

Mémoire pour l’Habilitation à Diriger des Recherches

Université des Sciences et Technologies de Lille 1

Génomique écologique de l’adaptation chez *Arabidopsis thaliana*

Fabrice ROUX

Présenté le 15 mai 2013 devant la commission d’examen :

Jacques DAVID, Professeur, SupAgro Montpellier

Rapporteur

Jean-Christophe SIMON, DR INRA, INRA Rennes

Rapporteur

Christophe THEBAUD, Professeur, Université Paul Sabatier Toulouse

Rapporteur

Maud TENAILLON, CR CNRS, Ferme du Moulon – Gif-sur-Yvette

Examineur

Joël CUGUEN, Professeur, Université de Lille 1

Examineur

Remerciements

Le temps des remerciements est venu et une angoisse m'envahit. Qui dois-je remercier ? Ou plutôt qui ne pas oublier de mettre dans les remerciements ? Depuis mes débuts en recherche, j'ai été accompagné, soutenu et guidé par de nombreuses personnes qui, par leur personnalité, m'ont autant apporté au niveau humain qu'au niveau scientifique. Je tiens donc à tous les remercier ; mais avant tout, je tiens à m'excuser. De quoi ? D'oublier des personnes ou de ne pas écrire les mots justes pour leur dire ce que je ressens.

Je tiens tout d'abord à remercier le laboratoire GEPV dans son intégralité pour sa très bonne ambiance. Je remercie aussi le laboratoire de m'avoir permis de mettre à profit de façon très satisfaisante les ressources et moyens mis à ma disposition. Je tiens notamment à remercier Joël Cuguen pour m'avoir demandé en 2005 (alors que j'étais ATER au laboratoire GEPV) de postuler au CNRS. Sortant de thèse, je n'avais pas forcément une idée très claire des postes auxquels je voulais postuler. A cette époque, partir effectuer un post-doc à l'étranger était ce qui m'importait avant tout. Postuler au CNRS m'a donc permis de prendre du recul sur mes recherches et de proposer un projet à partir de quelques idées qui me trottaient dans la tête depuis la thèse.

Merci aux thésards et post-docs Arabidopsiesques : Benjamin Brachi, Romain Villoutreix, Etienne Baron et Marilyne Debieu. Bon pas la peine de tourner autour du pot, je vous adore. Travailler et monter des projets avec vous est un réel plaisir. J'avoue qu'il n'est pas toujours facile de travailler avec une espèce modèle telle qu'Arabidopsis thaliana. Malgré tous les avantages que cette espèce présente, il faut quand même se coltiner pas mal de physiologie moléculaire aussi bien dans les papiers que dans les congrès... et vous savez de quoi je parle : vous avez tous assisté à un congrès international Arabidopsis dont certains titres de session suffisaient à eux seuls à nous filer une migraine. Vous m'avez aussi appris qu'il n'existe pas une seule manière d'encadrer et qu'un encadrant doit savoir s'adapter à la personnalité de chacun.

Il me faut aussi remercier les nombreux stagiaires que j'ai embarqués durant mes projets de recherche. J'espère ainsi avoir suscité un quelconque intérêt à l'étude de l'adaptation.

Merci au personnel de la serre pour toute l'aide si précieuse que vous avez su apporter à mes expériences : je promets que je ne lancerai plus de manip avec plus de 10000 plantes. Mention spéciale pour Nathalie Faure. Tu es arrivée en même temps que moi au laboratoire et tu as bien voulu m'accompagner dans cette aventure. Je ne te remercierai jamais assez pour ton aide, ton ouverture d'esprit, ta volonté de mettre en place des projets sur le long terme et nos discussions aussi bien en serre que sous la hotte de la salle bactério. Cédric Glorieux, je ne t'oublie pas. Que dire ? C'est juste super de bosser avec toi. Phénotypeur fou,

ton enthousiasme et ta sympathie n'ont pas d'égal. Nathalie, Cédric, vous allez énormément me manquer.

Merci à Cécile Godé, Adeline Courseaux et Anne-Catherine Holl pour m'avoir aidé dans tous les aspects liés à la biologie moléculaire, aussi bien d'un point de vue protocoles que commandes de consommables. Je n'arrive toujours pas à comprendre pourquoi vous n'avez jamais râlé alors que je faisais toujours mes manips de BM à l'arrache.

Merci à Sandrine Belingheri, Michèle Delecourt et Muriel Fikry pour m'avoir aidé à trouver mon chemin dans les méandres de l'administration (même si les voies sans issue ou les sens interdits sont encore trop nombreux à mon goût), surtout quand on doit jongler entre l'Université et le CNRS.

Merci à Claire-Lise Meyer, Nina Hautekète, Yves Piquot et Laurent Amsellem pour toutes ces discussions scientifiques, notamment au niveau de l'écologie, domaine que je découvre petit à petit et que j'apprécie de plus en plus. Un grand merci à ma collègue de bureau Hélène Frérot : comment as-tu fait pour ne pas me mettre à la porte du bureau, après toutes ces questions incessantes que je te posais ? C'est un vrai plaisir d'interagir avec toi et j'espère qu'on pourra continuer à le faire.

Je tiens aussi à remercier 3 personnes qui m'ont beaucoup influencé dans ma carrière scientifique. Tout d'abord, Xavier Reboud, mon encadrant de thèse. Nous avons passé énormément d'heures à remplir d'idées le tableau blanc de ton bureau, quitte parfois à oublier d'aller manger. Nous n'étions pas toujours d'accord et c'est tant mieux ! Ceci nous a amenés à des débats bien souvent au-delà du sujet de ma thèse, poussant les limites de notre imagination. Toutes ces discussions me manquent encore beaucoup ! Passons à Valérie Le Corre que je surnomme ma marraine de thèse. Tu m'as initié à la génétique quantitative et je ne te remercierai jamais assez pour ça. Je terminerai avec Joy Bergelson qui est devenue bien plus qu'une chef de post-doc. Je pense qu'il faut être honnête : il existe dans le monde des gens très intelligents et très doués, et je pense que tu en fais partie. Au tout début de ma thèse, je prenais plaisir à lire tous les articles sortant de ton laboratoire. Pour chaque article, je savais qu'une idée simple mais très originale y serait présente. Alors quand tu m'as proposé de faire un post-doc chez toi, j'étais évidemment plus que comblé. Plus que jamais, tu me supports dans les projets que j'entreprends et ça fait un bien fou. Je n'oublierai aussi jamais ce que tu m'as dit un jour : « quand on a des idées, il n'y a pas de compétition ».

Un grand merci à tous !

Fabrice, Arabidopsis addicted mais qui se soigne

Sommaire

<i>Curriculum vitae</i>	3
Publications et communications scientifiques	8
DOCUMENT DE SYNTHÈSE	15
Mieux comprendre l'adaptation des traits complexes dans la nature	16
I. <i>Arabidopsis thaliana</i> : une espèce modèle en génétique et en écologie	19
II. Etude de traits adaptatifs de plus en plus complexes	23
A. La résistance aux herbicides : un cas d'école	23
1. La balance coût-bénéfice	23
2. Impact de l'hétérogénéité de l'environnement sur la dynamique d'une adaptation ...	25
3. Conclusion	26
B. La résistance aux pathogènes : qualitatif vs. quantitatif	27
1. Valeur adaptative d'une résistance qualitative en fonction de la concentration initiale en pathogènes	29
2. Identification des bases génétiques de la résistance quantitative aux bactéries pathogènes de la communauté foliaire d' <i>A. thaliana</i>	30
2.1. <i>Pseudomonas viridiflava</i>	30
2.2. <i>Xanthomonas</i> sp.	31
2.2.1. <i>Xanthomonas campestris</i> pv. <i>campestris</i>	31
2.2.2. Souches isolées dans des populations naturelles d' <i>A. thaliana</i>	34
3. Conclusion.....	36
C. La phénologie : date de floraison, période de floraison et période reproductive .	37
1. Génétique et épigénétique	37
1.1. Identification des bases génétiques associées à la variation phénologique.....	37
1.2. Contribution de l'épigénétique dans la variation naturelle phénologique.....	40
1.3. Reconstruction de la première marche de la marche adaptative vers la précocité de floraison	41
2. Ecologie.....	43
2.1. Adaptation au climat à l'échelle mondiale	43
2.2. Identification à différentes échelles spatiales des agents sélectifs agissant sur la phénologie	45
3. Evolution	46
3.1. Evolution expérimentale dans des conditions environnementales contrôlées.	46

3.2. Lien entre fitness et date de floraison : serre vs. nature	47
4. Conclusion.....	48
III. Phénotype global et normes de réaction.....	49
A. Cohorte de germination et effet année	50
B. Enfouissement des graines x cohorte de germination.....	52
C. Interactions plante - plante.....	52
D. Environnement cytoplasmique	54
IV. Perspectives : génomique écologique de l'adaptation dans les communautés végétales.....	55
A. Interactions plante - pathogène.....	56
B. Interactions plante - plante.....	58
Références bibliographiques	60

CURRICULUM VITAE

Né le 11 Juillet 1977 à Luzy (58).

Nationalité : Française.

Adresse : UMR CNRS 8198, Laboratoire de Génétique et Evolution des Populations Végétales, Université des Sciences et Technologies de Lille 1, 59655 Villeneuve d'Ascq FRANCE

E-mail : fabrice.roux@univ-lille1.fr

Téléphone : 03 20 43 40 33

Fax : 03 20 43 69 79

PARCOURS PROFESSIONNEL

Janvier 2012 – Juillet 2012 : Lauréat Fulbright, laboratoire de Joy Bergelson, Department of Ecology and Evolution, University of Chicago, USA

Depuis janvier 2012 : Co-responsable avec Yves Piquot (MCF) de la thématique de recherche CGENECO 'Changements globaux : de la génomique écologique à l'écologie des communautés'

Depuis janvier 2011 : **Chargé de Recherche 1^{ère} classe CNRS**

Lab. de Génétique et Evolution des Populations Végétales, Université Lille1

Responsable de la thématique de recherche 'Génomique évolutive et écologique de l'adaptation chez l'espèce modèle *Arabidopsis thaliana*'.

2007 - 2010 : **Chargé de Recherche 2^{ème} classe CNRS**

Lab. de Génétique et Evolution des Populations Végétales, Université Lille1

Responsable de la thématique de recherche 'Génomique évolutive et écologique de l'adaptation chez l'espèce modèle *Arabidopsis thaliana*'.

2006 : **Chercheur CDD Post-doc**

Department of Ecology and Evolution, University of Chicago, USA

Direction: Joy Bergelson

Financement: Dropkin Foundation

- Genome-wide association mapping des bases génétiques de la résistance quantitative chez le système hôte-pathogène *Arabidopsis thaliana* – *Pseudomonas viridiflava*
- Modélisation et mise en évidence de la coévolution chez le système hôte-pathogène *Arabidopsis thaliana* – *Pseudomonas viridiflava*

2004 – 2005 : **ATER (1/2 service)**

UMR CNRS 8016, Lab. de Génétique et Evolution des Populations Végétales, Université Lille1

Direction : Joel Cuguen et Pascal Touzet

- Analyse bibliographique sur l'évolution de la précocité de floraison.

2001 – 2004 : **Doctorat en Biologie des Populations, mention Très Honorable avec les félicitations du jury**

UMR Biologie et Gestion des Adventices, INRA, Dijon

Direction : Xavier Reboud et Jacques Gasquez

Financement : Bayer Crop Sciences

Soutenance : 24 novembre 2004

- Gestion et prévention de la résistance aux herbicides : utilisation d'*Arabidopsis thaliana* comme modèle expérimental.

2000 – 2001 : **DSER en Génétique des Populations Végétales, mention Très Honorable**

UMR Biologie et Gestion des Adventices, INRA, Dijon

Direction : Xavier Reboud et Marie-Jeanne Perrot-Minot

- Analyse de la diversité génétique neutre et sélectionnée dans des métapopulations d'*Arabidopsis thaliana*.

1999 - 2000 : **DEA Biologie de l'Evolution et Ecologie, mention Assez Bien**

Université de Montpellier II

Directeur de stage : Xavier Reboud (INRA, Dijon)

- Evolution de la diversité génétique et du potentiel d'adaptation d'*Arabidopsis thaliana* dans un système en métapopulation avec ou sans extinction locale.

1998 - 1999 : **Maîtrise de Biologie des Populations et des Ecosystèmes, mention Bien**
Université de Bourgogne

ENSEIGNEMENT

- 2007-2012 : 3h cours magistral par année. M2R Master Recherche Génétique, Ecologie et Evolution, Université de Lille 1. Etude de l'adaptation chez les plantes.
- 2004-2005 : **Attaché Temporaire d'Enseignement et de Recherche, Université Lille1**
Génétique des populations : Licence 2^{ème} année, 48h ETD
Biologie évolutive : Master 1^{ère} année, 30h ETD
Biologie végétale : Licence 3^{ème} année, 6h ETD
Biodiversité : Licence 3^{ème} année, 12h ETD
- 2003-2004 : **Agent Temporaire Vacataire (ATV), Université de Bourgogne**
Biologie Végétale : DEUG, 20h ETD
- 2002-2003 : **(1) ATV, Université de Bourgogne**
Génétique Moléculaire et Cellulaire : Maîtrise, 35h ETD
Génétique formelle : Licence, 12h ETD
(2) ATV, Institut National Agronomique Paris-Grignon (INA-PG)
Option "Populations de mauvaises herbes" : deuxième année de l'INAPG, 2h ETD
- 2001-2002 : **ATV, Université de Bourgogne**
Biologie Végétale : DEUG, 27h ETD
Biologie Végétale : Licence, 16h ETD
Génétique Moléculaire et Cellulaire : Maîtrise, 48h ETD
- 2000-2001 : **ATV, Université de Bourgogne**
Biologie Végétale : DEUG, 40h ETD
Ecologie : DEUG, 16h ETD

RECAPITULATIF	2000-2001	2001-2002	2002-2003	2003-2004	2004-2005
Heures equivalent TD	56	91	49	20	96

COMPETENCES SCIENTIFIQUES

- **Expérimentation:** mise en place de protocoles et designs expérimentaux, expériences en serre et sur le terrain (travail du sol, semis, repiquage, notation de caractères phénotypiques, récolte des plantes), préparation de solutions herbicides et traitement herbicide de plantes en conditions contrôlées, inoculation de plantes par des bactéries pathogènes.
- **Echantillonnage de populations naturelles** (*Arabidopsis thaliana* à travers l'Europe et sur les îles du Cap Vert, betteraves sauvages sur les îles Canaries et Madère).
- **Biologie moléculaire:** extraction ADN, amplification PCR (RAPD, RFLP, microsatellites, PASA (PCR allèle-bispcifique)), électrophorèse, isoenzymes, clonage par vecteur bactérien, purification ADN, séquençage, génotypage haut-débit (Amplifluor, Taqman, KASPAR).
- **Culture in vitro** : semis de graines et repiquage de plantules sur milieu gélosé, culture de bactéries pathogènes.
- **Analyses de données** :
 - **Statistiques** : tests paramétriques et non-paramétriques, modèles linéaires et non-linéaires, analyse de variance, analyses multivariées, modèle mixte, ACP, AFC, Partial-Least Square Regression. Utilisation des logiciels SAS, R et SYSTAT pour analyse des données.

- **Génétique des Populations** : F-STAT pour analyses, METAPOPOP pour analyses et simulations.
- **Génétique Quantitative** : METAPOPOP pour analyses et simulations.
- **QTL mapping** : mixed-model composite interval mapping (MCIM) avec QTLnetwork.
- **GWA mapping**: tests de Wilcoxon ou Fisher (sans contrôle de la structure des populations) et modèle mixte (avec contrôle de la structure des populations, programme EMMAX).
- **Analyses de séquences et Evolution Moléculaire** : PubMed et TAIR pour recherche de séquences nucléotidiques ou protéiques ; Bioedit, Sequencher, Multalin pour vérification des données et alignement des séquences ; DnaSP et MEGA pour analyses moléculaires ; SNPMaster, SDS2.0 pour analyse de données issues de génotypage haut-débit.
- **Programmation sous R**

ACTIVITES RELEVANT DES MISSIONS AUTRES QUE LA RECHERCHE

- **Encadrements et co-encadrements:**

- **3 étudiants en thèse (directeur: Joël Cuguen)**

- Benjamin Brachi (2007-2010, Allocataire MESR). Etude de la variation naturelle de traits phénologiques chez *Arabidopsis thaliana* par une approche de génomique écologique.
 - Romain Villoutreix (2010-2013, Contrat doctoral de l'Université de Lille 1). Adaptation locale vs. plasticité phénotypique chez la plante modèle *Arabidopsis thaliana*.
 - Etienne Baron (2011-2014, Contrat doctoral de l'Université de Lille 1 – Région Nord-Pas-de-Calais, co-encadrement Laurent Amsellem). Suivi pérenne de populations expérimentales témoins d'*Arabidopsis thaliana* en relation avec le climat local du Nord-Pas-de-Calais : influence des interactions plante - plante.

- **Contrats post-doctoraux**

- Marilyne Debieu (post-doc, 2010-2013, financement : ANR Blanc, projet QUANTIREX)
 - Cécile Edelist (ATER, 2007)

8 étudiants de M2 : 1 Master Recherche Gène, Sélection, Adaptation, Université de Bourgogne ; 3 Master Recherche Génétique, Ecologie et Evolution, Université de Lille 1 ; 1 Master Recherche BGAE, Université de Montpellier 2 ; 1 Master BioSciences, Ecole Normale Supérieure Lyon ; 1 Master Pro Science et Technologies Biologie-Biochimie, Université d'Orléans ; 1 étudiante en DSR, Université de Lille 1.

13 étudiants de M1 : 3 Maîtrise de Biologie des Populations et des Ecosystèmes, Université de Bourgogne ; 1 Maîtrise de Sciences et Techniques du Végétal, Université de Bourgogne ; 1 Maîtrise de Biochimie, Université de Bourgogne ; 8 Master Biodiversité et Ecosystèmes Continentaux et Marins, Université de Lille 1.

10 étudiants de Licence.

- **Membre du bureau éditorial de Plant Biology** (2009-2012)

- **Activités Editoriales d'Arbitrage** : Acta Oecologica, American Naturalist, Ecology Letters, Genetics, Heredity, Journal of Applied Ecology, Molecular Biology and Evolution, Molecular Ecology, Nature Reviews Genetics, New Phytologist, Perspectives in Plant Ecology Evolution and Systematics, Plant Biology, PLoS Genetics, Weed Science.

- **Participant au GDR 'Génétique Quantitative dans les Populations Naturelles'** 2011-2014. En charge avec Hélène Frérot (MCF GEPU, Université de Lille 1) de la thématique 'Adaptation en milieu hétérogène'.

- **Membre de la Société Française d'Ecologie**

- **Membre du bureau du service commun des serres de l'Université de Lille 1.** Participation au projet « serre et extensions » dans le cadre du Plan campus Grand Lille.
- **Expert :**
 - Membre de jurys de concours :
 - ✓ CR2 INRA 2010 – Ecologie évolutive et biodiversité
 - ✓ CR2 INRA 2011 – Génétique et biologie moléculaire végétale
 - Membre de jurys de thèse :
 - ✓ Bertrand Jacquemin (examinateur, juin 2008, Université de Bourgogne)
 - ✓ Abdoul-Aziz Saïdou (examinateur, mars 2011, Montpellier SupAgro)
 - ✓ Eléonore Durand (examinateur, juin 2011, Université Paris XI Orsay)
 - ✓ Sylvain Poque (examinateur, décembre 2012, Université Bordeaux Segalen)
 - Membre de jury du concours des allocations de recherche 2008 de l'Ecole Doctorale Environnements, Santé, STIC (Dijon)
 - Membre du vivier MCF section CNU 67-68 (Université de Lille 1). 2008-2011, 2011-
 - Membre de 6 comités de thèse : Eléonore Durand (UMR de Génétique Végétale, Le Moulon), Coraline Caullet et Solène Bellanger (UMR Biologie et Gestion des Adventices, INRA Dijon), Juliette Doumayrou (UMR Génétique et Evolution des Maladies Infectieuses, Montpellier), Charlotte Trontin (Variation and Abiotic Stress Tolerance group, INRA Versailles), Pierre Nouhaud (UMR Biologie des Organismes et des Populations appliquée à la Protection des Plantes, INRA Rennes).
 - Outside faculty member of dissertation committee: Christopher Meyer (Joy Bergelson's lab, University of Chicago).
 - Co-organisateur de la session 'Intégration de l'interaction Génotype x Environnement dans l'étude des capacités d'adaptation'. Congrès Ecologie 2010 – Montpellier.
 - Expert pour projets NSF (2009, 2010), Netherlands Organisation for Scientific Research NOW (2011, 2012), European Plant Phenotyping Network EPPN (2012), German Research Foundation DFG (2012), FRB (2012).
- **Brevet européen PCT/EP2013/051490:** Resistance to black rot disease in plants from the Brassicaceae, Solanaceae and Cucurbitaceae families. Inventeur principal, responsable du projet: Dominique Roby. 6 inventeurs au total.
- Représentant des étudiants du Second Cycle à l'Université de Bourgogne (1998-2000)
- Gagnant d'un "IWSS Travel Award" de la "International Weed Science Society" comme récompense de la meilleure communication (Durban, South Africa, Juin 2004)

FINANCEMENTS

- 2007 : Bonus Qualité Recherche, Université des Sciences et Technologies de Lille 1. 20k€.
- 2009 - 2013 : ANR Blanc, projet QUANTIREX. PI : Dominique Roby (LIPM, Toulouse). 355 k€ (montant obtenu pour l'équipe 'Génomique évolutive et écologique de l'adaptation chez *Arabidopsis thaliana*' : 155k€).
- 2012 – 2016 : ANR BIOADAPT, projet CYTOPHENO (Cyto-nuclear co-adaptation and plant adaptive phenotypes). PI : Françoise Budar (INRA IJPB). 468 k€ (montant obtenu pour l'équipe CGENECO: 25 k€).
- Janvier 2012 – Juillet 2012 : Lauréat Fulbright. Determining the potential for adaptation of natural plant populations under climate change by identification of key genes underlying pathogen resistance. 15k\$.

COLLABORATIONS

1. Genome-wide association mapping chez *A. thaliana* : Joy Bergelson (University of Chicago, USA), Magnus Nordborg (Mendel Institute, Vienna, Austria), Valérie Le Corre (INRA, Dijon)
2. Séquençage de génomes: Joy Bergelson (University of Chicago, USA) et consortium international
3. Ecologie et échelle de l'adaptation chez *A. thaliana* : Joy Bergelson (University of Chicago, USA), Johanna Schmitt (Brown University, Providence, USA), Nina Hautekèete et Yves Piquot (Laboratoire GEVP, Université de Lille 1)
4. Suivi pérenne de populations expérimentales sur des sites ateliers du Nord de la France : Nina Hautekèete, Laurent Amsellem et Yves Piquot (Laboratoire GEVP, Université de Lille 1)
5. Résistance aux pathogènes : Dominique Roby (LIPM, CNRS-INRA, Toulouse), Joy Bergelson (University of Chicago, USA)
6. Interactions nucléo-cytoplasmiques : Françoise Budar (INRA Versailles)
7. Résistance herbicide chez les mauvaises herbes: Xavier Reboud, (INRA Dijon), Martin Vila-Aiub (Universidad de Buenos Aires, Argentina), Paul Neve (University of Warwick, UK).

MOBILITES

Afin de maintenir la très bonne collaboration établie en 2006 lors de mon séjour post-doctoral dans le laboratoire de Joy Bergelson (Department of Ecology and Evolution, University of Chicago, USA), j'essaie régulièrement de passer quelques semaines dans son laboratoire. Un point marquant a été un séjour de 7 mois (janvier – juillet 2012) dans le laboratoire de Joy Bergelson dans le cadre d'une bourse Fulbright. Ce séjour m'a permis d'avoir accès à des infrastructures de phénotypage à haut-débit de la résistance quantitative aux pathogènes, mais aussi de collaborer avec des chercheurs travaillant sur l'évolution de la virulence chez des bactéries pathogènes.

Grace au Collège Doctoral Européen de l'Université de Lille 1 et à l'Ecole Doctorale SMRE à laquelle est affilié le laboratoire GEVP, j'encourage mes étudiants en thèse à faire des séjours de courte durée dans des laboratoires étrangers. Ainsi, Benjamin Brachi (doctorant que j'ai co-encadré entre 2007 et 2010) est parti un mois (mai 2009) dans le laboratoire de Joy Bergelson. A la suite de ce séjour, Benjamin Brachi s'est vu proposé une bourse post-doctorale de 5 ans par Joy Bergelson. Benjamin Brachi est actuellement en post-doc à Chicago depuis janvier 2011. Romain Villoutreix (doctorant que j'encadre depuis octobre 2010) est parti un mois (juillet 2011) dans le laboratoire de Johanna Schmitt (Brown University, Providence, USA) et 6 semaines dans le laboratoire de Joy Bergelson (juin – juillet 2012).

PUBLICATIONS ET COMMUNICATIONS SCIENTIFIQUES

ARTICLES SCIENTIFIQUES

A. Publiés

Revues	Nombre	Facteur d'impact 2011
Nature Reviews Genetics	1	38.075
Nature	1	36.280
Science	1	31.201
Trends in Plant Science	1	11.047
PLoS Genetics	2	8.694
New Phytologist	2	6.645
The Plant Journal	1	6.160
Molecular Biology and Evolution	1	5.550
Evolution	1	5.146
Heredity	2	4.597
PLoS One	2	4.092
Genetics	5	4.007
Theoretical and Applied Genetics	1	3.297
Genetics, Selection, Evolution	1	2.885
Pest Management Science	2	2.251
Weed Research	1	1.924
Genetical Research	1	1.712
Crop Protection	1	1.402
Biofutur	1	0.023
Plant Signaling and Behavior	1	-

1. Cosson, P., V. Schurdi-Levraud, H. Le Quang, O. Sicard, **F. Roux**, O. Le Gall, T. Candresse and F. Revers. 2012. The RTM resistance to potyviruses in *Arabidopsis thaliana*: natural variation of the *RTM* genes and evidence for the implication of additional genes. **PLoS One** 7(6): e39169.
2. Brachi, B., C. Aimé, C. Glorieux, J. Cuguen and **F. Roux**. 2012. Adaptive value of phenological traits in stressful environments: predictions based on seed production and Laboratory Natural Selection. **PLoS One** 7(3):e32069.
3. Hancock, A.M., B. Brachi, N. Faure, M.W. Horton, L.B. Jarymowycz, F.G. Sperone, C. Toomajian, **F. Roux** and J. Bergelson. 2011. Adaptation to climate across the *Arabidopsis thaliana* genome. **Science** 334: 83-86.
4. Vila-Aiub, M., P. Neve and **F. Roux**. 2011. A unified approach to the estimation and interpretation of resistance costs in plants. **Heredity** 107(5): 386-394.
5. **Roux, F.***, M. Colomé-Tatché*, C. Edelist, R. Wardenaar, P. Guerche, F. Hospital, V. Colot, R.C. Jansen and F. Johannes. 2011. Genome-wide epigenetic perturbation jump-starts patterns of heritable variation found in nature. **Genetics** 188: 1015-1017. *Authors contributed equally to this work. *Highlighted by the Faculty of 1000 (June 2011) and The Scientist (October 2011 "Evolution, Tout de suite. Epigenetic perturbations could jump-start heritable variation.")*.
6. Budar, F. and **F. Roux**. 2011. The role of organelle genomes in plant adaptation: time to get to work! **Plant Signaling and Behavior** 6(5): 635-639.
7. Bergelson, J. and **F. Roux**. 2010. Towards identifying genes underlying ecologically relevant traits in *Arabidopsis thaliana*. **Nature Reviews Genetics** 11: 867-879.
8. Moison, M., **F. Roux**, M. Quadrado, R. Duval, M. Ekovich, D.-H. Lê, M. Verzaux and F. Budar. 2010. Cytoplasmic phylogeny and evidence of cyto-nuclear co-adaptation in *Arabidopsis thaliana*. **Plant Journal** 63: 728-738.

9. **Roux, F.**, L. Gao and J. Bergelson. 2010. Impact of initial pathogen density on resistance and tolerance in a polymorphic disease resistance system in *Arabidopsis thaliana*. **Genetics** 185: 283-291.
10. Brachi, B. , N. Faure, M. Horton, E. Flahauw, A. Vazquez, M. Nordborg, J. Bergelson, J. Cuguen and **F. Roux**. 2010. Linkage and association mapping of *Arabidopsis thaliana* flowering time in nature. **PLoS Genetics** 6(6):e1000940.
11. Platt, A., M. Horton, Y.S. Huang, Y. Li, A.E. Anastasio, N.W. Mulyati, J. Agren, O. Bossdorf, D. Byers, K. Donohue, M. Dunning, E.B. Holub, A. Hudson, V. Le Corre, O. Loudet, **F. Roux**, N. Warthmann, D. Weigel, L. Rivero, R. Scholl, M. Nordborg, J. Bergelson and J.O. Borevitz. 2010. The scale of population structure in *Arabidopsis thaliana*. **PLoS Genetics** 6: e1000843.
12. Atwell, S., Y. Huang, B.J. Vilhjalmsen, G. Willems, M. Horton, Y. L. D. Meng, A. Platt, A. Tarone, T.T. Hu, R. Jiang, N.W. Mulyati, X. Zhang, M. A. Amer, I. Baxter, B. Brachi, J. Chory, C. Dean, M. Debieu, J. de Meaux, J.R. Ecker, N. Faure, J.M. Kniskern, J.D.G. Jones, T. Michael, A. Nemri, **F. Roux**, D.E. Salt, C. Tang, M. Todesco, M.B. Traw, D. Weigel, P. Marjoram, J. Borevitz, J. Bergelson and M. Nordborg. 2010. Genome-wide association study of 107 phenotypes in a common set of *Arabidopsis thaliana* inbred lines. **Nature** 465(7298): 627-631.
13. Neve, P., M. Vila-Aiub and **F. Roux**. 2009. Evolutionary-thinking in agricultural weed management. **New Phytologist** 184: 783-793.
14. Gao, L., **F. Roux** and J. Bergelson. 2009. Quantitative fitness effects of infection in a gene-for-gene system. **New Phytologist** 184: 465-494.
15. Paris, M.*, **F. Roux***, A. Bérard and X. Reboud. 2008. The effects of genetic background on herbicide resistance fitness cost and its associated dominance in *Arabidopsis thaliana*. **Heredity** 101: 499-506. *Authors contributed equally to this work.
16. **Roux, F.**, M. Paris and X. Reboud. 2008. Delaying weed adaptation to herbicide by environmental heterogeneity: a simulation approach. **Pest Management Science** 64: 16-29.
17. Le Corre, V. and **F. Roux**. 2007. Gene research into the precocity of blossoming. **Biofutur** 282: 40-43.
18. **Roux, F.** and X. Reboud. 2007. Herbicide resistance dynamics in a spatially heterogeneous environment. **Crop Protection** 26: 335-341.
19. **Roux, F.**, P. Touzet, J. Cuguen and V. Le Corre. 2006. How to be early-flowering: an evolutionary perspective. **Trends in Plant Science** 11: 375-381.
20. **Roux, F.**, S. Giancola, S. Durand and X. Reboud. 2006. Building of an experimental cline with *Arabidopsis thaliana* to estimate herbicide fitness cost. **Genetics** 173: 1023-1031.
21. Giancola, S., H.I. McKhann, A. Bérard, C. Camilleri, S. Durand, P. Libeau, **F. Roux**, X. Reboud, I.G. Gut and D. Brunel. 2006. Utilization of the three high-throughput SNP genotyping, the GOOD assay, Amplifluor and TaqMan, in diploid and polyploid plants. **Theoretical and Applied Genetics** 112: 1115-1124.
22. **Roux, F.**, C. Camilleri, A. Bérard and X. Reboud. 2005. Multigenerational vs. single generation studies to estimate herbicide resistance fitness cost in *Arabidopsis thaliana*. **Evolution** 59(10): 2264-2269.
23. **Roux, F.**, C. Camilleri, S. Giancola, D. Brunel and X. Reboud. 2005. Epistatic interactions among herbicide resistances in *Arabidopsis thaliana*: the fitness cost of multi-resistance. **Genetics** 171: 1277-12788.
24. **Roux, F.**, A. Matějček, J. Gasquez and X. Reboud. 2005. Dominance variation across six herbicides of the *Arabidopsis thaliana* *csr1-1* and *csr1-2* resistance alleles. **Pest Management Science** 61: 1089-1095.
25. **Roux, F.** and X. Reboud. 2005. Is the cost of herbicide resistance expressed in the breakdown of the relationships between characters? A case study using synthetic auxin resistant *Arabidopsis thaliana* mutants. **Genetical Research** 85(2): 101-110.
26. **Roux, F.**, A. Matějček and X. Reboud. 2005. Response of *Arabidopsis thaliana* to 22 ALS-inhibitors: baseline toxicity and cross-resistance of *csr1-1* and *csr1-2* mutants. **Weed Research** 45: 220-227.
27. **Roux, F.**, J. Gasquez and X. Reboud, 2004. The dominance of the herbicide resistant cost in

several *Arabidopsis thaliana* mutant lines. **Genetics** 166: 449-460.

28. Le Corre, V., **F. Roux** and X. Reboud, 2002. DNA polymorphism at the *FRIGIDA* gene in *Arabidopsis thaliana*: extensive non-synonymous variation is consistent with local selection for flowering time. **Molecular Biology and Evolution** 19: 1261-1271.
29. Lavigne, C., X. Reboud, M. Lefranc, E. Porcher, **F. Roux**, I. Olivieri and B. Godelle, 2001. Evolution of genetic diversity in metapopulations: *Arabidopsis thaliana* as an experimental model. **Genetics, Selection, Evolution** 33 (Suppl.1): S399-S423.

B. Soumis

30. Huard-Chauveau, C., L. Perchepped, M. Debieu, S. Rivas, I. Kars, J. Bergelson, **F. Roux** and D. Roby. An atypical kinase under balancing selection confers broad-spectrum disease resistance in *Arabidopsis*. **Nature** (in revision).
31. Brachi, B., R. Villoutreix, N. Faure, N. Hautekèete, Y. Piquot, M. Pauwels, D. Roby, J. Cuguen, J. Bergelson and **F. Roux**. Investigation of the geographical scale of adaptive phenological variation and its underlying genetic bases in *Arabidopsis thaliana*. **Molecular Ecology** (in revision).
32. Brachi, B., N. Faure, J. Bergelson, J. Cuguen and **F. Roux**. Genome-wide association mapping of flowering time in *Arabidopsis thaliana* in nature: genetics for underlying components and reaction norms across two successive years. **Acta Botanica Gallica – Botany Letters** (in revision).

C. Actes de colloque, article de vulgarisation et chapitre de livre

33. **Roux, F.**, J. Gasquez and X. Reboud. 2004. Comment prévenir et gérer la résistance aux herbicides chez les mauvaises herbes ? **19ème conférence du COLUMA, Journées Internationales sur les Mauvaises Herbes**. Dijon, France (Texte intégrale : 8p).
34. **Roux, F.**, J. Gasquez and X. Reboud. 2004. A spatially based concept to manage herbicide resistance: the mosaic strategy. **Proceedings of the XIIth International Conference on Weed Biology**. Dijon, France. pp 611-618.
35. Reboud, X., H. McKhann, N. Scarcelli, V. Le Corre, **F. Roux**, J.L. David, T. Bataillon, C. Camilleri, A. Bérard and D. Brunel. 2004. 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.
36. Bataillon, T., I. Bonnin, J. David, C. Lavigne, V. Le Corre, S. Poirier, E. Porcher, X. Reboud, J. Ronfort, **F. Roux**, S. Santoni, and A.C. Thuillet. 2003. Dynamique de la diversité génétique dans les populations naturelles et artificielles de plantes autogames. **Les Actes du BRG 4**, 187-189.

COMMUNICATIONS ORALES – CONGRÈS & COLLOQUES

A. Sur invitation

1. **Roux, F.** September 2012. Ecological genomics of adaptation in *Arabidopsis thaliana*: cases of flowering time and pathogen resistance. Learning from nature to protect crops. Wageningen, The Netherlands.
2. **Roux, F.**, M. Debieu, C. Huard-Chauveau, L. Perchepped, B. Braci, C. Glorieux, N. Faure, J. Bergelson and D. Roby. July 2012. A novel mechanism of quantitative and broad-spectrum resistance to bacteria exhibits balancing selection in *Arabidopsis*. XXIIIth International Conference on *Arabidopsis* Research. Vienna, Austria.
3. **Roux, F.** May 2011. Genotype-Environment interactions : genetic architecture and genetic bases of phenological traits in *Arabidopsis thaliana*: greenhouse vs. nature. Ecole Chercheur "Génomique des caractères complexes". La Colle-sur-Loup, France.
4. Brachi, B., N. Faure, M. Horton, E. Flahauw, A. Vazquez, M. Nordborg, J. Bergelson, J. Cuguen and **F. Roux**. January 2011. Linkage and association mapping of *Arabidopsis thaliana* flowering time in nature. XIIIth Rencontres de virologie végétale, Aussois, France.

5. Brachi, B., N. Faure, M. Horton, E. Flahauw, A. Vazquez, M. Nordborg, J. Bergelson, J. Cuguen and **F. Roux**. July 2010. Linkage and association mapping of *Arabidopsis thaliana* flowering time in nature. 18th annual meeting of the Society for Molecular Biology and Evolution - SMBE 2010. Lyon, France.
6. **Roux F.**, M. Horton and J. Bergelson. January 2010. Robust associations in *Arabidopsis thaliana* for GWA mapping of quantitative resistance to the pathogen *Pseudomonas viridiflava*. IXth Rencontres Plantes-Bactéries, Aussois, France.

B. Sur abstracts

7. Baron, E., J. Richirt, L. Amsellem and **F. Roux**. March 2012. Mise en évidence des bases génétiques de l'adaptation à la compétition chez *Arabidopsis thaliana*. Journée des doctorants IRePSE, Villeneuve d'Ascq, France.
8. Villoutreix, R., B. Brachi, N. Faure, N. Hautekeete, Y. Piquot, D. Roby, J. Cuguen, **F. Roux**. January 2011. Identification des facteurs écologiques exerçant une pression de sélection sur des traits phénologiques clés chez l'espèce annuelle *Arabidopsis thaliana*. Journée des doctorants IRePSE, Villeneuve d'Ascq, France.
9. **Roux, F.** and D. Roby. November 2010. Quantitative resistance to *Xanthomonas campestris* pv. *campestris*: identification of key genes. Réunion annuelle du réseau national INRA « Xanthomonas ». Toulouse, France.
10. Brachi, B., N. Faure, J. Cuguen and **F. Roux**. August 2008. Valeur adaptative et bases génétiques de la durée du cycle de vie et de la date de floraison chez *Arabidopsis thaliana*. XXX^{ème} Réunion annuelle du Groupe de Biologie et Génétique des Populations : Le Petit Pois Dérivé 2008. Rennes, France, Agrocampus Ouest - Rennes.
11. Brachi, B., N. Faure, J. Cuguen and **F. Roux**. July 2008. Adaptive value and genetic bases of duration of flowering period in *Arabidopsis thaliana*. XXth International Congress of Genetics. Germany, Berlin.
12. **Roux, F.**, M. Horton and J. Bergelson. July 2008. Genome-wide Association mapping in *Arabidopsis thaliana* to identify genes causing quantitative resistance to the pathogen *Pseudomonas viridiflava*. XXth International Congress of Genetics. Germany, Berlin.
13. **Roux, F.**, M. Horton and J. Bergelson. August 2007. Genome-wide Association mapping in *Arabidopsis thaliana* to identify genes causing quantitative resistance to the pathogen *Pseudomonas viridiflava*. XIth Congress of the European Society for Evolutionary Biology. Sweden, Uppsala.
14. **Roux, F.**, J. Gasquez and X. Reboud. December 2004. Comment prévenir et gérer la résistance aux herbicides chez les mauvaises herbes? 19^{ème} conférence du COLUMA – Journées internationales sur la Lutte contre les Mauvaises Herbes. Dijon, France.
15. **Roux, F.**, J. Gasquez and X. Reboud. September 2004. A spatially based concept to manage herbicide resistance: the mosaic strategy. XIIth International Conference on Weed Biology. Dijon, France.
16. **Roux, F.** and X. Reboud. June 2004. Dominance of the herbicide resistance in treated and untreated areas. IVth International Weed Science Congress. Rep. of South Africa, Durban.
17. **Roux, F.**, X. Reboud et V. Le Corre. June 2004. Relation entre évolution moléculaire et adaptation phénotypique chez *Arabidopsis thaliana* : le gène *LUMINIDEPENDENS* et la date de floraison. X^{ème} Forum Jeunes Chercheurs. Besançon, France, Faculté Médecine.
18. **Roux, F.**, J. Gasquez et X. Reboud. May 2004. Réponse de l'espèce modèle *Arabidopsis thaliana* à 22 herbicides inhibiteurs de l'acétolactate synthase. XXXIV^{ème} Congrès du Groupe Français des Pesticides. Dijon, France, INRA.
19. Reboud, X., **F. Roux** and J. Gasquez. May 2004. Grand nettoyage au rayon des herbicides. XXXIV^{ème} Congrès du Groupe Français des Pesticides. Dijon, France, INRA.
20. **Roux, F.**, C. Camilleri, S. Durand et V. Le Corre. August 2003. Relation entre évolution moléculaire et adaptation phénotypique chez *Arabidopsis thaliana* : le gène *LUMINIDEPENDENS* et la date de floraison. XXV^{ème} Réunion annuelle du Groupe de Biologie et Génétique des Populations : Le Petit Pois Dérivé 2003. Toulouse, France, Université Paul Sabatier Toulouse II.

21. Reboud, X., **F. Roux**, J. Doyle and S. Powles. July 2003. *Chlamydomonas* : a model organism for insights into herbicide resistance. XIXth International Congress of Genetics. Melbourne, Australia.
22. **Roux, F.** et X. Reboud. August 2000. Evolution expérimentale dans un système en métapopulation d'*Arabidopsis thaliana*. XXII^{ème} Réunion annuelle Groupe de Biologie et Génétique des Populations : Le Petit Pois Dérivé. Dijon, France, Université Bourgogne.

COMMUNICATIONS ORALES - SÉMINAIRES

1. **Roux, F.**, M. Debieu, C. Huard-Chauveau, L. Perchevied, B. Braci, C. Glorieux, N. Faure, J. Bergelson and D. Roby. December 2012. A novel mechanism of quantitative and broad-spectrum resistance to bacteria exhibits balancing selection in *Arabidopsis*. Séminaire du laboratoire GEV. Université de Lille1, France.
2. **Roux, F.** October 2012. Ecological genomics of adaptation in *Arabidopsis thaliana*: cases of flowering time and pathogen resistance. Institute of Plant Sciences, Bern, Switzerland.
3. **Roux, F.** April 2012. Identification of key genes underlying quantitative resistance to bacterial pathogens in *Arabidopsis thaliana* by GWA mapping. Brown University, Providence (USA).
4. **Roux, F.** March 2012. Ecological genomics of adaptation in *Arabidopsis thaliana*. Laboratoire des Interactions Plantes-Microorganismes (LIPM), CNRS-INRA, Toulouse.
5. Hancock, A.M., B. Brachi, N. Faure, M.W. Horton, L.B. Jarymowycz, F.G. Sperone, C. Toomajian, **F. Roux** and J. Bergelson. November 2011. Adaptation to climate across the *Arabidopsis thaliana* genome. Séminaire du laboratoire GEV. Université de Lille1, France.
6. Brachi, B., N. Faure, M. Horton, E. Flahauw, A. Vazquez, M. Nordborg, J. Bergelson, J. Cuguen and **F. Roux**. June 2011. Linkage and association mapping of *Arabidopsis thaliana* flowering time in nature. SupAgro Montpellier.
7. Brachi, B., N. Faure, M. Horton, E. Flahauw, A. Vazquez, M. Nordborg, J. Bergelson, J. Cuguen and **F. Roux**. November 2010. Linkage and association mapping of *Arabidopsis thaliana* flowering time in nature. INRA Bordeaux.
8. Brachi, B., N. Faure, M. Horton, E. Flahauw, A. Vazquez, M. Nordborg, J. Bergelson, J. Cuguen and **F. Roux**. March 2010. Linkage and association mapping of *Arabidopsis thaliana* flowering time in nature. Warwick HRI, University of Warwick, UK.
9. Brachi, B., N. Faure, M. Horton, E. Flahauw, A. Vazquez, M. Nordborg, J. Bergelson, J. Cuguen and **F. Roux**. February 2010. Linkage and association mapping of *Arabidopsis thaliana* flowering time in nature. Atelier thématique "Genome Wide Association mapping", UMR Stress abiotiques et différenciation des Végétaux cultivés, Université de Lille 1, Villeneuve d'Ascq.
10. **Roux, F.**, M. Horton and J. Bergelson. February 2010. Robust associations in *Arabidopsis thaliana* for GWA mapping of quantitative resistance to the pathogen *Pseudomonas viridiflava*. Atelier thématique "Genome Wide Association mapping", UMR Stress abiotiques et différenciation des Végétaux cultivés, Université de Lille 1, Villeneuve d'Ascq.
11. **Roux, F.**, M. Horton and J. Bergelson. February 2010. Le Genome-wide association mapping permet-il de faire autre chose que trouver des gènes (candidats) ? Le cas de la résistance quantitative à *Pseudomonas viridiflava* chez *Arabidopsis thaliana*. UMR ECOBIO, Université de Rennes 1, Rennes.
12. **Roux, F.**, M. Horton and J. Bergelson. January 2010. Le Genome-wide association mapping permet-il de faire autre chose que trouver des gènes (candidats) ? Le cas de la résistance quantitative à *Pseudomonas viridiflava* chez *Arabidopsis thaliana*. Ferme du Moulon, Gif-sur-Yvette.
13. **Roux, F.**, M. Horton and J. Bergelson. January 2009. Genome-wide association mapping chez *Arabidopsis thaliana* : identification des bases génétiques de la résistance quantitative au pathogène *Pseudomonas viridiflava*. INRA Station de Génétique et d'Amélioration des Plantes, Mauguio.
14. **Roux, F.**, M. Horton and J. Bergelson. September 2008. Genome-wide association mapping chez *Arabidopsis thaliana* : identification des bases génétiques de la résistance quantitative au

pathogène *Pseudomonas viridiflava*. Laboratoire d'Ecologie Alpine, Université Joseph Fourier, Grenoble.

15. **Roux, F.**, M. Horton and J. Bergelson. January 2008. Genome-wide Association mapping in *Arabidopsis thaliana* to identify genes causing quantitative resistance to the pathogen *Pseudomonas viridiflava*. SFB 680. Germany, Köln.
16. **Roux, F.**, M. Horton and J. Bergelson. December 2007. Genome-wide association mapping chez *Arabidopsis thaliana* : identification des bases génétiques de la résistance quantitative au pathogène *Pseudomonas viridiflava*. Laboratoire des Interactions Plantes-Microorganismes (LIPM), CNRS-INRA, Toulouse.
17. **Roux, F.** November 2006. The genetic basis of quantitative resistance in *Arabidopsis* to *Pseudomonas viridiflava*. Genetics of Model Organisms Research Meetings. Chicago, USA.
18. **Roux, F.** October 2004. La date de floraison chez *Arabidopsis thaliana*. Séminaire du laboratoire GEVP. Université de Lille1, France.
19. **Roux, F.**, J. Gasquez and X. Reboud. March 2004. The dominance of the herbicide resistance cost in several *Arabidopsis thaliana* mutants lines. Western Australian Herbicide Resistance Initiative, University of Western Australia, Perth, Australia.
20. **Roux, F.** December 2000. DNA polymorphism at flowering time genes in *Arabidopsis thaliana*. Molecular Weed Sciences Laboratory, Department of Crop Sciences, Champaign-Urbana, Illinois, USA.

POSTERS

1. **Baron, E.**, J. Richirt, L. Amsellem and **F. Roux**. July 2012. Identification of genomic bases associated with adaptation to competition in *Arabidopsis thaliana*. XXIIIth International Conference on Arabidopsis Research. Vienna, Austria.
2. **Schurdi-Levraud, V.**, P. Cosson, Z. Nehr, M. Caballero, **F. Roux** and F. Revers. April 2012. Adaptation to biotic stress: towards genome-wide association genetics to unravel the genetic basis of responses to Potyviruses in *Arabidopsis thaliana*. Conférences Jacques Monod: Theoretical and empirical advances in evolutionary genomics. France, Roscoff.
3. **Brachi, B.**, Z. Qingrun, N. Faure, C. Glorieux, A. Vazquez, H. Guillaume, M. Nordborg, J. Bergelson and **F. Roux**. June 2011. Combining Nested Association Mapping and correlated genome associations to a quantitative trait network to unravel the genetic basis of fitness in *Arabidopsis thaliana*. XXIIth International Conference on Arabidopsis Research. Madison, USA.
4. **Faure, N.**, B. Brachi, R. Villoutreix, Z. Qingrun, A. Darracq, C. Camilleri, M. Simon, J. Cuguen, J. Bergelson, M. Nordborg and **F. Roux**. June 2011. Combining Nested Association Mapping and ecology – phenotype relationships to identify ecological and evolutionary forces acting on vegetative growth in *Arabidopsis thaliana*. XXIIth International Conference on Arabidopsis Research. Madison, USA.
5. **R. Villoutreix, B. Brachi**, N. Faure, N. Hautekeete, Y. Piquot, D. Roby, J. Cuguen and **F. Roux**. June 2011. The scale of adaptation in *Arabidopsis thaliana*: identifying the ecological factors that act on phenology. XXIIth International Conference on Arabidopsis Research. Madison, USA.
6. **Debieu, M.**, A. Genissel, C. Glorieux, N. Faure, C. Chauveau, J. Bergelson, D. Roby and **F. Roux**. June 2011. Identification of key genes underlying quantitative resistance to *Xanthomonas campestris* in *Arabidopsis thaliana* by Genome Wide Association mapping. XXIIth International Conference on Arabidopsis Research. Madison, USA.
7. **Schurdi-Levraud, V.**, P. Cosson, Z. Nehr, M. Debieu, M. Caballero, **F. Roux** and F. Revers. June 2011. Towards genome-wide association genetics to identify loci involved in responses to Potyviruses in plants. XXIIth International Conference on Arabidopsis Research. Madison, USA.
8. **Cosson, P.**, L. Sofer, H. Le, V. Schurdi-Levraud, **F. Roux**, O. Le Gall, T. Candresse and F. Revers. January 2011. Variabilité génétique naturelle des genes de résistance RTM chez *Arabidopsis thaliana*. XIIIth Rencontres de virologie végétale, Aussois, France.

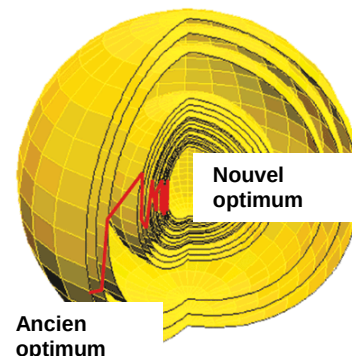
9. Schurdi-Levraud, V., Z. Nehr, P. Cosson, **F. Roux**, C. Caranta and F. Revers. January 2011. Genome-wide association genetics to identify loci involved in responses to Potyviruses in plants. XIIIth Rencontres de virologie végétale, Aussois, France.
10. Schurdi-Levraud, V., Z. Nehr, P. Cosson, **F. Roux**, C. Caranta and F. Revers. October 2010. Genome-wide association genetics to identify loci involved in responses to Potyviruses in plants. V^{ème} colloque du réseau Evolution Virale. France, Montpellier.
11. Brachi, B., N. Faure, M. Horton, J. Bergelson, J. Cuguen and **F. Roux**. July 2009. Combining Genome Wide Association mapping and QTL mapping in *Arabidopsis thaliana* for detecting genes underlying flowering time natural variation measured in ecologically realistic conditions. XXth International Conference on Arabidopsis Research. Edinburgh, Scotland.
12. **Roux, F.**, M. Horton and J. Bergelson. July 2009. Robust associations in *Arabidopsis thaliana* for GWA mapping of quantitative resistance to the pathogen *Pseudomonas viridiflava*. XXth International Conference on Arabidopsis Research. Edinburgh, Scotland.
13. Vilhjalmsson, B.J., S. Atwell, G. Willems, Y. Huang, A. Tarone, R. Jiang, C. Toomajian, T. Hu, D. Meng, Y. Li, X. Zhang, R.M. Clarck, M.J. Aranzana, K. Zhao, C.-C. Lee, H. Jung, B. Brachi, N. Faure, I. Baxter, D.E. Salt, J. Jones, **F. Roux**, M. Horton, C. Shindo, C. Lister, A. Platt, C. Dean, J.R. Ecker, D. Weigel, P. Marjoram, J. Borevitz, J. Bergelson and M. Nordborg. July 2008. Genome-Wide Association study in *Arabidopsis thaliana*. XIXth International Conference on Arabidopsis Research. Montreal, Canada.
14. **Roux, F.**, M. Horton and J. Bergelson. May 2007. Genome-Wide Association mapping in *Arabidopsis thaliana* to identify genes underlying quantitative resistance to the pathogen *Pseudomonas viridiflava*. Conférences Jacques Monod: Evolutionary Genomics. France, Roscoff.
15. **Roux, F.** and X. Reboud. June 2005. How to measure a cost of adaptation? XVIth International Conference on Arabidopsis Research – American Genetic Association “Plant Evolution: Genes and Phenotypes”. USA, Madison.
16. Reboud, X. et **F. Roux**. December 2004. Evolution de la résistance aux herbicides : apport du génotypage haut-débit. ACI ‘Séquençage végétal et Post-séquençage’. France, Paris, INRA.
17. **Roux, F.**, J. Gasquez and X. Reboud. June 2004. The dominance of the herbicide resistance cost in several *Arabidopsis thaliana* mutant lines. IVth International Weed Science Congress. Republic of South Africa, Durban.
18. **Roux, F.**, J. Gasquez et X. Reboud. May 2004. Gestion et prévention de la résistance aux herbicides: utilisation d'*Arabidopsis thaliana* comme modèle expérimental. XXXIV^{ème} Congrès du Groupe Français des Pesticides. France, Dijon, INRA.
19. **Roux, F.**, J. Gasquez and X. Reboud. August 2003. The dominance of the herbicide resistance cost in several *Arabidopsis thaliana* mutant lines. IXth Congress of the European Society for Evolutionary Biology. United Kingdom, Leeds.
20. Le Corre, V., L. Jacquens and **F. Roux**. August 2003. Molecular evolution at the within and between population levels in *Arabidopsis thaliana*: the *FRIGIDA* gene. IXth Congress of the European Society for Evolutionary Biology. United Kingdom, Leeds.
21. **Roux, F.**, J. Gasquez et X. Reboud. June 2003. Gestion et prévention de la résistance aux herbicides: utilisation d'*Arabidopsis thaliana* comme modèle expérimental. IX^{ème} Forum des jeunes chercheurs 2003. France, Dijon, Faculté de Médecine.

Document de synthèse

Génomique écologique de l'adaptation chez *Arabidopsis thaliana*

Box 1. Théorie génétique de l'adaptation.

La théorie récente de l'adaptation tire son origine du modèle géométrique de Fisher (Orr 1998, Barton and Keightley 2002, Orr 2005). Dans ce modèle, chaque organisme est un point dans un espace à n-dimensions, où chaque dimension représente un trait phénotypique. Un point donné dans cet espace phénotypique représente un optimum de fitness. Un changement environnemental induit un nouvel optimum qui se trouve éloignée de la population. Le modèle Fisher - Orr décrit l'ensemble du chemin adaptatif parcouru par une population vers ce nouvel optimum. Bien que la



variation génétique initialement présente dans une population puisse être une source d'allèles bénéfiques, le modèle Fisher – Orr considère l'adaptation à partir de nouvelles mutations. Le nouvel optimum phénotypique est supposé fixe durant l'adaptation et la fitness est une fonction continue du phénotype. Les principales prédictions du modèle sont :

- à chaque pas parcouru sur le chemin adaptatif correspond la fixation d'une mutation bénéfique. La réduction proportionnelle de la distance à l'optimum est similaire entre les pas successifs.
- l'effet de la première mutation fixée est attendu très important (de grands sauts adaptatifs sont ainsi attendus au début du chemin adaptatif)
- les effets bénéfiques fixés s'approchent d'une séquence géométrique avec les mutations à effet large fixées les premières et celles à effet plus faible à la fin du chemin adaptatif.

Ce modèle a été par la suite modifié pour tenir compte de la base moléculaire de la variation phénotypique (Orr 2002). La théorie de l'adaptation des séquences ADN ou « modèle de l'espace mutationnel » est basée sur les mêmes suppositions que la théorie de l'adaptation phénotypique, mais admet aussi une forte sélection couplée à un faible taux de mutation, de telle manière qu'une population est essentiellement fixée pour une séquence unique à n'importe quel moment. La probabilité qu'une séquence mutante bénéfique particulière est la prochaine séquence substituée dans le chemin adaptatif est proportionnelle à l'effet de cette séquence sur la fitness. Les prédictions de ce modèle sont très similaires à ceux du modèle phénotypique, la prédiction principale étant que la distribution des effets fixés sur la fitness durant l'adaptation est une exponentielle décroissante.

Quand l'adaptation est basée sur de la variation génétique initialement présente dans la population, la probabilité de fixation d'un allèle dépend de ses effets délétères et bénéfiques avant et après le changement environnemental, respectivement (Hermisson and Pennings 2005). Si un allèle bénéfique est neutre ou faiblement délétère avant le changement environnemental, sa probabilité de fixation est faiblement dépendante de son coefficient de sélection après le changement environnemental. La probabilité de fixation d'un allèle va notamment dépendre de sa fréquence initiale avant le changement environnemental. Ainsi, l'adaptation à partir de variation génétique standard pourrait mener à la sélection d'allèles avec de plus petits effets que l'adaptation à partir de nouvelles mutations.

Mieux comprendre l'adaptation des traits complexes dans la nature

Depuis Darwin (1859), le paradigme de l'évolution sous l'influence majeure de la **sélection naturelle** est devenu le modèle dominant de la biologie de l'évolution. Lorsqu'une population est confrontée à un changement de l'environnement, le phénotype moyen de cette population se trouve éloigné du phénotype optimal qui permettrait de maximiser la valeur sélective des individus dans ce nouvel environnement. La sélection naturelle favorise les phénotypes ayant les caractéristiques répondant au mieux aux critères imposés par les contraintes de l'environnement (Tonsor *et al.* 2005). Face à une pression de sélection donnée, une caractéristique phénotypique devient alors prévalente dans une population à cause de l'avantage sélectif que lui confère cette caractéristique par rapport aux autres phénotypes. L'adaptation correspond donc à un trait ou un ensemble de traits qui augmente la valeur sélective (fitness) des individus le(s) portant. L'adaptation peut aussi être considérée comme la résultante d'un processus selon lequel le phénotype moyen d'une population va, au fil des générations, tendre à rejoindre un optimum phénotypique par fixations successives d'allèles avantageux aux loci déterminant le phénotype. Ce processus, initialement formalisé par Fisher (1930), est appelé **la marche adaptative** (Box 1 ; Orr 2005). L'adaptation est ainsi considérée comme un processus en deux étapes : (1) la mutation est à l'origine des allèles ayant différents effets sur la fitness, et (2) les allèles qui améliorent la fitness augmentent en fréquence par sélection naturelle. De nombreuses études théoriques ont porté sur la deuxième étape de ce processus, incluant le calcul de la probabilité de fixation d'une nouvelle mutation favorable (Haldane 1927) et son taux d'accroissement (Haldane 1924).

Les formalisations mathématiques de la marche adaptative se sont généralement focalisées sur l'étude de l'adaptation au sein d'une population théorique. Toutefois, il ne faut pas oublier que les espèces évoluent dans un environnement hétérogène à la fois dans l'espace et dans le temps et qui exerce des pressions de sélection multiples sur les organismes, de type abiotique (climat, ressources,...) ou biotique (compétition, prédation, attaque de pathogènes...). Dans l'aire de répartition d'une espèce, l'optimum phénotypique peut donc varier selon les conditions environnementales locales, et ainsi mener à des phénomènes de sélection divergente entre les populations (Haldane 1948, Kawecki and Ebert 2004). Il en résulte des **adaptations locales**, qui sont à l'origine d'une part de la variation phénotypique intra-spécifique. Cependant, d'autres processus peuvent également contribuer à la variation intra-

Box 2. Questions soulevées lors de mon parcours de recherche. Différentes disciplines étant utilisées pour étudier l'adaptation, les questions énoncées ci-dessous ne sont pas totalement indépendantes les unes des autres.

1. EVOLUTION

- Quelle est la part adaptative de la variation naturelle du trait étudié ?
- Quel type de sélection (directionnelle, stabilisante, disruptive...) agit sur le trait étudié ?
 - Le type et l'intensité de la sélection dépendent-ils de la population considérée ?
- L'adaptation est-elle basée sur des variants génétiques déjà présents dans la population ou bien à partir de nouvelles mutations ?
- Quels sont les patterns de sélection des gènes liés à la variation phénotypique naturelle ?
 - Pourquoi tel gène a été recruté par la sélection naturelle et pas un autre ?
 - L'évolution génétique est-elle prévisible ?
- Quelle est l'importance de la variation naturelle des normes de réaction dans les capacités d'adaptation ?
 - Les gènes de plasticité sont-ils plus sous sélection que les gènes liés à l'adaptation locale ?

2. ECOLOGIE

- Quels sont les agents sélectifs du trait étudié ?
- Quelles sont les échelles spatiale et temporelle auxquelles l'adaptation s'est effectuée ?
 - La taille du grain de l'environnement est-elle la même suivant les agents sélectifs ?
 - Dans un environnement hétérogène, quel est l'effet du grain ou de plusieurs grains superposés (1) sur la distribution des optima phénotypiques dans un paysage adaptatif et (2) sur la dynamique d'une mutation adaptative ?

3. GENETIQUE

- Quelle est l'architecture génétique (nombre et effet des QTLs) associée à la variation naturelle du trait étudié ?
 - L'architecture génétique est-elle dépendante du trait étudié ?
 - Si oui, existe-t-il une relation entre l'architecture génétique et le type de sélection agissant sur le trait ?
 - L'architecture génétique pour un trait étudié dépend-elle de l'environnement où ce trait a été phénotypé ?
- Quelle est la part relative de l'épigénétique dans la variation naturelle phénotypique ?
- Quelle est l'identité moléculaire sous-jacente aux QTLs ?
 - Quel type de polymorphisme (cis-regulation, mutations dans régions codantes, méthylation...) est associé à la variation phénotypique naturelle ?
 - Existe-il un lien entre type de mutation (ponctuelle, indels, duplication...) et effet phénotypique ?

spécifique. C'est le cas notamment de l'équilibre mutation/dérive (Kimura 1955), ou de la sélection indirecte, c'est-à-dire lorsque la sélection agit indirectement sur un trait via sa corrélation génétique avec celui sous sélection (Lande and Arnold 1983).

Afin de mieux comprendre les mécanismes sous-jacents à l'adaptation de traits phénotypiques dans les populations naturelles, deux approches sont donc nécessaires. La première nécessite de mettre en évidence la part de la variation phénotypique qui est adaptative, et à quelle échelle géographique se réalise l'adaptation, par une étude approfondie des patrons de variations des traits phénotypiques et des facteurs écologiques auxquels les populations sont confrontées. La seconde doit s'atteler à étudier l'architecture génétique et les bases génétiques de la variation phénotypique adaptative dans la nature, afin de mieux comprendre les marches adaptatives suivies par les populations naturelles. Combiner ces deux aspects de l'étude de l'adaptation correspond au champ de la **génomique écologique** qui cherche à comprendre les mécanismes génétiques sous-jacents aux réponses des populations à leurs environnements naturels (Ungerer *et al.* 2008). La génomique écologique peut aussi être définie comme l'intégration de l'écologie dans les analyses d'association 'variation génétique - variation phénotypique adaptative', permettant de retracer la trajectoire évolutive des caractères adaptatifs dans les populations naturelles (Bergelson and Roux 2010). La génomique écologique est donc une discipline interdisciplinaire visant à réaliser une synthèse entre génétique, évolution et écologie.

L'objectif de mes activités de recherches est d'établir un lien entre la biologie moléculaire et l'écologie évolutive pour étudier la réponse adaptative à la sélection naturelle au sein des espèces végétales. Durant mes activités de recherche, j'ai donc toujours eu la volonté d'aborder l'étude de l'adaptation au niveau intra-spécifique par 3 disciplines complémentaires (chacune avec plusieurs méthodes d'investigation, Figure 1) en m'intéressant aux points suivants (voir Box 2 pour une liste plus exhaustive de questions):

- **Ecologie** : identification des agents sélectifs du trait étudié, détermination des échelles spatiale et temporelle auxquelles l'adaptation s'est effectuée.
- **Génétique** : nombre et effet des gènes impliqués dans la variation naturelle phénotypique, identification des bases génétiques et validation fonctionnelle.
- **Evolution** : estimation de la valeur adaptative du trait étudié, reconstruction de la marche adaptive vers un optimum phénotypique, étude des patterns de sélection des gènes liés à la variation naturelle phénotypique.

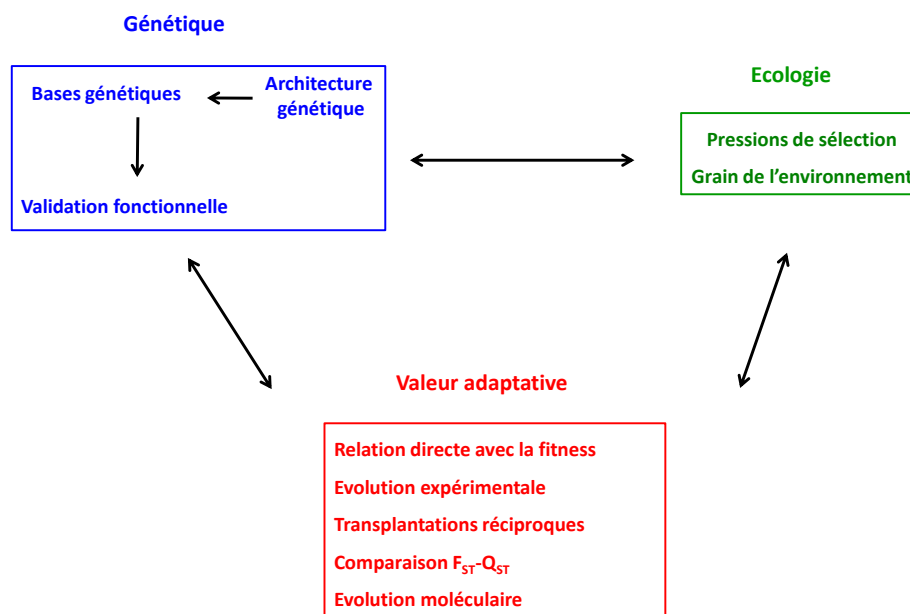


Figure 1. Génomique écologique de l'adaptation.

Le modèle biologique commun à mes travaux est *Arabidopsis thaliana*, espèce modèle en génétique végétale, mais qui apparaît aussi depuis quelques années comme une espèce modèle en écologie évolutive (Gaut 2012). En effet, bien qu'étant trop souvent considérée comme une espèce de « laboratoire », *A. thaliana* est avant tout une espèce sauvage dont l'évolution des traits phénotypiques se réalise dans des environnements naturels et contrastés. Je me suis donc focalisé sur cette espèce modèle pour les nombreux avantages qu'elle présente et que je détaillerai dans un premier temps.

Au fil de mon parcours professionnel, je me suis intéressé à des traits adaptatifs de plus en plus complexes et pour lesquels peu ou pas de réponses aux questions citées ci-dessus sont disponibles (Tableau 1). Pour chacun de ces traits, j'ai donc essayé de me focaliser sur des questions peu abordées dans la littérature correspondante. Ces traits constitueront le fil conducteur de la deuxième partie de ce document.

Tableau 1. Traits adaptatifs étudiés. 'X' : connaissances acquises, ' ? ' : peu ou pas de connaissances.

Période	Trait adaptatif	Ecologie	Génétique	Evolution
Thèse	Résistance aux herbicides	X	X	?
Post-doc et CR CNRS	Résistance aux bactéries pathogènes	X	?	?
ATER et CR CNRS	Phénologie	?	?	?

L'étude individuelle de ces traits adaptatifs m'a tout naturellement amené à reconsidérer la marche adaptative dans un milieu hétérogène en ne travaillant plus sur un seul trait dans un environnement unique, mais sur un ensemble de traits dans plusieurs environnements ; et ainsi à prendre en compte la variation naturelle des normes de réactions dans l'étude des capacités d'adaptation. Cette troisième partie de ce document correspondra plus à un état des lieux de mes projets de recherche en cours, qu'à une série de résultats déjà publiés.

Finalement, dans la quatrième et dernière partie de ce document, je présenterai le projet de recherche que je souhaite mener au cours des prochaines années. Cette fois-ci, mes recherches en génomique écologique ne porteront pas uniquement sur *A. thaliana*, mais sur les communautés végétales.

I. *Arabidopsis thaliana* : une espèce modèle en génétique et en écologie

Bergelson and Roux 2010 Nature Reviews Genetics, Brachi et al. Molecular Ecology in revision.

Originnaire d'Eurasie, *A. thaliana* (Figure 2) est une Brassicacée annuelle présentant une répartition mondiale (Shindo *et al.* 2007). *A. thaliana* est décrite comme une espèce colonisatrice souvent trouvée dans des milieux pauvres ou perturbés, rarement en compétition avec d'autres espèces. A l'heure actuelle, *A. thaliana* demeure l'espèce modèle principale de la génomique chez les plantes. Elle est très utilisée dans le cadre de la génétique fonctionnelle en raison de sa facilité de culture, son cycle de vie court (en conditions de serre) ainsi que sa capacité à s'autoféconder, permettant ainsi de maintenir des lignées homozygotes et de les phénotyper un nombre de fois infini (Weigel &



Figure 2. *Arabidopsis thaliana* cultivée en serre. (Photographie par Benjamin Brachi).

Nordborg 2005). Ces caractéristiques, combinées à la petite taille de son génome (5 chromosomes, ~125 Mb), ont conduit au séquençage complet de son génome, le premier chez les plantes supérieures, achevé en 2000 (accession Col-0, The Arabidopsis Genome Initiative 2000). La séquence de référence a permis d'annoter tous les gènes. Actuellement, la base de

données TAIR 9 (The Arabidopsis Information Resource ; <http://www.arabidopsis.org/>) compte 33,518 gènes dont 27,379 codant pour des protéines (Swarbreck *et al.* 2008).

Depuis plus d'une dizaine d'années, *A. thaliana* apparaît aussi comme une espèce modèle en écologie évolutive (Gaut 2012). Sur son aire de répartition mondiale, *A. thaliana* est présente dans une grande diversité d'habitats aussi bien d'un point de vue abiotique que biotique (Jakob *et al.* 2002, Mitchell-Olds and Schmitt 2006, Shindo *et al.* 2007). Cette diversité d'habitats s'observe même à une échelle géographique de l'ordre de quelques kilomètres (Figure 3).



Figure 3. Diversité des habitats d'*A. thaliana*. *A. thaliana* peut vivre dans des habitats contrastés, même à une échelle géographique fine. Sur ces photos, des populations naturelles observées en France représentent différents niveaux de perturbation, depuis les quais de gare et les champs cultivés aux prairies permanentes et garrigues. a. bord de route (Bretagne) ; b. prairie permanente (Bourgogne) ; c. quai de gare (Bourgogne) ; d. garrigue (Languedoc) ; e. champ de colza (Bretagne) ; f. talus (Nord). (photographies a, d, e et f par Benjamin Brachi et Nathalie Faure ; photographies b et c par Fabrice Roux).

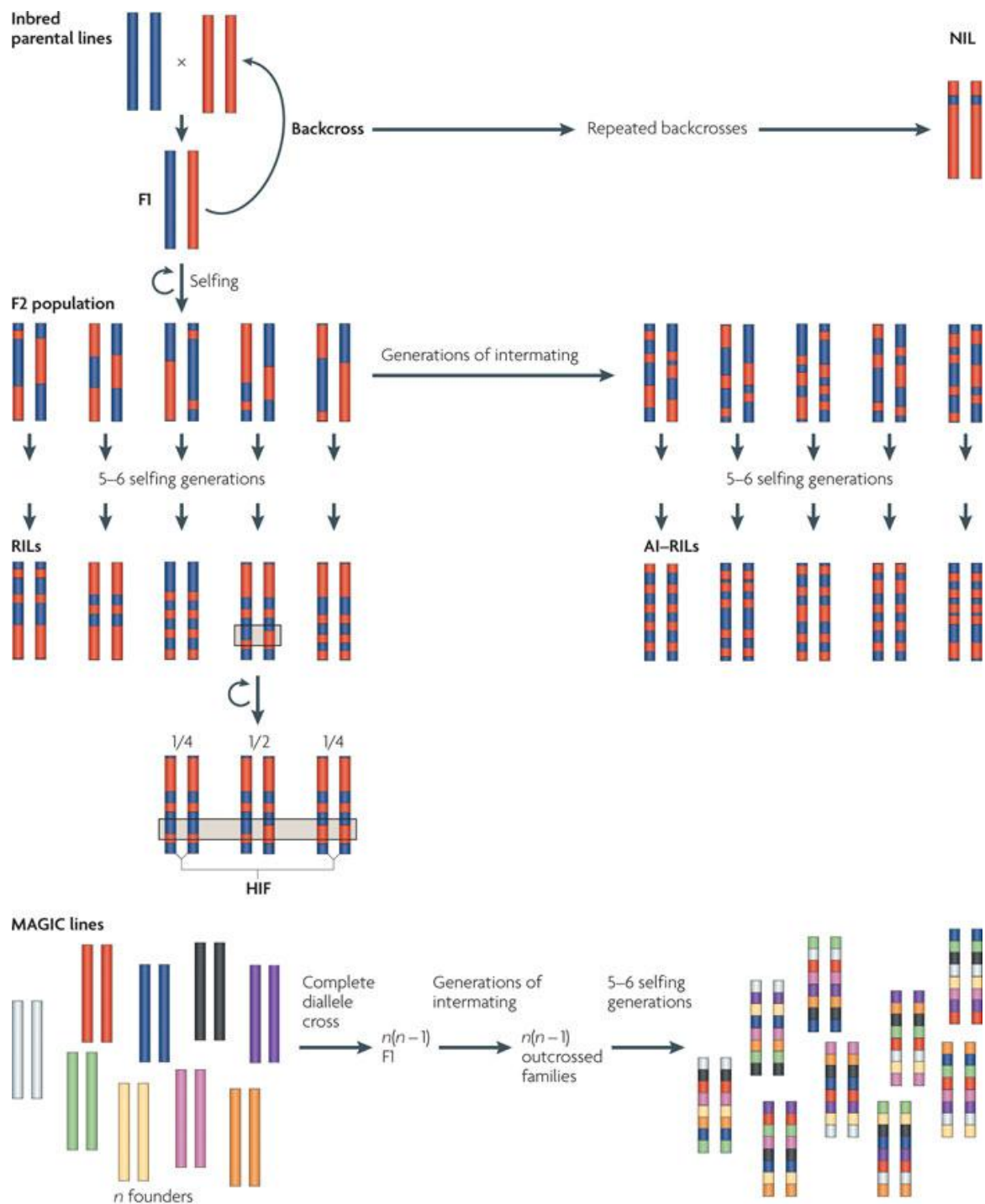


Figure 4. Différentes populations de cartographie QTLs chez *A. thaliana* (Bergelson and Roux 2010). RILs, Recombinant Inbred Lines. AI-RILs, Advanced Intercross – Recombinant Inbred Lines. NIL, Near-Isogenic Line. HIF, Heterogeneous Inbred Family. MAGIC lines, Multiparent Advanced Generation Inter-Cross lines. La résolution de la cartographie et les diversité génétique des populations de cartographie QTLs dépendent du nombre (1) de lignées génétiques fondatrices, (2) de générations d'inter-croisements, et (3) de générations d'autogamie.

L'effort commun de plusieurs laboratoires travaillant sur *A. thaliana* a permis de créer des ressources génétiques importantes. En rendant ces ressources publiques et disponibles *via* des centres de ressources, la communauté scientifique travaillant sur *A. thaliana* en a fait un modèle unique pour étudier la variation phénotypique et ses bases génétiques :

- **Lignées artificielles** : des ressources génétiques importantes pour détecter les bases génétiques de la variation phénotypique artificielle sont publiquement disponibles (altération ou perturbation aléatoire des gènes : mutagenèse EMS, T-DNA mutants ; altération ou perturbation spécifique de gènes : gene silencing par amiRNA) (Alonso and Ecker 2006, O'Malley and Ecker 2010), rendant la validation fonctionnelle par complémentation quantitative ou silencing quantitatif beaucoup plus aisée.
- **Population expérimentales** (Figure 4): de la population F2 classique jusqu'aux lignées Multiparent Advanced Generation Inter-Cross (MAGIC) les plus récemment développées, le but commun de ces populations expérimentales est de cartographier les QTLs associés à la variation phénotypique naturelle présente entre 2 ou plusieurs accessions naturelles. Etant donné la complète homozygotie des lignées recombinantes consanguines (Recombinant Inbred Lines), ces populations RILs permettent d'avoir des réplicats d'un même génotype au sein d'une expérience et/ou entre plusieurs conditions environnementales ; et demeurent donc le type de population le plus couramment utilisé chez *A. thaliana*. Plus de 60 familles RILs ont été développées chez *A. thaliana*.
- **Accessions naturelles** : plusieurs milliers d'accessions mondiales sont publiquement disponibles. Depuis quelques années, un effort considérable a été mis en place pour caractériser génétiquement ces accessions naturelles. Le séquençage complet du génome de 19 accessions mondiales a permis d'estimer que le déséquilibre de liaison moyen chez *A. thaliana* était en moyenne de 10kb (Kim *et al.* 2007), permettant ainsi d'envisager une résolution très fine pour cartographier le long du génome les régions génomiques associées à la variation naturelle phénotypique (méthode du Genome Wide Association mapping). En se basant sur ces données de séquençage complet de génomes, des jeux de 149 à 384 SNPs polymorphes à l'échelle mondiale ont été mis en place afin d'obtenir une caractérisation génétique rapide des accessions (>8000 accessions génotypées à ce jour ; Platt *et al.* 2010, Bomblies *et al.* 2010, Brachi *et al.* in revision). Afin d'identifier les régions génomiques associées à la variation naturelle phénotypique par GWA mapping, une puce de génotypage de 250k SNPs (soit en moyenne 1 SNP tous les 500bp) a été créée et utilisée pour génotyper >1300

accessions mondiales (Horton *et al.* 2012). A l'heure actuelle, afin d'obtenir une meilleure précision dans la localisation des bases génétiques associées à la variation naturelle phénotypique (*i.e.* obtenir un accès direct aux mutations causales), le séquençage complet du génome d'environ 1200 accessions naturelles d'*A. thaliana* est en cours *via* un consortium international (<http://www.1001genomes.org/>). Ces ~1200 accessions sont aussi en cours de caractérisation pour les niveaux de méthylation (base par base) sur l'ensemble du génome. Cette caractérisation épigénétique permettra prochainement de réaliser des analyses de GWA mapping en utilisant les niveaux de méthylation comme marqueurs polymorphes.

- **Populations naturelles** : 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. Pourtant, des populations génétiquement et phénotypiquement polymorphes ont été mises en évidence (Le Corre 2005) et des taux d'allogamie allant parfois jusqu'à 20 % ont été récemment décrits dans des populations naturelles d'*A. thaliana* (Bomblies *et al.* 2010, Platt *et al.* 2010). Il est donc possible, en utilisant le modèle biologique *A. thaliana*, de s'intéresser aux patrons d'évolution de traits phénotypiques et de leurs bases génétiques à l'échelle de l'espèce, mais aussi à une échelle très locale.

Depuis le début de ma thèse, ma contribution à ces ressources génétiques a été d'échantillonner plus de 200 accessions françaises, européennes ou africaines, *via* des projets scientifiques ou bien de simples vacances. Parmi ces deux cents accessions, 69 ont été génotypées pour 250k SNPs. Plus récemment, des projets développés au sein de mon équipe ont amené à la récolte de graines de plus de 1200 plantes issues de 51 populations françaises. Suivant la question biologique posée, la caractérisation génétique de ces lignées a été faite (ou est en cours) soit par un génotypage avec 149 SNPs ($n = 800$), soit par séquençage individuel complet du génome ($n = 179$ dont 34 par l'intermédiaire du projet 1001 genomes), soit par séquençage complet du génome de bulk d'ADN (chaque bulk d'ADN correspondant à une population, $n = 49$). Nous reviendrons sur l'utilisation de ces lignées dans la troisième partie de ce document.

II. Etude de traits adaptatifs de plus en plus complexes

A. La résistance aux herbicides : un cas d'école

Neve et al. New Phytologist 2009

Assimilée à une pression de sélection, l'utilisation de pesticides a inévitablement entraîné une réponse adaptative des espèces nuisibles, à savoir un développement de mécanismes de résistance aux pesticides. La résistance à un pesticide est l'aptitude héréditaire d'un individu à survivre et à se reproduire à une dose de pesticide qui serait normalement létale pour un individu de type sauvage (Friesen *et al.* 2000). L'apparition et la propagation de la résistance aux herbicides chez les adventices augmentent de manière exponentielle à l'heure actuelle, suivant les tendances déjà observées chez les insectes et les champignons. Le problème de la résistance aux herbicides est donc majeur et pourrait être d'un impact économique et/ou écologique considérable. Il est alors important que des stratégies de prévention et de gestion de la résistance chez les adventices soient développées et enfin adoptées afin de limiter ce phénomène.

Bien qu'elle pose avant tout un problème économique pour les agriculteurs, la résistance aux herbicides s'avère être aussi un cas extrême d'adaptation à une pression de sélection. L'architecture génétique est souvent simple (un seul gène impliqué dans la résistance), les pressions de sélection sous la dépendance de l'agriculteur sont connues et restent manipulables, les effets phénotypiques forts... Ainsi, pour répondre à la question « Comment prévenir et gérer la résistance aux herbicides ? », j'ai tiré profit des principes et concepts développés théoriquement pour comprendre les trajectoires évolutives d'un trait adaptatif, ceci afin de développer des stratégies de non-adaptation de la résistance. Dans l'optique de développer des stratégies efficaces de lutte contre la résistance, j'ai privilégié cette approche par deux points :

- la recherche de paramètres pertinents qui m'a conduit à étudier **la balance coût-bénéfice** de huit mutations à l'origine de résistance aux herbicides.
- **l'impact de l'hétérogénéité de l'environnement** (mode et alternance de zones traitées ou non traitées) sur l'évolution d'une résistance herbicide.

1. La balance coût-bénéfice

Roux et al. Genetics 2004, Roux et al. Evolution 2005, Roux et al. 2005 Genetical Research, Roux et al. Genetics 2005, Roux et al. Pest Management Science 2005, Roux et al. Weed Research 2005, Giancola et al. Theoretical and Applied Genetics

2006, Roux *et al.* *Genetics* 2006, Paris *et al.* *Heredity* 2008, Vila-Aiub *et al.* *Heredity* 2011.

Comme toute autre adaptation, l'évolution d'une résistance résulte de l'interaction entre 4 forces, à savoir la sélection avec une balance coût (en absence d'herbicide) – bénéfice (en présence d'herbicide), la migration entre les habitats favorables et défavorables à la résistance, la dérive génétique et la mutation. La très bonne caractérisation moléculaire et physiologique de 8 mutations liées à la résistance aux herbicides chez *A. thaliana* m'a permis d'étudier plus précisément les avantages et coûts liés aux mutations qui confèrent la résistance, permettant par la même occasion de réfuter quelques 'vérités' largement admises (mais rarement vérifiées) chez les mauvaises herbes :

- **avantage de la résistance en présence d'herbicide** : le facteur de résistance associée à une mutation ainsi que la dominance de cette mutation en présence d'herbicide peuvent varier suivant la dose d'herbicide appliquée sur les plantes. La valeur sélective d'une mutation adaptative est donc dépendante de l'intensité de la pression de sélection exercée sur les plantes.
- **coût de la résistance sur la fitness** : pour la première fois chez une espèce, j'ai testé et comparé expérimentalement 3 méthodes d'estimation du coût d'une adaptation (Bourguet *et al.* 2004), dans le cadre de la résistance aux herbicides.
 - **comparaison directe** : comparaison des composantes de la fitness, à savoir la reproduction, la croissance et la phénologie entre des plantes homozygotes sensibles (SS) et des plantes homozygotes résistantes (RR).
 - **évolution des fréquences**: étude sur plusieurs générations de l'évolution de la fréquence d'un allèle de résistance dans des populations isolées non traitées avec des herbicides.
 - **cline artificiel**: étude de l'évolution de la fréquence d'un allèle de résistance dans des populations situées le long d'un transect entre des zones traitées et non traitées connectées entre elles par de la migration. J'ai tiré parti de la mise en place expérimental d'un cline le long d'un environnement hétérogène, à savoir une zone traitée avec un herbicide et une zone non traitée.

L'estimation du coût sur la fitness de plusieurs résistances aux herbicides par ces trois méthodes ont permis de mettre en évidence les avantages et inconvénients liés à chaque méthode. Bien que la comparaison de caractères entre des plantes SS et RR permette de connaître l'origine du coût (i.e. changements phénotypiques liés à une diminution de la

fitness), seule l'étude de l'évolution des fréquences alléliques au cours du temps (permise par des outils de génotypage haut-débit) est à même de prédire la valeur globale du coût sur la fitness (Tableau 2). Par l'étude de dynamiques micro-évolutives dans les méthodes « Evolution des fréquences » et « Cline artificiel », il a été mis en évidence que le contrôle expérimental d'une force évolutive telle que la migration permet de disséquer le poids relatif de deux autres forces évolutives à savoir la sélection et la dérive génétique (Tableau 2).

Tableau 2. Comparaison des trois méthodes d'estimation du coût.

	Comparaison directe	Evolution des fréquences	Cline artificiel
estimation réelle du coût ¹	non	oui	oui
origine du coût	oui	non	non
durée de l'expérience	courte	longue	courte
influence de la dérive génétique	-	oui	non
prise en compte de la fréquence dépendance	non	oui	oui
évaluation de la fréquence allélique	-	lourde	lourde
place en serre ou en champ	faible	importante	importante

¹ L'estimation du coût reste cependant fonction des conditions expérimentales.

A travers le semis de plus de 1,3 millions de graines de *A. thaliana* et plusieurs approches combinant analyses morphologique et/ou génétique sur différents pas de temps (évolution expérimentale), j'ai montré que les valeurs des paramètres sont très variables entre les résistances, suivant les herbicides utilisés (cas des facteurs de résistance et de la dominance en présence d'herbicide), et le contexte (cas du coût sur la fitness variable selon la méthode d'estimation employée, le fond génétique et l'environnement biotique et/ou abiotique). **Les valeurs de paramètres génétiques semblent donc imprédictibles car inféodées aux mutations qui confèrent la résistance.** C'est une conclusion qui sera décevante pour tous les tenants de la généralité des modèles mais pas très surprenantes pour ceux qui se sont frottés à la question.

2. Impact de l'hétérogénéité de l'environnement sur la dynamique d'une adaptation

Roux and Reboud Crop Protection 2007, Roux et al. Pest Management Science 2008

Un des autres principes et concepts développés théoriquement pour comprendre les trajectoires évolutives d'un trait adaptatif est **le grain de l'environnement** (Kassen 2002). Il a été démontré que le grain de l'environnement pouvait influencer la dynamique d'une mutation adaptative (Via and Lande 1985, Kassen 2002). Ce résultat dépend de l'interaction entre

l'intensité des flux géniques entre les habitats favorables et défavorables au gène adaptatif ; et l'intensité de la sélection et de la contre-sélection du gène adaptatif dans les habitats favorables et défavorables, respectivement. Un environnement avec un grain fin conduit à d'importants flux géniques entre les habitats favorables et défavorables, favorisant ainsi la diffusion du gène adaptatif à travers les grains de l'environnement ; et ceci d'autant plus que la contre-sélection du gène adaptatif dans les habitats défavorables est faible. A l'opposé, un environnement avec un grain grossier limite considérablement les flux géniques entre les habitats favorables et défavorables, favorisant ainsi l'adaptation locale du gène adaptatif dans les seuls habitats favorables (Slatkin 1973, Nagylaki 1975). Il existe donc un grain intermédiaire où le flux génique du gène adaptatif à partir de l'habitat favorable est plus fortement contre balancé par la contre-sélection du gène adaptatif dans l'habitat défavorable et par le flux génique à partir de l'habitat défavorable. Dans ce cas, le gène adaptatif ne disperse pas à travers à l'environnement et reste inféodé à l'habitat favorable où il est apparu (Caprio 2001, Ronce and Kirkpatrick 2001).

Bien que l'hétérogénéité dans l'espace n'ait jamais été abordée dans le cas de la résistance aux herbicides, j'ai mis en évidence par modélisation et simulations que l'alternance spatiale de deux herbicides pouvait s'avérer une piste potentiellement efficace dans le ralentissement de la diffusion de la résistance, à condition qu'une taille optimale des zones traitées soit adoptée (grain de l'environnement). Ce résultat qui peut paraître évident pour des généticiens des populations ou des écologistes constitue une nouvelle approche dans l'application des herbicides sur le terrain (Lenormand and Raymond 1998, Vacher *et al.* 2003) et rejoint les développements théoriques des modèles de gestion de l'évolution des résistances aux insecticides des plantes génétiquement modifiées (exemple: stratégies haute dose / refuge pour le Maïs Bt).

En intégrant dans les modèles les valeurs expérimentales de la balance coût-bénéfice obtenues chez *A. thaliana* ainsi que des traits d'histoire de vie (comme le régime de reproduction), j'ai montré que l'impact du grain de l'environnement sur la propagation d'une résistance dépend dans un premier temps de l'interaction entre le régime de reproduction et la dominance de l'allèle de résistance en présence d'herbicide, puis du coût de la résistance sur la fitness, et enfin de la dominance du coût.

3. Conclusion

Bien qu'aucune résistance aux herbicides n'ait été décrite dans des populations naturelles d'*A. thaliana*, l'approche simultanée par l'expérimentation avec un organisme modèle et la

modélisation des concepts développés pour prédire les trajectoires évolutives d'une adaptation m'a permis de décrire des paramètres génétiques peu étudiés chez les espèces adventices et/ou de montrer leur importance pour une meilleure prévention et gestion de la résistance aux adventices. En effet, par une approche couplant **modélisation et démarche expérimentale tirant bénéfice des avantages d'un organisme modèle tel que *A. thaliana***, je me suis employé à **utiliser des concepts théoriques développés en écologie et en évolution pour répondre à une problématique appliquée**. Considérant la résistance aux herbicides comme un cas extrême d'adaptation, j'ai ainsi pu montrer que (1) les valeurs de paramètres semblent imprédictibles car propres aux mutations qui confèrent la résistance, et (2) l'impact du grain de l'environnement sur la propagation d'une résistance dépend principalement de l'interaction entre les facteurs génétiques associés aux résistances et des traits d'histoire de vie associés aux espèces.

La validation expérimentale de stratégies développées par modélisation permettrait de faire passer les prédictions des modèles du stade de présomption au stade d'affirmation. L'utilisation d'organismes modèles ne peut que renforcer cette validité : (1) le cycle de vie généralement court des organismes modèles permet d'apprécier une dynamique évolutive de l'adaptation sur un nombre de générations élevé (Elena and Lenski 2003, Ungerer *et al.* 2003), (2) le contrôle et/ou l'estimation des forces évolutives dans une expérimentation permettrait une compréhension fine de l'évolution d'une résistance aux herbicides dans un environnement hétérogène où le grain peut aussi être contrôlé expérimentalement. L'évolution expérimentale constitue donc une opportunité difficilement remplaçable dans la prévention et la gestion de la résistance, et plus largement dans l'étude de l'adaptation.

B. La résistance aux bactéries pathogènes : qualitatif vs. quantitatif

Les bactéries pathogènes représentent de sévères menaces pour les populations végétales. Un des objectifs aussi bien d'un point de vue appliqué (*i.e.* sélection de variétés résistantes) que d'un point de vue fondamental (*i.e.* comprendre les trajectoires évolutives des populations en présence de pathogènes) est d'identifier les bases génétiques associées à la variation naturelle de la résistance des plantes aux bactéries pathogènes. A ce stade, je dois faire la distinction entre deux types de résistance. **Le premier type de résistance correspond à la résistance qualitative** où l'inoculation de l'agent pathogène entraîne une réponse binaire chez les plantes : résistance complète ou sensibilité complète (Figure 5a). Les bases génétiques de cette résistance qualitative sont de mieux en mieux connues chez *A. thaliana*,

avec notamment la description des gènes *R* (*R-genes* qui comprend la classe majoritaire des gènes NB-LRR *R* genes (nucleotide binding domain (NB) and a leucine rich repeat (LRR) domains)). Il faut toutefois noter que les concentrations initiales en bactéries utilisées pour mettre en évidence ce phénotype binaire n'existent pratiquement pas dans la nature. Durant mon séjour post-doctoral dans le laboratoire de Joy Bergelson (University of Chicago), je me suis donc intéressé aux bénéfices de la résistance qualitative en fonction de la concentration initiale en pathogènes.



Figure 5. Résistance qualitative vs. résistance quantitative. **a.** Résistance qualitative : Feuilles inoculées avec une solution bactérienne de *Pseudomonas syringae* DC3000 contenant le gène d'avirulence *avrPph3 avr* (10^8 bactéries/mL). Les feuilles inoculées d'une accession résistante et d'une accession sensible sont entourées en rouge. **b.** Résistance quantitative : Symptômes associés à l'attaque de la bactérie pathogène *Pseudomonas viridiflava* sur des plantes d'*A. thaliana* dans des populations naturelles.

Le deuxième type de résistance correspond à la résistance quantitative et est beaucoup plus prévalente dans les populations naturelles que la résistance qualitative (Poland *et al.* 2009). Dans le cas de la résistance quantitative, nous observons une distribution continue des symptômes (Figure 5b). A ce jour, seulement 5 gènes de résistance quantitative ont été clonés et ne concernent que la résistance à des champignons ou à des nématodes (Krattinger *et al.* 2009, Fu *et al.* 2009, Fukuoka *et al.* 2009, Cook *et al.* 2012, Liu *et al.* 2012). Aucun gène de résistance quantitative à des bactéries pathogènes n'a été à ce jour cloné. Par ailleurs, bien que le séquençage de nombreux gènes de résistance qualitative ont permis de mettre en évidence des patterns de sélection très diversifiés (du balayage sélectif récent à la sélection balancée sur plusieurs millions d'années ; Bakker *et al.* 2006), aucun pattern de sélection n'a été décrit pour un gène de résistance quantitative.

Projet initié lors de mon séjour post-doctoral, je me suis lancé dans l'identification des bases génétiques associées à la variation naturelle de la résistance quantitative à différentes souches du pathogène *Pseudomonas viridiflava* collectionnées dans des populations naturelles nord-américaines d'*A. thaliana*. A cette époque, la méthode de GWA mapping chez *A. thaliana* en était à ses prémices et commençait à être développée entre autres par le

laboratoire de Joy Bergelson. J'ai ainsi eu la chance d'être un des tous premiers à tester cette méthode sur un jeu de données phénotypiques de résistance quantitative.

De retour en France, en collaboration avec l'équipe de Dominique Roby (LIPM, CNRS-INRA Toulouse), j'ai monté un projet sur : 1) l'identification par GWA mapping des bases génétiques associées à la variation naturelle de la résistance quantitative à la bactérie pathogène la plus prévalente dans les populations nord-américaines d'*A. thaliana*, à savoir *Xanthomonas* sp., 2) la validation fonctionnelle de ces gènes de résistance quantitative, et 3) la détection de traces de sélection associées à ces gènes.

1. Valeur adaptative d'une résistance qualitative en fonction de la concentration initiale en pathogènes

Gao et al. New Phytologist 2010, Roux et al. Genetics 2010

Bien que l'intensité de la sélection imposée par des agents pathogènes varie fortement d'une population à l'autre, peu de données sont disponibles sur les normes de réaction de génotypes soumis à un gradient de pression de sélection. Dans une expérience, je me suis focalisé sur l'interaction 'gene-for-gene' entre le gène de résistance *R Rpm1* chez *A. thaliana* et le gène d'avirulence *avr AvrRpm1* chez la bactérie pathogène foliaire *Pseudomonas syringae*. Je me suis notamment intéressé à l'effet quantitatif d'un gradient d'infection sur la fitness des plantes résistantes et sensibles (*i.e.* présence ou absence du gène de résistance *Rpm1* ; Figure 6).

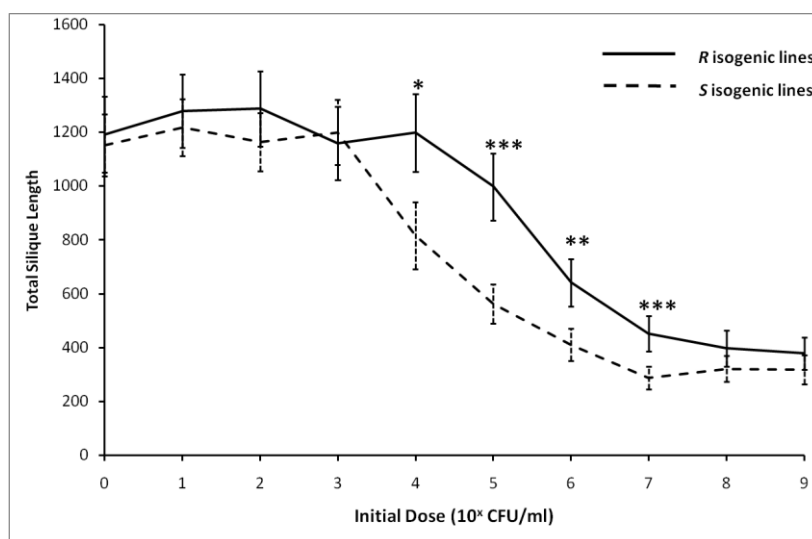


Figure 6. Effet du gène de résistance *Rpm1* sur la fitness des plantes, mesurée comme la longueur total des fruits (Total Silique Length exprimé en mm). CFU : Colony- Forming Units. R and S isogenic lines : lignées isogéniques ne différant que par la présence (R) ou l'absence (S) du gène *Rpm1*.

En utilisant à la fois des lignées transgéniques isogéniques ne différant que par le gène de résistance *Rpm1* et des accessions naturelles polymorphes pour ce même gène de résistance, j'ai démontré que les bénéfices d'une résistance qualitative étaient fonction de la concentration initiale en pathogènes (Figure 6) ; et pour la première fois chez *A. thaliana*, que des gènes de résistance qualitative pouvaient aussi être des gènes de tolérance à des bactéries pathogènes (*i.e.* capacité d'un individu à soutenir une infection sans une perte de fitness).

2. Identification des bases génétiques de la résistance quantitative aux bactéries pathogènes de la communauté foliaire d'*A. thaliana*

2.1. *Pseudomonas viridiflava*

Atwell et al. Nature 2010

J'ai phénotypé la résistance quantitative chez 192 accessions naturelles d'*A. thaliana* génotypées pour 214k SNPs, et ceci pour 5 souches de *P. viridiflava*, isolées à partir de populations naturelles nord-américaines d'*A. thaliana* (Figure 7).



Figure 7. Echelle de sévérité des symptômes liés à *Pseudomonas viridiflava*.

L'héritabilité du phénotype de résistance quantitative était proche de 0.6 quel que soit la souche de *P. viridiflava* considérée, démontrant une forte transmission du caractère de résistance quantitative d'une génération à l'autre. Pour chacune des souches de *P. viridiflava*, je me suis dans un premier temps focalisé sur les 500 SNPs les plus associés au phénotype de résistance quantitative (*i.e.* top SNPs). Il s'avère que seulement 15% des top SNPs sont communs à au moins 2 souches et aucun top SNP n'était commun aux 5 souches testées. Ce résultat indique que **les gènes de résistance quantitative à large spectre seraient peu nombreux au sein d'*A. thaliana* vis-à-vis de la bactérie pathogène *P. viridiflava*.** Dans 61% des cas, les top SNPs communs à au moins 2 souches sont localisés à moins de 20 kb de

gènes candidats (*i.e.* *R* genes; homologues à des protéines de résistance aux maladies, synthèse acide salicylique, acide jasmonique, éthylène; transporteurs d'ions; WRKY (régulateur de transcription)...). 85% des top SNPs sont spécifiques à une seule souche de *P. viridiflava*. En comparaison avec les top SNPs communs à au moins 2 souches, seulement 38% des top SNPs spécifiques sont localisés à moins de 20 kb de gènes candidats.

2.2. *Xanthomonas* sp

2.2.1. *Xanthomonas campestris* pv. *campestris*

Huard-Chauveau et al. Nature in revision, ANR Blanc QUANTIREX 2009-2013

Ce projet sur l'identification des bases génétiques de la résistance quantitative s'est poursuivi avec la bactérie pathogène *Xanthomonas campestris* pv. *campestris* (*Xcc*). *Xanthomonas campestris* est l'espèce de bactérie pathogène la plus prévalente au sein de la communauté foliaire dans les populations naturelles d'*A. thaliana* (Kniskern *et al.* 2007). C'est aussi un pathogène entraînant de fortes baisses de rendement des cultures à travers le monde. Ce projet s'effectue en collaboration avec Dominique Roby (LIPM, Toulouse) et est soutenu par un financement ANR Blanc (2009-2013).

Avant le commencement du projet, le laboratoire de Dominique Roby était déjà en train de cloner le premier gène de résistance quantitative à la souche *Xcc568* à partir de la famille RIL Col-5 x Kas-1. Ce clonage s'est terminé au début du projet avec l'identification du gène *RKS1* expliquant 53.7% de la variance phénotypique de la résistance quantitative à la souche *Xcc568*. Le gène *RKS1* correspond à une kinase atypique avec aucune homologie de structure aux gènes de résistance qualitative connus à ce jour, et correspondrait à un nouveau mécanisme de résistance encore inconnu chez les plantes. Deux transcrits de taille différente sont produits à partir de *RKS1* : un transcrit long (*RKS1-L*) et un transcrit court (*RKS1-S*). Le niveau de résistance quantitative est uniquement fonction de l'expression relative de *RKS1-L* : plus il est exprimé, plus la plante est résistante.

Afin de tester si *RKS1* était aussi à l'origine de la variation naturelle de la résistance quantitative à *Xcc568* à l'échelle de l'espèce (et n'était donc pas spécifique au croisement entre les accessions Col-5 x Kas-1), 381 accessions naturelles d'*A. thaliana* (toutes génotypées pour 214k SNPs) ont été phénotypées pour la résistance quantitative à *Xcc568* (Figure 8).

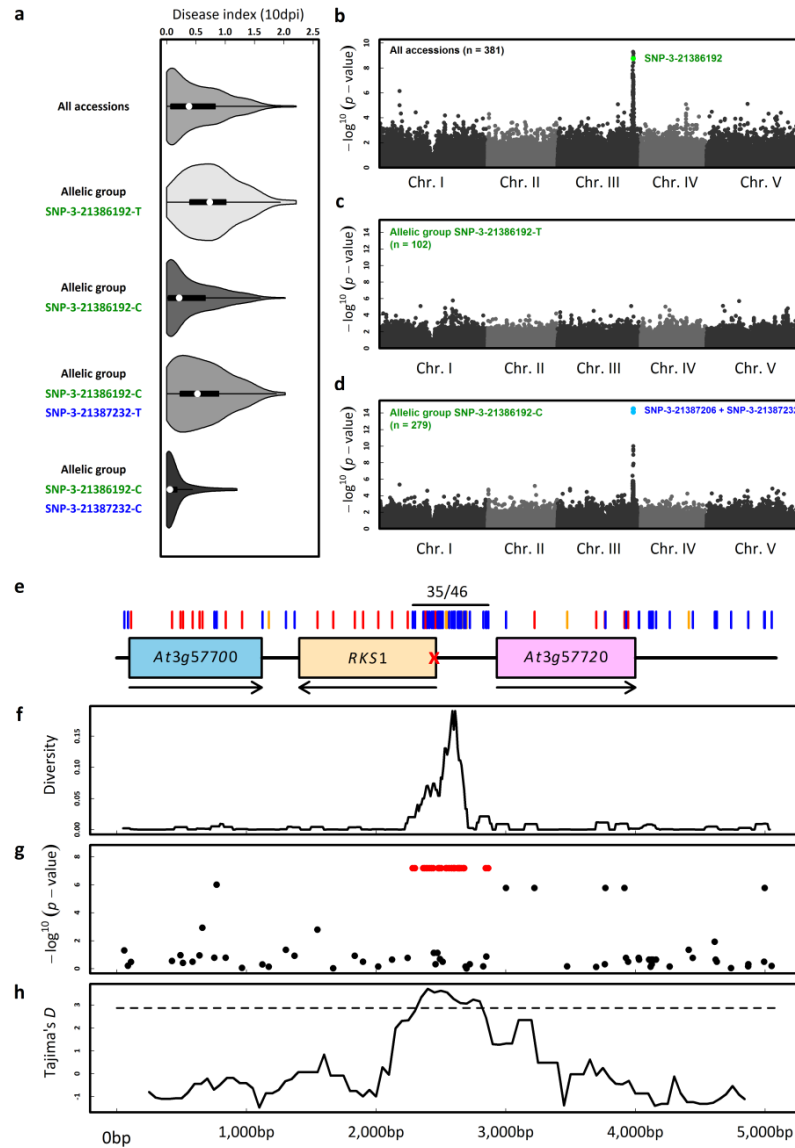


Figure 9. Variation naturelle de la résistance quantitative à *Xcc568* et du gène *RKS1* à l'échelle de l'espèce. (a). Violin plots représentant la distribution phénotypique des symptômes (voir Figure 8). Scan génomique des 214k SNPs pour l'association avec les symptômes (b) en utilisant l'ensemble des 381 accessions, (c) à l'intérieur du groupe allélique SNP-3-21386192-T et (d) à l'intérieur du groupe allélique SNP-3-21386192-C. (e) Diversité de séquences dans une région génomique de 5kb centrée sur *RKS1*. Les barres rouges, bleues et oranges indiquent les mutations non-synonymes, les mutations silencieuses et les indels, respectivement. '35/46' indique la localisation des 35 polymorphismes en déséquilibre de liaison complet. La croix rouge indique la position du codon stop au 4^{ème} acide aminé. Les flèches indiquent l'orientation des gènes le long du génome. (f) Diversité nucléotidique observée entre les deux haplogroupes intergénomiques. (g) Scan pour une association avec l'expression génique relative de *RKS1* dans la région génomique centrée sur *RKS1*. Les points en rouge indiquent les 35 polymorphismes en LD complet. (h) Valeurs du *D* de Tajima le long de la région génomique de 5kb centrée sur *RKS1*. La ligne en pointillés correspond au seuil de significativité de 1% d'une distribution empirique basée sur 876 fragments de 500bp séquencés sur le même jeu de 95 accessions utilisées dans cette étude pour séquencer la région génomique de 5kb centrée sur *RKS1* (Nordborg *et al.* 2005). Pour (b), (c), (d) et (g), l'axe des ordonnées correspond au $-\log_{10}$ de la probabilité d'association entre un SNP et la résistance quantitative à *Xcc568*. Cette probabilité d'association est obtenue à partir d'un modèle linéaire mixte incluant une matrice d'identité génétique entre les accessions afin de contrôler pour l'effet de la structure des populations, et ainsi diminuer le taux de faux positifs (méthode EMMAX, Kang *et al.* 2010).



Figure 8. Echelle de sévérité des symptômes liés à *Xanthomonas campestris* pv. *campestris* dix jours après inoculation. *Xcc* étant une bactérie pathogène du système vasculaire des plantes, l'inoculation s'effectue en trempant une aiguille dans une solution bactérienne puis en perçant la veine centrale en trois positions équidistantes.

Les principaux résultats sont les suivants :

- Nous avons observé une variation naturelle importante de la résistance quantitative à *Xcc568*, avec **une prévalence d'accessions résistantes** (Figure 9a).
- Les analyses de GWA mapping ont révélé un pic unique d'association à la fin du chromosome 3 (Figure 9b). Ce pic est centré sur le gène *RKS1* avec le 4^{ème} top SNP (SNP-3-21386192) localisé dans la région 5' de *RKS1*. **Ce résultat est une belle démonstration de la puissance du GWA mapping pour détecter précisément les régions génomiques associées à la variation phénotypique naturelle chez *A. thaliana*.** Pour information, la cartographie fine de la région génomique contenant *RKS1* à partir de la famille RIL Col-5 x Kas-1 a pris plus de 3 années ; alors que la détection de cette même région a pris 2 mois par GWA mapping (incluant phénotypage + analyses statistiques).
- Après avoir séparé les 381 accessions selon le polymorphisme au 4^{ème} top SNP (SNP-3-21386192), une variation phénotypique substantielle était toujours observée au sein des deux groupes alléliques. Une analyse GWA mapping a été effectuée à l'intérieur de chacun des deux groupes alléliques. Alors qu'aucun pic d'association n'a été identifié pour le groupe allélique SNP-3-21386192-T le plus sensible (Figure 9c), un pic unique d'association a été trouvé pour le groupe allélique SNP-3-21386192-C relativement plus résistant (Figure 9d). Les deux top SNPs de ce nouveau pic sont situés au début de la région codante de *RKS1*. **Cette nouvelle approche de nested GWA mapping** que j'ai mise en place **suggère une série allélique au gène *RKS1*** expliquant entre 43.8% - 47.1% de la variation naturelle dans les populations locales.

Nous avons par la suite séquencé 95 accessions pour une région génomique de 5kb centrée sur *RKS1* :

- 99 polymorphismes ont été détectés (Figure 9e). De manière intéressante, la région comprenant les 186 premières paires de bases de *RKS1* et la région intergénique entre *RKS1* et le gène *AT3G57720* définit deux haplogroupes (*i.e.* un haplogroupe intergénique résistant et un haplogroupe intergénique sensible) très divergent et distingués par 35 polymorphismes en déséquilibre de liaison complet (Figures 9e et 9f). L'analyse des séquences a par ailleurs mis en évidence que notre groupe allélique initial SNP-3-21386192-T le plus sensible étaient composé deux allèles sensibles indépendants situés à l'intérieur de l'haplogroupe intergénique résistant (Figure 10).
- L'haplotype le plus divergent d'une espèce sœur d'*A. thaliana* (*i.e.* *A. lyrata*) est fortement associé à fort niveau d'expression de *RKS1-L* (Figures 9g et 10). Ce résultat suggère encore que **la régulation de l'expression de *RKS1* est une composante majeure de la résistance quantitative à *Xcc568*.**
- Comme pour un certain nombre de *R-genes*, l'importance écologique et évolutive de *RKS1* est supportée par **une signature de sélection balancée agissant sur les deux haplotypes intergéniques** (Figure 9h), et la présence de populations polymorphes à travers l'Europe.

Les 381 accessions naturelles utilisées pour le phénotypage de la résistance quantitative à *Xcc568* ont aussi phénotypées pour la résistance quantitative à 4 autres souches de *Xcc* appartenant chacune à une race différente de *Xcc*. Pour deux souches de *Xcc* autres que *Xcc568*, nous avons aussi trouvé un pic unique d'association sur le gène *RKS1* suggérant que ce gène confère une résistance quantitative à un large spectre de races de *Xcc*. Pour les deux dernières souches étudiées, des pics d'association uniques et spécifiques à chacune des souches ont été trouvés sur le chromosome 5 (pour l'une des deux souches, là encore, l'approche de nested GWA mapping a suggéré une série allélique). La validation fonctionnelle des gènes de résistance quantitative à ces deux dernières souches ainsi que la détection de traces de sélection associées à ces gènes sont en cours. Il faut noter que, comme pour la bactérie pathogène *P. viridiflava*, aucun pic d'association n'a été trouvé en commun entre les 5 souches de *Xcc* testées.

2.2.2. Souches isolées dans des populations naturelles d'*A. thaliana* (projet Fulbright)

Les 5 souches de *Xcc* utilisées lors du projet ANR QUANTIREX ont été isolées sur des plantes hôtes autres qu'*A. thaliana*. Dans un souci de replacer la résistance aux pathogènes chez *A. thaliana* dans un contexte écologiquement réaliste, j'ai mis en place un projet visant non seulement à identifier les bases génétiques de la résistance quantitative en utilisant des souches de *Xanthomonas sp.* isolées dans des populations naturelles d'*A. thaliana* ; mais aussi visant à identifier les bases génétiques de la virulence chez ces même souches bactériennes. Ce projet vise donc à identifier les interacteurs géniques dans un système naturel hôte – pathogène. Pour débiter ce projet, je suis parti 7 mois dans le laboratoire de Joy Bergelson (University of Chicago, USA) dans le cadre d'une bourse Fulbright. En collaboration avec Chris Meyer (doctorant dans le laboratoire de Joy Bergelson), 174 accessions mondiales d'*A. thaliana* pour lesquelles la séquence génomique complète sera bientôt obtenue (<http://www.1001genomes.org/>) ont été phénotypées pour la résistance quantitative à 24 souches de *Xanthomonas sp.* pour lesquelles le séquençage complet du génome est en cours (12 souches isolées à partir de populations françaises d'*A. thaliana* et 12 souches isolées à partir de populations nord-américaines d'*A. thaliana*).

Comme pour les 5 souches de *Xcc* utilisées lors du projet ANR QUANTIREX, une prévalence d'accessions résistantes est observée (Figure 11). Pour différentes souches isolées à partir d'une même population d'*A. thaliana*, la réponse des accessions d'*A. thaliana* peuvent être très contrastées d'une souche à l'autre (voir par exemple les souches isolées à partir de la population PLY, Figure 11). Bien que les séquences complètes des 174 accessions ne soient pas encore disponibles, des analyses de GWA mapping basées sur 95 accessions déjà génotypées pour 214k SNPs ont été effectuées. Pour plus de 90% des souches, un pic unique d'association a été détecté suggérant une architecture génétique avec un QTL majeur. De manière surprenante, aucun pic d'association n'a été trouvé en commun entre au moins deux souches. On peut alors se poser la question suivante : à partir de combien de souches testées, un pic commun d'association entre au moins deux souches sera détecté ? Il paraît difficilement envisageable que pour chaque nouvelle souche de *Xanthomonas sp.* corresponde un nouveau gène de résistance quantitative.

Les gènes NB-LRR *R* associés à la résistance qualitative constituent une des familles de gènes les plus polymorphes génétiquement (très forte diversité nucléotidique, présence de nombreuses mutations décrites comme drastique (mutation stop, décalage du cadre de

lecture...), très fort polymorphisme de présence-absence entre les accessions ; Cao *et al.* 2011). En est-il de même pour les gènes de résistance quantitative ?

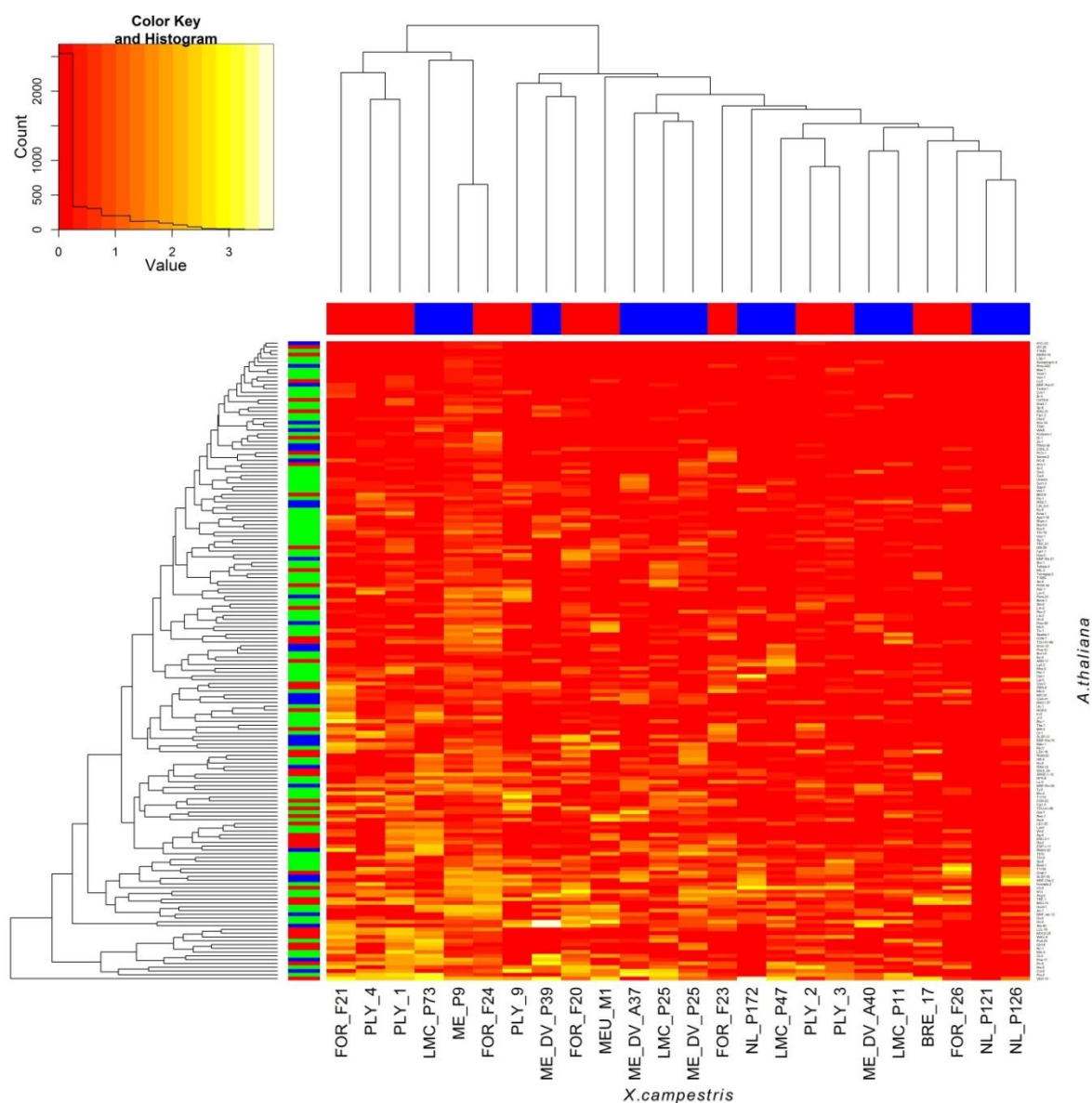


Figure 11. Heatmap de la résistance quantitative résultante des interactions entre *A. thaliana* et *Xanthomonas sp.* Sur l'axe des abscisses, les couleurs rouge et bleue indiquent les souches de *Xanthomonas sp.* françaises et nord-américaines, respectivement. Sur l'axe des ordonnées, les couleurs rouge, bleue et verte indiquent les accessions d'*A. thaliana* françaises, nord-américaines et provenant d'autres pays que la France ou les USA, respectivement. 'Value' sur l'axe des abscisses de la clé de couleur correspond à la sévérité des symptômes (voir Figure 8). Chaque case correspond à la valeur génotypique prédite de résistance quantitative de chaque accession.

A partir de la matrice phénotypique de résistance quantitative 24 x 174 obtenue (Figure 11), l'objectif est de faire une analyse de GWA mapping conjointe (joint GWA mapping) entre les génomes des deux espèces de ce système naturel hôte – pathogène afin d'identifier

les interacteurs géniques entre les deux partenaires. Le développement d'une telle méthode est en cours *via* une collaboration avec Mary Sara McPeck (Department of Biostatistics, University of Chicago, USA). Les objectifs finaux sont :

- de valider fonctionnellement plusieurs paires de gènes en interaction entre *A. thaliana* et *Xanthomonas sp.*, permettant ainsi de pouvoir dégager des généralités quant aux gènes impliqués dans les interactions (classe de gènes, homologie de structure...).
- d'estimer l'âge de sélection des interacteurs géniques et de tester si l'âge de sélection des gènes de résistance quantitative est corrélé à l'âge de sélection des gènes de virulence. Une corrélation significative pourrait suggérer une dynamique coévolutive entre *A. thaliana* et *Xanthomonas sp.*

3. Conclusion

Mes travaux de recherche sur la résistance aux pathogènes ont naturellement suivi les progrès technologiques en génomique (appliqués notamment chez *A. thaliana*), permettant ainsi d'aborder des questions encore difficilement abordables chez des espèces non-modèles :

- L'intérêt du GWA mapping quant à une identification précise et rapide des gènes liés à la variation naturelle de traits phénotypiques a été démontré dans le cas de la résistance quantitative à la *Xcc568*. Cette démonstration a été rendue possible car le gène *RKS1* avait été validé fonctionnellement par ailleurs. Même si l'utilisation du GWA mapping permet une identification plus rapide et précise des régions génomiques associées à la variation naturelle que les méthodes traditionnelles de QTL mapping, il ne faut pas oublier que les résultats du GWA mapping demeurent de simples hypothèses quant à l'identité du gène sous-jacent au QTL concerné : **la validation fonctionnelle demeure une étape certes fastidieuse mais incontournable dans la validation des bases génétiques et moléculaires.**
- Le fait d'avoir identifié peu de pics d'association communs à plusieurs souches appartenant à la même espèce bactérienne pathogène soulève plus de questions que n'apporte de réponses : les différents pics d'association trouvés pour une même espèce de bactérie pathogène sont-ils communs entre plusieurs espèces de bactérie pathogène ? si c'est le cas, l'évolution d'un gène de résistance quantitatif ne dépendrait-il pas plus de la fréquence du gène de virulence sur l'ensemble de la communauté de bactéries pathogènes que de la fréquence d'une bactérie pathogène particulière ? L'identification des gènes de virulence chez *Xanthomonas sp.* devrait permettre d'apporter des éléments de réponse.

C. La phénologie : date de floraison, période de floraison et période reproductive

Chez les plantes annuelles, les facteurs environnementaux définissent une période favorable durant laquelle les plantes doivent compléter leur cycle de vie afin de maximiser leur fitness. Il est donc admis que la durée du cycle de vie doit s'ajuster à la durée de la période favorable, et donc résulter d'une adaptation aux conditions environnementales. Le cycle de vie chez les espèces annuelles est composé de deux périodes : la période végétative avec accumulation des ressources et la période reproductive qui commence à l'initiation de la floraison. Cette dernière période peut elle-même être divisée en 2 périodes plus ou moins chevauchantes: la période de floraison (depuis l'apparition de la première fleur jusqu'à la sénescence de la dernière fleur) et la période reproductive (depuis l'apparition de la première fleur jusqu'à la maturation du dernier fruit). Le ratio (période de floraison / période reproductive) est supposé être un indicateur d'une trade-off entre qualité (ratio faible) et quantité de graines (ratio élevé).

Bien que les bases génétiques de la variation naturelle de la date de floraison commencent à être bien connues chez *A. thaliana* (Andrés and Coupland 2012), il faut toutefois rappeler que l'identification de ces bases génétiques a été toujours été réalisée à partir d'expériences menées dans des conditions contrôlées, souvent très éloignées des conditions environnementales rencontrées dans les populations naturelles d'*A. thaliana* à travers le monde. Savoir si les mêmes gènes seraient retrouvés dans des conditions écologiquement plus réalistes restait une question ouverte ?

La connaissance des bases génétiques de la variation naturelle de la période de floraison et de la période reproductive aussi bien chez *A. thaliana* que chez d'autres plantes annuelles sont quasi-inexistantes. De même, la valeur adaptative ainsi que les agents sélectifs de la date de floraison, période de floraison et période reproductive restent largement inconnus chez *A. thaliana*. Pour combler ces lacunes et mieux comprendre la dynamique adaptative de ces traits phénologiques dans un environnement hétérogène, j'ai aussi abordé cette thématique de recherche par la complémentarité écologie – génétique – évolution.

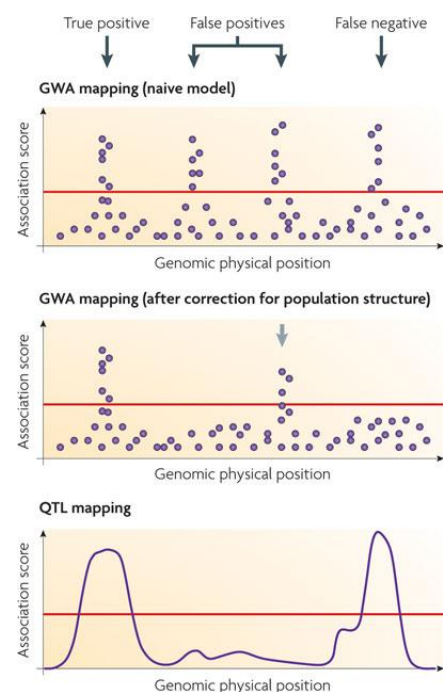
1. Génétique et épigénétique

1.1. Identification des bases génétiques associées à la variation phénologique

Thèse Benjamin Brachi 2007-2010, Atwell et al. 2010 *Nature*, Bergelson & Roux 2010 *Nature Reviews Genetics*, Brachi et al. 2010 *PLoS Genetics*, Platt et al. 2010 *PLoS Genetics*, Brachi et al. *Acta Botanica Gallica – Botany Letters in revision*.

Comme pour la résistance quantitative aux pathogènes, je n'aurais pu utiliser que la méthode du GWA mapping pour identifier les régions génomiques associées à la variation phénologique chez *A. thaliana*. Cependant, une étude portant sur la date de floraison chez *A. thaliana* avait montré que de nombreuses associations SNP – variation naturelle de la date de floraison pouvaient correspondre à des faux positifs (Zhao *et al.* 2007). En effet, par rapport à la résistance aux pathogènes, la variation naturelle de la date de floraison chez *A. thaliana* se recoupe fortement avec la structure des populations présente au sein de cette espèce (Aranzana *et al.* 2005). Différentes méthodes statistiques ont été proposées pour diminuer le taux de faux positifs, notamment un modèle linéaire mixte incluant une matrice d'identité génétique entre les accessions afin de contrôler pour l'effet de la structure des populations (méthode EMMAX Efficient Mixed Model Association eXpedited, Kang *et al.* 2010). Cependant, ces méthodes n'éliminent pas complètement les faux positifs. Même si elles le faisaient, on pourrait alors se trouver en présence de faux négatifs, c'est-à-dire de vrais positifs qui se recourent avec la structure des populations et qui deviennent donc non significatifs après correction pour la structure des populations. Une solution pour diminuer à la fois le taux de faux positifs et le taux de faux négatifs est de combiner les avantages du GWA mapping (résolution fine) et des méthodes traditionnelles de QTL mapping (pas d'effet de la structure des populations ; Figure 12).

Figure 12. Avantages de combiner GWA mapping et méthodes traditionnelles de QTL mapping (Bergelson and Roux 2010). Cette combinaison permet aux vrais positifs et aux faux négatifs d'être distingués des faux positifs. Les vrais positifs sont des SNPs causaux détectés par GWA mapping et recouverts par des régions QTL. Les corrections pour la structure des populations mettent en évidence les faux positifs qui correspondent à de fausses associations phénotype – génotype. Parce que les méthodes statistiques contrôlant pour la structure des populations réduisent seulement (et n'éliminent pas) tous les faux positifs, des faux positifs peuvent toujours être présents (flèche grise). Dans ce cas, les faux positifs restant ne sont pas validés par les régions QTL, démontrant ainsi la valeur ajoutée du QTL mapping dans la détection des vrais positifs. Les faux négatifs sont des SNPs causaux qui sont perdus comme un artefact des corrections pour la structure des populations, mais peuvent être validés par des régions QTL. La ligne rouge horizontale indique le seuil de significativité d'une association phénotype – génotype.



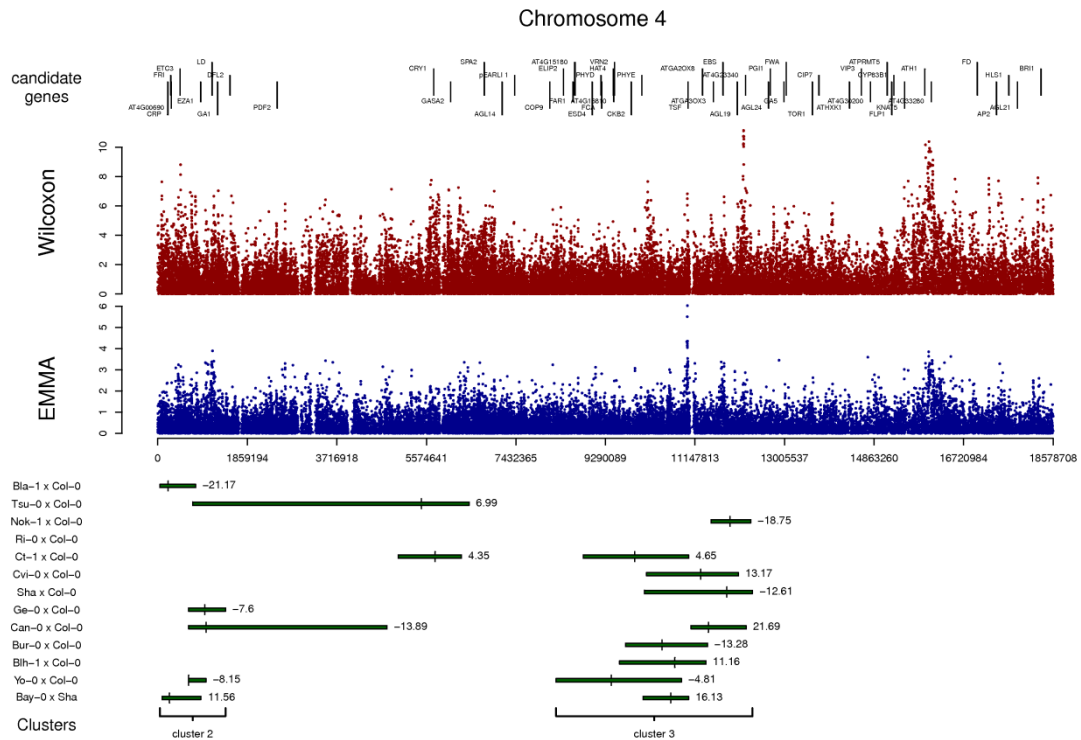


Figure 14. Comparaison des résultats du GWA et QTL mapping (QTLs additifs) sur le chromosome 4 pour la date de floraison mesurée sur un terrain expérimental. L'axe des abscisses indique la position physique des SNPs le long du chromosome. Top panel : position des 52 gènes candidats de la date de floraison localisés sur le chromosome 4. Mid-panel : l'axe des ordonnées correspond à la valeur d'association entre chaque SNP et la date de floraison selon un modèle naïf (Wilcoxon) et un modèle statistique tenant compte de la structure des populations (EMMA). Bottom panel : régions QTL pour chacune des 13 familles RILs. Pour chaque famille RIL, les barres vertes représentent les intervalles de confiance à 95% de la position du QTL (trait vertical). Parmi les 5 clusters de QTLs trouvés pour la date de floraison, deux se retrouvent sur le chromosome 4. Le cluster 3 confirme le pic d'association trouvé par la méthode EMMA sur le gène *TSF*.

Afin de combiner les avantages du GWA mapping et des méthodes traditionnelles de QTL mapping, une expérience de phénotypage de grande envergure a été menée dans des conditions naturelles (terrain expérimental de l'Université de Lille 1), avec la mesure de traits phénologiques sur plus de 18 800 plantes (Figure 13). Ces plantes correspondent à 200 accessions mondiales génotypées pour 214k SNPs (soit en moyenne 1 SNP tous les 500bp) et 13 familles de lignées recombinantes consanguines (pour un total de 4366 RILs génotypées en moyenne pour 85 marqueurs).



Figure 13. Expérience de phénotypage de différents traits phénologiques sur le terrain expérimental de l'Université de Lille 1. La photo de droite est caractéristique des conditions hivernales rencontrées dans le Nord de la France, conditions climatiques difficilement reproductible dans des chambres de culture.

Outre une très bonne description de l'architecture génétique de la date de floraison par les 13 familles RILs, ces expériences ont permis pour la première fois de montrer que la combinaison du GWA mapping et du QTL mapping (1) augmentait de manière significative la détection fine des gènes candidats de la date de floraison par rapport à chacune des méthodes utilisée indépendamment, et (2) diminuait le nombre de faux positifs liés à la structure des populations (jusqu'à 40% des top SNPs) et de faux négatifs apparaissant après correction pour la structure des populations (jusqu'à 24% des top SNPs ; Figure 14). Par ailleurs, **cette étude constitue la première étude de GWA mapping réalisée (toutes espèces confondues) en conditions naturelles**, permettant de mettre en évidence que de nombreux gènes de l'horloge circadienne ainsi que le florigène *TSF* étaient liés à la variation naturelle de la date de floraison. Ces gènes candidats n'avaient jamais été cités dans la littérature basée sur des expériences réalisées dans des conditions contrôlées.

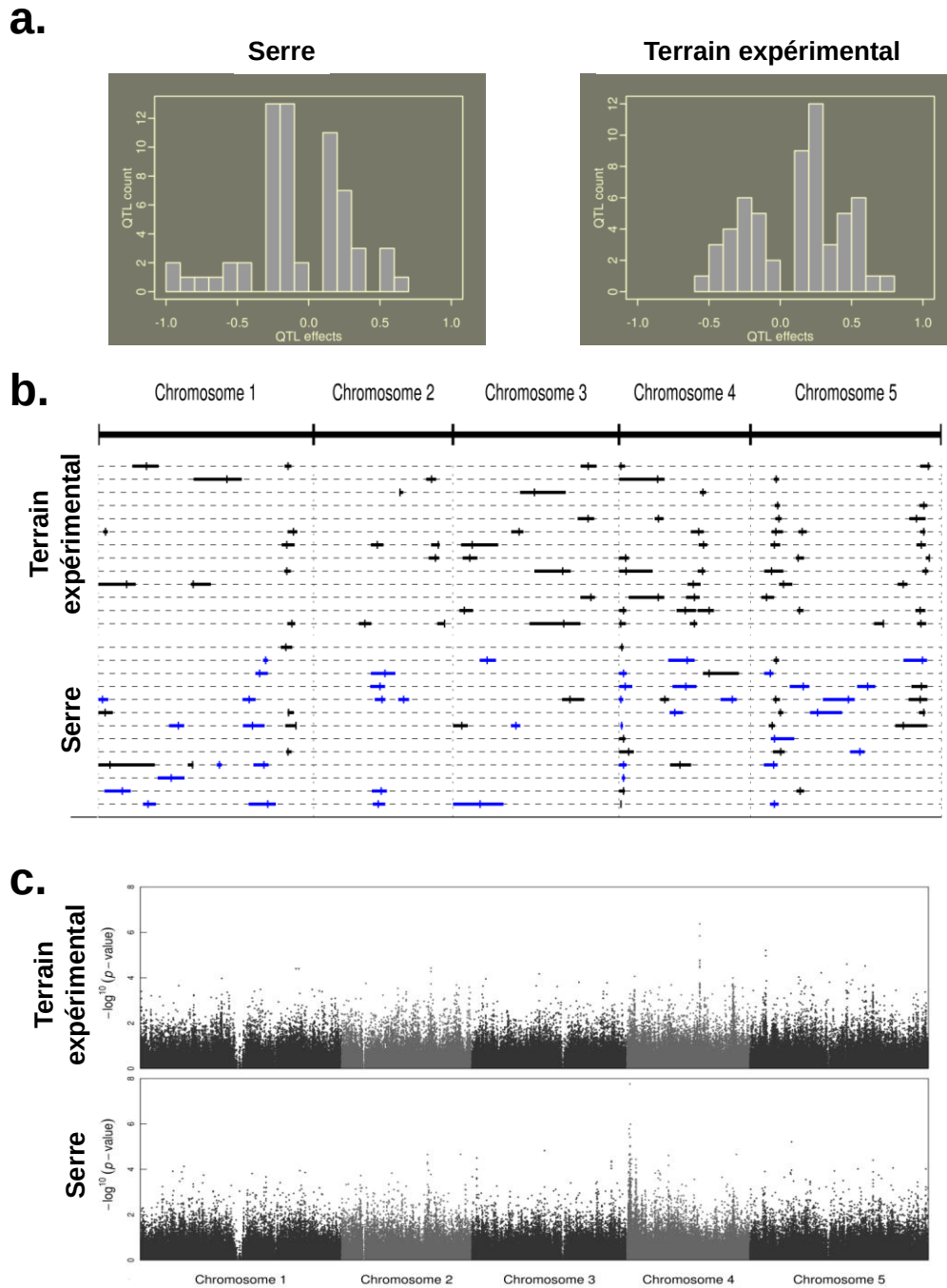


Figure 15. Comparaison des résultats entre serre et terrain expérimental pour la date de floraison. **a. Architecture génétique.** Pour chaque environnement, les effets des QTLs ont été standardisés en divisant l'effet du QTL par l'écart-type phénotypique. **b. QTL mapping.** Pour chaque environnement, chaque ligne correspond à un des 13 familles RIL utilisées dans l'étude. Pour chaque famille RIL, les barres représentent les intervalles de confiance à 95% de la position du QTL (trait vertical). Les barres bleues représentent les QTLs spécifiques aux conditions de serre. **c. GWA mapping.** L'axe des abscisses indique la position physique des SNPs le long du chromosome. Les 5 chromosomes sont représentés avec différentes nuances de gris. L'axe des ordonnées correspond à la valeur d'association entre chaque SNP et la date de floraison selon un modèle statistique tenant compte de la structure des populations (EMMA).

Afin de vérifier que cette distinction entre terrain expérimental et conditions contrôlées ne provenait pas de l'utilisation d'un matériel génétique (accessions naturelles et familles RIL) différent entre notre étude et d'autres études publiées, nous avons mené en serre exactement la même expérience menée sur le terrain expérimental (même design expérimental, même matériel génétique, $n = 18\,800$ plantes, 20°C / 16h photopériode / pas de vernalisation).

Les principaux résultats pour la date de floraison sont les suivants :

- Alors qu'une distribution normale de la date de floraison est observée sur le terrain expérimental, une distribution L-shaped avec un excès de lignées génétiques précoces est observée en serre.
- Bien que le nombre de QTLs additifs soit à peu près le même entre les deux environnements (68 QTLs pour la serre, 66 QTLs pour le terrain expérimental), **l'architecture génétique est très différente entre la serre et le terrain expérimental** (Figure 15a). Alors que la majorité des QTLs détectés sur le terrain expérimental ont un effet faible à moyen, de nombreux QTLs à effet fort ont été détectés en serre.
- 41 des 68 QTLs (=60%) détectés en serre sont spécifiques à cet environnement (Figure 15b). En combinant les 500 top SNPs identifiés par GWA mapping dans chaque environnement, **seulement 1.2% de ces top SNPs sont communs aux deux environnements** (Figure 15c). Alors que les gènes liés à l'horloge circadienne et le florigène *TSF* ont été identifiés sur le terrain expérimental, les gènes candidats identifiés en serre correspondent aux gènes de vernalisation *FRI* et *FLC* ainsi qu'à des gènes de la voie des hormones.

Ces résultats très contrastés tant au niveau de l'architecture génétique, des régions génomiques associées à la variation naturelle de la date de floraison que des gènes candidats identifiés soulignent **l'importance de réaliser aussi le phénotypage dans des milieux proches de ceux rencontrés par l'espèce étudiée.**

1.2. Contribution de l'épigénétique dans la variation naturelle phénologique

Roux et al. 2011 Genetics

Comme la variation phénotypique naturelle peut être aussi d'origine épigénétique, l'identification des bases épigénétiques de la date de floraison a aussi été abordée dans cette

thématique de recherche. Afin d'établir la contribution de l'épigénétique dans la variation naturelle de la date de floraison chez *A. thaliana*, une expérience de phénotypage sur un terrain expérimental de l'Université de Lille 1 a été effectuée en utilisant une famille de 500 epiRILs. Cette famille est issue d'un croisement entre 2 plantes identiques au niveau génétique mais différentes au niveau de la méthylation de l'ADN sur l'ensemble du génome (Johannes *et al.* 2009). Ainsi, seuls les niveaux de méthylation de l'ADN ségrégent entre les 500 epiRILs. Dans cette même expérience, des accessions naturelles d'*A. thaliana* ont aussi été utilisées. Le phénotypage de ces epiRILs et accessions naturelles non seulement pour la date de floraison mais aussi pour 9 autres traits (morphologiques, architecturaux) ont permis de montrer que :

- L'héritabilité d'origine épigénétique dans les epiRILs était de 30% pour la date de floraison.
- Les patterns d'héritabilité d'origine épigénétique étaient fortement corrélés à ceux observés dans les populations naturelles.

Ces travaux indiquent que la variation épigénétique induite ou naturelle des traits complexes pourraient partager une architecture polygénique et suggèrent de nouveaux mécanismes d'adaptation.

1.3. Reconstruction de la première marche de la marche adaptative vers la précocité de floraison

Le Corre et al. Molecular Biology and Evolution 2002, Roux et al. Trends in Plant Science 2006, Le Corre and Roux Biofutur 2007.

Comme annoncé précédemment, le GWA mapping permet une identification plus rapide et plus précise des bases génétiques liées à la variation naturelle de la date de floraison chez *A. thaliana*. Cependant, une question reste en suspens : pourquoi certains gènes semblent être les cibles majeures de la sélection naturelle pour la précocité de floraison, alors que de plus de 280 gènes de la date de floraison pourraient être la cible de la sélection naturelle pour une variation de la date de floraison ? Une des hypothèses concernant la précocité de floraison chez *A. thaliana* est que le nombre de mutations adaptatives (*i.e.* conférant la précocité) est plus important pour certains gènes de floraison que pour d'autres gènes de la date de floraison.

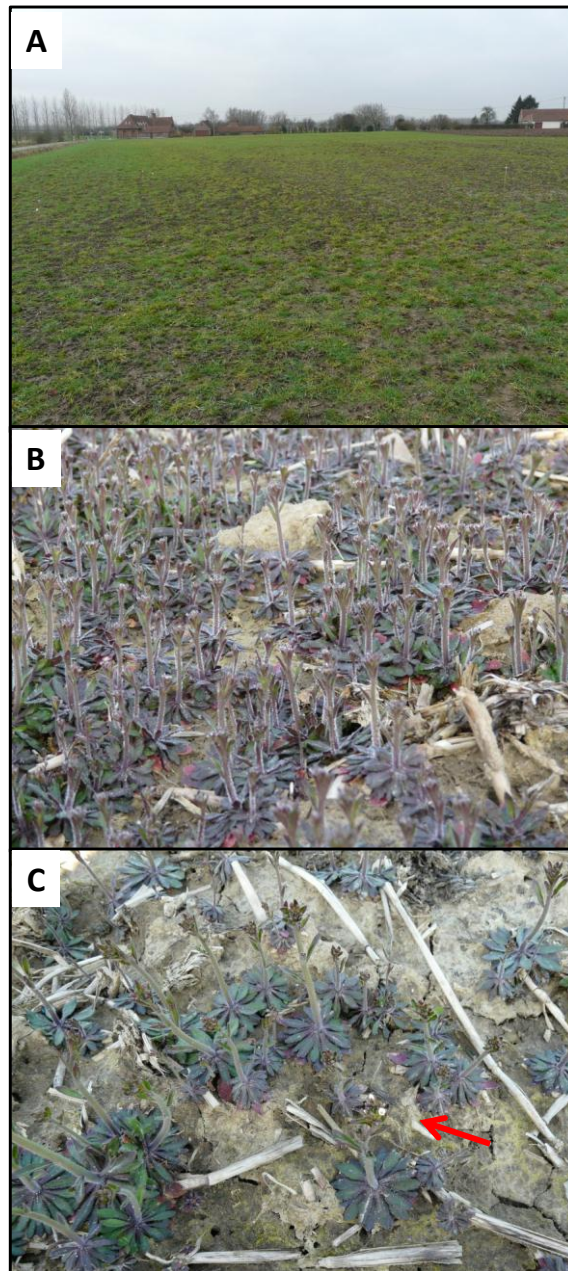


Figure 16. Sélection de mutants spontanés associés à une précocité de floraison. A. Parcelle agricole utilisée pour le semis de 23 millions de graines. **B.** Plantes toutes issues de la même plante grand-mère de l'accession suédoise Lov-5 très tardive pour la date de floraison. **C.** Identification d'une plante fleurissant 3-4 jours avant les autres plantes se trouvant à proximité.

Par exemple, intéressons-nous aux gènes *FRI* et *APETALA 1 (AP1)*. Le gène *AP1* est un gène intégrateur situé en aval des voies de la date de floraison. La protéine *AP1* est impliquée dans la formation des fleurs. Une mutation dans la région codante du gène entraîne la non-formation des fleurs. Durant la phase végétative, le gène *FLC* (gène clé de la floraison) réprime théoriquement l'expression du gène *AP1* en se fixant sur la CA₂G-box (8 paires de base) située dans la région promotrice du gène *AP1*. La précocité de floraison ne peut être obtenue que par mutation dans cette région promotrice de 8 paires de base, comme précédemment observé chez le blé (Yan *et al.* 2003). En revanche, le gène *FRI* code pour une protéine comprenant deux domaines coiled-coil lui permettant de se fixer dans la région promotrice du gène *FLC* (Johanson *et al.* 2000). Une perte de fonction de la protéine *FRI* provoque une précocité de floraison : l'ensemble de la région codante (plus de 1500 paires de base) et la région promotrice peuvent donc être la cible de la sélection naturelle pour la précocité de floraison, augmentant ainsi théoriquement le nombre potentiel de mutations adaptatives: toute mutation de perte de fonction semblant pouvoir être positivement sélectionnée. Dans une population théorique avec tous les mutants potentiels conférant une précocité de floraison, la sélection d'une mutation au gène *FRI* serait le résultat d'une balance mutation/dérive/sélection, avec comme gagnant le gène ayant le taux de mutation efficace le plus élevé, à coefficient de sélection identique. A l'heure actuelle, aucune donnée expérimentale n'est disponible pour valider ou réfuter cette hypothèse.

Pour tester cette hypothèse, j'ai mis en place une expérience pour identifier les mutations spontanées retenues au cours d'une expérience de sélection sur une génération. Pour conférer une réalité écologique à la reconstruction du premier pas de la marche adaptative vers la précocité de floraison, plus de 23 millions de graines (toutes issues de la même plante grand-mère de l'accession suédoise Lov-5 très tardive pour la date de floraison) ont été semées sur une parcelle agricole de 1ha proche de Lille le 6 septembre 2010 (correspondant ainsi à la période de la principale cohorte de germination chez *A. thaliana* dans le Nord de la France). Sur les 23 millions de graines, seules 2,5 millions ont germées et sont arrivées au stade de floraison. Sur ces 2,5 millions de plantes qui ont fleuri, 64 lignées ayant une floraison plus précoce ont été identifiées (Figure 16). L'estimation de la date de floraison de ces 64 lignées en comparaison à la lignée Lov-5 d'origine ainsi que la sélection de lignées homozygotes sont en cours sur un terrain expérimental de l'Université de Lille 1. Outre l'estimation directe (et rare) d'un taux de mutation potentiellement adaptative, le séquençage des génomes des mutants spontanés précoces permettra d'estimer un nombre de mutations potentiellement

adaptatives par gène de floraison, et ainsi valider ou non l'hypothèse selon laquelle le nombre de mutations précoces est plus important pour certains gènes de floraison que pour d'autres gènes de floraison.

2. Ecologie

2.1. Adaptation au climat à l'échelle mondiale

Thèse Benjamin Brachi 2007-2010, Hancock et al. 2011 Science

En collaboration avec Joy Bergelson (University of Chicago), nous voulions tester si les facteurs climatiques sont des agents sélectifs importants chez *A. thaliana*. Nous avons alors réalisé un scan génomique des relations polymorphismes génétiques – facteurs climatiques à partir de 948 accessions mondiales, toutes génotypées pour 214k SNPs et pour lesquelles 13 variables climatiques étaient disponibles dans les localités d'origine. Nos prédictions étaient les suivantes :

- Si *A. thaliana* est adapté au climat, on devrait être capable de prédire dans une localité donnée (et donc un climat donné) les accessions qui devraient avoir la fitness la plus importante et ceci à partir des SNPs les plus associés aux 13 facteurs climatiques.
- La date de floraison étant fortement corrélée à la latitude chez *A. thaliana*, il a été suggéré que les facteurs climatiques devaient être des agents sélectifs importants pour ce trait (Stinchcombe *et al.* 2004). On s'attend donc à retrouver un enrichissement en gènes de floraison parmi les gènes les plus fortement associés aux facteurs climatiques.

Les principaux résultats de cette étude sont les suivants :

- les SNPs associés à des changements d'acide-aminés (contrairement aux SNPs dans les régions intergéniques) étaient significativement enrichis parmi les SNPs les plus associés au climat, suggérant que notre scan génomique a efficacement détecté des allèles adaptatifs au climat.
- notre première prédiction a été validée. En se basant sur les SNPs les plus associés aux 13 facteurs climatiques, nous avons tout d'abord prédit les allèles les plus avantageux pour le climat à Lille. Puis, dans une expérience mise en place sur le terrain expérimental de l'Université de Lille 1, nous avons montré que les accessions mondiales qui ont ces allèles avantageux prédits ont effectivement eu une fitness

relative plus importante que les accessions mondiales qui n'avaient pas ces allèles avantageux prédits (Figure 17).

- les SNPs les plus associés au climat sont 3 fois plus associés à des balayages sélectifs récents que des SNPs choisis au hasard sur le génome. Ce résultat suggère qu'une espèce comme *A. thaliana* pourrait rapidement atteindre une limite adaptative dans le contexte d'un changement climatique rapide, ceci à cause des contraintes imposées par l'attente de nouvelles mutations.

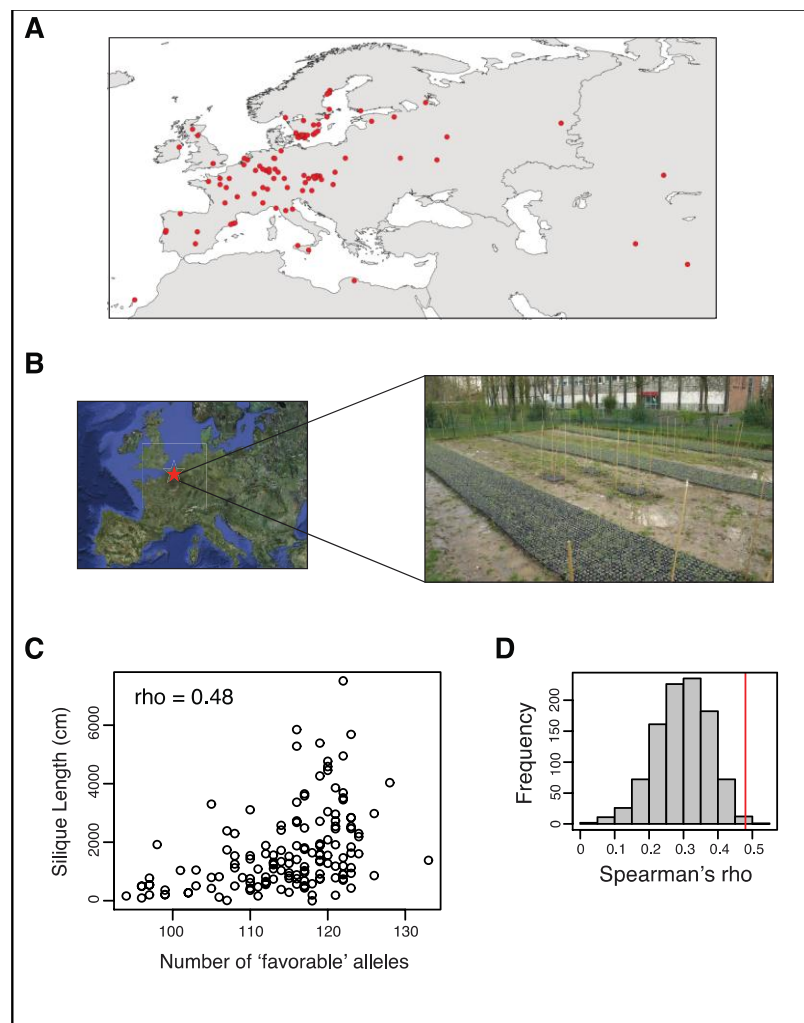


Figure 17. Les SNPs les plus associés au climat prédisent le rang du succès reproducteur à Lille (France). **A.** Localisation géographique des accessions utilisées lors de l'étude. **B.** Localisation géographique et photo du terrain expérimental de l'Université de Lille 1. **C.** Relation entre la longueur totale des siliques (i.e. une mesure de la fitness) et le nombre d'allèles associés au climat prédits avantageux à Lille. **D.** Coefficient de corrélation de Spearman entre la longueur total des fruits (=siliques) et le nombre d'allèles favorables (ligne rouge) en comparaison avec une distribution de coefficients de corrélation issus de permutations.

Il faut cependant noter 1) qu'une partie de la variance de la fitness sur le terrain expérimental de l'Université de Lille 1 demeure inexpliquée par les allèles les plus avantageux prédits, et 2) que peu de SNPs les plus associés au climat sont localisés dans ou à proximité de gènes de floraison. Ces deux points suggèrent que d'autres agents sélectifs agissant sur *A. thaliana* restent à être identifiés et que la valeur adaptative de la date de floraison doit peut-être être étudiée à d'autres échelles géographiques que l'échelle mondiale.

2.2. Identification à différentes échelles spatiales des agents sélectifs agissant sur la phénologie

Thèse Benjamin Brachi 2007-2010, Brachi et al. Molecular Ecology in revision

Afin d'identifier les facteurs écologiques susceptibles d'agir comme agents sélectifs sur la phénologie chez *A. thaliana*, 49 populations françaises naturelles d'*A. thaliana* représentant 800 lignées génétiques et provenant de 4 régions françaises (Bourgogne, Bretagne, Languedoc et Nord ; entre 11 et 16 populations par région et 200 lignées par région) ont été caractérisées écologiquement pour 42 variables écologiques (climat ($n = 25$), sol ($n = 14$) et compétition intra- et inter-spécifique ($n = 3$)) ; phénologiquement pour 6 traits phénologiques selon deux cohortes de germination (automne et printemps) simulées dans des conditions contrôlées de serre ; et génétiquement (135 SNPs supposés neutre). Les principaux résultats sont les suivants :

- l'échelle de structuration de la variation naturelle est différente entre les traits phénologiques. Par exemple, la variation naturelle pour la période de floraison et la période reproductive s'observe notamment au niveau intra-population ; alors que la variation naturelle pour la date de floraison se répartit de manière équitable entre les régions, entre les populations au sein des régions et à l'intérieur des populations.
- jusqu'à 63% de la variation phénologique peut être expliquée par la diversité génétique neutre, suggérant qu'il est important de contrôler pour des processus non-sélectifs dans la détection des associations phénotype – écologie, ceci afin d'éviter des faux positifs.
- les relations entre variables écologiques et traits phénologiques sont plus fortes au niveau régional qu'à l'échelle de la France, suggérant une adaptation locale des populations.
- bien que les facteurs climatiques aient souvent été suggérés comme agents sélectifs principaux de la phénologie chez *A. thaliana*, les conditions édaphiques et la compétition

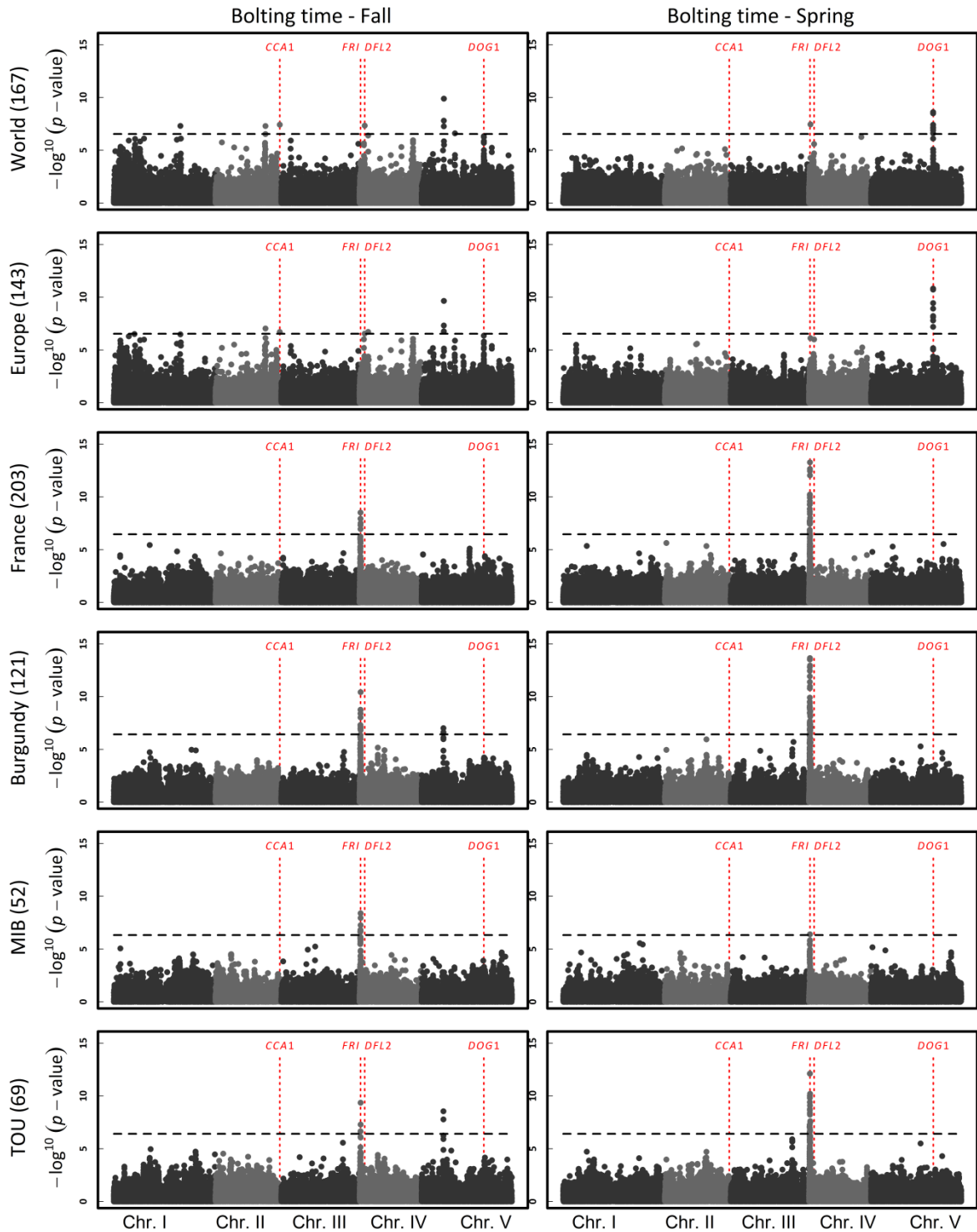


Figure 18. Identification de régions génomiques associées à la variation naturelle de la date de floraison à différentes échelles géographiques et selon deux cohortes simulées de germination. L'axe des abscisses indique la position physique des SNPs le long du chromosome. Les 5 chromosomes sont représentés avec différentes nuances de gris. L'axe des ordonnées correspond à la valeur d'association entre chaque SNP et la date de floraison selon un modèle statistique tenant compte de la structure des populations (EMMAX). La ligne en pointillés correspond au seuil de Bonferroni. Les lignes verticales rouges en pointillés indiquent la position physique de 4 gènes candidats de la date de floraison identifiés dans cette étude.

interspécifique apparaissent comme des agents sélectifs tout aussi, voire plus, importants dans certaines régions.

Cette étude représente la caractérisation écologique la plus complète de populations naturelles d'*A. thaliana* et suggère une échelle géographique restreinte à laquelle l'adaptation phénologique s'est potentiellement effectuée. La détermination de cette échelle géographique devrait permettre une meilleure détection et compréhension des bases génétiques associées à la variation naturelle de la phénologie. Ainsi, pour compléter cette histoire, nous avons réalisé des analyses de GWA mapping à différentes échelles spatiales (d'une échelle mondiale à l'échelle de la population locale) à partir de 400 accessions mondiales et/ou françaises, toutes génotypées pour 214k SNPs et phénotypées en même temps que les 49 populations françaises. Ces analyses ont mis en évidence que la détection des bases génétiques associées à la variation phénologique chez *A. thaliana* dépend clairement de l'échelle géographique considérée, avec la détection de pics d'association spécifiques à certaines populations locales (Figure 18). Par exemple, pour la date de floraison mesurée dans la cohorte de printemps, le gène *DOG1* a été identifié uniquement à l'échelle mondiale alors que le gène *FRI* a été identifié uniquement aux échelles française et locale (*DOG1* et *FRI* sont deux gènes qui ont été validés fonctionnellement par d'autres études comme associés à la variation naturelle de la date de floraison ; Andrés and Coupland 2012). De même, comme précédemment décrit pour la comparaison entre la serre et le terrain expérimental, les régions génomiques associées à la variation phénologique diffèrent entre les deux cohortes simulées de germination (Figure 18).

3. Evolution

3.1. Evolution expérimentale dans des conditions environnementales contrôlées

Thèse Benjamin Brachi 2007-2010, Brachi et al. 2012 PLoS One

En combinant estimation directe de la fitness et expériences d'évolution expérimentale, nous avons mis en évidence que les coefficients de sélection associés à la date de floraison, durée de floraison et durée de reproduction étaient très variables suivant la pression de sélection imposée aux plantes (stress hydrique vs. compétition interspécifique par exemple), mais aussi suivant l'intensité de la pression de sélection. Ces expériences en conditions environnementales contrôlées ont aussi permis de valider des associations phénologie – écologie décrites ci-dessus (partie II.C.2.2), suggérant que le stress hydrique et la compétition interspécifique sont certainement de vrais agents sélectifs de la phénologie chez *A. thaliana*.

3.2. Lien entre fitness et date de floraison : serre vs. nature

Thèse Benjamin Brachi 2007-2010

L'étude décrite juste avant a été effectuée dans des conditions contrôlées de serre. Cependant, nous avons précédemment vu que l'architecture et les régions génomiques associées à la variation naturelle de la date de floraison pouvaient être très différentes entre des conditions contrôlées de serre et des conditions écologiquement plus réalistes du terrain expérimental. Qu'en est-il de la relation entre fitness et date de floraison ? Pour répondre à cette question, la longueur totale des fruits (nombre total de fruits x longueur moyenne d'un fruit) comme proxy de la fitness a été mesurée pour toutes les plantes décrites dans la partie II.C.1.1 (200 accessions mondiales et 4366 RILs) aussi bien en serre que sur le terrain expérimental.

En se basant sur l'ensemble des lignées génétiques, la variation de la date de floraison explique 39.6% de la variation de la fitness en serre ; alors que la variation de la date de floraison explique seulement 5.4% de la variation de la fitness sur le terrain expérimental, (Figure 19). En serre, une forte sélection directionnelle non-linéaire significative a été trouvée pour la précocité de floraison (fitness relative = $-0.40 \times \text{floraison} + 0.09 \times \text{floraison}^2$). Le pattern de sélection sur le terrain expérimental est assez différent : une sélection stabilisante significative combinée à une faible sélection directionnelle significative pour la précocité de floraison ont été détectées (fitness relative = $-0.06 \times \text{floraison} - 0.25 \times \text{floraison}^2$).

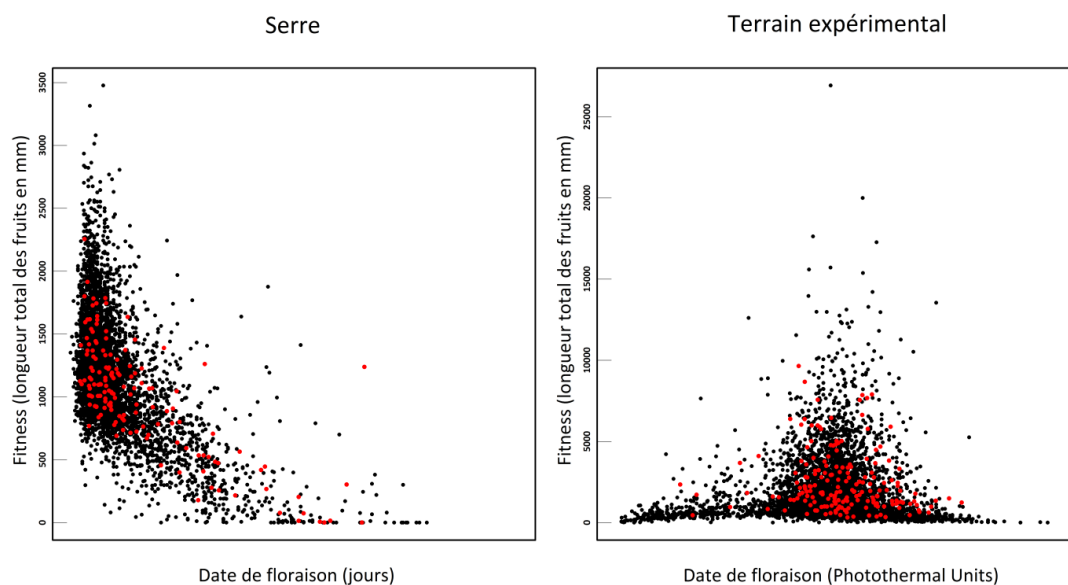


Figure 19. Relation entre fitness et date de floraison suivant deux environnements : serre vs. terrain expérimental. Les points en noir correspondent aux moyennes génotypiques par RIL. Les points en rouge correspondent aux moyennes génotypiques par accessions naturelle.

4. Conclusion

En y regardant de plus près, je pense que j'ai abordé cette thématique du mauvais côté. La date de floraison étant décrite comme un trait adaptatif à l'échelle mondiale chez *A. thaliana*, je me suis déjà focalisé sur l'architecture génétique et l'identification des régions génomiques associées à la variation de ce trait ; et ceci **en m'appuyant sur la combinaison innovante GWA mapping – QTL mapping et avec la volonté de replacer *A. thaliana* dans un contexte écologique réaliste.** J'avais cependant omis un détail important : le caractère supposé adaptatif de la date de floraison avait été mise en évidence dans des conditions contrôles de serre. Et en effet, des différences nettes ont été observées entre la serre et le terrain expérimental. On pourrait penser que ces deux environnements sont complémentaires dans l'étude génétique de la date de floraison. Mais au final quel est l'intérêt de phénotyper des plantes dans des conditions environnementales qui n'existent pas sur l'aire de distribution d'*A. thaliana* ? Je conçois très bien que ces conditions contrôlées soient importantes pour l'identification des différentes voies physiologiques sous-jacentes à la date de floraison. Mais dans le cas de l'étude de l'adaptation de la date de floraison, ne vaudrait-il pas mieux se focaliser sur un phénotypage dans différents environnements naturels ?

Combiner à la fois l'étude d'association écologie – génétique et des mesures de fitness dans des conditions écologiques réalistes a permis de mettre en évidence une adaptation au climat chez *A. thaliana*. Cependant l'absence d'un enrichissement de gènes de floraison associés aux facteurs climatiques a suggéré que la date de floraison n'était pas un trait fortement lié à la fitness à l'échelle mondiale. Cette conclusion semble confirmée par le faible pourcentage (5.4%) de variation de fitness expliquée par la date de floraison sur le terrain expérimental. Deux hypothèses peuvent être alors émises :

- **soit l'étude de la valeur adaptative de la date de floraison doit être étudiée à d'autres échelles géographiques que l'échelle mondiale.** En se basant sur un échantillonnage hiérarchique de populations françaises d'*A. thaliana*, les fortes corrélations phénologie – écologie observées à une échelle régionale semble suggérer qu'une étude de l'adaptation de la phénologie à cette échelle serait plus appropriée. Les résultats du GWA mapping à différentes échelles spatiales vont dans ce sens. Mais il ne faut pas oublier que cette étude a été réalisée dans des conditions de serre.
- **soit une grande proportion de la variation de la date de floraison est neutre.** Là aussi, en se basant sur un échantillonnage hiérarchique de populations françaises d'*A. thaliana*, nous avons pu mettre en évidence qu'une grande partie de la variation

phénologique pouvait être expliquée par la diversité génétique neutre. Mais dans ce cas, quels sont les traits phénotypiques sous sélection chez *A. thaliana* ?

Bien que le phénotypage apparaisse de plus en plus comme le facteur limitant en comparaison avec les développements technologiques en génomique, la première étape de l'étude de l'adaptation chez *A. thaliana* devrait être le phénotypage de nombreux traits phénotypiques (y compris la fitness), et ceci dans plusieurs environnements écologiquement réalistes à partir d'un échantillonnage hiérarchique de populations naturelles. Une fois les traits adaptatifs et leur échelle de variation naturelle correspondante identifiés, des projets visant à déterminer l'architecture génétique et à identifier les régions génomiques associées à cette variation naturelle peuvent être envisagés.

III. Phénotype global et normes de réaction

Je poursuis mes recherches sur l'adaptation chez *A. thaliana* (1) en étudiant cette fois-ci plusieurs traits (y compris la fitness) de manière simultanée ; et (2) en phénotypant des lignées génétiques non plus dans un seul environnement, mais dans plusieurs environnements.

En ce qui concerne le premier point, outre des traits phénologiques, de nouveaux traits supposés adaptatifs sont pratiquement systématiquement mesurés lors des expériences: croissance végétative et forme de la rosette (accumulation des ressources), hauteur de la plante et nombre de branches (allocation des ressources et dispersion des graines), herbivorie, attaque par des pathogènes (champignons et bactéries).

Sur le deuxième point, je m'intéresse à l'effet de différents facteurs abiotiques et biotiques (pouvant être rencontrés par *A. thaliana* dans son milieu naturel) sur l'expression des traits phénotypiques et de leur relation avec la fitness. Phénotyper des plantes dans plusieurs environnements écologiquement réalistes amène tout naturellement à **étudier les interactions 'génotype x environnement', i.e. la variation naturelle génétique des normes de réaction** (NB : si nous définissons la plasticité phénotypique comme la capacité d'un génotype à produire différents phénotypes selon différents environnements, cette propriété est graphiquement représentée par la relation entre le phénotype exprimé par un génotype et la gamme de valeurs d'un facteur environnemental, relation qu'on appelle communément norme de réaction). Donc, très récemment, mes recherches en génomique écologique visent aussi à

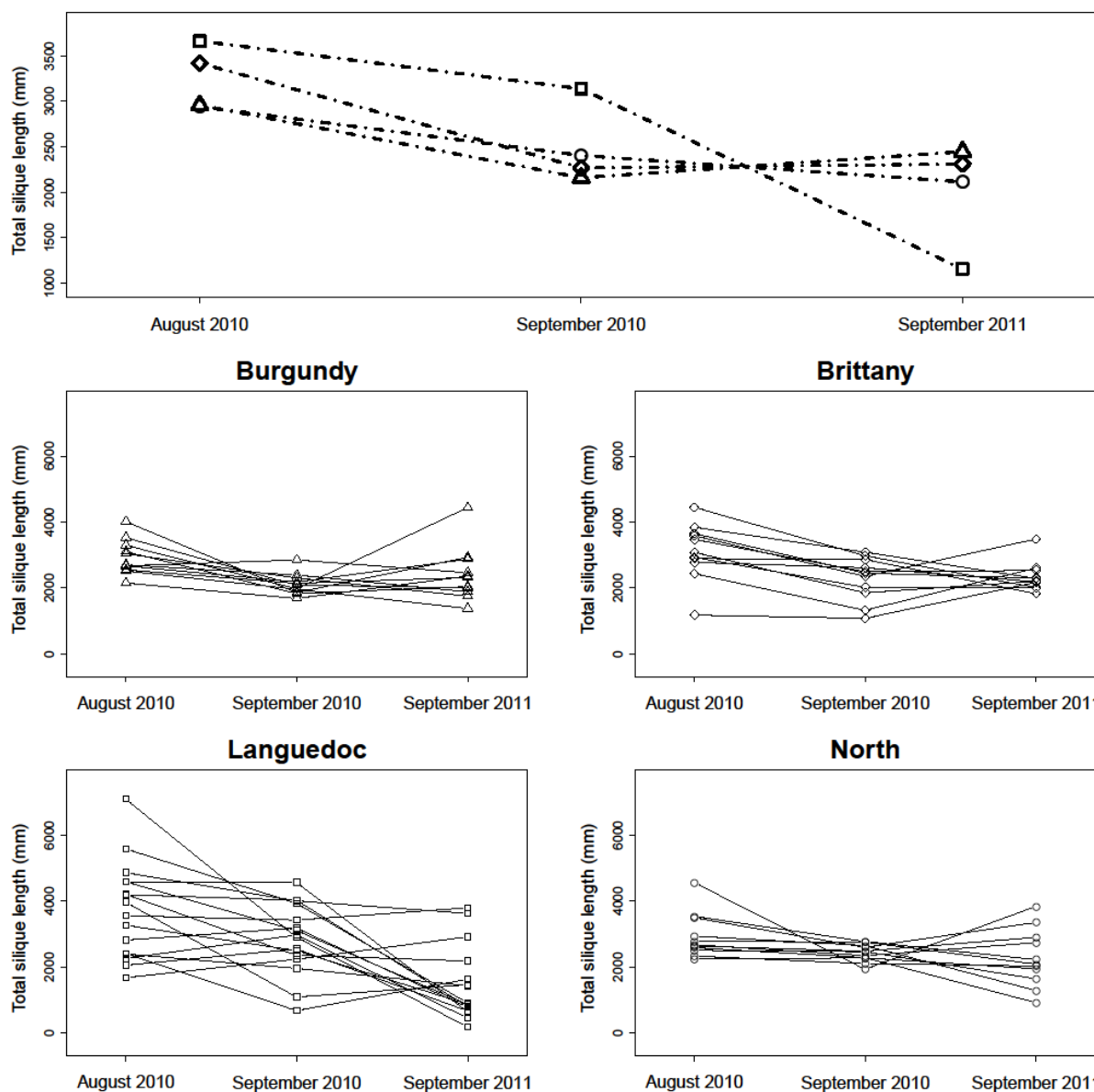


Figure 20. Norme de réaction de la fitness à différentes échelles géographiques. Pour la figure du haut, chaque point représente la moyenne entre populations pour une région donnée : Δ : Bourgogne, $\diamond$: Bretagne, $\square$: Languedoc, $\circ$: Nord. Pour les 4 figures du bas, chaque point correspond à la moyenne entre lignées pour une population donnée.

étudier l'échelle géographique et la valeur adaptative de la variation génétique naturelle des normes de réaction ; et par la suite à identifier les régions génomiques associées à cette variation naturelle.

Comme énoncé dans l'introduction, cette troisième partie de ce document correspondra plus à un état des lieux de mes projets de recherche en cours, qu'à une série de résultats déjà publiés. Par ailleurs, bien que les différents facteurs biotiques et abiotiques testés peuvent apparaître nombreux et sans lien évident, ils permettront (je l'espère) de tirer des généralités quant à l'importance de la variation naturelle des normes de réaction dans les capacités d'adaptation chez *A. thaliana*.

A. Cohorte de germination et effet année

Thèse Romain Villoutreix 2010-2013

Quelques études ont démontré l'importance de la date de germination sur l'expression des traits post-germination chez *A. thaliana* (Donohue *et al.* 2005, Wilczek *et al.* 2009). De même, dans certaines régions géographiques où l'on trouve *A. thaliana*, le climat peut être très variable d'une année sur l'autre (Hancock *et al.* 2011). Afin d'identifier les bases génétiques et tester la valeur adaptative de la variation génétique naturelle des réponses phénotypiques à ces environnements rencontrés par *A. thaliana*, nous avons établi une expérience sur un terrain expérimental de l'Université de Lille 1 où 800 lignées génétiques issues des 49 populations collectées en France (partie II.C.2.2) ont été semées fin août et fin septembre 2010 (correspondant au début et à la fin de la principale cohorte de germination principale dans le Nord de la France) et fin septembre 2011. Les deux années ont été très contrastées au niveau climatique avec un automne dans les normales saisonnières, un hiver froid et très enneigé, et un printemps chaud et sec pour l'année 2010-2011 ; un automne chaud et sec, un hiver froid mais peu enneigé, et un printemps froid et pluvieux pour l'année 2011-2012.

Les premières analyses statistiques ont révélé une forte variation naturelle des normes de réaction pour tous les traits mesurés y compris la fitness (Figure 20), aussi bien entre les régions qu'entre les populations au sein des régions et qu'entre les lignées génétiques au sein des populations. Les analyses statistiques pour identifier les traits sous sélection suivant les 4 régions et les 3 environnements testés sont en cours ; ainsi que les analyses statistiques visant

à déterminer si les normes de réaction entre les deux cohortes de germination ou entre les deux années sont adaptatives, neutres ou maladaptatives.

Afin d'identifier les régions génomiques associées à la fitness mais aussi à des facteurs écologiques et présentant des traces de sélection (Figure 21), nous avons entrepris en collaboration avec Joy Bergelson (University of Chicago, USA) le séquençage des 49 populations en bulkant les ADN par population (couverture attendu = 30X par population). Le séquençage est presque achevé et l'analyse des données de séquençage permettra d'obtenir les fréquences alléliques des polymorphismes génétiques le long du génome. Un scan génomique sera alors effectué pour d'identifier les gènes liés à la fitness, déterminer le type de sélection liée à ces gènes mais aussi identifier les facteurs écologiques sous-jacents à la sélection naturelle. Une comparaison entre les 4 régions permettra de déterminer si ce sont les mêmes gènes liés à la fitness, les mêmes types de sélection et enfin les mêmes facteurs écologiques.

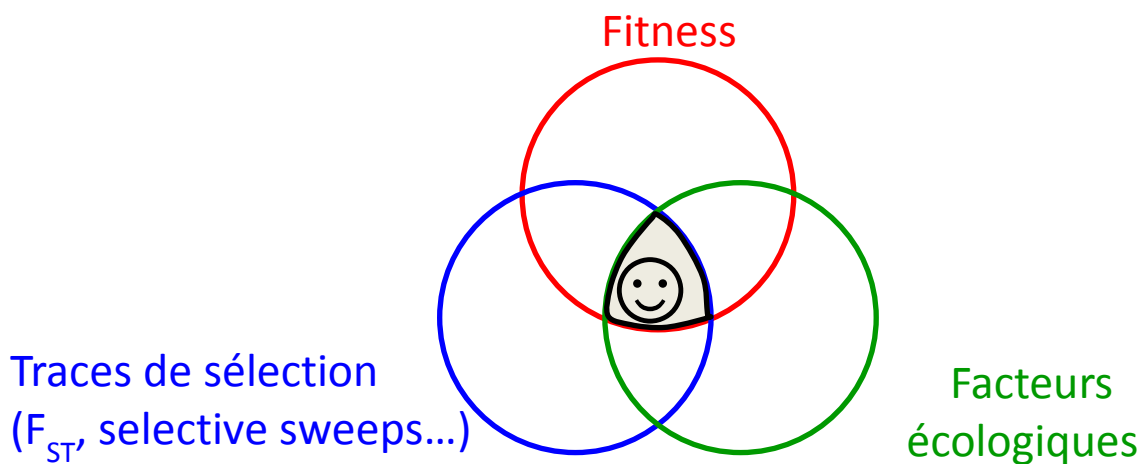


Figure 21. Connections entre différentes approches pour étudier l'adaptation.

Ce type d'analyses pourra être réalisé non seulement sur la fitness mais aussi sur tous les traits phénotypiques et leurs normes de réaction associées. Nous pourrions ainsi tester si les gènes de plasticité sont plus (ou moins) sous sélection que les gènes liés à l'adaptation locale.

B. Enfouissement des graines x cohorte de germination

Projet Nathalie Faure, IE Service commun des serres de l'Université de Lille 1

Des lots de graines de 200 accessions mondiales génotypées pour 214k SNPs ont été enfouis à 5-10 cm de profondeur dans le sol au mois de juin 2010, simulant ainsi un enfouissement naturel. En parallèle, des lots de graines ont été conservés à température ambiante dans un bureau (traitement de 'non enfouissement'). La sortie et le semis de ces lots de graines à 3 périodes différentes de l'année (fin août et fin septembre correspondant au début et à la fin de la principale cohorte de germination principale dans le Nord de la France + début mars correspondant à la deuxième cohorte de germination dans le Nord de la France) ont permis de tester l'effet de l'interaction 'enfouissement x cohorte de germination' sur l'expression des traits post-germination chez les 200 accessions mondiales. Les premiers résultats sur la date de floraison indiquent de fortes interactions 'accession x enfouissement x cohorte de germination'. Par exemple, semées fin août – fin septembre, des accessions du nord de la Suède ne fleurissent qu'au printemps prochain quand les graines ont été stockées à température ambiante. Quand les graines de ces mêmes accessions sont enfouies, certaines de ces accessions fleurissent dès le mois d'octobre. L'estimation de la fitness de ces plantes devrait nous indiquer si les normes de réaction à l'enfouissement et/ou la cohorte de germination sont adaptatives, neutres ou maladaptatives. Les analyses de GWA mapping devraient nous permettre d'identifier les régions génomiques associées à la variation naturelle des normes de réaction à l'enfouissement et/ou la cohorte de germination.

C. Interactions plante – plante

Thèse Etienne Baron 2011-2014

Bien qu'étant considérée comme une espèce pionnière, *A. thaliana* cohabite avec de nombreuses autres espèces végétales dans certaines populations naturelles, et ceci sur plusieurs générations (observations personnelles). Afin d'identifier les régions génomiques associées à la réponse à la compétition par *A. thaliana* face à divers compétiteurs, 51 lignées d'*A. thaliana* issues de la population naturelle polymorphe bourguignonne TOU-A et génotypées pour 214k SNPs ont été mises en situation de compétition intraspécifique et de situation de compétition interspécifique avec 4 espèces compétitrices que l'on retrouve dans

la communauté naturelle de la pop TOU-A : *Poa annua* (port droit), *Stellaria media* et *Veronica arvensis* (port rampant), *Trifolium repens* (port intermédiaire).

Les premiers résultats indiquent des interactions ‘lignée x traitement de compétition’ très importantes pour la croissance végétative et la date de floraison. Les analyses de GWA mapping réalisées sur les 51 lignées ont révélé des régions génomiques différentes suivant l’espèce compétitrice. Par exemple, pour la date de floraison, le pic d’association localisé au début du chromosome 5 en absence de compétition disparaît dans le traitement de compétition interspécifique avec *Stellaria media* (Figure 22). Là encore, l’estimation de la fitness de ces plantes devrait nous indiquer si les normes de réaction de réponse à la compétition intra- ou interspécifique sont adaptatives, neutres ou maladaptatives.

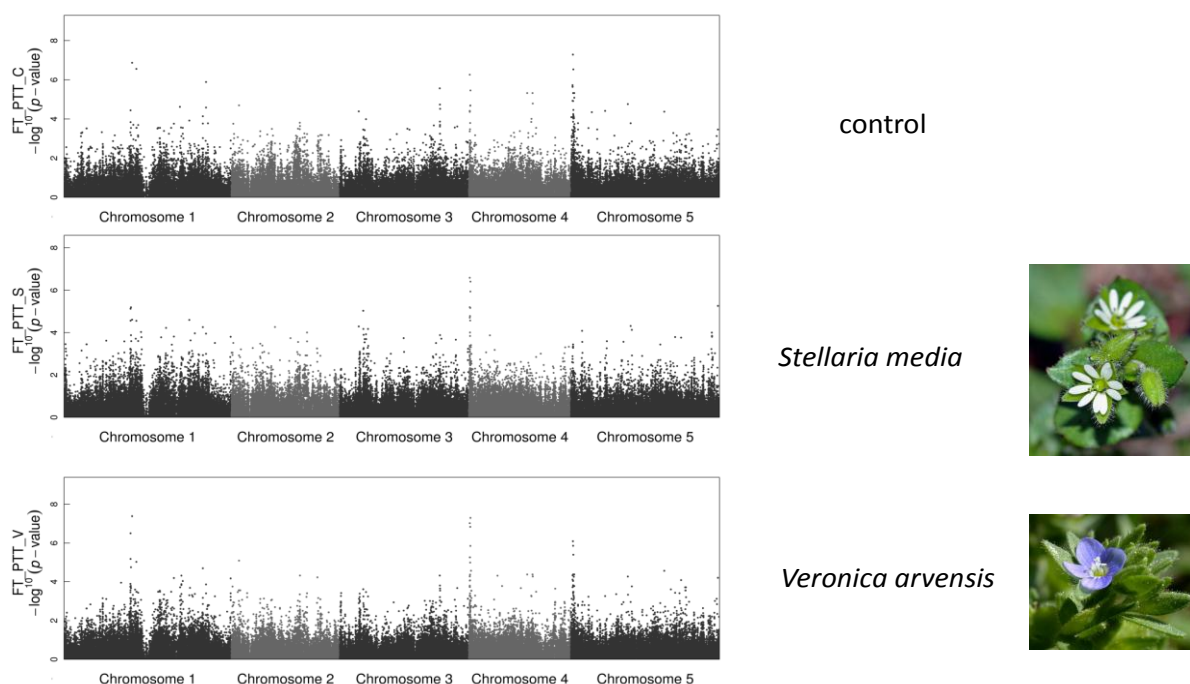


Figure 22. Identification des régions génomiques associées à la variation de la date de floraison selon différents traitements de compétition (control : absence de compétition, compétition interspécifique avec *Stellaria media* et *Veronica arvensis*). L’axe des abscisses indique la position physique des SNPs le long du chromosome. Les 5 chromosomes sont représentés avec différentes nuances de gris. L’axe des ordonnées correspond à la valeur d’association entre chaque SNP et la date de floraison selon un modèle statistique tenant compte de la structure des populations (EMMAX). FT_PTT : date de floraison exprimée en unités photothermales.

D. Environnement cytoplasmique

Moison et al. 2010 Plant Journal, Budar & Roux 2011 Plant Signaling and Behavior, ANR BIOADAPT CYTOPHENO 2012 – 2016

Dans un projet dirigé par Françoise Budar (INRA Versailles), nous avons récemment suggéré une co-adaptation cyto-nucléaire au sein d'*A. thaliana*. En effet, en se basant sur des croisements réciproques entre accessions naturelles, un effet du cytoplasme sur la capacité de germination des graines a été détecté. Pour continuer à étudier cette co-adaptation cyto-nucléaire, 56 cytolignées ont été créées à partir de 8 accessions mondiales: une cytolignée correspond à une accession dont le cytoplasme d'origine a été remplacé par le cytoplasme d'une autre accession. Pour une accession donnée, nous avons donc 7 cytolignées qui ont le même génome nucléaire que l'accession étudiée mais chacune avec le cytoplasme d'une autre des 7 accessions. L'étude de ces cytolignées a été financée par un projet ANR BIOADAPT où ma participation consistera à phénotyper ces 56 cytolignées ainsi que les 8 accessions parentales dans un contexte écologique réaliste. Comme les séquences génomiques nucléaires et cytoplasmiques seront disponibles pour ces 8 accessions, un des objectifs finaux sera de d'identifier les régions génomiques en interaction entre les deux génomes et qui affectent un phénotype donné.

IV. Perspectives : génomique écologique de l'adaptation dans les communautés végétales

Dans le contexte des changements globaux, un des défis en écologie évolutive est de déterminer le type et la structure de la variation génétique qui sous-tend les traits phénotypiques pour évaluer le potentiel et la dynamique d'une réponse adaptative à la sélection naturelle et / ou aux changements environnementaux d'origine anthropique. Un éventuel manque de réponse des espèces à la sélection face aux changements globaux (comme le réchauffement climatique) pourrait causer une érosion de la biodiversité, ce qui pourrait perturber durablement les écosystèmes. Ceci est d'autant plus vrai pour les espèces végétales dont la dispersion est en moyenne limitée par rapport aux agents pathogènes et aux animaux.

Précédemment dans ce document, une étude sur l'adaptation au climat a été décrite avec le résultat principal suivant : les polymorphismes les plus associés au climat sont 3 fois plus associés à des balayages sélectifs récents que des polymorphismes choisis au hasard sur le génome. Ce résultat suggère qu'une espèce comme *A. thaliana* pourrait rapidement atteindre une limite adaptative dans le contexte d'un changement climatique rapide, ceci à cause des contraintes imposées par l'attente de nouvelles mutations. En revanche, une autre étude sur l'adaptation au climat chez *A. thaliana* suggère que la sélection a surtout été basée sur la variation génétique déjà existante (Fournier-Level *et al.* 2011).

Bien que ces différents résultats entre ces deux études restent à expliquer, nous devons garder à l'esprit que **la réponse adaptative potentielle d'une espèce aux changements globaux ne dépendra pas seulement de ses interactions directes avec l'environnement abiotique, mais aussi des effets de l'environnement sur ses interactions biotiques** (interactions avec les pathogènes, compétition intra- et inter-spécifique ...). Cependant, nos connaissances sur la dynamique adaptative des systèmes interactifs organisme – organisme, et plus précisément sur la co-évolution entre deux espèces en interaction, reste étonnamment très limitée. Ceci est particulièrement pertinent lorsque l'on étudie le recrutement des mutations chez les deux partenaires biotiques conférant les variants phénotypiques susceptibles d'être retenues par la sélection naturelle. Parce que le paysage moléculaire des interactions organisme – organisme (interactions hôte-pathogène, interactions plante-plante) ne peut être approché qu'en tenant compte de leur spécificité génétique, la caractérisation des bases moléculaires sous-jacentes aux spécificités hôte-pathogène et aux spécificités plante-plante

permettra d'augmenter considérablement notre compréhension et les prédictions sur l'émergence des maladies et l'évolution des communautés végétales, respectivement.

L'objectif de ce nouveau projet est toujours d'établir un lien entre la biologie moléculaire et écologie évolutive, mais ceci afin de prédire les trajectoires évolutives des communautés de plantes et des agents pathogènes face aux changements globaux. Les technologies NGS et le génotypage par séquençage (Genotyping-by-Sequencing, GBS) facilitent déjà l'accès à des polymorphismes causaux associés à la variation naturelle des traits complexes. Le coût sans cesse décroissant du séquençage est une incroyable opportunité pour cartographier les bases génétiques de caractères complexes chez les espèces non-modèles et les replacer dans un contexte écologique.

Pour mener à bien ce projet, je pense l'aborder selon 3 axes:

- je vais tout d'abord continuer à travailler sur l'identification des bases génétiques de l'interaction dans le pathosystème *A. thaliana* – *Xanthomonas* sp. J'ai d'ores et déjà abordé cette problématique dans la partie II.B.2.2.2. (page 35, « Souches isolées dans des populations naturelles d'*A. thaliana* ») et je ne reviendrais donc pas dessus.
- ensuite, je souhaite prolonger l'identification des bases génétiques des interactions entre deux partenaires dans d'autres pathosystèmes naturels, impliquant onze espèces de plantes communément associées à *A. thaliana* dans des communautés végétales.
- enfin, je souhaite commencer à élucider les bases génétiques des interactions entre deux partenaires dans un système plante - plante, toujours sur la base d'*A. thaliana* et des onze espèces de plantes communément associées à *A. thaliana* dans des communautés végétales.

A. Interactions plante – pathogène

Combiner génomique des populations et études écologiques et phénotypiques chez de multiples espèces qui occupent les mêmes habitats devrait permettre **des études comparatives de la variation génétique adaptative entre des espèces qui ont potentiellement évolué en parallèle en présence des mêmes pressions de sélection** (Brachi *et al.* 2011). J'ai l'intention de collecter 100 communautés végétales françaises, toutes incluant *A. thaliana*. En France, les populations d'*A. thaliana* se trouvent dans des habitats naturels ou dans des habitats d'origine anthropique. Cinquante populations d'*A. thaliana* seront

échantillonnés dans des écosystèmes naturels, tandis que les cinquante autres populations d'*A. thaliana* seront échantillonnées dans des agro-systèmes. Un tel échantillonnage permettra de tester si le potentiel adaptatif face aux changements globaux peut différer entre les écosystèmes naturels et les agro-systèmes. Onze autres espèces communément associées à *A. thaliana* dans les communautés végétales en France (dont certaines espèces ayant différents traits d'histoire de vie) seront collectées en même temps que *A. thaliana* (Tableau 3). Dans chaque communauté, les graines de 10 à 30 individus seront récoltées par espèce végétale et constitueront ainsi des familles génétiques.

Tableau 3. Espèces communément associées à *A. thaliana* dans les communautés végétales.

Name	Family	Life-cycle	Mating system
<i>Arenaria serpyllifolia</i>	Caryophyllaceae	Annual	Entomogamy
<i>Capsella bursa-pastoris</i>	Brassicaceae	Annual-Bisannual	Autogamy
<i>Cardamine hirsuta</i>	Brassicaceae	Annual	Autogamy
<i>Cerastium glomeratum</i>	Caryophyllaceae	Annual	Autogamy
<i>Erodium cicutarium</i>	Geraniaceae	Annual	Entomogamy and autogamy
<i>Erophila verna</i>	Brassicaceae	Annual	Entomogamy
<i>Poa annua</i>	Poaceae	Annual-Bisannual	Autogamy and wind pollination
<i>Stellaria media</i>	Caryophyllaceae	Annual	Entomogamy
<i>Trifolium arvense</i>	Fabaceae	Annual-Bisannual	Entomogamy and autogamy
<i>Veronica arvensis</i>	Plantaginaceae	Annual	Entomogamy and autogamy
<i>Vicia sativa</i>	Fabaceae	Annual	Entomogamy and autogamy

Dans chaque communauté, chaque espèce végétale sera caractérisée au niveau génomique par séquençage du génome complet des bulks d'ADN de tous les individus échantillonnés. Cette étape permettra d'obtenir les fréquences alléliques des polymorphismes génétiques le long du génome pour chaque espèce végétale et dans chaque communauté.

Dans chaque communauté, la description des communautés de pathogènes potentiels (bactéries et champignons) sera effectuée pour chaque espèce végétale par séquençage ciblé des régions 16S et ITS afin de classer ces pathogènes suivant différents OTUs (Operational Taxonomic Units).

La description génomique conjointe des communautés de plantes et de pathogènes permettra de tester si (par scan génomique chez les plantes) les régions génomiques associées à des descripteurs de communautés de pathogènes (comme la présence-absence d'une espèce pathogène) sont les mêmes entre des espèces végétales soumises aux mêmes pressions de sélection.

Si la présence d'*A. thaliana* se trouve être corrélée dans les populations naturelles à la présence d'une espèce particulière de pathogène, une expérience sera mis en place pour inoculer une accession d'*A. thaliana* de chaque population naturelle avec une souche du pathogène de chaque population naturelle. Un tel design expérimental permettra de tester à la fois l'adaptation locale et une potentielle coévolution dans un système hôte-pathogène. La caractérisation génomique des deux partenaires aidera à l'identification des mécanismes moléculaires de cette spécificité plante-pathogène génétique.

Ce projet devrait permettre d'apporter des éléments de réponse aux questions suivantes :

- Est-ce que les gènes associés à la variation naturelle de la résistance aux pathogènes sont identiques entre plusieurs espèces végétales qui ont potentiellement fait face aux mêmes pressions de sélection exercées par les pathogènes?
- Observe-t-on un effet des traits d'histoire de vie (comme le système de reproduction) sur le type de variantes génétiques qui pourraient être recrutés par la sélection naturelle?
- La détection des gènes en interaction entre les deux partenaires plante – pathogène permet-elle de mettre en évidence une potentielle coévolution dans des pathosystèmes autres que *A. thaliana* – *Xanthomonas* sp.?
- Quels sont les mécanismes moléculaires qui sous-tendent cette interaction plante - pathogène?

B. Interactions plante - plante

Bien que la spécialisation biotique entre génotypes voisins de deux espèces végétales ait été suggérée de plus de 25 ans (Aarssen and Turkington 1985), les bases génétiques de cette interaction restent à élucider. Profitant de l'échantillonnage et de la caractérisation génomique de 100 communautés végétales (décrit ci-dessus), je vais me concentrer sur l'espèce végétale la plus associée à *A. thaliana* à travers les 100 communautés végétales. Bien qu'il ait souvent été supposé que la compétition avait peu d'importance dans l'élaboration des stratégies de vie chez des plantes annuelles comme *A. thaliana*, une variation génétique pour les stratégies de compétition a déjà été mis en évidence chez *A. thaliana* (Mutic and Wolf 2007). Par ailleurs,

les régions génomiques associées à la réponse à la compétition interspécifique chez *A. thaliana* ont été trouvées différentes entre plusieurs espèces compétitrices (Figure 22).

Si la présence d'*A. thaliana* se trouve être corrélée à la présence d'une espèce végétale particulière à travers les 100 communautés végétales, une expérience sera mise en place pour mettre en compétition une accessions d'*A. thaliana* de chaque population naturelle avec une accessions de chaque population naturelle de l'autre espèce végétale. A partir de la matrice phénotypique de la réponse à la compétition obtenue, nous pourrions tester à la fois l'adaptation locale et une potentielle coévolution dans un système interactif plante-plante. La caractérisation génomique des deux partenaires facilitera l'identification des mécanismes moléculaires de cette spécificité génétique plante - plante.

Ce projet devrait permettre d'apporter des éléments de réponse aux questions suivantes :

- Dans le cas de la compétition interspécifique, le potentiel pour la plasticité phénotypique et les interactions génotype x environnement varient-ils selon les espèces végétales?
- La détection des gènes en interaction entre les deux partenaires plante – plante permet-elle de mettre en évidence une potentielle coévolution ?
- Quels sont les mécanismes moléculaires qui sous-tendent cette interaction plante - plante?

Références bibliographiques

- Aarssen, L.W. and R. Turkington 1985 Biotic specialization between neighbouring genotypes in *Lolium perenne* and *Trifolium repens* from a permanent pasture. *J. Ecol.* 73: 605-614.
- Andrés, F. and G. Coupland 2012 The genetic bases of flowering responses to seasonal cues. *Nat. Rev. Genet.* 13: 627-639.
- Alonso, J.M. and J.R. ECKER 2006 Moving forward in reverse: genetic technologies to enable genome-wide phenomic screens in *Arabidopsis*. *Nat. Rev. Genet.* 7: 524-536.
- Aranzana, M.J., S. Kim, K. Zhao, E. Bakker, M. Horton *et al.* Genome-Wide Association mapping in *Arabidopsis* identifies previously known flowering time and pathogen resistance genes. *PLoS Genet.* 1(5): e60.
- Bakker, E.G., C. Toomajian, M. Kreitman and J. Bergelson 2006 A genome-wide survey of *R* gene polymorphisms in *Arabidopsis*. *Plant Cell* 18: 1803-1818
- Barton, N.H and P.D. Keightley 2002 Understanding quantitative genetic variation. *Nat. Rev. Genet.* 3: 11-21.
- Bomblies, K., L. Yant, R.A. Laitinen, S.-T. Kim, J.D. Hollister *et al.* 2010 Local-scale patterns of genetic variability, outcrossing, and spatial structure in natural stands of *Arabidopsis thaliana*. *PLoS Genet.* 6: e1000890.
- Bourguet, D., T. Guillemaud, C. Chevillon and M. Raymond 2004 Fitness costs of insecticide resistance in natural breeding sites of the mosquito *Culex pipiens*. *Evolution* 58: 128-135.
- Cao, J., K. Schneeberger, S. Ossowski, T. Günther, S. Bender *et al.* 2011 Whole-genome sequencing of multiple *Arabidopsis thaliana* populations. *Nat. Genet.* 43(10): 956-963.
- Caprio, M.A. 2001 Source-sink dynamics between transgenic and non-transgenic habitats and their role in the evolution of resistance. *J. Econ. Entomol.* 94: 698-705.
- Cook, D.E., T.G. Lee, X. Guo, S. Melito, K. Wang *et al.* 2012 Copy number variation of multiple genes at *Rgh1* mediates nematode resistance in soybean. *Science* 338: 1206-1209.
- Darwin, C.R. 1859 The Origin of Species. J. Murray, London.
- Donohue, K., L. Dorn, C. Griffith, E. Kim, A. Aguilera *et al.* 2005 Niche construction through germination cueing: life-history responses to timing of germination in *Arabidopsis thaliana*. *Evolution* 59: 771-785.
- Elena, S.F. and R.E. Lenski R.E. 2003 Evolution experiments with microorganisms: the dynamics and genetic bases of adaptation. *Nat. Rev. Genet.* 4: 457-469.
- Fisher, R.A. 1930 The Genetical Theory of Natural Selection. Oxford Univ. Press, Oxford.

- Fournier-Level, A., A. Korte, M. Cooper, M. Nordborg, J. Schmitt *et al.* 2011 A map of local adaptation in *Arabidopsis thaliana*. *Science* 334: 86-95.
- Friesen, L.J.S., G.M. Ferguson and J.C. Hall 2000 Management strategies for attenuating herbicide resistance: untoward consequences of their promotion. *Crop Prot.* 19: 891-895.
- Fu, D., C. Uauy, A. Distelfeld, A. Blechl, L. Epstein *et al.* 2009 A kinase-START gene confers temperature-dependent resistance to wheat stripe rust. *Science* 323: 1357-1360.
- Fukuoka, S., N. Saka, H. Koga, K. Ono, T. Shimizu *et al.* 2009 Loss of function of a proline-containing protein confers durable disease resistance in rice. *Science* 325: 998-1001.
- Gaut, B. 2012 *Arabidopsis thaliana* as a model for the genetics of local adaptation. *Nat. Genet.* 44: 115-121.
- Haldane, J.B.S. 1924 A mathematical theory of natural and artificial selection, part I. *Trans. Camb. Philo. Soc.* 23: 19-41.
- Haldane, J.B.S. 1927 A mathematical theory of natural and artificial selection, part V: selection and mutation. *Proc. Camb. Philo. Soc.* 28: 838-844.
- Haldane, J. B.S. 1948 The theory of cline. *J. Genet.* 48: 277-284.
- Hermisson, J. and Pennings, P.S. 2005 Soft sweeps: molecular population genetics of adaptation from standing genetic variation. *Genetics* 169: 2335–2352
- Horton, M., A. Hancock, Y. Huang, C. Toomajian, S. Atwell *et al.* 2012 Genome-wide patterns of genetic variation in worldwide *Arabidopsis thaliana* accessions from the RegMap panel. *Nat. Genet* 44: 212-218.
- Jakob, K., E.M. Goss, H. Araki, T. Van, M. Kreitman *et al.* 2002 *Pseudomonas viridiflava* and *P. syringae*—natural pathogens of *Arabidopsis thaliana*. *Mol. Plant Microbe In.* 15: 1195–1203.
- Johannes, F. E. Porcher, F.K. Teixeira, V. Saliba-Colombani, M. Simon *et al.* 2009 Assessing the impact of transgenerational epigenetic variation on complex traits. *PLoS Genet.* 5: e10000530.
- Johanson, U., J. West, C. Lister, M. Scott, R. Amasino *et al.* 2000 Molecular analysis of *FRI*, a major determinant of natural variation in *Arabidopsis* flowering time. *Science* 290: 344-347.
- Kassen, R. 2002. The experimental evolution of specialists, generalists, and the maintenance of diversity. *J. Evolution. Biol.* 15: 173-190.
- Kang, H.M., J.H. Sul, S.K. Service, N.A. Zaitlen, S.-Y. Kong *et al.* 2010 Variance component model to account for sample structure in genome-wide association studies. *Nat. Genet.* 42: 348-402.

- Kawecki, T.J. and D. Ebert 2004 Conceptual issues in local adaptation. *Ecol. Lett.* 7: 1225-1241.
- Kim, S., V. Plagnol, T.T. Hu, C. Toomajian, R.M. Clark *et al.* 2007 Recombination and linkage disequilibrium in *Arabidopsis thaliana*. *Nat. Genet.* 39: 1151-1155.
- Kimura, M. 1955 Random genetic drift in multi-allelic locus. *Evolution* 9: 419-435.
- Kniskern, J.M., B.M. Traw and J. Bergelson 2007 Salicylic acid and jasmonic acid signaling defense pathways reduce natural bacterial diversity on *Arabidopsis thaliana*. *Mol. Plant Microbe In.* 20(12): 1512-1522.
- Krattinger, S.G., E.S. Lagudah, W. Spielmeyer, R.P. Singh, J. Huerta-Espino *et al.* 2009 A putative ABC transporter confers durable resistance to multiple fungal pathogens in wheat. *Science* 323: 1360-1363.
- Lande, R. and S.J. Arnold 1983 The measurement of selection on correlated characters. *Evolution* 37: 1210-1226.
- Le Corre, V. 2005 Variation at two flowering time genes within and among populations of *Arabidopsis thaliana*: comparison with markers and traits. *Mol. Ecol.* 14: 4181-4192.
- Lenormand, T. and M. Raymond 1998 Resistance management: the stable zone strategy. *P. Roy. Soc. B – Biol. Sci.* 265: 1985-1990.
- Liu, S., P.K. Kandath, S.D. Warren, G. Yeckel, R. Heinz *et al.* 2012 A soybean cyst nematode resistance gene points to a new mechanism of plant resistance to pathogens. *Nature* 492: 256-260.
- Mitchell-Olds, T. and J. Schmitt 2006 Genetic mechanisms and evolutionary significance of natural variation in *Arabidopsis*. *Nature* 441: 947–952.
- Mutic, J.J. and J.B. Wolf 2007 Indirect genetic effects from ecological interactions in *Arabidopsis thaliana*. *Mol. Ecol.* 16: 2371-2381.
- Nagylaki T. 1975 Conditions for the existence of clines. *Genetics* 80: 595-615.
- Nordborg, M., T.T. Hu, Y. Ishino, J. Jhaverii, C. Toomajian *et al.* 2005 The pattern of polymorphism in *Arabidopsis thaliana*. *PLoS Biol.* 3: e196.
- O'Malley, R.C. and J.R. Ecker 2010 Linking genotype to phenotype using the *Arabidopsis* unimutant collection. *Plant J.* 61: 928-940.
- Orr, H.A. 1998 The population genetics of adaptation: the distribution of factors fixed during adaptive evolution. *Evolution* 52: 935-949.
- Orr, H.A. 2002 The population genetics of adaptation: the adaptation of DNA sequences. *Evolution* 56: 1317-1330.
- Orr, H.A. 2005 The genetic theory of adaptation: a brief history. *Nat. Rev. Genet.* 6: 119-127.

- Platt, A., M. Horton, Y. Huang, Y. Li, A.E. Anastasio *et al.* 2010 The scale of population structure in *Arabidopsis thaliana*. *PLoS Genet.* 6: e10000843.
- Poland, J.A., P.J. Balint-Kurti, R.J. Wisser, R.C. Pratt and R.J. Nelson 2009 Shades of gray: the world of quantitative disease resistance. *Trends Plant Sci.* 14(1): 21-29.
- Ronce, O. and M. Kirkpatrick M. 2001 When sources become sinks: migrational meltdown in heterogeneous habitats. *Evolution* 55: 1520-1531.
- Sindo, C., G. Bernasconi and C.S. Hardtke 2007 Natural genetic variation in *Arabidopsis*: tools, traits and prospects for evolutionary ecology. *Ann. Bot.* 99: 1043-1054.
- Slatkin M. 1973 Gene flow and selection in a cline. *Genetics* 75: 733-756.
- Stinchcombe, J.R., C. Weinig, M. Ungerer, K.M. Olsen, C. Mays *et al.* 2004 A latitudinal cline in flowering time in *Arabidopsis thaliana* modulated by the flowering time gene *FRIGIDA*. *Proc. Natl. Acad. Sci. USA* 101: 4712-4717.
- Swarbreck, D., C. Wilks, P. Lamesch, T.Z. Berardini, M. Garcia-Hernandez *et al.* 2008 The Arabidopsis Information Resource (TAIR): gene structure and function annotation. *Nucleic Acids Res.* 36: D1009-D1014.
- Tonsor, S.J., C. Alonso-Blanco and M. Koornneef 2005 Genes function beyond the single trait: natural variation, gene effects, and evolutionary ecology in *Arabidopsis thaliana*. *Plant Cell Environ.* 28: 2-20.
- Ungerer, M., L.C. Johnson and M.A. Herman 2008 Ecological genomics: understanding gene and genome function in the natural environment. *Heredity* 100: 178-183
- Ungerer, M.C., C.R. Linder C.R. and L.H. Rieseberg 2003 Effects of genetic background on response to selection in experimental populations of *Arabidopsis thaliana*. *Genetics* 163: 277-286.
- Vacher C., D. Bourguet, F. Rousset, C. Chevillon and M.E. Hochberg 2003 Modelling the spatial configuration of refuges for a sustainable control of pests: a case study of *Bt* cotton. *J. Evolution. Biol.* 16: 378-387.
- Via, S. and R. Lande 1985 Genotype-environment interaction and the evolution of phenotypic plasticity. *Evolution* 39: 505-522
- Weigel, D. and M. Nordborg 2005 Natural variation in Arabidopsis. How do we find the causal genes? *Plant Physiol.* 138 : 567-568.
- Wilczek, A.M., J.L. Roe, M.C. Knapp, M.D. Cooper, C. Lopez-Gallego *et al.* 2009 Effects of genetic perturbation on seasonal life history plasticity. *Science* 323: 930-934.
- Yan, L., A. Loukoianov, G. Tranquilli, M. Helguera, T. Fahima *et al.* 2003 Positional cloning of the wheat vernalization gene *VRN1*. *Proc. Natl. Acad. Sci. U.S.A.* 100: 6263-6268.

Zhao, K., M. Aranzana, S. Kim, C. Lister, C. Shindo *et al.* 2007 An *Arabidopsis* example of association mapping in structured samples. *PLoS Genet.* 3: e4.