le QTL1 et un régulateur de mécanismes de réponse au stress chez des jeunes plantules d'*A. thaliana* (*ABI Five Binding Protein 1*, *AFP1,* Garcia *et al.* 2008) pour le QTL2. Le gène candidat *PERK13/RHS10* code un récepteur de type kinase (RLK) appartenant à la famille des Proline Rich Extensin-like Kinase (PERK). Pour rappel, les RLK correspondaient à l'un des processus biologiques surreprésentés en condition plurispécifique, mis en évidence dans le chapitre précédent.

Considérant l'ensemble de ces critères, il a été choisi de se focaliser sur l'identification du/des gène(s) sous-jacent(s) au QTL1 qui présente donc plusieurs avantages: (i) un pic d'association fin (~2 kb, donc potentiellement peu de gènes candidats à tester); (ii) un type d'interaction plante-plante simple (interaction monospécifique avec *P. annua*), et (iii) un gène candidat codant pour un récepteur kinase (*PERK13/RHS10*) et présentant un rôle potentiellement intéressant dans les interactions plante-plante .

Dans ce troisième chapitre, je me suis donc attaché à identifier le/les gènes causaux sous-jacents au QTL1, QTL majeur identifié dans le second chapitre en conditions d'interaction monospécifique avec *P. annua*. En combinant (i) l'analyse de mutants insertionnels correspondant aux gènes candidats, (ii) l'étude de la diversité nucléotidique du gène candidat au sein de la population locale TOU-A, et (iii) la production de lignées complémentées par différents haplotypes naturels, j'ai cherché à répondre aux questions suivantes:

i. Quel est le gène causal sous-jacent au QTL1 ?
ii. Quelle est la fonction, en lien avec les interactions plante-plante, codée par ce gène ?
iii. Quelle est la diversité nucléotidique observée pour le gène identifié ?
iv. Ce gène confère-t-il une réponse compétitive à d'autres espèces que *P. annua* ?

NB : dans ce chapitre, mon travail a consisté (i) à effectuer la caractérisation phénotypique et moléculaire de mutants affectés dans les gènes candidats, (ii) à produire différentes lignées complémentées ainsi qu'à les caractériser aux niveaux phénotypique et moléculaire, (iii) à effectuer les analyses statistiques, (iv) à étudier la diversité nucléotidique du gène candidat présente au sein de la population TOU-A, et (v) à déterminer la spécificité de la réponse à la compétition à diverses espèces de graminées ainsi qu'aux deux espèces herbacées utilisées lors du chapitre précédent.

B) Manuscrit: An Arabidopsis receptor-like kinase mediates competitive plant-plant interactions.

Cyril Libourel[1], Fabrice Roux[1, 2*], Dominique Roby[1, 2*]

Affiliations :

[1] LIPM, Université de Toulouse, INRA, CNRS, Castanet-Tolosan, France

[2] These authors contributed equally to this work.

* To whom correspondence should be addressed. E-mail: dominique.roby@inra.fr & fabrice.roux@inra.fr

Manuscrit

“An Arabidopsis receptor-like kinase mediates competitive plant-plant interactions”

Cyril Libourel, Fabrice Roux, Dominique Roby

Summary

While competition between plant species are recognized as a major factor responsible for crop yield and plant community dynamics, the genetic and molecular mechanisms underlying such biotic interactions remained poorly characterized. By considering a Genome Wide Association study that investigated different plant assemblages interacting with *A. thaliana*, we identified a specific QTL related to plant competition response. Here, we report that a proline rich, extensin like receptor like kinase (PERK13/RHS10) drives in *A. thaliana* an escape strategy (ratio height / rosette diameter, HD ratio) in response to the presence of *Poa annua*, a species that interacts with *A. thaliana* in natural populations. While the wild-type and over-expressor lines displayed an increased escape strategy in response to *P. annua*, this response was abolished in the *perk13-1* loss-of-function mutant. The complementation of *perk13-1* by the wild-type *PERK13* gene showed a restored response to *P. annua*. In addition, lines complemented with two highly differentiated natural haplotypes of *PERK13* showed contrasted competitive phenotypes in response to the presence of *P. annua*. Interestingly, in a preliminary experiment, we observed that the presence of *P. annua* resulted in reduced *PERK13* expression in aboveground organs. Together, these results demonstrate that *PERK13* is the causal gene underlying the QTL conferring the response to *P. annua* and acts as a positive regulator of HD ratio. Our work provides the first functional validation of a gene involved in natural variation of competitive response to the presence of a neighboring species.

Key words: *Arabidopsis thaliana*, plant-plant interactions, receptor-like kinase, local population, competitive response.

Introduction

In natural environments, the structure, diversity and dynamics of plant communities are largely mediated by competitive interactions at the interspecific level (Tilman 1985, Goldberg & Barton 1992, Chesson 2000, Martorell & Freckelton 2014). Similarly, in absence of pesticides, negative plant-plant interactions between weeds and crop species represent the main biotic factor reducing crop biomass and grain yield (~36%), in comparison with animal pests (~18%) and pathogens (~16%)(Oerke *et al.* 2006, Neve *et al.* 2009). Despite the importance of interspecific competition in the functioning of natural plant communities and crop fields, our understanding of the genetic and molecular bases associated with natural variation of interspecific competition is largely limited in comparison with other types of biotic interactions such as plant-pathogen interactions (Roux & Bergelson 2016). Yet, it may help to predict the dynamics of natural plant communities in ecological time (Pierik *et al.* 2013, Frachon *et al.* 2017) and to accelerate breeding programs for the selection of crop varieties with enhanced competitive ability against weeds (Worthington & Reberg-Horton 2013). This is even more relevant in the context of current climate change that leads to (i) modifications of plant assemblages, mainly resulting from geographic range shifts that differ among plant species (Gilman *et al.* 2010, Singer *et al.* 2013) and (ii) increased deleterious impact of weeds on crop yield (Basu *et al.* 2004, Peters *et al.* 2014).

So far, only five traditional Quantitative Trait Loci (QTL) mapping studies (based on F2 or Recombinant Inbred Lines mapping populations) and three Genome Wide Association mapping studies (GWAS) reported the genetic architecture of competitive response of a focal species to the presence of neighboring species (reviewed in Subrahmaniam *et al.* 2018, Onishi *et al.* 2018, Libourel *et al.* 2019). In all these studies, the genetic architecture was polygenic,

ranging from the identification of few QTLs with medium effect to the identification of up to tens of QTLs with small effect (reviewed in Subrahmaniam *et al.* 2018). In addition, the genetic architecture was highly dependent on the identity of the neighboring species, the number of species surrounding the focal species and the plant assemblage structure (Libourel *et al.* 2019). In the three GWAS, all using *Arabidopsis thaliana* as a focal species, the fine mapping of genomic regions associated with natural variation of interspecific competition revealed that most candidate genes belong to plant functional categories that have been rarely identified in artificial environments simulating plant-plant interactions, such as shade or root spatial constraints (Subrahmaniam *et al.* 2018). For instance, in the context of monospecific interactions, the candidate genes were mainly related to either cell wall modification, histone modification or meristem identity/life history traits (Baron *et al.* 2015, Frachon *et al.* 2017). On the other hand, receptor-like kinases and transporters were significantly enriched in plurispecific neighborhoods (Libourel *et al.* 2019). Studies reporting the cloning of QTLs associated with interspecific competition are however scarce, not to say absent.

In this study, we therefore aimed at identifying the causative gene underlying a QTL that we previously fine mapped in a local GWA mapping population for the competitive response of *A. thaliana* to the presence of the bluegrass *Poa annua* (Libourel *et al.* 2019). *P. annua* is a common weed in cultivated fields (Warwick 1979) and one of the main grasses co-occurring with *A. thaliana* in natural plant communities and permanent meadows (Frachon *et al.* 2017, Frachon *et al.* 2019). By adopting multiple approaches, we demonstrated that *PROLINE RICH, EXTENSIN-LIKE RECEPTOR KINASE 13* also named *ROOT HAIR SPECIFIC 10* (*PERK13/RHS10*) mediates natural variation of an escape strategy of *A. thaliana* in presence of *P. annua* and common wheat.

Material and Methods

Plant materials

The T-DNA mutant lines of *A. thaliana* used in this study were ordered to the Nottingham Arabidopsis Stock Centre (NASC, http://arabidopsis.info/BasicForm) and are in the Col-0 background (*SI Appendix,* Table S1). One T-DNA mutant line is a GABI-Kat line (GK-345C10) whereas the remaining T-DNA mutants were identified in the SALK library (http://signal.salk.edu). The position of the T-DNA insertion was confirmed by polymerase chain reaction (PCR) using LP and RP primers designed using the online T-DNA Primer Design tools (http://signal.salk.edu/tdnaprimers.2.html) and the specific left border primer T-DNA insertion LBb1.3 (*SI Appendix,* Table S2). Amplicons were sequenced using specific LP, RP and LBb1.3 primers and assembled using Phred, Phrap and Consed software. The results from sequencing were on-line BLAST using the web interface provided by NCBI (https://blast.ncbi.nlm.nih.gov/Blast.cgi) to identify the T-DNA insertional position in the genome. Seeds of the overexpressor line of *PER13/RHS10*, named *RHS10ox*, were kindly provided by the Hyung-Taeg Cho lab. This line, generated in the Col-0 background, contains the *PERK13/RHS10* coding sequence under the control of the *EXPANSIN A7* promoter (Cho & Cosgrove 2002). In order to reduce maternal effects, seeds of all these lines were produced under the same greenhouse conditions.

In this study, we also used the annual bluegrass *Poa annua* (Poaceae) as a neighboring species and six other species, namely the chickweed *Stellaria media* (Caryophyllaceae), the speedwell *Veronica arvensis* (Plantaginaceae), the Kentucky bluegrass *Poa prantensis* (Poaceae), the cat grass *Dactylis glomerata* (Poaceae), the oat *Avena sativa* (Poaceae), the common wheat

Triticum aestivum (Poaceae) (*SI Appendix,* Table S3). Seeds for the first four species were obtained from the Arbiotech company (http://www.arbiotech.com). Seeds for *A. sativa* and *T. aestivum* were kindly provided by Etienne-Pascal Journet (AGIR, INRA, Castanet-Tolosan, France).

***PERK13/RHS10* sequencing, plasmid constructions and transgenic plant generation**

The *PERK13/RHS10* gene and its flanking regions (~5.6kb) from the Col-0 accession and from eight accessions of the local TOU-A mapping population, was sequenced after amplification with the RHS10_LR_Fwd and RHS10_LR_Rv primers using the PrimeSTAR® GXL DNA Polymerase (Takara) (*SI Appendix,* Table S2). Sequencing was performed by the Sanger technology using 23 primers (RHS10_LR_X) to cover the 5.6kb region (*SI Appendix,* Table S2). Sequences were assembled using the Phred, Phrap and Consed softwares. The results were on-line BLAST using the web interface provided by NCBI (Madden 2013, https://blast.ncbi.nlm.nih.gov/Blast.cgi).

To generate the constructs for complementation experiments, amplicons obtained for Col-0, TOU-A1-124 and TOU-A6-61 with the primers attB1_primer and attB2R_primer (*SI Appendix,* Table S2) were cloned into the donor vector pDONR207, using multisite Gateway technology (Life Technologies). Subsequently, the respective constructs were cloned into the pEG301 vector and introduced in *Agrobacterium tumefaciens* (strain GV3101) by electroporation. Three week-old *perk13-1* loss-of-function plants were transformed by floral dip (Clough & Bent 1998). For each construct, at least three independent homozygous lines were selected for phenotyping and molecular characterization.

Measurement of aboveground traits

Experimental design. Phenotyping experiments were replicated three times for each line used in this study (Table S4). For phenotyping the T-DNA mutant lines and the complemented lines, we used for each replicate a split-plot design arranged as a randomized complete block design (RCBD) with two competition treatments (i.e. absence and presence of *P. annua*) nested within blocks (*SI Appendix,* Table S4). We included Col-0 as a control in each experiment (*SI Appendix,* Table S4). For testing the specificity of the competitive response mediated by *PERK13/RHS10* towards other plant species than *P. annua*, we used a split-plot design arranged as a randomized complete block design (RCBD) with eight competition treatments (i.e. absence and presence of seven different species, *SI Appendix,* Table S3) nested within blocks (*SI Appendix,* Table S4).

Growth conditions. The experiments were conducted in the same growth chamber of the Toulouse Plant-Microbe Phenotyping Platform (TPMP, https://www6.toulouse.inra.fr/tpmp). Pots (7cm x 7cm x 6cm) were filled with damp standard culture soil (PROVEEN MOTTE 20, Soprimex). In presence of neighboring species, each *A. thaliana* plant was surrounded by three neighboring plants. Seeds for neighboring plants were evenly spaced, 2 cm away from the *A. thaliana* central position. During the experiments, plants were grown at 20 °C under artificial light to provide a 16-hr photoperiod and were bottom watered without supplemental nutrients. *A. thaliana* focal seedlings and neighbor seedlings were thinned to one per pot 6 to 12 days after seed sowing. Germination date of *A. thaliana* target seedlings was daily monitored.

Phenotypic traits. Two raw phenotypic traits were measured on each focal plant of *A. thaliana* at the time of their flowering, which was measured as the number of days between germination and flowering date. The first trait corresponds to the height from the soil to the first flower on

the main stem (H1F expressed in mm). H1F is related to seed dispersal (Wender *et al.* 2005) and shade avoidance (Dorn *et al.* 2000) in *A. thaliana*. The second trait corresponds to the maximum diameter of the rosette, which was measured at the nearest millimeter (DIAM; Weinig *et al.* 2006). This trait is a proxy of the growth of the rosette of the focal plant from germination to flowering. These traits allowed us to estimate the HD ratio as the height from the soil to the first flower on the rosette diameter (i.e. H1F/DIAM)**.**

Statistical analysis. The following mixed model (PROC MIXED procedure, REML method, SAS 9.3, SAS Institute Inc) was used to explore the phenotypic differences among the different lines:

$$Y_{iikl} = \mu_{trait} + Replicate_j + Block_i\,(Replicate_j) + Treatment_k + Line_l + Block_i \times Treatment_k + Treatment_k \times Line_l + Treatment_k \times Replicate_j + Replicate_j \times Treatment_k \times Line_l + \varepsilon_{iikl} \quad (1)$$

where '*Y*' is the HD ratio scored on focal *A. thaliana* plants, 'μ' is the overall phenotypic mean; 'Replicate' accounts for differences among the temporal replicates; 'Block' accounts for environmental variation among experimental blocks within each replicate; 'Treatment' corresponds to effect of the presence of a neighboring species (absence of competitor *vs* presence of *P. annua, V. arvensis, S. media, P. pratensis, D. glomerata, A. sativa* and *T. aestivum*); 'Line' measures the effect of the different genetic lines; the interaction term 'Treatment' x Line' accounts for variation among genetic lines for their reaction norms across the treatments. All factors were treated as fixed effects. For the calculation of *F*-values, terms were tested over their appropriate denominators. Given the split-split-plot design used in this study, the variance associated with 'Block x Treatment' was used as the error term for testing the 'Block' and 'Treatment' effects. Least Square means (LSmeans) of HD ratio were obtained for each 'Treatment x Line' combination following the model (1).

Measurement of root hair relative traits

Experimental design. Phenotyping of root hair phenotypes of Col and the *perk13-1* loss-of-function mutant were replicated three times. For each replicate, we used a split-plot design arranged as a RCBD with two competition treatments (i.e. absence and presence of one *P. annua* plant) nested within three blocks (*SI Appendix,* Table S4). In each 'Block x Treatment' combination, three plants of Col-0 wild type and the mutant *perk13-1* were randomized for a total of 36 plants per replicate and 108 plants for the full experiment.

Growth conditions. *A. thaliana s*eeds were surface sterilized with 2.4% bleach during two minutes (10 min for *P. annua* seeds) and rinsed three times with sterile water. Seeds were then stored at 4°C in water for stratification and sown on full-strength Murashige and Skoog agar medium (MS Duchefa, M0221) supplemented with 3% sucrose (w/v) and 1‰ B5 vitamin (v/v). Plates were placed vertically in a growth chamber at 20°C with a photoperiod of 16h light /8h dark.

Phenotypic traits. Fourteen days after germination, three root hair related traits were measured on each plant of *A. thaliana*. The first trait corresponds to the root hair density (i.e. total number of root hairs divided by the length of the section investigated). The second trait corresponds to the sum of the total root hair length divided by the length of the section investigated. The third trait corresponds to the mean root hair length. Root hairs were observed under a stereomicroscope (V.16 Zeiss axiozoom, magnification 50x). All root hairs of each side of the main root of each plant were measured with the NeuronJ plugin (Meijering *et al.* 2004) of the ImageJ software. At the end of experiment, 101 plants were phenotyped (~94% of the total number of plants).

Statistical analysis. The following mixed model (PROC MIXED procedure, REML method, SAS 9.3, SAS Institute Inc) was used to explore differences of root hair related traits between Col-0 and the *perk13-1* loss-of-function mutant:

$$Y_{iikl} = \mu_{trait} + Replicate_j + Block_i\,(Replicate_j) + Treatment_k + Line_l + Block_i \times Treatment_k + Treatment_k \times Line_l + \varepsilon_{iikl} \quad (2)$$

where '*Y*' is one of the phenotypic traits scored on focal *A. thaliana* plants, 'μ' is the overall phenotypic mean; 'Experiment' accounts for differences among the three temporal replicates; 'Block' accounts for environmental variation among experimental blocks within each replicate; 'Treatment' corresponds to effect of the presence of *P. annua*; 'Line' measures mean differences between Col-0 and *perk13-1*; the interaction term 'Treatment' x Line' accounts for variation between Col-0 and *perk13-1* in their reaction norms across the two treatments. All factors were treated as fixed effects. For the calculation of *F*-values, terms were tested over their appropriate denominators. Given the split-split-plot design used in this study, the variance associated with 'Block x Treatment' was used as the error term for testing the 'Block' and 'Treatment' effects. LSmeans were obtained for all 'Treatment x Line' combinations following the model (2).

1) <u>Quantitative reverse transcription PCR</u>

Quantitative reverse transcription PCR were conducted to estimate the relative expression of (i) the four genes underlying the QTL investigated, (ii) *PERK13/RHS10* in the complemented lines, and (iii) *PERK13/RHS10* in Col-0 at different time points after sowing (*SI Appendix,* Fig. S1 & 2). In *in vitro* conditions, total RNA was extracted from 14 day-old plant roots from eight to nine biological replicates per genotype from three independent experiments, with the NucleoSpin® RNA kit (Macherey Nagel). The same protocol was used to estimate the

PERK13/RHS10 expression in different organs (leaves, roots, inflorescences, flowers) of Col-0 plants grown in *in vitro* and growth chamber conditions from 4 to 28 day-old plants. 500ng of total RNA were used for cDNA synthesis with the reverse transcriptase Transcriptor according the manufacturer's protocol (Roche E1372). Quantitative RT-PCR reactions were performed in 10µL using SYBR ® Green II master mix (Brilliant II SYBR Green QPCR Master Mix, Sigma-Aldrich). Gene expression was normalized using the housekeeping genes *ACT2* (At3g18780) and *MON1* (At2g28390) *(*Czechowski *et al.* 2005).

Results

In a previous GWAS performed with 91 local French accessions of *A. thaliana* (TOU-A population) grown in absence and presence of *P. annua* (Figure 1*A* and *B)*, we identified a QTL explaining ~15% of genetic variation of both the height from the soil to the first flower on the main stem (H1F) and the HD ratio (i.e. height of the first flower on the rosette diameter, H1F/DIAM) (Libourel *et al.* 2019). A close up indicates that this QTL corresponds to a neat association peak covering a short genomic region of ~20kb (Fig. 1*C*), which includes four genes (Fig. 1*D*). Because the HD ratio quantifies the degree of the escape strategy adopted by *A. thaliana* in response to the presence of neighboring plants (Baron *et al.* 2015) and was under selection in a local population of *A. thaliana* inhabiting a highly competitive environment (Frachon *et al.* 2017), we decided to identify the causal gene underlying this QTL.

PERK13/RHS10* mediates response to *Poa annua

To identify the causal gene associated with natural variation of HD ratio, we phenotyped eight T-DNA mutant lines (Col-0 genetic background) located in the 20kb region underlying the QTL (Figure 1*D*, *SI Appendix*, Table S1) in absence and presence of *P. annua*. We also included in this screen a *PERK13/RHS10* over-expressor line (*RHS10ox*, Hwang *et al.* 2016).

The wild-type Col-0 accession exhibits a significant increase (+40%) of HD ratio in response to the presence of *P. annua* (Fig. 1*D* and *SI Appendix*, Table S5). Six T-DNA mutants showed a similar response than Col-0, whereas no significant increase of HD ratio was observed for the loss-of-function *perk13-1* and *108* mutants (Fig. 1*D* and *SI Appendix*, Table S5). Interestingly, the *RHS10ox* line showed an HD ratio ~38% and ~26% significantly higher than Col-0 in absence and presence of *P. annua*, respectively (Fig. 1*D* and *SI Appendix*, Table S5). To validate *PERK13/RHS10* as the causal gene, we complemented the *perk13-1* loss-of-function mutant with a ~5.5kb DNA region including the Col-0 *PERK13* allele (Fig. 2*A*). Three complemented lines (*perk13-1*|*PERK13*-Col-0) were tested and showed a restored response to the presence of *P. annua* with a significant increase of 57%, 63% and 59% of HD ratio (Fig. 2*B* and *SI Appendix*, Table S6). Together, these results demonstrate that *PERK13/RHS10* is the causal gene underlying the QTL conferring a competitive response of *A. thaliana* to the presence of *P. annua*.

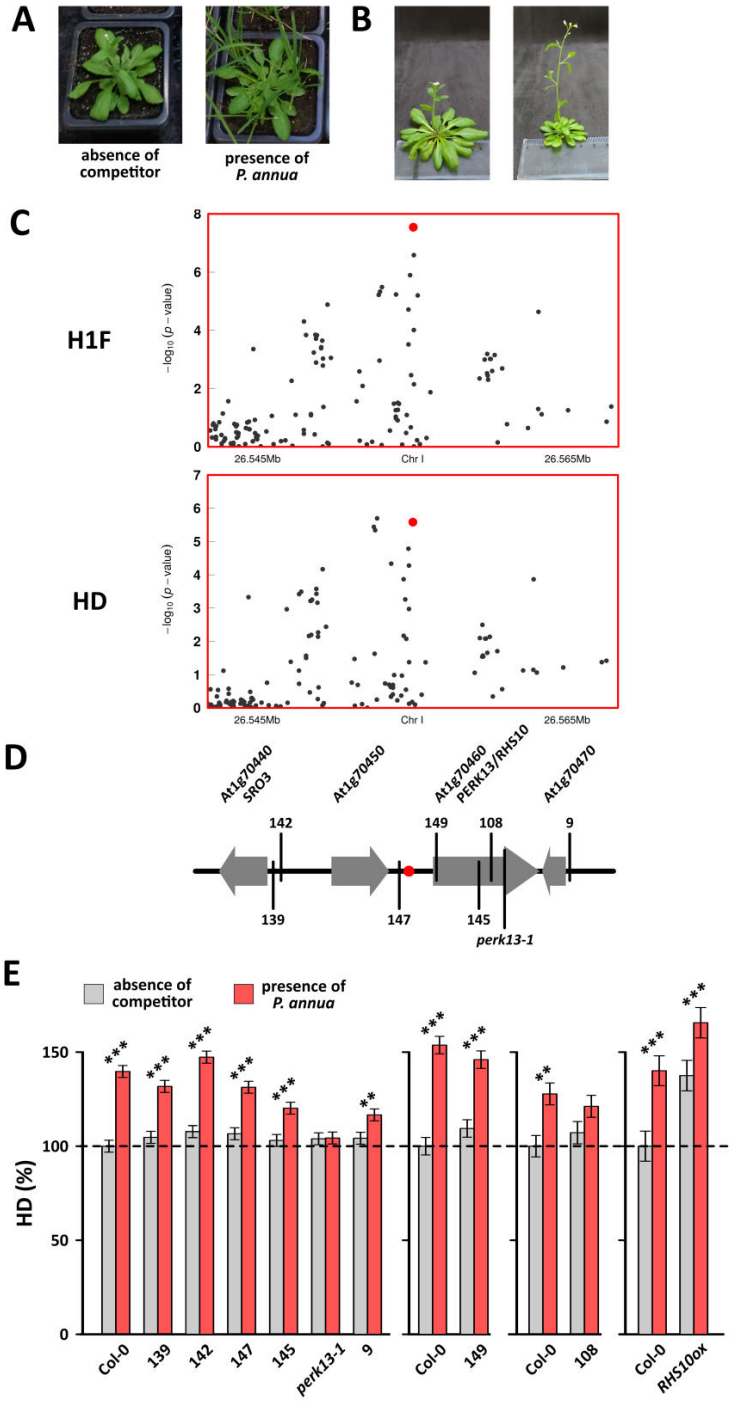

Fig. 1. Identification of *PERK13*/*RHS10* as the gene associated with natural variation of HD ratio in a local population of *A. thaliana* in presence of *P. annua*. (*A*) *A. thaliana plants* growing in absence or in presence of *P. annua.* (*B*) Natural variation of HD ratio. (*C*) Close up of the association peak identified for H1F and HD ratio. (*D*) Schematic representation of the genes underlying the QTL identified in response to the presence of *P. annua*. The red dot indicates the position of the top associated SNP (1_26.555.224).The black vertical lines and numbers indicate the location of the mutations considered in this study. (*E*) Barplots of HD ratio in absence (grey bars) and presence (red bars) of *P. annua* expressed in percentage of HD ratio measured on Col-0 in the control treatment. Data were collected from at least three independent experiments. FDR corrected *p*-values: *0.05 > *P* > 0.01, **0.01 > *P* > 0.001, ***$P < 0.001$, absence of symbols: non-significant.

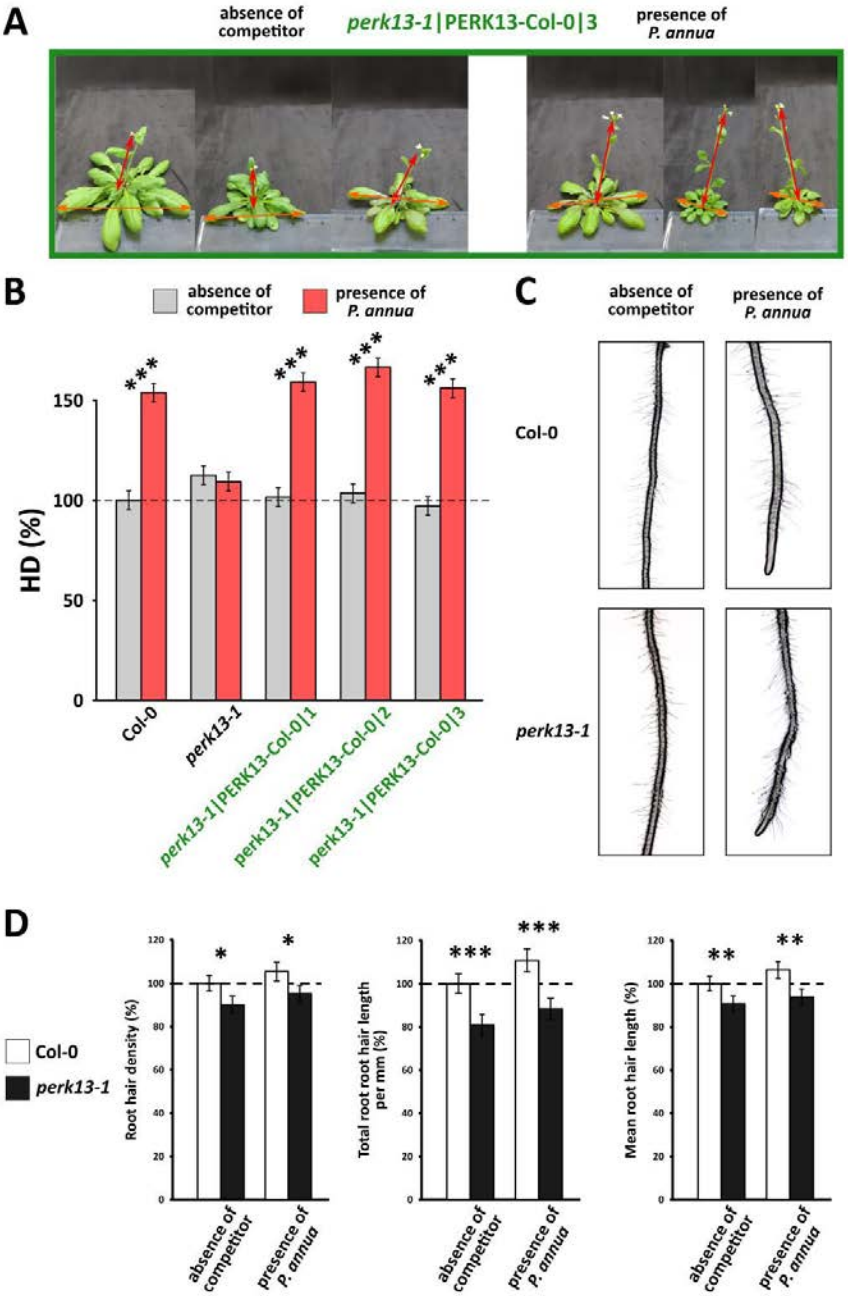

Fig. 2. *PERK13/RHS10* is the causal gene controlling competitive response of *A. thaliana* to the presence of *Poa annua.* (*A*) Illustration of HD ratio for one complemented line in absence and presence of *P. annua*. (*B*) Barplots of HD ratio in absence and presence of *P. annua* expressed in percentage of Col-0 in absence of *P. annua*. FDR corrected *p*-values between the absence and presence of *P. annua* for each genotype: *0.05 > *P* > 0.01, **0.01 > *P* > 0.001, *** *P* < 0.001, absence of symbols: non-significant. (*C*) Representative images of wild type and *perk13-1* mutant plant roots. (*D*) Barplots of root hair density, total root hair length per mm and mean root hair length in Col-0 and the *perk13-1* mutant line in absence and presence of *P. annua*. Data represent LSmeans ±SE from 1916 (29 seedlings, Col-0 in absence of *P. annua*), 1059 (21 seedlings, Col-0 in presence of *P. annua*), 1549 (26 seedlings, *perk13-1* in absence of *P. annua*) and 1172 (25 seedlings, *perk13-1* in presence of *P. annua*) root hairs. Data were collected from three independent experiments. FDR corrected *p*-values between the two genotypes (i.e. 'Line' effect, *SI Appendix,* Table S7): *0.05 > *P* > 0.01, **0.01 > *P* > 0.001, *** *P* < 0.001, absence of symbols: non-significant.

The *RHS10* gene encodes a receptor like kinase whose expression is differentially regulated in leaves and roots in the absence/presence of *P. annua*

PERK13/RHS10 encodes a proline rich extensin-like receptor kinase (PERK) that has been reported to exert a negative control on root hair growth (Hwang *et al.* 2016). In this study, the loss-of-function *rhs10.1* mutant (same mutant as *perk-13-1*) exhibits longer root hairs than the wild-type Col-0, when root hair growth was measured on 3-day-old seedlings under *in vitro* growth conditions (Hwang *et al.* 2016). Accordingly, the *RHS10ox* line over-expressing *PERK13/RHS10* showed a strongly reduced root hair growth (Hwang *et al.* 2016).

To investigate the putative role of belowground interactions between *A. thaliana* and *P. annua*, we measured *in vitro* root hair related traits on 14-day-old Col-0 and *perk13-1* mutant seedlings, in absence and presence of *P. annua* (Fig. 2*C* and *D*). Surprisingly, the mean root hair length is significantly smaller in the *perk13-1* mutant, both in absence and presence of *P. annua* (Fig. 2*D*, *SI Appendix*, Table S6). Similar results were observed for the root hair density and the total root hair length per mm (Fig. 2*D*, *SI Appendix*, Table S6). For these three traits, no significant effect of the presence of *P. annua* was detected (*SI Appendix*, Table S6).

We also measured *PERK13/RHS10* gene expression in Col-0 seedlings grown under *in vitro* or growth chamber conditions, in different organs at different time points in presence and absence of *P. annua* (Fig. 3). Based on only one replicate, we observed a stronger expression level of *PERK13/RHS10* in the root compartment under *in vitro* growth conditions in both 4-day-old and 10-day-old seedlings than in the leaf compartment, as previously reported (Hwang *et al.* 2016)(Fig. 3*A*). In agreement with our root hair phenotypic data, we did not observe any significant impact of the presence of *P. annua* on the relative expression of *PERK13/RHS10* (Fig. 3*A*). In the leaf compartment, we observed an increase of *PERK13/RHS10* relative

expression over time (from 10 to 28 days after sowing) under growth chamber conditions, and *PERK13/RHS10* gene expression was reduced in presence of *P. annua* (Fig. 3*B* and *C*). Interestingly, *PERK13/RHS10* expression was found higher in aboveground organs in absence of *P. annua* 28 days after sowing than in roots at earlier stages (Fig. 3*A*, *B* and *C*). All these results reveal a more complex role and expression pattern for *PERK13/RHS10* than previously described.

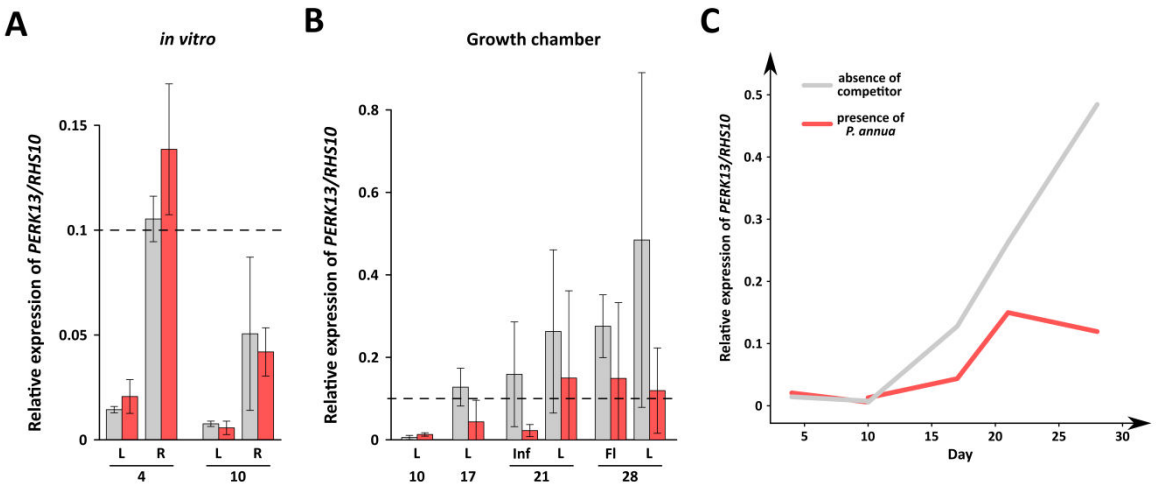

Fig. 3. *PERK13/RHS10* expression in different organs of Col-0 plants grown under *in vitro* and growth chamber conditions, in presence (red bars) or absence (grey bars) of *Poa annua*. (*A*) *PERK13/RHS10* normalized expression level in *in vitro* conditions in presence (red bars) /absence (grey bars) of one *P. annua* plant. (*B*) *PERK13/RHS10* normalized expression level in growth chamber condition in presence/absence of three *P. annua* plants. (*C*) *PERK13/RHS10* normalized expression level in leaves in both *in vitro* and growth chamber conditions in presence/absence of *P. annua* plants. L: leaves; R: roots; Inf: inflorescence; F: flowers.

PERK13/RHS10 sequence polymorphisms are associated with response to the presence of *Poa annua* in natural accessions

To get insight into a putative relationship between *PERK13/RHS10* natural diversity and its role in response to the presence of *P. annua*, we selected eight TOU-A accessions used in the initial

GWAS, based on their allele for the most associated SNP at the position 26,555,224 on chromosome 1, with four accessions chosen for each allele. Among the nine accessions (Col-0 and eight TOU-A accessions) sequenced for a ~5.6kb region encompassing the promoter and coding regions of *PERK13/RHS10*, we detected 93 indels and 88 SNPs (181 polymorphisms, Fig. 4*A*). None of the polymorphisms was present in the transmembrane domain or in the extra-cellular domain of PERK13/RHS10, most of the polymorphisms being located in the kinase domain (49%) and in the promoter region (38%) (Fig. 4*A*). Noteworthy is the absence of non-synonymous mutation or indel among the 17 polymorphisms located in the exons.

Among the eight TOU-A accessions, we identified three haplotypes (Fig. 4*A*). While Haplotype 1 (accessions A1-115, A1-79 and A6-104) and Haplotype 2 (accession A1-124) are closely related to the haplotype of Col-0, Haplotype 3 strongly differs from Col-0 and the two other haplotypes, especially in the kinase domain in which the 88 polymorphisms are in complete linkage disequilibrium. Although a number of polymorphisms are located in the promoter, *PERK13/RHS10* expression level in 14 days plant roots was similar among the haplotypes and was not significantly affected by the presence of *P. annua* (Fig. 4*B*, *SI Appendix*, Table S7).

To test whether these contrasted haplotypes mediate different competitive response to the presence of *P. annua*, we complemented the mutant *perk13-1* with *PERK13*-Haplotype 2 from the TOU-A1-124 accession and with *PERK13*-Haplotype 3 from the TOU-A6-61 accession (Fig. 5*A* and *B*). As previously observed for Col-0, the three complemented lines *perk13-1*|PERK13-A1-124|haplotype2 showed a significant increase of HD ratio in response to the presence of *P. annua* (Fig. 5*B* and *C*, *SI Appendix*, Table S8). In contrast, the three complemented lines *perk13-1*|PERK13-A6-61|haplotype3 showed a similar HD ratio in absence and presence of *P. annua* (Fig. 5*B* and *C*, *SI Appendix*, Table S8).

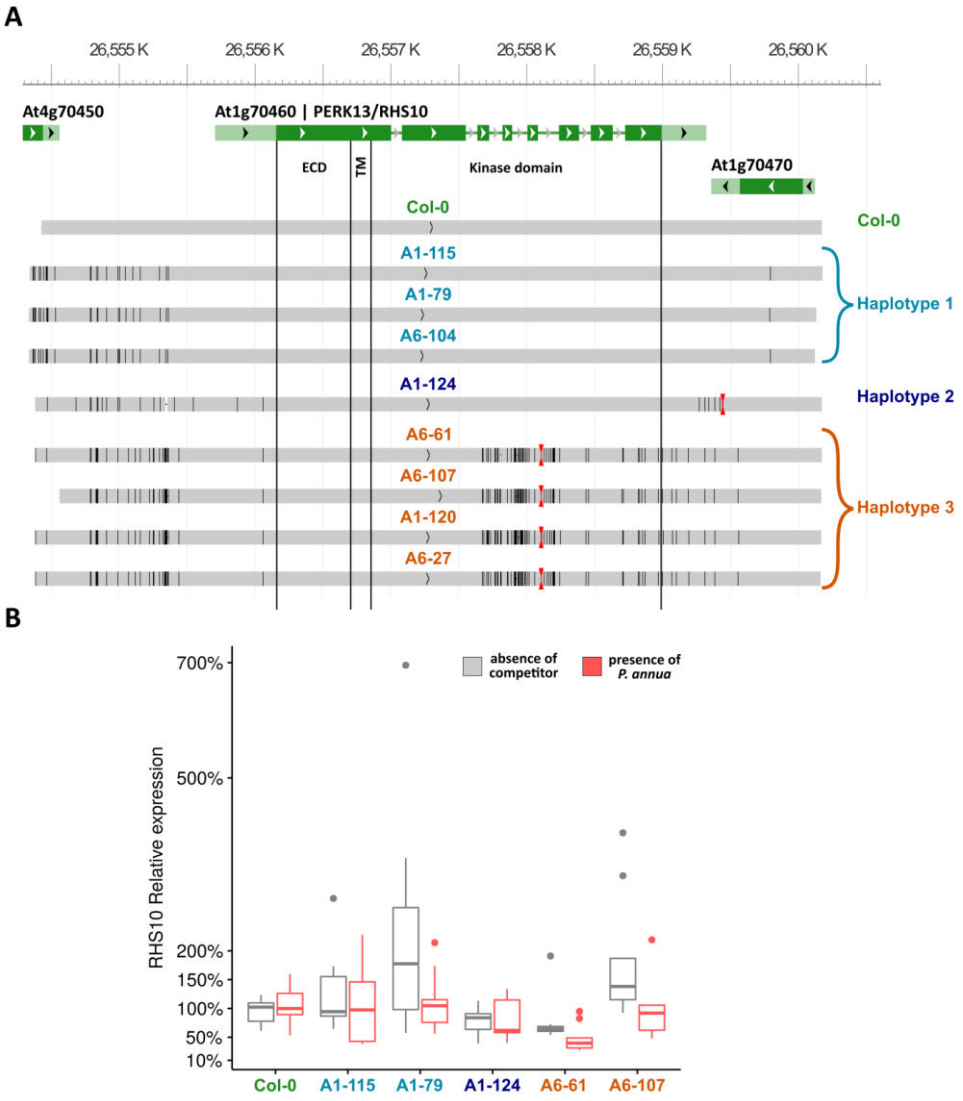

Fig. 4. Natural diversity of *PERK13/RHS10*. (*A*) Sequence diversity observed in a ~5.6kb region centered on *PERK13/RHS10* in Col-0 and eight TOU-A accessions. Black vertical lines indicate mismatches and white vertical lines indicate gaps. Insertions are represented by red hourglasses with a bar on both top and bottom. (*B*) *PERK13/RHS10* normalized expression level in 14 day-old plant roots in Col-0 and five TOU-A accessions in absence or presence of *P. annua* in *in vitro* conditions.

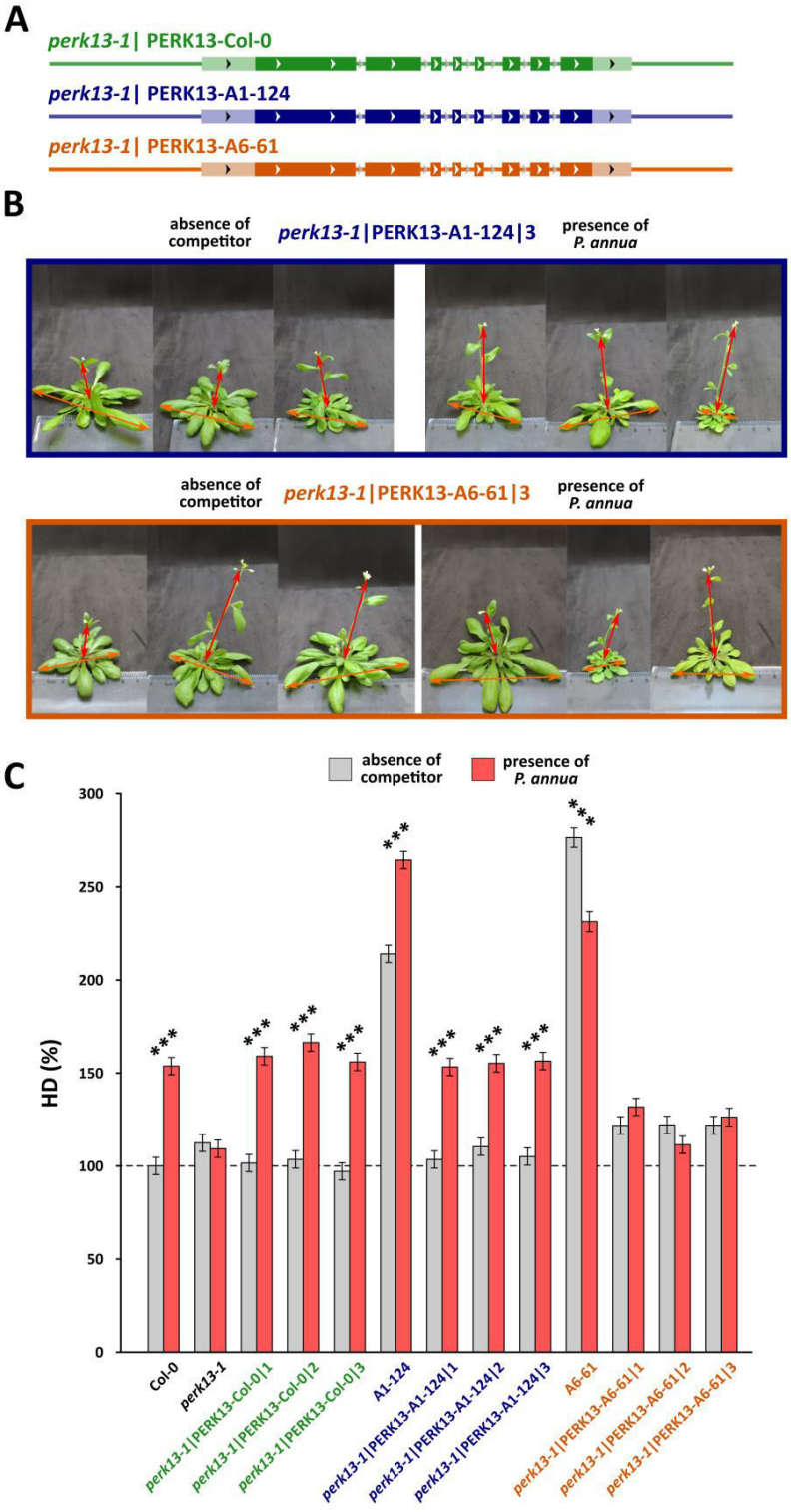

Fig. 5. Complementation of *perk13-1* by three haplotypes of *PERK13/RHS10*. (*A*) Representation of the three haplotypes used for complementation. (*B*) Illustration of HD ratio for two complemented lines in absence and presence of *P. annua*. (*C*) Barplots of HD ratio in absence and presence of *P. annua* expressed in percentage of Col-0 in absence of *P. annua*. FDR corrected *p*-values between absence and presence of *P. annua* for each genotype: *0.05 > *P* > 0.01, **0.01 > *P* > 0.001, *** *P* < 0.001, absence of symbols: non-significant.

Specificity of *PERK13/RHS10* in competitive response

We explored the specificity of *PERK13/RHS10* by growing Col-0, *perk13-1* and *RHS10ox* lines in presence of *P. annua*, four other grass species and two herb species commonly associated with *A. thaliana* in natural plant communities (Fig. 6, *SI Appendix*, Table S3).

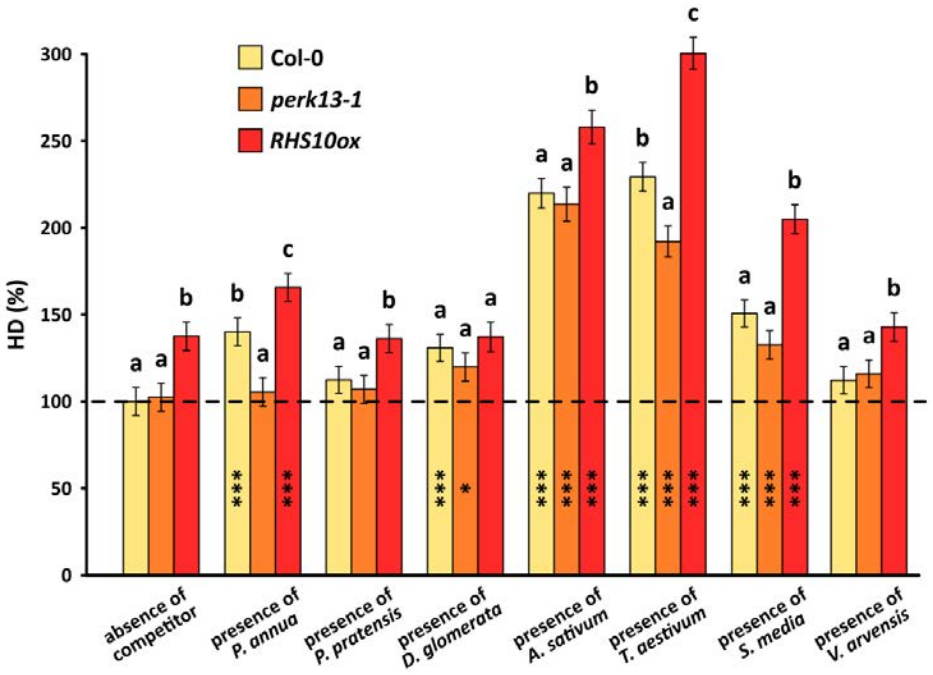

Fig. 6. Barplots illustrating the specificity of *PERK13/RHS10* towards other plant species than *P. annua*. HD ratio in absence and presence of seven species expressed in percentage of HD ratio measured on Col-0 in absence of *P. annua*. FDR corrected *p*-values for each genetic line between each treatment of interspecific competition and absence of competitor: *0.05 > *P* > 0.01, **0.01 > *P* > 0.001, *** *P* < 0.001, absence of symbols: non-significant. For each treatment, different letters indicate different groups according to the genetic lines after a FDR correction.

In absence of any neighboring species, the *perk13-1* showed a similar HD ratio in comparison with Col-0, in contrast with the *RHS10ox* line that exhibited a higher HD ratio (Fig. 6, *SI Appendix*, Table S9). Similar responses were observed when plants were grown in presence of *S. media* and *V. arvensis,* or in presence of the two grass species *A. sativum* and *P. pratensis* (Fig. 6, *SI Appendix*, Table S9). In the opposite, in presence of *T. aestivum*, the three lines

expressed the same relative response than in presence of *P. annua* (Fig. 6, *SI Appendix*, Table S9), with HD ratio being smaller for *perk13-1* and higher for *RHS10ox* in comparison with Col-0 (Fig. 6). Finally, in presence of *D. glomerata*, we observed a different pattern, with the three lines expressing a similar HD ratio (Fig. 6, *SI Appendix*, Table S9).

Discussion

In nature, a wide range of plant-plant interactions can be observed either at the intraspecific or interspecific level, and can be competitive, commensal, cooperative/mutualistic or asymmetric (Subrahmaniam *et al.* 2018). Our understanding of the genetic and molecular bases underlying natural variation of plant–plant interactions is largely limited in comparison with other types of biotic interactions. Only three studies currently led to the identification and functional validation of four genes underlying a specific type of plant-plant interactions, i.e. parasitism. All of them confer resistance to the parasitic plant *Striga* sp. in three crops. A coiled-coil nucleotide-binding site leucine-rich repeat (CC-NBS-LRR) confers resistance to *S. gesneriodes* in cowpea (Li & Timko 2009). Two cytochrome P450 genes (*SBL1* and *SBL2* involved in the biosynthesis of strigolactones) were found responsible for resistance to *S. hermonthia* in rice (Cardoso *et al.* 2014). Another gene, the sulfotransferase *LGS1* (*LOW GERMINATION STIMULANT 1*), underlying a major QTL of resistance to both *S. asiatica* and *S. hermonthica* was identified in sorghum (Gobena *et al.* 2017).

***PERK13/RHS10* | phenotypes.** In a recent study focusing on plant competitive response in *A. thaliana*, we highlighted a putative predominant role of transport proteins and receptor like kinases in response to plurispecific interactions (Libourel *et al.* 2019). In line with these findings, we identified and functionally validated in this study the gene *PERK13/RHS10* as the

gene underlying a QTL identified in monospecific plant-plant interactions (Libourel *et al.* 2019). *PERK13/RHS10* mediates an escape strategy (HD ratio) adopted by natural accessions in response to the presence of *P. annua* (Baron *et al.* 2015, Frachon *et al.* 2017). Interestingly, *PERK13/RHS10* was first characterized as a negative regulator of root hair growth (Won *et al.* 2009, Hwang *et al.* 2016). However in our study, the *perk13-1* mutant had smaller root hairs and a slightly higher root hair density as compared to the wild type. These contrasted results may be due to potentially different plant growth conditions. In our experimental design, *P. annua* did not affect root hair length and density, which is consistent with no change in relative expression of *PERK13/RHS10* in the root compartment. Conversely, we observed a negative effect of *P. annua* on *PERK13/RHS10* relative expression in above-ground organs (i.e. leaves, green inflorescence and flowers) of Col-0, which exhibited an increase of HD ratio. The relationships between *PERK13/RHS10* gene expression/plant compartment (root/aerial part) and the phenotype in response to *P. annua* remain complex and require further studies in the different *PERK13/RHS10* transgenic lines, in the different compartments and in the same environmental conditions. At the opposite, the *RHS10ox* line, which presents an over-expression of *PERK13/RHS10* supposed to be observed mainly in roots (i.e. *AtEXP7* gene promoter, *ProE7*; Cho & Cosgrove 2002), showed a constitutive higher HD ratio in absence of *P. annua* and an increase of HD in response to *P. annua* as Col-0. We cannot exclude however that a certain level of expression might be detectable in aerial organs, previous studies on the root specificity of *PERK13/RHS10* overexpression being based only on Northern analysis (Cho & Cosgrove 2002). Based on these results, we proposed that *PERK13/RHS10* is a positive regulator of the competitive response phenotype (HD ratio). We proposed that a high *PERK13/RHS10* expression in leaves might induce a low HD ratio in absence of *P. annua*. At

the opposite, *PERK13/RHS10* expression is repressed by the presence of *P. annua*, thereby leading to an increase of the HD ratio and then promoting an escape strategy.

***PERK13/RHS10* | molecular functions.** *RHS10* encodes a PERK (proline-rich, extensin-like receptor-like kinase, PERK13) and among the 15 members of this gene family, is closely related to the pollen-specific PERK11 and -12 (Humphrey *et al.* 2015). PERKs are thought to act as sensors/receptors at the cell wall due to their extracellular proline rich, extension like domains (Nakhamchik *et al.* 2004, Humphrey *et al.* 2007) and seem to act as regulators of plant growth and response to stress. *PERK8*, *-9*, and *-10* were shown as *RHS10/PERK13*, to be involved in the control of root growth (Humphrey *et al.* 2015). In addition, *perk4* mutants displayed increased root elongation as compared to the wild type, in response to ABA (Bai *et al.* 2009). Ectopic expression of *BnPERK1* in Arabidopsis resulted in hypocotyl length changes in dark-grown seedlings (Haffani *et al.* 2006). Interestingly, AtPERK1/NsAK was reported to be involved in biotic interactions since it interacts with the geminivirus nuclear shuttle protein (NSP) and acts as a positive regulator of viral infection (Florentino *et al.* 2006). Taken together, these findings suggest multiple functions for PERKs, PERK13/RHS10 being a particularly interesting player since it seems to exert functions both in plant development and biotic interactions. A better understanding of PERK13/RHS10 multiple functions would require the identification of the signaling components operating downstream the receptor like kinase. In this context, an RNase (RNS2) was also identified as a putative downstream target of PERK13/RHS10 which was shown to regulate the accumulation of reactive oxygen species (ROS) in the root (Hwang *et al.* 2016). Interactors of PERK8, 9, 10, which are close homologs of PERK13 and whose functions are unknown, were identified by yeast 2 hybrid screening: KIPK1 and -2 interact with the cytosolic kinase domain and are members of *Arabidopsis* AGC VIII kinases (Humphrey *et al.* 2015). KIPK1 and KIPK-2 belong to the AGC1 subgroup

including D6 PROTEIN KINASE members that are implicated in PIN protein regulation in relation to phototropic responses (Willige *et al.* 2013). A search for interacting partners of PERK13/RHS10, together with a transcriptomic approach, in the context of the competitive interaction between Arabidopsis and *Poa annua* would help to decipher the signaling response initiated by PERK13/RHS10. Identification of the ligand directly or indirectly produced by *P. annua* and *T. aestivum* that both induce the escape strategy observed in *A. thaliana*, would also help to better understand *PERK13/RHS10* functions.

***PERK13/RHS10* | genetic variation.** As previously observed for most *A. thaliana* genes associated with natural variation of plant-pathogen interactions (Roux & Bergelson 2016), a contrasted pattern of nucleotide diversity was observed across *PERK13/RHS10*. No polymorphisms was identified in the extracellular (ECD) and transmembrane domains (TM), suggesting a purifying selection acting on these two domains. The absence of polymorphisms in ECD is in line with the putative function of this domain in PERKs, i.e. binding cell wall compounds, such as extensin proteins (Hwang *et al.* 2016). On the other hand, we identified highly differentiated haplotypes with tens of polymorphisms in complete (LD), suggesting the maintenance of long-lived polymorphisms associated with plant-plant interactions through balancing selection. Such a signature of selection has already been reported in *A. thaliana* for a handful of genes associated with natural variation of resistance to pathogenic bacteria (Roux & Bergelson 2016). In most cases, signatures of balancing selection were observed on gene presence/absence polymorphisms, such as the *R* genes *RPM1* and *RPS5* (Stahl *et al.* 1999, Tian *et al.* 2002). More recently, signatures of selection acting on a promoter region has been reported for the *RKS1* gene conferring quantitative resistance to the bacterial pathogen *Xanthomonas campestris* (Huard-Chauveau *et al.* 2013). In our study, intriguingly, balancing selection seems to act mainly on the kinase domain in which none of the 88 polymorphisms in

complete LD is non-synonymous. However, we cannot rule out that some promoter polymorphisms in complete LD with the kinase domain polymorphisms also contribute to natural variation of response to the presence of *P. annua*. A deeper functional analysis of both the promoter region and the kinase domain is clearly needed to identify the causal polymorphism(s) that are target(s) of balancing selection.

ACKNOWLEDGEMENTS

We are also grateful to the staff of the LIPM greenhouse for their assistance during the growth chamber experiments. This work was funded by a PhD fellowship from the University of Paul Sabatier Toulouse to CL. This study was also supported by the LABEX TULIP (ANR-10-LABX-41; ANR-11-IDEX-0002-02). Part of this work was carried out on the Toulouse Plant-Microbe Phenotyping facility (http://tpmp.inra.fr), which is part of the LIPM – UMR INRA441/CNRS2594.

Data Accessibility

Raw phenotypic data will be available in the Dryad database: doi:10.5061/dryad.XXX.

Author contributions

F.R. and D.R. supervised the project. F.R., D.R. and C.L. designed the experiments. C.L. and F.R. conducted the greenhouse experiments. C.L. and F.R. measured the phenotypic traits. C.L. analyzed the phenotypic traits. C.L., F.R. and D.R. wrote the manuscript.

References

Bai, L., Zhang, G., Zhou, Y., Zhang, Z., Wang, W., Du, Y., ... & Song, C. P. (2009). Plasma membrane-associated proline-rich extensin-like receptor kinase 4, a novel regulator of Ca2+ signalling, is required for abscisic acid responses in *Arabidopsis thaliana*. *The Plant Journal*, *60*(2), 314-327.

Baron, E., Richirt, J., Villoutreix, R., Amsellem, L., & Roux, F. (2015). The genetics of intra-and interspecific competitive response and effect in a local population of an annual plant species. *Functional Ecology*, *29*(10), 1361-1370.

Bartoli, C., & Roux, F. (2017). Genome-wide association studies in plant pathosystems: toward an ecological genomics approach. *Frontiers in plant science*, *8*, 763.

Basu, C., Halfhill, M. D., Mueller, T. C., & Stewart Jr, C. N. (2004). Weed genomics: new tools to understand weed biology. *Trends in plant science*, *9*(8), 391-398.

Cardoso, C., Zhang, Y., Jamil, M., Hepworth, J., Charnikhova, T., Dimkpa, S. O., ... & Wang, Y. (2014). Natural variation of rice strigolactone biosynthesis is associated with the deletion of two MAX1 orthologs. *Proceedings of the National Academy of Sciences*, *111*(6), 2379-2384.

Chesson, P. (2000). Mechanisms of maintenance of species diversity. *Annual review of Ecology and Systematics*, *31*(1), 343-366.

Cho, H. T. (2016). Arabinogalactan protein motif-containing receptor-like kinases are likely to play the negative feedback factor to maintain proper root hair length. *Plant signaling & behavior*, *11*(9), 2007-22.

Cho, H. T., & Cosgrove, D. J. (2002). Regulation of root hair initiation and expansin gene expression in Arabidopsis. *The Plant Cell*, *14*(12), 3237-3253.

Clough, S. J., & Bent, A. F. (1998). Floral dip: a simplified method for Agrobacterium-mediated transformation of *Arabidopsis thaliana*. *The plant journal*, *16*(6), 735-743.

Czechowski, T., Stitt, M., Altmann, T., Udvardi, M. K., & Scheible, W. R. (2005). Genome-wide identification and testing of superior reference genes for transcript normalization in Arabidopsis. *Plant physiology*, *139*(1), 5-17.

Dorn, L. A., Pyle, E. H., & Schmitt, J. (2000). Plasticity to light cues and resources in *Arabidopsis thaliana*: testing for adaptive value and costs. *Evolution*, *54*(6), 1982-1994.

Feng, L., Gao, Z., Xiao, G., Huang, R., & Zhang, H. (2014). Leucine-rich repeat receptor-like kinase FON1 regulates drought stress and seed germination by activating the expression of ABA-responsive genes in rice. *Plant molecular biology reporter*, *32*(6), 1158-1168.

Florentino, L. H., Santos, A. A., Fontenelle, M. R., Pinheiro, G. L., Zerbini, F. M., Baracat-Pereira, M. C., & Fontes, E. P. (2006). A PERK-like receptor kinase interacts with the geminivirus nuclear shuttle protein and potentiates viral infection. *Journal of virology*, *80*(13), 6648-6656.

Frachon, L., C. Bartoli, B. Mayjonade, N.C. Hautekèete and F. Roux. Adaptation to plant communities across the genome of Arabidopsis thaliana. *Molecular Biology and Evolution* (in revision).

Frachon, L., Libourel, C., Villoutreix, R., Carrère, S., Glorieux, C., Huard-Chauveau, C., ... & Roux, F. (2017). Intermediate degrees of synergistic pleiotropy drive adaptive evolution in ecological time. *Nature ecology & evolution*, *1*(10), 1551.

Gilman, S. E., Urban, M. C., Tewksbury, J., Gilchrist, G. W., & Holt, R. D. (2010). A framework for community interactions under climate change. *Trends in ecology & evolution*, *25*(6), 325-331.

Gobena, D., Shimels, M., Rich, P. J., Ruyter-Spira, C., Bouwmeester, H., Kanuganti, S., ... & Ejeta, G. (2017). Mutation in sorghum LOW GERMINATION STIMULANT 1 alters strigolactones and causes *Striga* resistance. *Proceedings of the National Academy of Sciences*, *114*(17), 4471-4476.

Goldberg, D. E., & Barton, A. M. (1992). Patterns and consequences of interspecific competition in natural communities: a review of field experiments with plants. *The American Naturalist*, *139*(4), 771-801.

Haffani, Y., Silva-Gagliardi, N., Sewter, S., Aldea, M. G., Zhao, Z., Nakhamchik, A., ... & Goring, D. (2006). Altered expression of PERK receptor kinases in Arabidopsis leads to changes in growth and floral organ formation. *Plant signaling & behavior*, *1*(5), 251-260.

Huard-Chauveau, C., Perchepied, L., Debieu, M., Rivas, S., Kroj, T., Kars, I., ... & Roby, D. (2013). An atypical kinase under balancing selection confers broad-spectrum disease resistance in Arabidopsis. *PLoS genetics*, *9*(9), e1003766.

Humphrey, T. V., Bonetta, D. T., & Goring, D. R. (2007). Sentinels at the wall: cell wall receptors and sensors. *New Phytologist*, *176*(1), 7-21.

Humphrey, T. V., Haasen, K. E., Aldea-Brydges, M. G., Sun, H., Zayed, Y., Indriolo, E., & Goring, D. R. (2015). PERK–KIPK–KCBP signalling negatively regulates root growth in *Arabidopsis thaliana*. *Journal of experimental botany*, *66*(1), 71-83.

Hwang, Y., Lee, H., Lee, Y. S., & Cho, H. T. (2016). Cell wall-associated ROOT HAIR SPECIFIC 10, a proline-rich receptor-like kinase, is a negative modulator of Arabidopsis root hair growth. *Journal of experimental botany*, *67*(6), 2007-2022.

Libourel, C., Baron, E.,Lenglet, J., Amsellem, L., Roby, D., & Roux F. The genomic architecture of competitive response of *Arabidopsis thaliana* is highly flexible between monospecific and plurispecific neighborhoods. (In preparation).

Madden, T. (2013). The BLAST sequence analysis tool.

Martorell, C., & Freckleton, R. P. (2014). Testing the roles of competition, facilitation and stochasticity on community structure in a species-rich assemblage. *Journal of Ecology*, *102*(1), 74-85.

Meijering, E., Jacob, M., Sarria, J. C., Steiner, P., Hirling, H., & Unser, M. (2004). Design and validation of a tool for neurite tracing and analysis in fluorescence microscopy images. *Cytometry Part A: the journal of the International Society for Analytical Cytology*, *58*(2), 167-176.

Nakhamchik, A., Zhao, Z., Provart, N. J., Shiu, S. H., Keatley, S. K., Cameron, R. K., & Goring, D. R. (2004). A comprehensive expression analysis of the Arabidopsis proline-rich extensin-like receptor kinase gene family using bioinformatic and experimental approaches. *Plant and Cell Physiology*, *45*(12), 1875-1881.

Neve, P., Vila-Aiub, M., & Roux, F. (2009). Evolutionary-thinking in agricultural weed management. *New Phytologist*, *184*(4), 783-793.

Ninkovic, V. (2003). Volatile communication between barley plants affects biomass allocation. *Journal of Experimental Botany*, *54*(389), 1931-1939.

Oerke, E. C. (2006). Crop losses to pests. *The Journal of Agricultural Science*, *144*(1), 31-43.

Onishi, K., Ichikawa, N., Horiuchi, Y., Kohara, H., & Sano, Y. (2018). Genetic architecture underlying the evolutionary change of competitive ability in Asian cultivated and wild rice. *Journal of Plant Interactions*, *13*(1), 442-449.

Peters, K., Breitsameter, L., & Gerowitt, B. (2014). Impact of climate change on weeds in agriculture: a review. *Agronomy for sustainable development*, *34*(4), 707-721.

Pierik, R., Mommer, L., & Voesenek, L. A. (2013). Molecular mechanisms of plant competition: neighbour detection and response strategies. *Functional Ecology*, *27*(4), 841-853.

Platt, A., Horton, M., Huang, Y. S., Li, Y., Anastasio, A. E., Mulyati, N. W., ... & Dunning, M. (2010). The scale of population structure in *Arabidopsis thaliana*. *PLoS genetics*, *6*(2), e1000843.

Roux, F., & Bergelson, J. (2016). The genetics underlying natural variation in the biotic interactions of Arabidopsis thaliana: The challenges of linking evolutionary genetics and community ecology. In *Current topics in developmental biology* (Vol. 119, pp. 111-156). Academic Press.

Singer, A., Travis, J. M., & Johst, K. (2013). Interspecific interactions affect species and community responses to climate shifts. *Oikos*, *122*(3), 358-366.

Stahl, E. A., Dwyer, G., Mauricio, R., Kreitman, M., & Bergelson, J. (1999). Dynamics of disease resistance polymorphism at the Rpm1 locus of Arabidopsis. *Nature*, *400*(6745), 667.

Subrahmaniam, H. J., Libourel, C., Journet, E. P., Morel, J. B., Muños, S., Niebel, A., ... & Roux, F. (2018). The genetics underlying natural variation of plant–plant interactions, a beloved but forgotten member of the family of biotic interactions. *The Plant Journal*, *93*(4), 747-770.

Tian, D., Araki, H., Stahl, E., Bergelson, J., & Kreitman, M. (2002). Signature of balancing selection in Arabidopsis. *Proceedings of the National Academy of Sciences*, *99*(17), 11525-11530.

Tilman, D. (1985). The resource-ratio hypothesis of plant succession. *The American Naturalist*, *125*(6), 827-852.

Warwick, S. I. (1979). THE BIOLOGY OF CANADIAN WEEDS.: 37 *Poa annua* L. *Canadian Journal of Plant Science*, *59*(4), 1053-1066.

Weinig, C., Johnston, J., German, Z. M., & Demink, L. M. (2006). Local and global costs of adaptive plasticity to density in *Arabidopsis thaliana*. *The American Naturalist*, *167*(6), 826-836.

Wender, N. J., Polisetty, C. R., & Donohue, K. (2005). Density-dependent processes influencing the evolutionary dynamics of dispersal: a functional analysis of seed dispersal in *Arabidopsis thaliana* (Brassicaceae). *American Journal of Botany*, *92*(6), 960-971.

Willige, B. C., Ahlers, S., Zourelidou, M., Barbosa, I. C., Demarsy, E., Trevisan, M., .. & Schwechheimer, C. (2013). D6PK AGCVIII kinases are required for auxin transport and phototropic hypocotyl bending in Arabidopsis. *The Plant Cell*, tpc-113.

Won, S. K., Lee, Y. J., Lee, H. Y., Heo, Y. K., Cho, M., & Cho, H. T. (2009). Cis-element-and transcriptome-based screening of root hair-specific genes and their functional characterization in Arabidopsis. *Plant physiology*, *150*(3), 1459-1473.

Worthington, M., & Reberg-Horton, C. (2013). Breeding cereal crops for enhanced weed suppression: optimizing allelopathy and competitive ability. *Journal of Chemical Ecology*, *39*(2), 213-231.

Supplementary Information

An Arabidopsis receptor-like kinase mediates competitive plant-plant interactions.

Cyril Libourel, Fabrice Roux*, Dominique Roby*

* Correspondence: Corresponding Author: dominique.roby@inra.fr & fabrice.roux@inra.fr

2 Supplementary Figures

9 Supplementary Tables

Figure S1. Relative expressions of the four genes underlying the QTL in the root compartment of 14 day-old seedlings in *in vitro* conditions. For each gene, relative expressions are displayed for T-DNA mutant lines with the insertional sites inside the target gene or in the flanking region.

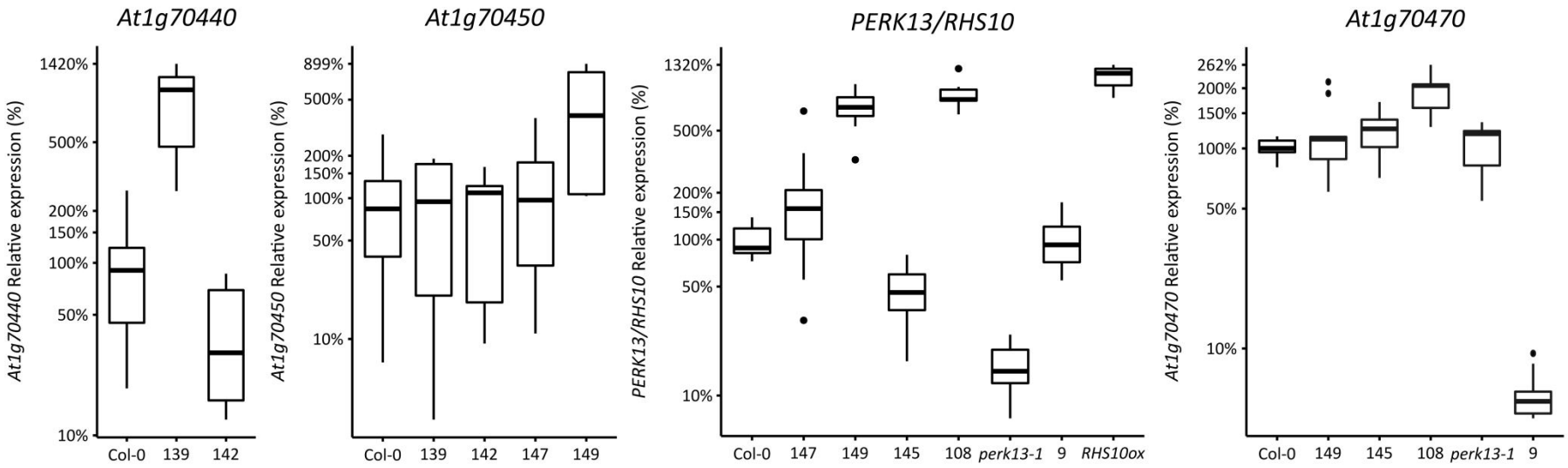

Figure S2. Relative expression of *RHS10* in root compartment of 14 day-old seedlings in *in vitro* conditions.

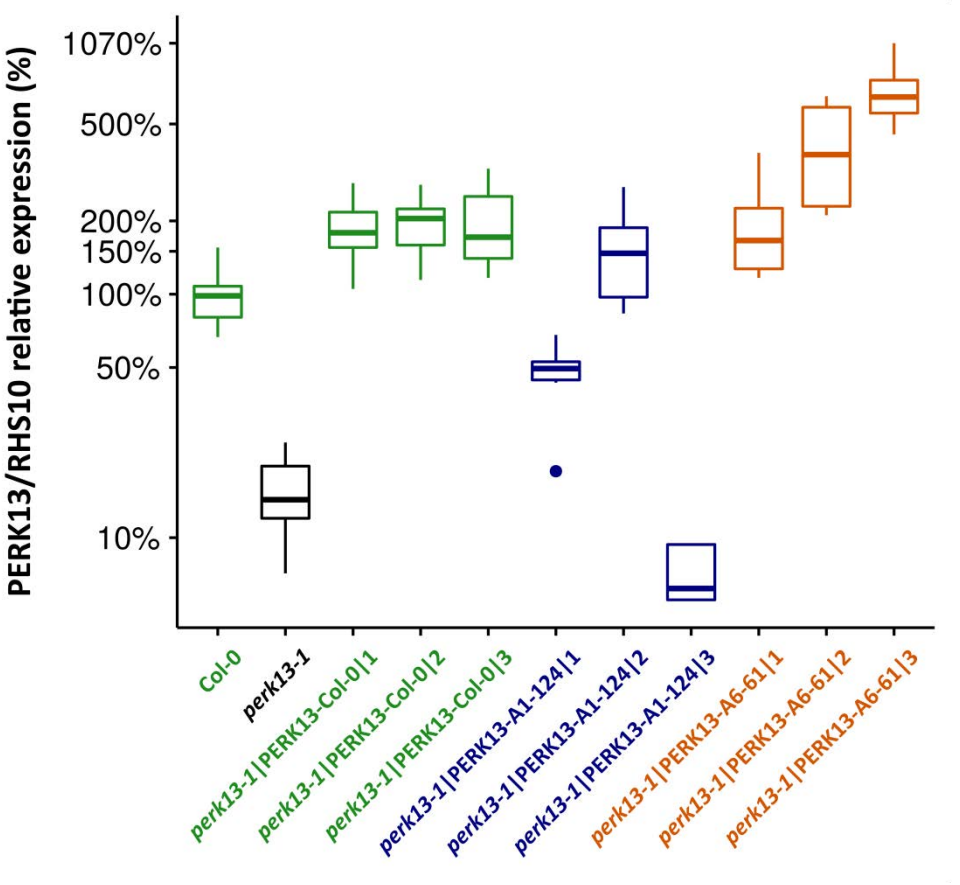

Table S1. List of T-DNA mutant lines used to identify the gene underlying the QTL.

id	T-DNA mutant
9	SALK_075882C
108	GK-345C10
139	SALK_125515C
142	SALK_124026
145	SALK_043490
147	SALK_045625
149	SALK_079932
perk13-1	SALK_075892

Table S2. Oligonucleotides used in this study.

Gene	Oligonucleotides	Sequence
PERK13/RHS10	RHS10_LR_Fwd	ACAAAGCAGTACTTGTGACGACTC
PERK13/RHS10	RHS10_LR_F1	TTGAAATACTTAAAAAACTAATGCGTGT
PERK13/RHS10	RHS10_LR_F2	GGATAAAACTTGAAGATATGTCAAACA
PERK13/RHS10	RHS10_LR_F3	AAAAGAAGTTTCACTAATATGGTGCC
PERK13/RHS10	RHS10_LR_F4	TCATCACAACAATCACTAAATTCAAA
PERK13/RHS10	RHS10_LR_F5	TCTGTTATACAGAAACAAATTCAAACAA
PERK13/RHS10	RHS10_LR_F6	ACCTCCAGAGGTGTTTGAACC
PERK13/RHS10	RHS10_LR_F7	GGCCGTTGTGTTCTTAGTCAG
PERK13/RHS10	RHS10_LR_F8	ACATACTTGGAGAAGGAGGTTTTG
PERK13/RHS10	RHS10_LR_F9	TGGCGTATTTACACGAAGACTG
PERK13/RHS10	RHS10_LR_F10	TTATGGGAACCTTCGGGTAAG
PERK13/RHS10	RHS10_LR_F11	CTCGTCCTCTGCTTCACAAAG
PERK13/RHS10	RHS10_LR_F12	CAGTGGAGACTACTCTGTCCAAGA
PERK13/RHS10	RHS10_LR_F13	CGGATTTTTTGTTGAGTGTGATAA
PERK13/RHS10	RHS10_LR_F14	ATTAAGACAATGTTCTGACCGGA
PERK13/RHS10	RHS10_LR_Rv	GATCAAGGAAAGGAGTGAGTGCTA
PERK13/RHS10	RHS10_LR_F9_bis	TAGCATATTTGCATGAAGACTG
PERK13/RHS10	RHS10_LR_R1	CCCACCTCCTTACACTGTAGA
PERK13/RHS10	RHS10_LR_R2	GCCGTGAGATTAGACCTGTG
PERK13/RHS10	RHS10_LR_R3	TAGCTGGAGAAAGATGAGAAAT
PERK13/RHS10	RHS10_LR_R4	CCATATATTGAGATGTCAAGACAGA
PERK13/RHS10	RHS10_LR_R5	GGAGGTGGAGGAAAAGGA
PERK13/RHS10	RHS10_LR_F3_bis	CATTTAAAGAACAAGAGCAATCAAGA
PERK13/RHS10	attB1_primer	GGGGACAAGTTTGTACAAAAAAGCAGGCT
PERK13/RHS10	attB2R_primer	GGGGACCACTTTGTACAAGAAAGCTGGGT
T-DNA	LBb1.3	ATTTTGCCGATTTCGGAAC
At3g18780	R1_Fwd	TTCCGCTCTTTCTTTCCAAGCTC

At3g18780	R1_Rv	CGGTACCATTGTCACACACGATTG
At1g70440	G1_Fwd	AGTCACTTGACGCTGCAAGAATG
At1g70440	G1_Rv	GCTCTCTCCTCAGTTTCCGTTTCC
At1g70450	G2_Fwd	ACATTGGAGCATCATTTGCATGGG
At1g70450	G2_Rv	GACTCTCCTAGCCCATTCAAGCAC
PERK13/RHS10	G3_Fwd	AACGTCCACGTATGGTTCAGGTTG
PERK13/RHS10	G3_Rv	ACTGCTTTGTCCCACTTTGTTACC
At2g28390	R2_Fwd	TTGATCCACTTGCAGACAAGGC
At2g28390	R2_Rv	TACCCTTTGGCACACCTGATTG
At1g70470	G4_Fwd	CGCGGTGGAAAGCAACACAAAG
At1g70470	G4_Rv	TCACGGCTCTTCTTTGACGATGG
PERK13/RHS10	G5_Fwd	TGATGCCTACAGTGACTCACAA
PERK13/RHS10	G5_Rv	TTGACCGTATAAGAATCCATCTGA

Table S3. List of species used to test the specificity of *PERK13/RHS10*.

Latin name		Family	Common name
Poa annua	*P.annua*	Poaceae	meadow grass
Poa prantensis	*P. prantensis*	Poaceae	blue grass
Dactylis glomerata	*D. glomerata*	Poaceae	cat grass
Veronica arvensis	*V. arvensis*	Plantaginaceae	corn speedwell
Stellaria media	*S. media*	Caryophyllaceae	chickweed
Avena sativa	*A. sativa*	Poaceae	oat
Triticum aestivum	*T. aestivum*	Poaceae	common wheat

Table S4. Characteristics of the experiments performed in this study.

Experiment	Number of plant per Block x Treatment	Number of blocks per experiment	Col-0	*9*	*139*	*142*	*145*	*147*	*perk13-1*	*149*	*108*	*RHS10ox*	*perk13-1*|PERK13-Col-0|1	*perk13-1*|PERK13-Col-0|2	*perk13-1*|PERK13-Col-0|3	A1-124	*perk13-1*|PERK13-A1-124|1	*perk13-1*|PERK13-A1-124|2	*perk13-1*|PERK13-A1-124|3	A6-61	*perk13-1*|PERK13-A6-61|1	*perk13-1*|PERK13-A6-61|2	*perk13-1*|PERK13-A6-61|3
A	2	37	X	X	X	X	X	X	X														
B	4	15	X								X												
C	2	28	X						X	X			X	X	X	X	X	X	X	X	X	X	X
D	7 | 9 | 9	4	X						X			X											

Table S5. Treatment effect (absence *vs* presence of *Poa annua*) for each line. Bold *p*-values indicate significant effects after FDR correction.

Experiment	Effect	Line	F Value	*p*-val
A	Treatment	Col-0	78.66	**2.02E-17**
A	Treatment	9	7.66	**6.73E-03**
A	Treatment	139	35.84	**8.18E-09**
A	Treatment	142	76.62	**2.74E-17**
A	Treatment	145	14.68	**2.14E-04**
A	Treatment	147	30.29	**9.07E-08**
A	Treatment	*perk13-1*	0.01	9.23E-01
B	Treatment	108	2.83	1.00E-01
B	Treatment	Col-0	11.68	**8.39E-04**
C	Treatment	149	30.93	**8.19E-08**
C	Treatment	Col-0	66.77	**2.74E-15**
D	Treatment	*RHS10ox*	12.98	**8.09E-04**
D	Treatment	Col-0	27.35	**1.88E-06**

Table S6. Variation of the three root hair related traits between Col-0 and *perk13-1* in presence and absence of *P. annua*. Bold *p*-values indicate significant effects after FDR correction.

Trait	Effect	F Value	*p*-val
Root hair density	**Block**	7.05381	**4.39E-06**
	Treatment	1.75726	3.22E-01
	Line	6.36948	**3.22E-02**
	Treatment*Line	0.00189	9.65E-01
Total root hair length per root length	**Block**	3.52592	**4.50E-03**
	Treatment	3.53846	1.27E-01
	Line	17.75527	**3.64E-04**
	Treatment*Line	0.11420	8.03E-01
Mean root hair length	**Block**	0.87	7.60E-01
	Treatment	1.5	3.90E-01
	Line	10.84	**4.50E-03**
	Treatment*Line	0.23	7.60E-01

Table S7. Pairwise differences of *PERK13/RHS10* normalized expression level in 14 day-old plant roots in Col-0 and five TOU-A accessions in absence or presence of *P. annua* in *in vitro* conditions. *P-values* obtain from the pairwise.wilcox.test in R are displayed for all pairwise differences**.** Bold *p*-values indicate significant effects after FDR correction. Green boxes highlight *p*-values representing the effect of the presence of *P. annua* on the expression of *PERK13/RHS10* of a given genotype.

	A1-115_presence_of_*P. annua*	A1-115_absence_of_competitor	A1-124_presence_of_*P. annua*	A1-124_absence_of_competitor	A1-79_presence_of_*P. annua*	A1-79_absence_of_competitor	A6-107_presence_of_*P. annua*	A6-107_absence_of_competitor	A6-61_presence_of_*P. annua*	A6-61_absence_of_competitor	Col-0_presence_of_*P. annua*
A1-115_absence_of_competitor	8.27E-01	-	-	-	-	-	-	-	-	-	-
A1-124_presence_of_*P. annua*	7.35E-01	2.35E-01	-	-	-	-	-	-	-	-	-
A1-124_absence_of_competitor	5.43E-01	3.32E-01	9.77E-01	-	-	-	-	-	-	-	-
A1-79_presence_of_*P. annua*	9.50E-01	9.50E-01	3.39E-01	4.56E-01	-	-	-	-	-	-	-
A1-79_absence_of_competitor	2.14E-01	4.56E-01	9.93E-02	**3.38E-02**	2.35E-01	-	-	-	-	-	-
A6-107_presence_of_*P. annua*	9.61E-01	8.24E-01	7.69E-01	7.57E-01	6.87E-01	1.65E-01	-	-	-	-	-
A6-107_absence_of_competitor	2.72E-01	2.81E-01	**3.38E-02**	**1.27E-02**	1.65E-01	9.76E-01	9.73E-02	-	-	-	-
A6-61_presence_of_*P. annua*	1.20E-01	**3.38E-02**	1.04E-01	1.13E-01	**2.60E-02**	**5.40E-03**	**3.38E-02**	**5.40E-03**	-	-	-
A6-61_absence_of_competitor	6.87E-01	9.93E-02	9.50E-01	4.93E-01	1.04E-01	**3.95E-02**	4.93E-01	**2.44E-02**	1.04E-01	-	-
Col-0_presence_of_*P. annua*	9.76E-01	9.77E-01	3.05E-01	2.35E-01	1.00E+00	2.14E-01	5.43E-01	1.88E-01	**2.04E-02**	1.65E-01	-
Col-0_absence_of_competitor	9.50E-01	8.27E-01	4.79E-01	4.56E-01	9.50E-01	1.88E-01	7.69E-01	**3.95E-02**	**2.60E-02**	9.93E-02	7.35E-01

Table S8. Treatment effect (absence *vs* presence of *Poa annua*) for the wild-type, *perk13-1* and complemented lines and two natural accessions (TOU-A1-124 and TOU-A6-61). Bold *p*-values indicate significant effects after FDR correction.

Effect	Line	F Value	*p*-val
Treatment	Col-0	66.77	**2.05E-15**
Treatment	*perk13-1*	0.23	6.34E-01
Treatment	*perk13-1*|PERK13-Col-0|1	75.74	**3.55E-17**
Treatment	*perk13-1*|PERK13-Col-0|2	90.00	**1.14E-19**
Treatment	*perk13-1*|PERK13-Col-0|3	79.52	**8.60E-18**
Treatment	A1-124	58.54	**7.57E-14**
Treatment	*perk13-1*|PERK13-A1-124|1	57.39	**1.14E-13**
Treatment	*perk13-1*|PERK13-A1-124|2	45.64	**3.26E-11**
Treatment	*perk13-1*|PERK13-A1-124|3	60.51	**3.47E-14**
Treatment	A6-61	36.24	**3.15E-09**
Treatment	*perk13-1*|PERK13-A6-61|1	2.29	1.54E-01
Treatment	*perk13-1*|PERK13-A6-61|2	2.65	1.35E-01
Treatment	*perk13-1*|PERK13-A6-61|3	0.42	5.60E-01

Table S9. Variation among Col-0, *perk13-1* and *RHS10ox* lines for the response of HD ratio to the presence of seven neighboring species. Bold *p*-values indicate significant effects after FDR correction.

	Treatments						
	P.annua	*P.pratensis*	*D.glomerata*	*A.sativa*	*T.aestivum*	*S.media*	*V.arvensis*
Experiment	7.04E-01	8.73E-01	8.06E-01	**2.17E-02**	3.52E-01	**7.81E-05**	3.55E-01
Block(Rep)	**2.68E-02**	6.81E-02	**2.68E-02**	3.98E-01	2.86E-01	6.93E-01	4.07E-01
Treatment	**1.35E-06**	3.42E-01	**6.67E-04**	**2.45E-31**	**2.45E-31**	**1.10E-15**	5.82E-02
Line	**1.38E-14**	**5.75E-10**	**6.37E-06**	**5.41E-04**	**1.12E-12**	**1.29E-13**	**1.46E-08**
Treatment x Line	**8.10E-03**	5.61E-01	**2.68E-02**	8.75E-01	**1.13E-03**	5.28E-02	8.06E-01
Treatment x Experiment	4.07E-01	2.91E-01	7.04E-01	**2.68E-02**	2.70E-01	**2.85E-04**	8.73E-01
Treatment x Line x Experiment	1.50E-01	7.32E-01	8.06E-01	8.57E-01	**5.50E-05**	6.81E-02	8.06E-01

C) Conclusion

Dans ce chapitre, j'ai combiné des approches complémentaires pour :

i. Identifier le gène causal sous-jacent au QTL1.
ii. Etudier la relation entre expression et phénotype.
iii. Déterminer la diversité et la spécificité écologique de ce gène.

Par des analyses de lignées mutantes, j'ai pu déterminer que le récepteur de type kinase *PERK13/RHS10* était bel et bien le gène sous-jacent au QTL1, associé à une stratégie d'évitement en réponse à la compétition avec *P. annua*. En effet, le mutant knock-out correspondant au gène *PERK13/RHS10* (*perk13-1*) n'a plus la capacité de répondre à la présence de *P. annua* au niveau du ratio HD et sa complémentation par la forme sauvage (Col-0) de *PERK13/RHS10* restaure totalement le phénotype observé initialement. Par ailleurs, nous avons observé dans la lignée qui sur-exprime *PERK13/RHS10* une valeur de ratio HD en absence de *P. annua* similaire à celle que l'on observe chez la lignée sauvage Col-0 en présence de *P. annua*. Nous avons ainsi pu montrer que la population locale TOU-A était un outil puissant pour l'identification des bases génétiques sous-jacentes aux interactions plante-plante. En effet, le LD court estimé dans cette population a permis de cartographier de façon fine le gène *PERK13/RHS10* pressenti dans le chapitre précédent comme un gène candidat d'intérêt, sous-jacent au QTL1.

Par un séquençage de différentes accessions, il nous a été possible d'identifier deux haplotypes fortement différentiés pour le gène *PERK13/RHS10*. Ces deux haplotypes se trouvent être :

(i) identiques au niveau de la séquence nucléotidique du domaine extracellulaire de la protéine correspondante, suggérant une sélection purifiante. Ce type de sélection est attendu car ce domaine pourrait être potentiellement impliqué dans la détection par *PERK13/RHS10* d'un ligand qui semble commun entre *P. annua* et le blé.

(ii) fortement différenciés au niveau de la région promotrice du gène, ainsi que dans le domaine kinase de la protéine, suggérant une sélection balancée sur ces deux régions de

RHS10. L'impact de ces polymorphismes sur la transcription et la traduction de *PERK13/RHS10* serait sans aucun doute intéressant à étudier. Cela est d'autant plus pertinent qu'aucune mutation dans la région kinase ne semble entraîner une quelconque modification de la séquence protéique (décalage du cadre de lecture, épissage alternatif ou mutation non synonyme), donc une altération possible de l'activité kinase du RLK. Il serait néanmoins nécessaire de séquencer un plus grand nombre d'accessions issues de la population TOU-A afin d'explorer pleinement la diversité nucléotidique de *PERK13/RHS10*. Ce séquençage pourrait être étendu à d'autres accessions issues d'autres populations, afin de déterminer si les différents haplotypes identifiés au sein de la population TOU-A sont aussi présents dans d'autres populations.

Du fait de l'existence de polymorphismes dans la région promotrice du gène *PERK13/RHS10*, des approches ont été développées pour tenter de mettre en relation le niveau d'expression de *PERK13/RHS10* et le phénotype ratio HD. Une cinétique du niveau d'expression de *PERK13/RHS10* a été réalisée dans différents organes de l'accession Col-0 en absence et en présence de *P. annua*. Nous avons pu montrer, lors d'une expérience préliminaire, que la présence de *P. annua* ne semblait pas modifier l'expression de *PERK13/RHS10* dans les racines à des stades précoces, mais semblait plutôt conduire à une réduction de son expression dans les parties aériennes (feuilles et inflorescences) à des stades tardifs. Ces données d'expression semblent pointer du doigt un effet de *P. annua* sur l'expression de l'haplotype Col-0 de *PERK13/RHS10* dans les parties aériennes, ce qui est assez intriguant car ce gène a été décrit comme étant spécifique des poils racinaires (Hwang *et al.* 2016)!

L'effet de la présence de *P. annua*, non détecté au niveau de l'expression racinaire de *PERK13/RHS10*, mais significatif au niveau des organes aériens, pourrait-il, au moins en partie, expliquer les phénotypes contrastés (ratio HD) observés entre les 3 haplotypes identifiés? Bien que nos données ne soient que préliminaires, nous pouvons émettre l'hypothèse que la présence de *P. annua* entraîne une diminution de l'expression de *PERK13/RHS10* chez les haplotypes Col-0, 1 et 2 (proches de Col-0) de *PERK13/RHS10* dans les parties aériennes, diminution qui s'accompagnerait d'une élongation plus grande de la tige principale. *A contrario*, la présence

de *P. annua* n'altèrerait pas l'expression de *PERK13/RHS10* chez les individus porteurs de l'haplotype 3.

Afin de progresser sur cette question, il serait nécessaire de répéter les mesures de niveau d'expression dans les différents organes (racines, feuilles et inflorescence) en conditions dans lequel le phénotype aérien est observé, c'est-à-dire en chambre de culture. Afin de confirmer ou non notre hypothèse, il serait nécessaire d'effectuer aussi ces mesures sur des lignées complémentées ainsi que sur des accessions naturelles. Ces dernières expériences pourraient nous permettre d'établir un lien fonctionnel potentiel entre la diversité naturelle que l'on observe au sein de la population locale TOU-A et la réponse à la présence de *P. annua* au niveau du ratio HD.

Bien que l'ensemble de l'analyse fonctionnelle du gène *PERK13/RHS10* reste à réaliser (voir Chapitre Perspectives ci-dessous), l'identification et la validation fonctionnelle du gène *PERK13/RHS10* représente une première étape dans la compréhension des mécanismes moléculaires qui sous-tendent la réponse compétitive et plus largement les interactions plante-plante, et ce grâce à l'utilisation d'une population locale d'*A. thaliana*.

Discussion générale et perspectives

En me plaçant à l'interface entre écologie évolutive et biologie moléculaire, l'objectif principal de ma thèse était de comprendre et de caractériser l'architecture génétique sous-jacente à la réponse à la compétition interspécifique chez *A. thaliana*. Pour cela, il m'a fallu dans un premier temps choisir une population d'*A. thaliana* adaptée à cet objectif et la caractériser aux niveaux génomique et phénotypique, afin de déterminer l'architecture génétique sous-jacente à la variation génétique naturelle de la réponse d'*A. thaliana* à la présence de la graminée *P.* annua (et ceci pour trois types de sol). J'ai également pu mettre en avant le caractère adaptatif de cette architecture génétique sur une courte échelle de temps. Par la suite, j'ai pu identifier les principaux processus biologiques associés aux QTL de réponse à la compétition dans différents contextes d'interactions plante-plante mono- et plurispécifiques. Pour finir, j'ai pu identifier et valider fonctionnellement le premier gène sous-jacent à la réponse compétitive chez *A. thaliana* (Figure d.1.).

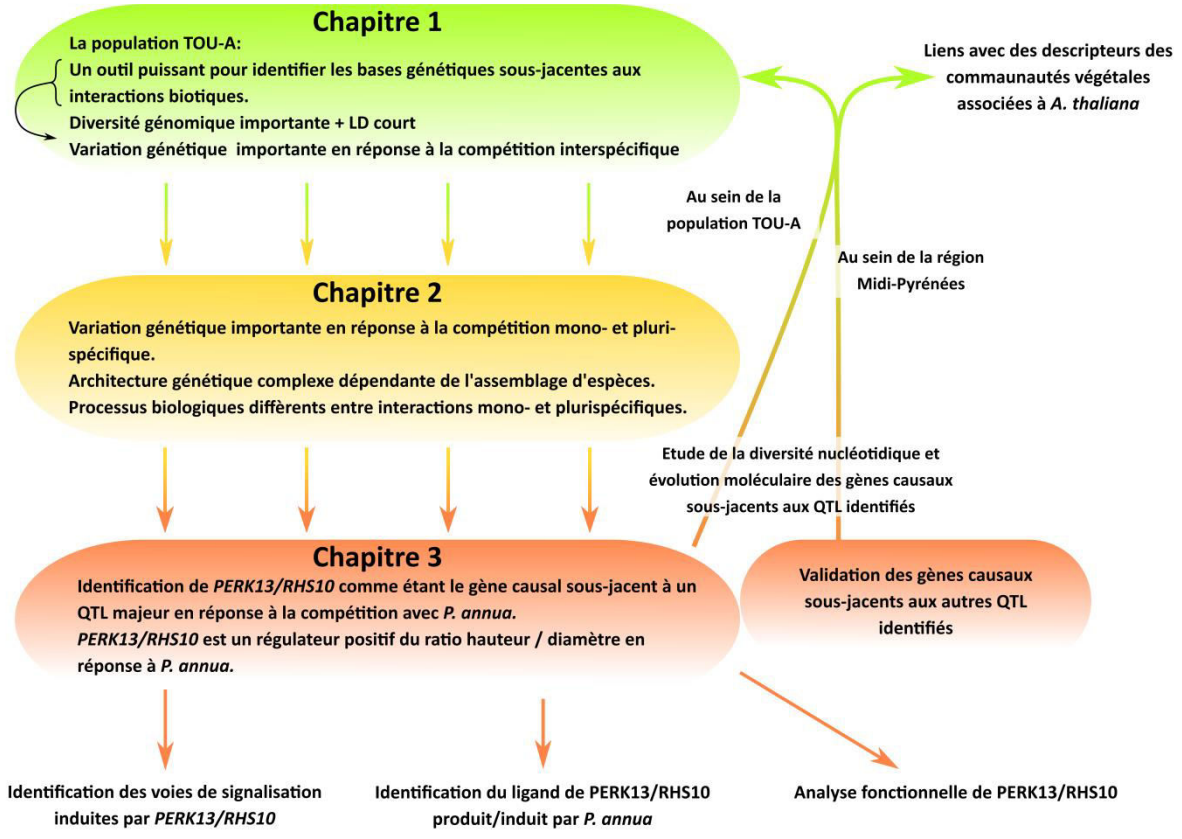

Figure d.1. Schéma représentant les principales conclusions des 3 chapitres de ma thèse ainsi que les perspectives de recherche associées.

I. Discussion générale

A) Variation génétique des interactions plante-plante : spécificité de l'interaction biotique

Lors d'une thèse réalisée au sein de l'équipe par Etienne Baron entre 2011 et 2014, une expérience avait été réalisée sur un terrain expérimental à partir de 48 accessions de la population locale TOU-A. Ces accessions avaient été cultivées en présence de 4 espèces communément associées à *A. thaliana* dans les communautés végétales (*Poa annua*, *Stellaria media*, *Trifolium arvense* et *Veronica arvensis*). Dans cette étude, des normes de réaction fortement croisées avaient été observées pour une série de neuf traits phénologiques, architecturaux et reliés à la production de graines (Baron *et al.* 2015). En effet, pour un trait donné, le rang des 48 accessions était fortement dépendant de l'identité des espèces voisines, suggérant une spécificité de réponse des accessions vis-à-vis de l'espèce avec laquelle elles interagissent. Durant ma thèse, j'ai non seulement confirmé ce phénomène de « spécialisation biotique » dans des conditions contrôlées de serre avec un panel de 96 accessions de la population TOU-A (chapitre 2), mais aussi démontré que cette spécialisation biotique était dépendante de l'environnement abiotique de phénotypage (i.e. type de sol dans le chapitre 1) et de l'assemblage des espèces voisines (chapitre 2).

Intégrer une telle variation génétique intraspécifique de la plasticité phénotypique pourrait permettre de mieux interpréter les assemblages de communautés. En effet, bien que la théorie traditionnelle de l'assemblage des communautés ignore largement cette variation génétique intraspécifique (Violle *et al.* 2012), l'intégrer semble primordial car elle est bien souvent essentielle à la survie, à la croissance et à la reproduction des espèces dans de nouvelles conditions environnementales (Joshi *et al.* 2001). Elle influencerait donc la réponse et l'évolution des populations face à des changements environnementaux et ainsi, affecterait les communautés biotiques qui leur sont associées (Albert *et al.* 2010). En effet, il y a de plus en plus de preuves empiriques que cette variabilité intraspécifique peut avoir des effets significatifs sur la dynamique d'une communauté, en modulant des mécanismes de différenciation de niches (Courbaud *et al.* 2012) et/ou en pouvant provoquer des chevauchements de performances

individuelles entre les espèces (Clark *et al.* 2004, Lichstein *et al.* 2007). Un nombre limité de génotypes spécialisés et donc une variation intraspécifique faible modifierait profondément la structure des communautés végétales et en réduirait leur dynamique.

Une telle spécialisation biotique est-elle aussi marquée pour d'autres types d'interactions biotiques ? Il semble que ce soit le cas pour les interactions plante-bactérie. Par exemple, quelques études ont montré que le classement d'accessions d'*A. thaliana* pour leur niveau de résistance quantitative à des bactéries phytopathogènes pouvait être très différent entre *Pseudomonas syringae* et *Pseudomonas viridiflava* (Atwell *et al.* 2010, Bartoli *et al.* 2018). De même, l'impact de facteurs abiotiques sur l'issue de l'interaction plante-pathogène peut être très dépendant du génotype de la plante. Par exemple, en utilisant une collection mondiale de 176 accessions d'*A. thaliana*, il a été observé que des accessions résistantes à la bactérie pathogène *Ralstonia solanacearum* à 27°C devenaient plus sensibles à 30°C, et vice-versa (Aoun *et al.* 2017). La question de la variabilité génétique de la réponse d'une espèce végétale à différents assemblages d'espèces bactériennes pathogènes reste par contre ouverte (Bartoli & Roux 2017). En revanche, la spécificité des interactions entre génotypes de plantes et génotypes d'espèces bactériennes est assez bien documenté (Bartoli & Roux 2017). Par exemple, dans une expérience où 130 accessions mondiales d'*A. thaliana* ont été phénotypées pour la résistance quantitative à 22 souches de *X. arboricola* isolées à partir de populations naturelles d'*A. thaliana*, le classement des accessions pour leur niveau de résistance quantitative était très dépendant de l'identité de la souche bactérienne et aucune accession n'était soit sensible, soit résistante à toutes les souches de *X. arboricola* (Wang *et al.* 2018).

Etonnamment, ces interactions génotype x génotype semblent avoir été moins étudiées dans le cadre des interactions plante-plante (Aarssen & Turkington 1985, Genung *et al.* 2012, Wuest & Niklaus 2018). En utilisant la population TOU-A, il serait pourtant intéressant de tester si le niveau de diversité génétique de plasticité phénotypique est différent entre le type d'interactions 'accessions d'*A. thaliana* x différentes espèces voisines' et le type d'interactions 'accessions d'*A. thaliana* x différents génotypes d'une même espèce voisine'. Pour mener une telle expérience, il serait judicieux d'utiliser différents génotypes de différentes espèces cohabitant avec *A. thaliana* dans la population TOU-A. Cela permettrait ainsi d'augmenter le

réalisme écologique de nos expériences, jusqu'à présent basées sur l'utilisation de lignées commerciales de *P. annua*, *S. media*, et *V. arvensis*.

B) Flexibilité de l'architecture génétique sous-jacente aux interactions plantes-plantes : maintien de la diversité génomique intra-population

La grande spécificité des interactions biotiques observée au niveau phénotypique suggère que l'architecture génétique sous-jacente à ces interactions est, elle aussi, belle et bien dépendante de la composition et de la structure du voisinage d'*A. thaliana*, elles-mêmes dépendantes des conditions de l'environnement abiotique. J'ai pu confirmer cette hypothèse grâce aux caractéristiques génomiques de la population TOU-A (diversité génomique importante, déséquilibre de liaison très court), favorables à une très bonne description de l'architecture génétique de la variation naturelle phénotypique par GWA mapping.

Dans le premier chapitre, en combinant une expérience de résurrection dans des conditions écologiquement réalistes avec des analyses d'association pangénomique, nous avons identifié une architecture génétique originale sous-jacente à l'évolution d'une population locale d'*A. thaliana* vers un nouvel optimum phénotypique. En effet, l'architecture génétique des 29 traits phénotypiques mesurés était très variable entre les six micro-habitats testés, avec plus de 78% des SNP les plus associés aux variations phénotypiques qui étaient spécifiques à un seul micro-habitat. Nous avons également trouvé que ces QTL spécifiques étaient faiblement sous sélection sur une période de 8 générations. A l'opposé, nous avons identifié une fraction très faible de SNP avec un degré de pléiotropie intermédiaire (liées à 3-5 éco-phénotypes) et fortement sous sélection. Ce type d'architecture génétique non encore décrite chez d'autres espèces (à notre connaissance), permettrait ainsi à une population naturelle de répondre rapidement à un changement global de l'environnement *via* la sélection de combinaisons phénotypiques optimales médiées par des QTL pléiotropes, tout en ajustant la réponse selon les micro-variations environnementales observées au sein de la population TOU-A *via* des QTL spécifiques. Par ailleurs, cette architecture génétique combinant (i) de rares QTL avec un niveau

de pléiotropie intermédiaire et fortement sélectionnés, et (ii) de très nombreux QTL spécifiques des micro-habitats et faiblement sous sélection, permettrait le maintien d'une diversité génétique au sein de la population TOU-A, et donc d'un potentiel adaptatif face à de futurs changements environnementaux comme l'apparition de nouvelles espèces végétales dans la communauté végétale TOU-A. Il reste néanmoins à tester si une telle architecture génétique est spécifique à la population TOU-A ou si elle peut être aussi observée dans d'autres populations naturelles d'*A. thaliana* qui présentent également un fort degré de diversité génomique.

Dans le second chapitre, en combinant des analyses d'association pangénomique et différents traitements de compétition basés sur 91 accessions de la population locale TOU-A, j'ai pu, là aussi, mettre en évidence une architecture génétique flexible entre les différents traitements. En effet, en accord avec les données phénotypiques, nous avons constaté que les bases génétiques étaient très dépendantes de l'assemblage d'espèces voisines considéré. Malheureusement, il ne nous a pas été possible d'étudier l'évolution phénotypique entre 2002 et 2010 car il existe un biais dans le choix des accessions. En effet, les accessions ont été choisies pour couvrir au mieux l'aire de distribution de la population d'origine. Il n'était donc pas possible d'estimer correctement des taux d'évolution phénotypique (i.e. haldanes) comme ce fut le cas dans le premier chapitre. Néanmoins, il nous est possible de tester si les top SNP identifiés dans le chapitre 2 présentent des traces de sélection temporelle.

Ainsi, en considérant les 9,817 SNP présentant les valeurs de F_{ST} temporel les plus élevées (soit 1% du nombre total de SNP), j'ai identifié 476 SNP parmi l'ensemble des 200 top SNP associés à chacun des quatre traits mesurés (H1F : hauteur du sol à la première fleur sur la tige principale, DIAM : diamètre maximum de la rosette, BIOMASS : biomasse sèche aérienne, ratio HD = H1F/DIAM) dans les 10 traitements d'interaction hétérospécifique considérés. Sur ces 476 SNP, 284 (60%) sont associés à des conditions d'interaction monospécifique dont 181 pour le diamètre maximal de la rosette en présence de *V. arvensis* et *P. annua,* et 77 pour le ratio HD en majorité en présence de *V. arvensis*. Les 193 SNP restants ont été identifiés en conditions d'interaction plurispécifique dont 156 pour le ratio HD. De manière intéressante, 95% de ces 156 SNP sont associés au ratio HD en présence d'au moins un individu de *P. annua* (104 PPX et 44 PXX ; X = S ou V). Il semblerait que la présence de la véronique des champs

ainsi que du pâturin annuel joue un rôle sur la sélection de traits comme le diamètre et le ratio HD, se traduisant par des SNP sous-jacents évoluant fortement en seulement 8 générations. Ces résultats semblent être en accord avec des observations faites au sein de la population confirmant une réduction depuis 2002 de la présence de la stellaire (mouron des oiseaux) et une présence toujours constante de la véronique et du pâturin annuel au sein de la communauté végétale de la population TOU-A.

Bien que représentant une première étape dans l'étude des relations entre architecture génétique des interactions plante-plante selon différents niveaux de complexité et traces de sélection temporelle, il serait là aussi intéressant d'augmenter le réalisme écologique des interactions plante-plante testées expérimentalement. En effet, les assemblages mis en place au sein de l'expérience du chapitre 2 pourraient être considérés comme peu réalistes d'un point de vue écologique. Notamment, il est très peu probable que dans la nature, toutes les plantes d'*A. thaliana* soient toujours entourées de 3 autres plantes. Par ailleurs, au sein de la communauté TOU-A, les assemblages testés ont certainement des probabilités différentes d'exister, probabilités qui sont directement dépendantes des abondances absolues et relatives de chacune des espèces compétitrices étudiées. Ces probabilités de co-occurrence impactent directement le taux de sélection agissant sur les gènes impliqués dans la réponse compétitive d'*A. thaliana*. En d'autres termes, si *A. thaliana* est rarement en présence d'un assemblage d'espèces compétitrices donné au sein de la communauté végétale TOU-A, les QTL de réponse compétitive spécifiques de cet assemblage seront rarement soumis à la sélection. A l'inverse, un assemblage d'espèces compétitrices très commun au sein de la population TOU-A devrait entraîner une sélection plus importante sur les gènes de réponse à cet assemblage. Pour tester cette hypothèse, une description de la communauté végétale TOU-A (i.e. identité et abondance de chacune des espèces végétales) est prévue au printemps 2019, avec une attention particulière sur la description du voisinage d'une cinquantaine de plantes d'*A. thaliana*. Combinées à la récolte de graines de plusieurs plantes de chacune des espèces, ces descriptions *in situ* devraient permettre la mise en place de nouvelles expériences en conditions contrôlées, qui permettront d'explorer plus largement la relation entre degré de pléiotropie des top SNP identifiés entre les assemblages d'espèces les plus communs et les traces de sélection temporelle.

Au-delà de l'étude des interactions plante-plante, cette population TOU-A peut aussi être un outil puissant pour caractériser l'architecture génétique en réponse à d'autres interactions biotiques. En effet, depuis quelques années, la population TOU-A a fait l'objet de phénotypage pour la résistance au virus de la mosaïque du navet (TuMV, Turnip Mosaic Virus) et à différentes bactéries pathogènes (*Pseudomonas viridiflava, Ralstonia solanacearum, Xanthomonas campestris*). En ajoutant ces données aux nombreuses données déjà acquises au sein de l'équipe (traits phénologiques, traits architecturaux, réponses aux interactions plante-plante…), l'utilisation de la population TOU-A pourrait permettre de mettre en lumière des mécanismes moléculaires communs à ces différents types d'interactions biotiques, et d'appréhender les bases moléculaires des mécanismes d'intégration de la réponse des plantes à un nombre quasi-infini de combinaisons de stress biotiques.

C) Fonctions et gènes identifiés, sous-jacents aux interactions plante-plante

1) De nouveaux processus biologiques identifiés

En considérant l'ensemble des 369 gènes situés à moins de 1kb des 20 top SNP identifiés pour chacune des combinaisons 'trait phénotypique x traitement de compétition' dans le chapitre 2, nous avons observé une sous- et une surreprésentation de certains processus biologiques (Table 1), dont l'identité se recoupe très peu avec celle des processus biologiques déjà identifiés et décrits comme étant impliqués dans les interactions plante-plante, à savoir photosynthèse, hormones, paroi, défense, transporteurs, modification d'histone et identité méristématique (Subrahmaniam *et al.* 2018). Ce faible recoupement pourrait s'expliquer entre autres par des différences du degré de réalisme écologique des approches utilisées au cours de nos travaux comparativement à ceux déjà publiés. Historiquement, les études portant sur les interactions plante-plant ont considéré soit un environnement artificiel simulant des interactions plante-plante (i.e. ombre), soit des interactions monospécifiques impliquant bien souvent des espèces végétales dont la co-occurrence dans la nature est relativement faible, voire nulle.

Parmi les 7 principaux processus biologiques identifiés dans Subrahmaniam *et al.* (2018), deux sont surreprésentés dans notre expérience conduite en conditions contrôlées de serre

(Table 1) : photosynthèse et/ou la perception de la lumière, et transport *via* des transporteurs de type ABC.

Fréquence normalisée par rapport à la fréquence dans le génome d'*A. thaliana*	*p*-val	Class
9.04	1.04E-05	Synthèse de tétrapyrroles
6.27	3.64E-05	Transformation d'acides organiques
6.19	3.42E-03	Métabolisme C1
2.7	4.53E-03	Photosystème
1.68	7.24E-04	Perception et transduction de signaux
1.57	6.81E-03	Transport
1.44	4.01E-04	Régulation des ARN
0.75	3.78E-06	Non assigné
0.15	1.72E-14	Modification des strucutures de l'ADN

Table 1. Processus biologiques (classification MapMan) significativement sur- et sous- représentés pour tous les gènes identifiés.

Au-delà de ces deux processus biologiques, nous avons pu mettre en évidence une surreprésentation de gènes associés à la perception et à la transduction de signaux et notamment de récepteurs de type kinase. Ces récepteurs sont associés à de nombreux processus tels que l'immunité, le développement et la croissance des plantes (Tang *et al.* 2017), mais aussi à la réponse aux stress abiotiques (Ye *et al.* 2017). Considérant les nombreux gènes de type récepteurs de type kinase identifiés en réponse à différents stress biotiques et abiotiques, ainsi que leur rôle potentiellement prépondérant (sur la base de leur surreprésentation) dans la réponse aux interactions plante-plante, il semblait particulièrement intéressant de tenter d'explorer le rôle de l'un d'entre eux.

2) *PERK13/ RHS10*, un gène qui contrôle les interactions compétitives plante-plante

Etant donnée (i) l'abondance de *P. annua* au sein de la population TOU-A, (ii) l'importance de *P. annua* dans les processus sélectifs qui semblent agir sur la population TOU-A, et (iii) l'identification d'un pic d'association intéressant (fin et forte significativité) en condition monospécifique avec *P. annua*, nous nous sommes focalisés sur le QTL1. De façon intéressante, dans la région génomique couverte par ce QTL, se trouve un gène RLK

(*PERK13/RHS10*). Grâce à la combinaison (i) d'une analyse de mutants situés dans la région génomique couverte par le QTL retenu pour notre étude, et (ii) de la complémentation du mutant *perk13 -1* par l'haplotype du fond génétique Col-0 et de deux haplotypes provenant de la population TOU-A, le gène *PERK13/RHS10* a pu être identifié comme étant un gène causal sous-jacent à la variation naturelle observée pour une stratégie d'évitement de la compétition (ratio HD).

Bien qu'ayant été décrit initialement pour être un régulateur négatif de la croissance des poils racinaires (Hwang *et al.* 2016), *PERK13/RHS10* semble donc aussi jouer un rôle essentiel dans la réponse d'*A. thaliana* à la présence de *P. annua* en agissant sur un phénotype au niveau aérien. À noter que cette réponse est observée aussi en réponse au blé tendre *T. aestivum,* mais pas en réponse à *A. sativa, D. glomerata, S. media, V. arvensis* et plus étonnamment *P. pratensis* (pourtant plus proche phylogénétiquement de *P. annua* que toutes les autres espèces), suggérant l'existence d'une spécificité de la réponse médiée par PERK13/RHS10. Par ailleurs, des résultats préliminaires semblent montrer que *P. annua* réprime l'expression de *RHS10/PERK13* dans les organes aériens. Tous ces résultats ouvrent de nombreuses perspectives extrêmement intéressantes, (i) comprendre par quels mécanismes/molécules *P. annua* est perçu par *A. thaliana,* (ii) comment ce message est-il transduit par la plante, et (iii) quelle reprogrammation génétique est associée à la réponse d'évitement à la compétition, suite à la perception de *P. annua via* PERK13/RHS10. Une analyse fonctionnelle précise du récepteur RHS10/PERK13 devrait constituer un excellent point de départ à la compréhension de l'ensemble des mécanismes sous-jacents à la réponse à la compétition.

3) Variation nucléotidique de *RHS10* au sein de la population TOU-A

Afin de replacer *PERK13/RHS10* dans les processus évolutifs abordés lors du premier chapitre, j'ai séquencé quelques accessions avec la méthode Sanger, ce qui m'a permis d'avoir une information plus précise de la diversité nucléotidique présente au niveau de *PERK13/RHS10*. En effet, les données de séquençage Illumina en 'short reads' ne permettent pas d'identifier les polymorphismes dans des régions très polymorphes, comme c'est le cas de la région promotrice et du domaine kinase de *PERK13/RHS10*. Nous avons identifié des haplotypes très différenciés, suggérant une sélection balancée avec un maintien sur de longs

temps évolutifs (i.e. plusieurs centaines de milliers d'années) de ces différents haplotypes accompagné d'une absence quasi-totale de recombinaison. Chez *A. thaliana*, ce type de sélection a déjà été observé pour d'autres types d'interactions, notamment pour des gènes associés la variation naturelle de la résistance à des bactéries pathogènes (Roux & Bergelson 2016). Dans la plupart des cas, la sélection balancée s'observe sur des polymorphismes de type présence/absence d'un gène, comme pour les gènes de type *R* (i.e. *RPM1* et *RPS5*) (Roux & Bergelson 2016). Récemment, une signature de sélection balancée a été détectée sur la région promotrice et le début de la région codante du gène *RKS1*, qui est associé à la résistance quantitative à la bactérie pathogène *Xanthomonas campestris* (Huard-Chauveau *et al.* 2013). En ce qui concerne le gène *PERK13/RHS10*, il semble que la sélection balancée concerne surtout le domaine kinase. A notre connaissance, aucune étude n'a rapporté une signature de sélection balancée agissant sur un domaine kinase (toutes espèces confondues). Ce qui est néanmoins troublant dans notre étude est l'absence de mutations non-synonymes dans le domaine kinase. Si ce domaine est réellement sous sélection balancée, quels sont les polymorphismes ciblés ? Par ailleurs, nous ne pouvons pas non plus exclure que d'autres polymorphismes situés dans la région promotrice ne jouent pas aussi un rôle essentiel dans la fonction du gène *PERK13/RHS10*. Pour identifier le(s) polymorphisme(s) causal/causaux sous-jacent(s) au QTL1, il serait intéressant de créer dans un premier temps des lignées swap qui combineraient les domaines de *PERK10/RHS10* de différents haplotypes (e.g. promoteur de l'haplotype 1 + région codante de l'haplotype 3 *vs* promoteur de l'haplotype 3 + région codante de l'haplotype 1).

II. Perspectives

Bien que ce travail ne soit pas achevé, nous avons identifié et validé fonctionnellement un gène sous-jacent à la réponse compétitive dans le contexte d'interactions plante-plante, et ce grâce à l'utilisation d'une population locale d'*A. thaliana*. L'ensemble des résultats obtenus ouvrent des perspectives diverses et notamment trois principales en lien avec les différents chapitres abordés (Figure d.1.).

A) Dissection des mécanismes de perception et des voies de signalisation associées à *PERK13/RHS10*

L'identification du premier gène de réponse compétitive à la présence d'une autre espèce représente un point de départ pour la compréhension des évènements moléculaires sous-jacents à la réponse à la compétition interspécifique. En effet, de nombreuses questions et champs d'investigation s'ouvrent non seulement sur la caractérisation fonctionnelle de PERK13/RHS10, l'exploration des voies de signalisations activées par PERK13/RHS10, mais aussi sur l'exploration de l'identité du ligand potentiel du récepteur (Figure d.1.).

1) Analyse fonctionnelle de *PERK13/RHS10*

Plusieurs approches peuvent être envisagées pour mieux comprendre le fonctionnement de ce récepteur putatif :

i. Analyser l'expression du gène dans le contexte des interactions plante-plante et dans les compartiments aériens ou racinaires, afin de préciser son patron d'expression, au cours du développement mais aussi en réponse à l'interaction avec *P. annua.*
ii. Générer des lignées transgéniques pour lesquelles l'expression de *PERK13/RHS10* est abolie ou réduite spécifiquement dans les racines, ou dans les parties aériennes, afin d'évaluer le rôle de ce récepteur dans chacun de ces compartiments.
iii. Faire une analyse structure-fonction de PERK13/RHS10 afin de déterminer le rôle des différents domaines dans l'interaction. Cette étude fournira de plus des outils nécessaires aux approches visant à élucider la nature du ligand et les voies de signalisation. Une attention particulière sera portée au domaine ECD (extracellular proline-rich domain) comportant des résidus proline ainsi que des motifs à arabinogalactane (AGP), qui se sont révélés nécessaires pour l'inhibition de la croissance des poils racinaires (Hwang *et al.* 2016, Cho 2016). Le domaine kinase probablement impliqué dans la transduction du signal et potentiellement sous sélection balancée sera également ciblé.
iv. Déterminer la localisation subcellulaire au niveau aérien en regard de sa localisation dans la membrane plasmique au niveau racinaire (Hwang *et al.* 2016)

v. Déterminer le rôle du gène 2 localisé dans la région génomique sous-jacente au QTL1 (*At1g70450*) dans la réponse à *P. annua*. En effet, dans notre analyse, les mutants disponibles pour ce gène ne sont que très partiellement affectés dans le niveau d'expression du gène. Et même si la complémentation du mutant *perk13-1* par le gène *PERK13/RHS10* conduit à la restauration complète du phénotype, on ne peut exclure à ce stade, que le gène *At1g70450* ne puisse pas aussi jouer un rôle dans la réponse. Ainsi, la génération de lignées amiRNA pour l'un et/ou l'autre des gènes (*PERK13/RHS10* et *At1g70450*) nous permettrait de déterminer si *At1g70450* est impliqué dans le phénotype de réponse à la compétition. Par ailleurs, *At1g70450* a pu être identifié comme étant un récepteur kinase cytoplasmique (RLCK, Shiu & Bleecker 2001). *At1g70450* ne présente pas de domaine extracellulaire riche en proline ni de domaine transmembranaire mais présente une certaine similarité de la région kinase avec *PERK13/RHS10* (Nakhamchik *et al.* 2004). Il serait donc intéressant de tester si At1g70450 interagit directement ou indirectement avec PERK13/RHS10.

vi. Déterminer le rôle des homologues les plus proches de PERK13/RHS10 (i.e. PERK12, 11, 1, 5, 8 et 10) dans la réponse à *P. annua* par une analyse de mutants affectés dans ce/ces gènes, ceci afin d'évaluer la possibilité d'une redondance fonctionnelle en lien avec la réponse à la compétition, comme cela a pu être montré en partie sur la longueur des poils racinaires (PERK1, 5, 8 et 10, Hwang *et al.* 2016).

2) Quelles voies de signalisation en aval de PERK13/RHS10 ?

La compréhension du fonctionnement de ce récepteur et de son rôle dans l'interaction plante-plante passe par l'identification des cascades de signalisation contrôlées par *PERK13/RHS10*, et aboutissant à la génération d'un phénotype complexe et intégratif touchant à la fois à la hauteur et au diamètre de la plante. Afin de mener à bien cet objectif, deux types d'approches principales pourraient être envisagées, au moins dans un premier temps :

i. Recherche des interacteurs de PERK13/RHS10.

Une approche *a priori* sera ciblée dans un premier temps sur certains interacteurs déjà connus comme RNS2 qui interagit avec le domaine kinase de PERK13/RHS10 (Hwang *et al.* 2016), ou KIPK1 et -2 qui interagissent avec le domaine kinase de PERK8, 9 et 10 (Humphrey *et al.* 2015). Le domaine kinase semble être relativement conservé entre différentes espèces (par exemple : 97.4% d'homologie entre AtPERK13 et PERK13 chez *Noccaea caerulescens*) mais aussi entre PERK au sein d'*A. thaliana* (Nakhamchik *et al.* 2004). Un premier axe serait de tester ces interacteurs déjà connus par co-immunoprécipitation et/ou système double hybride, et *in vivo* par FRET-FLIM. En cas de validation de l'interaction, leur rôle dans la réponse compétitive sera examiné par approche génétique.

Une recherche sans *a priori* pourra en parallèle être entreprise *via* un crible double hybride en utilisant des banques générées dans le contexte de l'interaction *A. thaliana – P. annua*. Naturellement, le rôle de ces interacteurs pourra ensuite être testé par des approches génétiques, notamment par la caractérisation de mutants insertionnels dans les gènes correspondant à ces interacteurs.

ii. Analyse du transcriptome de lignées dérégulées pour PERK13/RHS10

Au-delà des partenaires de PERK13/RHS10 interagissant physiquement avec la protéine, il serait également intéressant de mettre en place des analyses transcriptomiques permettant d'explorer les voies de signalisation contrôlées par *PERK13/RHS10*. Ces analyses pourront être conduites par RNAseq, sur les différentes lignées dérégulées pour l'expression de *PERK13/RHS10*, (i) dans différents organes, (ii) à différents points

cinétiques, (iii) en présence/absence de *P. annua*. Ces expériences permettraient d'identifier le réseau de gènes associés à *PERK13/RHS10*, dont la validation devra être réalisée par analyse de mutants, afin d'identifier les régulateurs majeurs de cette réponse.

3) Quel ligand pour PERK13/RHS10 ?

Le mode d'action des récepteurs kinase de type PERK semble se faire par (i) l'association du domaine extracellulaire avec certains composants de la paroi cellulaire, et la perception potentielle de ligands au niveau pariétal ou de la membrane plasmique, et (ii) une transduction du signal par le domaine kinase dans le cytoplasme (Hwang *et al.* 2016). Dans le cadre des interactions plantes-plantes, il serait particulièrement intéressant d'identifier les molécules perçues par PERK13. Cela nous renseignerait sur les mécanismes par lesquels *P. annua* est perçu de manière différentielle par les accessions de la population TOU-A. Deux hypothèses sont envisageables :

i. la perception de signaux externes produits par *P. annua* et le blé (exsudats, composés volatiles organiques, autres…)
ii. la perception de signaux internes produits par *A. thaliana* (Damage-Associated Molecular Pattern, DAMP, ou autres molécules).

Afin d'identifier un tel ligand, plusieurs approches sont envisageables :

i. Si ce(s) ligand(s) est/sont produit(s) au niveau racinaire par *P. annua* (point qui devrait être élucidé en point 1), une première expérience serait de phénotyper le comportement des différentes lignées *PERK13/RHS10* en possession de l'équipe (WT, mutant et lignées complémentées) sur du sol ayant servi à la croissance de *P. annua*, ou en ajoutant dans le milieu de culture les éluats du sol ayant servi à la croissance de *P. annua*. Si le ligand est présent et influence le phénotype de réponse à la compétition (ratio HD), il resterait à (i) extraire, (ii) purifier et (iii) identifier les composés bioactifs produits par *P. annua* en utilisant des méthodes telles que la spectrométrie de masse (Altemimi *et al.* 2017) notamment.

ii. Si ces ligands sont produits/perçus au niveau aérien (voir point 1), des approches similaires pourraient être menées mais cette fois-ci, en collectant les composés volatiles aériens (VOC) émis par *P. annua* et en les testant sur les différentes lignées *PERK13/RHS10* en possession de l'équipe.

iii. Enfin, une autre possibilité pourrait être que ce récepteur perçoit des modifications pariétales ou un signal issu de la paroi, résultant de l'interaction directe avec *P. annua,* hypothèse envisageable si on considère la nature de ce récepteur. Il semblerait que le domaine riche en proline et en particulier les motifs à arabinogalactane de PERK13/RHS10 puisse percevoir les modifications de structure de la paroi cellulaire (Hwang *et al.* 2016). De manière intéressante, la famille des PERK a été identifiée pour la première fois chez *Brassica napus* par l'identification de *BnPERK1* qui semble être impliqué dans la perception et la réponse de blessures et, dans une moindre mesure, à l'infection par *Sclerotinia sclerotiorum* (Silva & Goring 2002). Le test de mutants de la protéine PERK13/RHS10, affectés notamment pour leur interaction avec la paroi (mutations/délétions dans le domaine extracellulaire) ainsi que des mutants affectés dans la synthèse d'arabinogalactane, pourrait apporter des éléments de réponse à cette hypothèse. De même, la recherche de composés pariétaux (en absence ou en présence de *P. annua*) et de leur effet sur le phénotype de réponse à la compétition (ratio HD) pourrait être envisagés.

Pour l'ensemble de ces approches, le fait de disposer d'espèces induisant (*P. annua* et *T. aestivum*) ou non (*P. pratensis, D. glomerata, V. arvensis, S. media* et *A. sativa*) une réponse compétitive d'*A. thaliana via* PERK13/RHS10 constitue une opportunité pour générer des comparatifs utiles en vue de l'identification du ligand.

B) Validation fonctionnelle d'autres QTL identifiés par GWA mapping

Tout comme nous avons pu le faire pour *PERK13/RHS10*, il serait pertinent de valider fonctionnellement d'autres QTL que nous avons identifiés dans le chapitre 2 (Figure d.1.). Malgré le fait que ces projets demandent un investissement important, ce type d'approche est indispensable si l'on souhaite obtenir une compréhension la plus large possible des bases

génétiques et des mécanismes moléculaires sous-jacents à la variation naturelle des interactions plante-plante. En effet, la validation et l'analyse fonctionnelle de plusieurs gènes causaux permettraient d'établir et d'explorer plus largement les fonctions moléculaires et les voies de signalisation associées.

Afin d'obtenir une plus grande compréhension des mécanismes complexes sous-jacents aux interactions plante-plante tout en répondant à un besoin grandissant d'amélioration des espèces cultivées dans un contexte toujours plus respectueux de l'environnement (i.e. diminution de l'application de produits phytosanitaires), il serait judicieux de s'intéresser à des QTL associés à des traits phénotypiques en lien avec l'accumulation de biomasse et/ou de production de graines. Ces gènes peuvent être reliés à une meilleure préemption des nutriments essentiels à la croissance ou bien reliés à une limitation de la croissance des compétiteurs comme par exemple la production de composés allélopathiques.

Ce type d'interaction asymétrique pourrait être un critère de choix pour la sélection du prochain QTL à valider fonctionnellement. En s'intéressant à la fois aux données de biomasse des plantes focales et aux données de biomasse des compétiteurs, on serait tenté de chercher un QTL discriminant des accessions maximisant leur biomasse tout en diminuant la biomasse des compétiteurs.

C) Valeur adaptative de *PERK13/RHS10*

1) Quid du type sélection agissant sur *PERK13/RHS10*?

L'identification d'haplotypes très différenciés entre 8 accessions TOU-A suggère que le gène *PERK13/RHS10* est sous sélection balancée. Pour compléter cette étude préliminaire, j'ai estimé la proportion de chacun des haplotypes de *PERK13/RHS10* entre les 195 accessions de la population TOU-A, grâce à un SNPcalling effectué à partir des données Illumina sur les différents haplotypes. J'ai identifié une majorité d'accessions proches de l'haplotype 1 (68%) et de l'haplotype 2 (24%) contre seulement 8% similaires à l'haplotype 3. Les 16 accessions similaires à l'haplotype 3 sont représentées à part égales dans chacune des cohortes de 2002 et 2010. Il semble que malgré une faible proportion d'accessions possédant *PERK13/RHS10*-haplotype 3, ce dernier semble se maintenir au sein de la population TOU-A sur au moins quelques générations. Le génome de 115 accessions TOU-A échantillonnées en 2007 a été séquencé à l'automne 2018 et permettra d'ajouter un point supplémentaire dans l'étude de la dynamique de cet haplotype au sein de la population TOU-A.

Comment se maintient l'haplotype 3 à une fréquence de 8% au sein de la population TOU-A ? L'hypothèse d'un trade-off coût/bénéfice en absence/présence de *P. annua* semble la plus simple, si l'on prend aussi en compte la probabilité de rencontre entre *A. thaliana* et *P. annua* au sein de la population TOU-A. En présence de *P. annua*, les accessions portant les haplotypes 1 ou 2 possèderaient un avantage sélectif vis-à-vis des accessions portant l'haplotype 3. En absence de *P. annua*, c'est l'opposé qui se produit : les accessions portant l'haplotype 3 sont sélectionnées, alors que les accessions portant les haplotypes 1|2 sont contre-sélectionnées. Si la probabilité qu'*A. thaliana* ne rencontre pas *P. annua* est bien plus faible que la probabilité qu'elle rencontre *P. annua* au sein de la communauté végétale TOU-A, on s'attend à maintenir l'haplotype 3 à une fréquence faible.

Pour valider cette hypothèse, il serait intéressant de mettre en place des expériences d'évolution expérimentale en utilisant les lignées complémentées avec les différents haplotypes identifiés au sein de la population TOU-A. Différents types de populations expérimentales pourraient être créés en modifiant la densité de *P. annua*. Le suivi au cours des générations de

la fréquence de chacun des haplotypes permettrait d'estimer leur valeur sélective respective, et de tester si elles dépendent de la densité de *P. annua*.

2) Quid des autres populations naturelles d'*A. thaliana* cohabitant avec *P. annua* ?

L'identification du rôle de *PERK13/RHS10* au sein de la population TOU-A (Bourgogne) ne permet pas de prédire si *PERK13/RHS10* joue un rôle similaire dans d'autres populations d'arabettes. Pour répondre à cette question, la caractérisation des communautés végétales de 168 populations naturelles d'*A. thaliana* de la région Midi-Pyrénées réalisée dans le cadre de la thèse de Léa Frachon au sein de l'équipe (2014-2017) pourrait nous permettre de déterminer s'il existe un lien entre diversité génétique de *PERK13/RHS10* et présence/absence de certaines espèces végétales. Cette approche, bien que corrélative, nous permettrait non seulement d'avoir une idée de la distribution géographique des différents haplotypes de *PERK13/RHS10*, mais aussi d'identifier d'autres partenaires écologiques associés à *PERK13/RHS10*.

Si *PERK13/RHS10* n'est pas écologiquement pertinent au sein de la région Midi-Pyrénées, il serait facilement envisageable de lancer des études de GWA mapping en phénotypant la réponse à la présence de *P. annua* des 168 populations naturelles (dont les génomes ont été séquencés, Frachon *et al.* 2018). Dans tous les cas, l'identification de nouveaux QTL en réponse à la présence de *P. annua* permettrait d'améliorer notre compréhension, initiée par la découverte de *PERK13/RHS10*, des mécanismes génétiques sous-jacents aux interactions Arabette – graminées.

Aarssen, L. W., & Turkington, R. (1985). Biotic specialization between neighbouring genotypes in *Lolium perenne* and *Trifolium repens* from a permanent pasture. *The Journal of Ecology*, 605-614.

Albert, C. H., Thuiller, W., Yoccoz, N. G., Douzet, R., Aubert, S., & Lavorel, S. (2010). A multi-trait approach reveals the structure and the relative importance of intra-vs. interspecific variability in plant traits. *Functional Ecology*, *24*(6), 1192-1201.

Alonso, J. M., & Ecker, J. R. (2006). Moving forward in reverse: genetic technologies to enable genome-wide phenomic screens in Arabidopsis. *Nature Reviews Genetics*, *7*(7), 524.

Alonso-Blanco, C., Andrade, J., Becker, C., Bemm, F., Bergelson, J., Borgwardt, K. M., ... & Ecker, J. R. (2016). 1,135 genomes reveal the global pattern of polymorphism in Arabidopsis thaliana. *Cell*, *166*(2), 481-491.

Altemimi, A., Lakhssassi, N., Baharlouei, A., Watson, D., & Lightfoot, D. (2017). Phytochemicals: Extraction, isolation, and identification of bioactive compounds from plant extracts. *Plants*, *6*(4), 42.

Aoun, N., Tauleigne, L., Lonjon, F., Deslandes, L., Vailleau, F., Roux, F., & Berthomé, R. (2017). Quantitative disease resistance under elevated temperature: genetic basis of new resistance mechanisms to *Ralstonia solanacearum*. *Frontiers in plant science*, *8*, 1387.

Arabidopsis Genome Initiative. (2000). Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *nature*, *408*(6814), 796.

Aranzana, M. J., Kim, S., Zhao, K., Bakker, E., Horton, M., Jakob, K., ... & Toomajian, C. (2005). Genome-wide association mapping in Arabidopsis identifies previously known flowering time and pathogen resistance genes. *PLoS genetics*, *1*(5), e60.

Atwell, S., Huang, Y. S., Vilhjálmsson, B. J., Willems, G., Horton, M., Li, Y., ... & Jiang, R. (2010). Genome-wide association study of 107 phenotypes in *Arabidopsis thaliana* inbred lines. *Nature*, *465*(7298), 627.

Bais, H. P., Vepachedu, R., Gilroy, S., Callaway, R. M., & Vivanco, J. M. (2003). Allelopathy and exotic plant invasion: from molecules and genes to species interactions. *Science*, *301*(5638), 1377-1380.

Baron, E., Richirt, J., Villoutreix, R., Amsellem, L., & Roux, F. (2015). The genetics of intra-and interspecific competitive response and effect in a local population of an annual plant species. *Functional Ecology*, *29*(10), 1361-1370.

Bartoli, C., & Roux, F. (2017). Genome-wide association studies in plant pathosystems: toward an ecological genomics approach. *Frontiers in plant science*, *8*, 763.

Bartoli, C., Frachon, L., Barret, M., Rigal, M., Huard-Chauveau, C., Mayjonade, B., ... & Roux, F. (2018). In situ relationships between microbiota and potential pathobiota in *Arabidopsis thaliana*. *The ISME journal*, 1.

Basu, C., Halfhill, M. D., Mueller, T. C., & Stewart Jr, C. N. (2004). Weed genomics: new tools to understand weed biology. *Trends in plant science*, *9*(8), 391-398.

Baxter, I., Brazelton, J. N., Yu, D., Huang, Y. S., Lahner, B., Yakubova, E., ... & Vitek, O. (2010). A coastal cline in sodium accumulation in *Arabidopsis thaliana* is driven by natural variation of the sodium transporter AtHKT1; 1. *PLoS genetics*, *6*(11), e1001193.

Bay, R. A., Rose, N., Barrett, R., Bernatchez, L., Ghalambor, C. K., Lasky, J. R., ... & Ralph, P. (2017). Predicting responses to contemporary environmental change using evolutionary response architectures. *The American Naturalist*, *189*(5), 463-473.

Bergelson, J., & Roux, F. (2010). Towards identifying genes underlying ecologically relevant traits in *Arabidopsis thaliana*. *Nature Reviews Genetics*, *11*(12), 867.

Bomblies, K., Yant, L., Laitinen, R. A., Kim, S. T., Hollister, J. D., Warthmann, N., ... & Weigel, D. (2010). Local-scale patterns of genetic variability, outcrossing, and spatial structure in natural stands of *Arabidopsis thaliana*. *PLoS Genetics*, *6*(3), e1000890.

Brachi, B., Aimé, C., Glorieux, C., Cuguen, J., & Roux, F. (2012). Adaptive value of phenological traits in stressful environments: predictions based on seed production and laboratory natural selection. *PLoS One*, *7*(3), e32069.

Brachi, B., Faure, N., Horton, M., Flahauw, E., Vazquez, A., Nordborg, M., ... & Roux, F. (2010). Linkage and association mapping of *Arabidopsis thaliana* flowering time in nature. *PLoS genetics*, *6*(5), e1000940.

Brachi, B., Filiault, D., Darme, P., Le Mentec, M., Kerdaffrec, E., Rabanal, F., ... & Novikova, P. (2017). Plant genes influence microbial hubs that shape beneficial leaf communities. *Biorxiv*, 181198.

Brachi, B., Meyer, C. G., Villoutreix, R., Platt, A., Morton, T. C., Roux, F., & Bergelson, J. (2015). Coselected genes determine adaptive variation in herbivore resistance throughout the native range of *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences*, 201421416.

Brachi, B., Villoutreix, R., Faure, N., Hautekèete, N., Piquot, Y., Pauwels, M., ... & Roux, F. (2013). Investigation of the geographical scale of adaptive phenological variation and its underlying genetics in A rabidopsis thaliana. *Molecular Ecology*, *22*(16), 4222-4240.

Brooker, R. W., Maestre, F. T., Callaway, R. M., Lortie, C. L., Cavieres, L. A., Kunstler, G., ... & Armas, C. (2008). Facilitation in plant communities: the past, the present, and the future. *Journal of Ecology*, *96*(1), 18-34.

Bukowski, A. R., & Petermann, J. S. (2014). Intraspecific plant–soil feedback and intraspecific overyielding in A rabidopsis thaliana. *Ecology and evolution*, *4*(12), 2533-2545.

Callaway, R. M., & Aschehoug, E. T. (2000). Invasive plants versus their new and old neighbors: a mechanism for exotic invasion. *Science*, *290*(5491), 521-523.

Callaway, R. M., & Walker, L. R. (1997). Competition and facilitation: a synthetic approach to interactions in plant communities. *Ecology*, *78*(7), 1958-1965.

Callaway, R. M., Brooker, R. W., Choler, P., Kikvidze, Z., Lortie, C. J., Michalet, R., ... & Armas, C. (2002). Positive interactions among alpine plants increase with stress. *Nature*, *417*(6891), 844.

Chaney, L., & Baucom, R. S. (2014). The costs and benefits of tolerance to competition in *Ipomoea purpurea*, the common morning glory. *Evolution*, *68*(6), 1698-1709.

Chapin Iii, F. S., Zavaleta, E. S., Eviner, V. T., Naylor, R. L., Vitousek, P. M., Reynolds, H. L., ... & Mack, M. C. (2000). Consequences of changing biodiversity. *Nature*, *405*(6783), 234.

Chesson, P. (2000). Mechanisms of maintenance of species diversity. *Annual review of Ecology and Systematics*, *31*(1), 343-366.

Cho, H. T. (2016). Arabinogalactan protein motif-containing receptor-like kinases are likely to play the negative feedback factor to maintain proper root hair length. *Plant signaling & behavior*, *11*(9), 2007-22.

Clark, J. S., LaDeau, S., & Ibanez, I. (2004). Fecundity of trees and the colonization–competition hypothesis. *Ecological Monographs*, *74*(3), 415-442.

Clements, F. E. (1916). *Plant succession: an analysis of the development of vegetation* (No. 242). Carnegie Institution of Washington.

Corre, V. L. (2005). Variation at two flowering time genes within and among populations of *Arabidopsis thaliana*: comparison with markers and traits. *Molecular ecology*, *14*(13), 4181-4192.

Courbaud, B., Vieilledent, G., & Kunstler, G. (2012). Intra-specific variability and the competition–colonisation trade-off: coexistence, abundance and stability patterns. *Theoretical Ecology*, *5*(1), 61-71.

Curtin, S. J., Tiffin, P., Guhlin, J., Trujillo, D. I., Burghardt, L. T., Atkins, P., ... & Young, N. D. (2017). Validating Genome-Wide Association candidates through quantitative variation in nodulation. *Plant physiology*, pp-01923.

Debieu, M., Huard-Chauveau, C., Genissel, A., Roux, F., & Roby, D. (2016). Quantitative disease resistance to the bacterial pathogen *Xanthomonas campestris* involves an Arabidopsis immune receptor pair and a gene of unknown function. *Molecular plant pathology*, *17*(4), 510-520.

Farooq, M., Bajwa, A. A., Cheema, S. A., & Cheema, Z. A. (2013). Application of allelopathy in crop production. *Int J Agric Biol*, *15*(6), 1367-1378.

Fournier-Level, A., Korte, A., Cooper, M. D., Nordborg, M., Schmitt, J., & Wilczek, A. M. (2011). A map of local adaptation in *Arabidopsis thaliana*. *Science*, *334*(6052), 86-89.

Frachon, L., Bartoli, C., Carrère, S., Bouchez, O., Chaubet, A., Gautier, M., ... & Roux, F. (2018). A genomic map of climate adaptation in *Arabidopsis thaliana* at a micro-geographic scale. *Frontiers in plant science*, *9*.

Frachon, L., C. Bartoli, B. Mayjonade, N.C. Hautekèete and F. Roux. Adaptation to plant communities across the genome of *Arabidopsis thaliana*. *Molecular Biology and Evolution* (in revision)

French, E., Kim, B. S., & Iyer-Pascuzzi, A. S. (2016, August). Mechanisms of quantitative disease resistance in plants. In *Seminars in cell & developmental biology* (Vol. 56, pp. 201-208). Academic Press.

Fusco, G., & Minelli, A. (2010). Phenotypic plasticity in development and evolution: facts and concepts.

Gaut, B. (2012). *Arabidopsis thaliana* as a model for the genetics of local adaptation. *Nature Genetics*, *44*(2), 115.

Genung, M. A., Bailey, J. K., & Schweitzer, J. A. (2012). Welcome to the neighbourhood: interspecific genotype by genotype interactions in Solidago influence above-and belowground biomass and associated communities. *Ecology Letters*, *15*(1), 65-73.

Goldberg, D. E., & Barton, A. M. (1992). Patterns and consequences of interspecific competition in natural communities: a review of field experiments with plants. *The American Naturalist*, *139*(4), 771-801.

Halty, V., Valdés, M., Tejera, M., Picasso, V., & Fort, H. (2017). Modeling plant interspecific interactions from experiments with perennial crop mixtures to predict optimal combinations. *Ecological Applications*, *27*(8), 2277-2289.

Hancock, A. M., Brachi, B., Faure, N., Horton, M. W., Jarymowycz, L. B., Sperone, F. G., ... & Bergelson, J. (2011). Adaptation to climate across the *Arabidopsis thaliana* genome. *Science*, *334*(6052), 83-86.

Hanikenne, M., Talke, I. N., Haydon, M. J., Lanz, C., Nolte, A., Motte, P., ... & Krämer, U. (2008). Evolution of metal hyperaccumulation required cis-regulatory changes and triplication of HMA4. *Nature*, *453*(7193), 391.

Hansen, M. M., Olivieri, I., Waller, D. M., Nielsen, E. E., & GeM Working Group. (2012). Monitoring adaptive genetic responses to environmental change. *Molecular ecology*, *21*(6), 1311-1329.

Hector, A., Schmid, B., Beierkuhnlein, C., Caldeira, M. C., Diemer, M., Dimitrakopoulos, P. G., ... & Harris, R. (1999). Plant diversity and productivity experiments in European grasslands. *science*, *286*(5442), 1123-1127.

Hierro, J. L., & Callaway, R. M. (2003). Allelopathy and exotic plant invasion. *Plant and soil*, *256*(1), 29-39.

Hoffmann, A. A., & Sgro, C. M. (2011). Climate change and evolutionary adaptation. *Nature*, *470*(7335), 479.

Horton, M. W., Hancock, A. M., Huang, Y. S., Toomajian, C., Atwell, S., Auton, A., ... & Nordborg, M. (2012). Genome-wide patterns of genetic variation in worldwide *Arabidopsis thaliana* accessions from the RegMap panel. *Nature genetics*, *44*(2), 212.

Huard-Chauveau, C., Perchepied, L., Debieu, M., Rivas, S., Kroj, T., Kars, I., ... & Roby, D. (2013). An atypical kinase under balancing selection confers broad-spectrum disease resistance in Arabidopsis. *PLoS genetics*, *9*(9), e1003766.

Humphrey, T. V., Haasen, K. E., Aldea-Brydges, M. G., Sun, H., Zayed, Y., Indriolo, E., & Goring, D. R. (2015). PERK–KIPK–KCBP signalling negatively regulates root growth in *Arabidopsis thaliana*. *Journal of experimental botany*, *66*(1), 71-83.

Hwang, Y., Lee, H., Lee, Y. S., & Cho, H. T. (2016). Cell wall-associated ROOT HAIR SPECIFIC 10, a proline-rich receptor-like kinase, is a negative modulator of Arabidopsis root hair growth. *Journal of experimental botany*, *67*(6), 2007-2022.

Jabran, K., Mahajan, G., Sardana, V., & Chauhan, B. S. (2015). Allelopathy for weed control in agricultural systems. *Crop Protection*, *72*, 57-65.

Joshi, J., Schmid, B., Caldeira, M. C., Dimitrakopoulos, P. G., Good, J., Harris, R., ... & Mulder, C. P. H. (2001). Local adaptation enhances performance of common plant species. *Ecology Letters*, *4*(6), 536-544.

Kang, H. M., Sul, J. H., Service, S. K., Zaitlen, N. A., Kong, S. Y., Freimer, N. B., ... & Eskin, E. (2010). Variance component model to account for sample structure in genome-wide association studies. *Nature genetics*, *42*(4), 348.

Kim, S., Plagnol, V., Hu, T. T., Toomajian, C., Clark, R. M., Ossowski, S., ... & Nordborg, M. (2007). Recombination and linkage disequilibrium in *Arabidopsis thaliana*. *Nature genetics*, *39*(9), 1151.

Koornneef, M., Alonso-Blanco, C., & Vreugdenhil, D. (2004). Naturally occurring genetic variation in *Arabidopsis thaliana*. *Annu. Rev. Plant Biol.*, *55*, 141-172.

Kronholm, I., Pico, F. X., Alonso-Blanco, C., Goudet, J., & Meaux, J. D. (2012). Genetic basis of adaptation in *Arabidopsis thaliana*: local adaptation at the seed dormancy QTL DOG1. *Evolution: International Journal of Organic Evolution*, *66*(7), 2287-2302.

Lasky, J. R., Upadhyaya, H. D., Ramu, P., Deshpande, S., Hash, C. T., Bonnette, J., ... & Buckler, E. S. (2015). Genome-environment associations in sorghum landraces predict adaptive traits. *Science Advances*, *1*(6), e1400218.

Lichstein, J. W., Dushoff, J., Levin, S. A., & Pacala, S. W. (2007). Intraspecific variation and species coexistence. *The American Naturalist*, *170*(6), 807-818.

Loreau, M., & Hector, A. (2001). Partitioning selection and complementarity in biodiversity experiments. *Nature*, *412*(6842), 72.

Maestre, F. T., Callaway, R. M., Valladares, F., & Lortie, C. J. (2009). Refining the stress-gradient hypothesis for competition and facilitation in plant communities. *Journal of Ecology*, *97*(2), 199-205.

Maestre, F. T., Valladares, F., & Reynolds, J. F. (2005). Is the change of plant–plant interactions with abiotic stress predictable? A meta-analysis of field results in arid environments. *Journal of Ecology*, *93*(4), 748-757.

Matthies, D. (1997). Parasite–host interactions in Castilleja and Orthocarpus. *Canadian Journal of Botany*, *75*(8), 1252-1260.

Millennium Ecocystem Assessment (2005) Ecosystem an Human Well-being, synthesis. *Island Press, Washington D.C.*

Mitchell-Olds, T., & Schmitt, J. (2006). Genetic mechanisms and evolutionary significance of natural variation in Arabidopsis. *Nature*, *441*(7096), 947.

Nakhamchik, A., Zhao, Z., Provart, N. J., Shiu, S. H., Keatley, S. K., Cameron, R. K., & Goring, D. R. (2004). A comprehensive expression analysis of the Arabidopsis proline-rich extensin-like receptor kinase gene family using bioinformatic and experimental approaches. *Plant and Cell Physiology*, *45*(12), 1875-1881.

Neve, P., Vila-Aiub, M., & Roux, F. (2009). Evolutionary-thinking in agricultural weed management. *New Phytologist*, *184*(4), 783-793.

Nordborg, M., Hu, T. T., Ishino, Y., Jhaveri, J., Toomajian, C., Zheng, H., ... & Jakobsson, M. (2005). The pattern of polymorphism in *Arabidopsis thaliana*. *PLoS biology*, *3*(7), e196.

O'Malley, R. C., & Ecker, J. R. (2010). Linking genotype to phenotype using the Arabidopsis unimutant collection. *The Plant Journal*, *61*(6), 928-940.

Oerke, E. C. (2006). Crop losses to pests. *The Journal of Agricultural Science*, *144*(1), 31-43.

Pecl, G. T., Araújo, M. B., Bell, J. D., Blanchard, J., Bonebrake, T. C., Chen, I. C., ... & Falconi, L. (2017). Biodiversity redistribution under climate change: Impacts on ecosystems and human well-being. *Science*, *355*(6332), eaai9214.

Phillips, F. J. (1909). A study of pinon pine. *Botanical Gazette*, *48*(3), 216-223.

Pierik, R., Mommer, L., & Voesenek, L. A. (2013). Molecular mechanisms of plant competition: neighbour detection and response strategies. *Functional Ecology*, *27*(4), 841-853.

Pisula, N. L., & Meiners, S. J. (2010). Relative allelopathic potential of invasive plant species in a young disturbed woodland. *The Journal of the Torrey Botanical Society*, *137*(1), 81-87.

Platt, A., Horton, M., Huang, Y. S., Li, Y., Anastasio, A. E., Mulyati, N. W., ... & Dunning, M. (2010). The scale of population structure in *Arabidopsis thaliana*. *PLoS genetics*, *6*(2), e1000843.

Pluess, A. R., Frank, A., Heiri, C., Lalagüe, H., Vendramin, G. G., & Oddou-Muratorio, S. (2016). Genome–environment association study suggests local adaptation to climate at the regional scale in Fagus sylvatica. *New Phytologist*, *210*(2), 589-601.

Prasad, K. V., Song, B. H., Olson-Manning, C., Anderson, J. T., Lee, C. R., Schranz, M. E., ... & Reichelt, M. (2012). A gain-of-function polymorphism controlling complex traits and fitness in nature. *science*, *337*(6098), 1081-1084.

Press, M. C. (1998). Dracula or Robin Hood? A functional role for root hemiparasites in nutrient poor ecosystems. *Oikos*, 609-611.

Pugnaire, F. I., & Luque, M. T. (2001). Changes in plant interactions along a gradient of environmental stress. *Oikos*, *93*(1), 42-49.

Rehman, S., Shahzad, B., Bajwa, A. A., Hussain, S., Rehman, A., Cheema, S. A., ... & Li, P. (2018). Utilizing the Allelopathic Potential of Brassica Species for Sustainable Crop Production: A Review. *Journal of Plant Growth Regulation*, 1-14.

Rellstab, C., Zoller, S., Walthert, L., Lesur, I., Pluess, A. R., Graf, R., ... & Gugerli, F. (2016). Signatures of local adaptation in candidate genes of oaks (*Quercus spp.*) with respect to present and future climatic conditions. *Molecular ecology*, *25*(23), 5907-5924.

Roux, F., & Bergelson, J. (2016). The genetics underlying natural variation in the biotic interactions of *Arabidopsis thaliana*: The challenges of linking evolutionary genetics and community ecology. In *Current topics in developmental biology* (Vol. 119, pp. 111-156). Academic Press.

Roux, F., Voisin, D., Badet, T., Balagué, C., Barlet, X., Huard-Chauveau, C., ... & Raffaele, S. (2014). Resistance to phytopathogens e tutti quanti: placing plant quantitative disease resistance on the map. *Molecular Plant Pathology*, *15*(5), 427-432.

Rubio, B., Cosson, P., Caballero, M., Revers, F., Bergelson, J., Roux, F., & Schurdi-Levraud, V. (2018). Genome-wide association study reveals new loci involved in *Arabidopsis thaliana* and *Turnip mosaic virus* (Tu MV) interactions in the field. *New Phytologist*.

Sakai, A. K., Allendorf, F. W., Holt, J. S., Lodge, D. M., Molofsky, J., With, K. A., ... & McCauley, D. E. (2001). The population biology of invasive species. *Annual review of ecology and systematics*, *32*(1), 305-332.

Sala, O. E., Chapin, F. S., Armesto, J. J., Berlow, E., Bloomfield, J., Dirzo, R., ... & Leemans, R. (2000). Global biodiversity scenarios for the year 2100. *science*, *287*(5459), 1770-1774.

Shindo, C., Bernasconi, G., & Hardtke, C. S. (2007). Natural genetic variation in Arabidopsis: tools, traits and prospects for evolutionary ecology. *Annals of Botany*, *99*(6), 1043-1054.

Shiu, S. H., & Bleecker, A. B. (2001). Receptor-like kinases from Arabidopsis form a monophyletic gene family related to animal receptor kinases. *Proceedings of the National Academy of Sciences*, *98*(19), 10763-10768.

Silva, N. F., & Goring, D. R. (2002). The proline-rich, extensin-like receptor kinase-1 (PERK1) gene is rapidly induced by wounding. *Plant molecular biology*, *50*(4-5), 667-685.

Tang, D., Wang, G., & Zhou, J. M. (2017). Receptor kinases in plant-pathogen interactions: more than pattern recognition. *The Plant Cell*, *29*(4), 618-637.

Tesio, F., & Ferrero, A. (2010). Allelopathy, a chance for sustainable weed management. *International Journal of Sustainable Development & World Ecology*, *17*(5), 377-389.

Tilman, D. (1985). The resource-ratio hypothesis of plant succession. *The American Naturalist*, *125*(6), 827-852.

Tilman, D., Lehman, C. L., & Thomson, K. T. (1997). Plant diversity and ecosystem productivity: theoretical considerations. *Proceedings of the national academy of sciences*, *94*(5), 1857-1861.

Tilman, D., Reich, P. B., Knops, J., Wedin, D., Mielke, T., & Lehman, C. (2001). Diversity and productivity in a long-term grassland experiment. *Science*, *294*(5543), 843-845.

Tilman, D., Wedin, D., & Knops, J. (1996). Productivity and sustainability influenced by biodiversity in grassland ecosystems. *Nature*, *379*(6567), 718.

Turkington, R., & Harper, J. L. (1979). The growth, distribution and neighbour relationships of *Trifolium repens* in a permanent pasture: IV. Fine-scale biotic differentiation. *The Journal of Ecology*, 245-254.

Turner, T. L., Bourne, E. C., Von Wettberg, E. J., Hu, T. T., & Nuzhdin, S. V. (2010). Population resequencing reveals local adaptation of *Arabidopsis lyrata* to serpentine soils. *Nature genetics*, *42*(3), 260.

Van Ruijven, J., & Berendse, F. (2003). Positive effects of plant species diversity on productivity in the absence of legumes. *Ecology Letters*, *6*(3), 170-175.

Violle, C., Enquist, B. J., McGill, B. J., Jiang, L. I. N., Albert, C. H., Hulshof, C., ... & Messier, J. (2012). The return of the variance: intraspecific variability in community ecology. *Trends in ecology & evolution*, *27*(4), 244-252.

Vitousek, P. M., D'antonio, C. M., Loope, L. L., Rejmanek, M., & Westbrooks, R. (1997). Introduced species: a significant component of human-caused global change. *New Zealand Journal of Ecology*, 1-16.

Wang, M., Roux, F., Bartoli, C., Huard-Chauveau, C., Meyer, C., Lee, H., ... & Bergelson, J. (2018). Two-way mixed-effects methods for joint association analysis using both host and pathogen genomes. *Proceedings of the National Academy of Sciences*, 201710980.

Wang, Q., Yang, S., Liu, J., Terecskei, K., Ábrahám, E., Gombár, A., ... & Fodor, L. (2017). Host-secreted antimicrobial peptide enforces symbiotic selectivity in *Medicago truncatula*. *Proceedings of the National Academy of Sciences*, *114*(26), 6854-6859.

Weigel, D., & Nordborg, M. (2005). Natural variation in Arabidopsis. How do we find the causal genes?. *Plant Physiology*, *138*(2), 567-568.

Wuest, S. E., & Niklaus, P. A. (2018). A plant biodiversity effect resolved to a single chromosomal region. *Nature ecology & evolution*, *2*(12), 1933.

Yang, S., Wang, Q., Fedorova, E., Liu, J., Qin, Q., Zheng, Q., ... & Bisseling, T. (2017). Microsymbiont discrimination mediated by a host-secreted peptide in *Medicago truncatula*. *Proceedings of the National Academy of Sciences*, *114*(26), 6848-6853.

Ye, Y., Ding, Y., Jiang, Q., Wang, F., Sun, J., & Zhu, C. (2017). The role of receptor-like protein kinases (RLKs) in abiotic stress response in plants. *Plant cell reports*, *36*(2), 235-242.

Zhang, Q., Yao, L. J., Yang, R. Y., Yang, X. Y., Tang, J. J., & Chen, X. (2007). Potential allelopathic effects of an invasive species *Solidago canadensis* on the mycorrhizae of native plant species. *Allelopathy Journal*, *20*(1), 71-78.

Zhao, K., Aranzana, M. J., Kim, S., Lister, C., Shindo, C., Tang, C., ... & Nordborg, M. (2007). An Arabidopsis example of association mapping in structured samples. *PLoS genetics*, *3*(1), e4.

Identification des bases génétiques associées à la variation naturelle des interactions plante-plante chez *Arabidopsis thaliana*

Les interactions biotiques sont déterminantes dans la réponse des communautés végétales face aux changements environnementaux. Parmi ces interactions, les interactions plante-plante jouent un rôle important dans la structure, la diversité et la dynamique des communautés végétales. Bien qu'il soit largement admis que l'identification des gènes associés aux interactions plante-plante soit une étape importante pour prédire et comprendre les dynamiques adaptatives des communautés végétales, les études portant sur l'identification des variants génétiques associés à la variation naturelle des interactions plante-plante restent étonnamment rares. L'objectif principal de cette thèse a été de caractériser et d'identifier les bases génétiques associées à la variation naturelle de la réponse compétitive d'une population locale (TOU-A) de l'espèce modèle *Arabidopsis thaliana* en interaction avec différentes espèces. Dans un premier chapitre, par une approche de résurrection couplée à des analyses de GWA mapping et de différentiation génétique temporelle, j'ai pu montrer que l'utilisation de la population TOU-A était adaptée pour l'identification fine des bases génétiques adaptatives associées aux interactions plante-plante en condition d'interaction monospécifique. Dans le second chapitre, afin de prendre en compte la complexité des interactions plante-plante dans la nature, je me suis donc intéressé à caractériser l'architecture génétique de la réponse compétitive d'*A. thaliana* dans différents contextes d'interactions mono- et pluri-spécifiques. J'ai pu mettre en évidence des phénomènes de spécialisation biotique de certaines accessions en réponse à certains assemblages, et ce, malgré une réponse parfois similaire de l'ensemble de la population en réponse aux différents traitements d'interaction. Par une approche de GWA mapping, j'ai pu mettre en évidence que les QTL de réponse à la compétition d'*A. thaliana* étaient majoritairement très différents entre les 12 traitements d'interaction. J'ai aussi pu montrer que les processus biologiques étaient différents entre conditions mono- et pluri-spécifiques et, plus spécifiquement, que les récepteurs de type kinase joueraient un rôle prépondérant dans les interactions plante-plante. Dans un troisième chapitre, j'ai cherché à valider fonctionnellement un gène sous-jacent à un QTL associé à la réponse compétitive d'*A. thaliana* vis-à-vis de *Poa annua*. Par le phénotypage de lignées mutantes et de lignées complémentées, j'ai pu identifier que le gène *PERK13/RHS10* était le gène causal sous-jacent à ce QTL. Cette validation fonctionnelle m'a permis de caractériser *PERK13/RHS10* comme étant un régulateur positif d'une stratégie d'évitement de la compétition en réponse à *P. annua* mais aussi au blé tendre (*Triticum aestivum*). Ces travaux s'inscrivent dans la lignée d'approches interdisciplinaires, qui ont pour but de mieux comprendre les déterminants génétiques qui sous-tendent la variation naturelle adaptative des interactions biotiques.

Mots-clés : génomique écologique, interactions plante-plante, population locale, variation naturelle, bases génétiques, GWA mapping, validation fonctionnelle, *Arabidopsis thaliana*.

Identification of genetic bases associated with natural variation of plant-plant interactions in *Arabidopsis thaliana*

Biotic interactions are crucial in the response of plant communities to environmental changes. Among these interactions, plant-plant interactions play an important role in the structure, diversity and dynamics of plant communities. Although it is widely accepted that the identification of genes associated with plant-plant interactions is an important step in predicting and understanding the adaptive dynamics of plant communities, studies on the identification of genetic variants associated with natural variation in plant-plant interactions remain scarce. The main objective of this thesis was to characterize and identify the genetic bases associated with natural variation of competitive response in a local population (TOU-A) of the model species *Arabidopsis thaliana* interacting with different species. In a first chapter, with a resurrection approach coupled with GWA mapping and a temporal genetic differentiation analysis, I was able to show that the TOU-A population was adapted for the identification of associated adaptive genetic bases underlying monospecific plant-plant interactions. In the second chapter, in order to take into account the complexity of plant-plant interactions observed in nature, I was interested in characterizing the genetic architecture of the competitive response of *A. thaliana* in different contexts of mono- and multi-specific interactions. I was able to highlight biotic specialization of some accessions in response to certain assemblages, and this, despite observing a similar response of the entire population in response to different interaction treatments. Through an approach of GWA mapping, I was able to highlight that the QTLs of competitive response of *A. thaliana* were mostly different between the 12 interaction treatments. I have also been able to show that biological processes differ between mono- and multi-specific conditions and, more specifically, that receptor like-kinase play a major role in plant-plant interactions. In a third chapter, I sought to functionally validate a gene underlying a QTL associated with the competitive response of *A. thaliana* to *Poa annua*. By phenotyping mutant lines and complemented lines, I was able to identify that the *PERK13/RHS10* gene was the causative gene underlying this QTL. This functional validation allowed me to characterize *PERK13/RHS10* as a positive regulator of a competition avoidance strategy in response *to P. annua* but also to wheat (*Triticum aestivum*). This work is in line with interdisciplinary approaches, which intends to improve our understanding of the genetic determinants that underlie the adaptive natural variation of biotic interactions.

Keywords: ecological genomics, plant-plant interactions, local population, natural variation, genetic bases, GWA mapping, functional validation, *Arabidopsis thaliana*.