Thèse

Présentée par

Romain VILLOUTREIX

En vue de l'obtention du grade de

Docteur en Sciences de l'Université de Lille1

Discipline : Biologie Evolutive et Ecologie

Numéro d'ordre : 41247

Echelle spatiale et temporelle de l'adaptation chez *Arabidopsis thaliana* : intégration de la plasticité phénotypique

Soutenance prévue le 16 décembre 2013, devant le jury suivant :

Irène Till-Bottraud, DR CNRS, Université Joseph Fourier……..Rapporteur

Jacqui Shykoff, DR CNRS, Université Paris XI ORSAY..…….....Rapporteur

Virginie Orgogozo, CR CNRS, Université Paris Diderot……...…Examinateur

Laurène Gay, CR INRA, SUPAGRO de Montpellier…….Examinateur

Joël Cuguen, Professeur, Université de Lille 1……...……..Directeur

Fabrice Roux, CR CNRS, Université de Lille 1……..…….Co-directeur

Remerciements

Déjà trois ans ! Je ne le réalise toujours pas, même à l'écriture de ces lignes. Quel crâneur diront certains, une thèse se doit d'être un sacerdoce, une expérience difficile, on doit « en chier » comme dirait l'expression consacrée. Je leur répondrai que le temps passe vite quand on est bien accompagné.

Ainsi je voudrais remercier Fabrice ROUX, mon codirecteur de thèse, qui m'a accompagné durant ces trois années. Il a toujours été disponible pour moi et a fait preuve d'une patience sans faille à mon égard. Un grand merci donc à toi Fabrice, faire de la recherche avec toi a toujours été un plaisir!

Je tiens également à remercier Joël CUGUEN, mon second codirecteur de thèse, pour m'avoir permis d'effectuer ma thèse au laboratoire GEPV et pour les discussions toujours intéressantes que j'ai eues avec lui. Je ne lui serai jamais assez reconnaissant de m'avoir montré l'importance de la sémantique dans les discussions scientifiques. De plus, l'évocation de son nom dans les bâtiments administratifs engendra souvent un déblocage rapide de ma situation puisqu'il exerce également la fonction de directeur de l'école doctorale de l'Université de Lille 1. J'avoue que mon esprit rancunier et un brin sadique m'ont fait abuser de ce stratagème. Un grand merci à toi Joël et toutes mes excuses au personnel administratif de l'Université.

Je tiens particulièrement à remercier le personnel de la serre et du laboratoire. Cette thèse ne serait pas ce qu'elle est actuellement sans l'aide précieuse et inestimable de Cédric GLORIEUX. Cet homme, en plus d'être très sympathique et toujours souriant, est le Lucky Luck du comptage de fruits ! Je tiens également à remercier Nathalie FAURE pour son aide précieuse lors de nombreux relevés ainsi que pour nos nombreuses discussions. Je tiens

également à remercier Stella HUYNH et Benoit BUCHER, deux étudiants de Master 1, qui m'ont particulièrement aidé durant de nombreux relevés. Enfin je souhaite remercier Cécile GODE pour son aide en biologie moléculaire et Elise BARILLOT pour de nombreux dépannages !

Je tiens également à remercier Benjamin BRACHI, ancien doctorant de l'équipe, qui m'a initié à R. Cette thèse lui doit beaucoup. Et pour rester dans le thème de la filiation doctorale, je tiens à remercier spécialement Etienne BARON, nouveau doctorant de l'équipe, pour sa bonne humeur et ses farces nombreuses et variées ! Son départ pour Toulouse va rendre triste de nombreuses personnes.

Je tiens également à remercier Antoine DORNIER, ancien ATER du laboratoire, qui m'a appris les ficelles et rouages dans l'art d'enseigner. Je tiens également à le remercier pour nos nombreuses discussions de recherche.

Je tiens également à remercier Laurent AMSELLEM, maître de conférences au laboratoire, pour son humour décapant à la pause de midi et les pas assez nombreuses après-midi guitare que nous avons eus (le week-end !). Je le remercie également d'avoir essayé de me mettre au sport, même si sa tentative ne fût pas fructueuse.

Mais la vie n'est pas faite que de recherche. Tout passionné doit tout de même, de temps en temps, relâcher son esprit. Ainsi je souhaite remercier Cyril DYMNY qui m'apprend la guitare depuis 3 ans bientôt. Il a fait, lui aussi, preuve d'un patience hors norme avec moi et a également grandement contribué, de manière indirecte, à cette thèse. J'en profite pour faire sa publicité, il a écrit de très bonnes chansons que trouverez sur internet ou en le contactant.

La vie n'est pas faite que de recherche mais en faisant de la recherche on peut faire sa vie. En effet, durant ma thèse au GEPV j'ai rencontré ma compagne, Lieselot NGUYEN. Je

souhaite la remercier pour toute l'aide et l'amour qu'elle m'apporte, cette thèse lui doit beaucoup. Etant elle aussi en thèse, je lui souhaite toute la réussite possible.

Je tiens enfin à remercier mes parents qui m'ont supporté durant toutes mes études. Je me souviendrai toujours de cette discussion dans la voiture où je vous annonçais que la formation de l'IUT ne me plaisait pas après seulement trois jours de cours. Un grand merci à mon père, qui m'a conseillé ce jour-là de ne pas perdre de temps et de rejoindre immédiatement le cursus universitaire général. Merci de manière globale à toute ma famille et mes amis.

Mes sincères excuses aux personnes que j'aurais oubliées de remercier.

SOMMAIRE :

Les articles marqués d'une étoile figurent dans ce manuscrit.

Article 1*: B. BRACHI[§], **R. VILLOUTREIX[§],** N. FAURE, N. HAUTEKEETE,Y. PIQUOT, M. PAUWELS, D. ROBY, J. CUGUEN, J. BERGELSON and F. ROUX. 2013. Investigation of the geographical scale of adaptive phenological variation and its underlying genetics in *Arabidopsis thaliana*. *Molecular Ecology* 22: 4222-4240. [§]**These authors contributed equally to this work.**

Article 2*: **R. VILLOUTREIX**, C. GLORIEUX, B. BRACHI, J. BERGELSON, J. CUGUEN and F. ROUX Evidence of selection acting on reproductive strategies in *Arabidopsis thaliana* at different spatial scales: when temporal heterogeneity in the field matters! *Soumis à Molecular Ecology.*

Article 3* : **R. VILLOUTREIX**, C. GLORIEUX, J. CUGUEN and F ROUX. Adaptive value and costs of phenotypic plasticity across germination cohorts and years in four regional sets of natural *Arabidopsis thaliana* families. *A soumettre un fois l'article n°2 accepté.*

Article 4* : **R. VILLOUTREIX**, V. LE CORRE, E. BARON, L. AMSELLEM and F. ROUX. Rapid phenotypic evolution and fine-grained spatial variation in a natural population of *Arabidopsis thaliana*: a resurrection study. *En préparation*.

Article 5 : **R. VILLOUTREIX** and F. ROUX. Genome-wide association mapping in *Arabidopsis thaliana*: a transition towards evolutionary ecological genomics. *En préparation*.

Article 6: B. BRACHI, C. MEYER, A. PLATT, **R. VILLOUTREIX**, F. ROUX and J. BERGELSON. GWA reveals a complex adaptive history for glucosinolate profiles in *Arabidopsis thaliana*. *En préparation*.

Article 7: E. BARON, J. RICHIRT, **R. VILLOUTREIX**, E. BARILLOT, L. AMSELLEM and F. ROUX. The genetics of response to interspecific competition in *Arabidopsis thaliana. En préparation*.

INTRODUCTION

I. INTRODUCTION :

1. Identification des bases génétiques associées à la variation phénotypique naturelle intra-spécifique

Un défi majeur de la biologie moderne reste l'identification des bases génétiques sous-jacentes à la variation phénotypique naturelle observée à l'intérieur d'une espèce. Cette identification a des implications sociétales importantes notamment en médecine pour la découverte des bases génétiques associées aux maladies, ou bien encore en agronomie pour l'augmentation du rendement chez les plantes cultivées ou les animaux domestiques.

Bien que de nombreuses méthodes existent depuis longtemps pour identifier les bases génétiques associées à la variation naturelle, le récent développement de technologies de séquençage haut débit de nouvelle génération (Next Generation Sequencing, NGS) semble annonciateur de nouvelles avancées dans ce domaine. En effet, ces nouvelles techniques permettent un accès à la quasi-totalité des polymorphismes présents le long du génome pour de nombreux individus. Cette caractérisation des polymorphismes génétiques couplée à des méthodes statistiques récemment développées d'association phénotype – génotype le long du génome (Genome-Wide Association mapping, GWA mapping) pourrait permettre d'effectuer une cartographie fine et très rapide des régions génomiques associées à la variation naturelle phénotypique, et ceci chez de nombreux organismes. Cette nouvelle combinaison NGS – GWA mapping a notamment permis de nombreuses avancées dans des programmes de découverte des bases génétiques liées à des maladies humaines (Cookson *et al*. 2009 ; Cirulli & Goldstein 2010) et à l'amélioration du rendement chez les animaux domestiques (Goddard & Hayes 2009) et les plantes cultivées (Rafalski *et al.* 2010).

Figure 1 : ***Arabidopsis thaliana*** **cultivée en serre.** (Photographie par Benjamin Brachi).

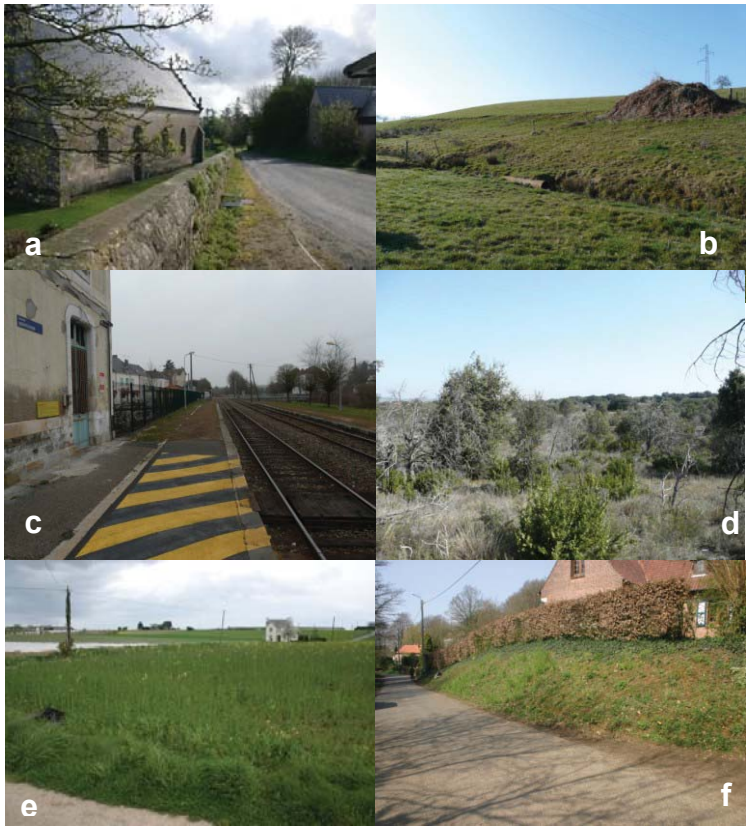


Figure 2 : Diversité des habitats où est présent *A. thaliana*. *A. thaliana* peut vivre dans des habitats contrastés, même à une échelle géographique fine. Sur ces photos, des populations naturelles observées en France représentent différents niveaux de perturbation, depuis les quais de gare et les champs cultivés aux prairies permanentes et garrigues. a. bord de route (Bretagne) ; b. prairie permanente (Bourgogne) ; c. quai de gare (Bourgogne) ; d. garrigue (Languedoc) ; 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).

L'apparition de ces récentes avancées technologiques a également favorisé l'émergence d'une nouvelle discipline : la génomique écologique évolutive qu'on peut définir comme l'identification des gènes associés à la variation phénotypique adaptative dans les environnements naturels et les populations (Feder & Mitchell-Olds 2003 ; Orsini *et al.* 2013). L'association des trois disciplines que sont la génomique, l'écologie et l'évolution reste cependant, à l'heure actuelle, l'apanage d'un nombre limité d'espèces modèles.

2. La plante modèle *Arabidopsis thaliana* : du GWA mapping à la génomique écologique évolutive

A. *Arabidopsis thaliana*

Originaire d'Eurasie, *Arabidopsis thaliana* (Figure 1), communément appelée l'arabette des dames, 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 encore la principale espèce modèle en 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 & Nordborg 2005). Ces caractéristiques combinées à la petite taille de son génome (5 chromosomes, ~125 MB) ont fait d'*A. thaliana* l'espèce indiquée pour le premier séquençage complet chez les plantes supérieures, achevé en 2000 (accession Col-0, The Arabidopsis Genome Initiative 2000). Cette séquence de référence a permis d'annoter un grand nombre de gènes. Actuellement, la base de données TAIR 9 (The

Box 1 : Historique du GWA mapping chez *A. thaliana*.

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 (Mitchell-Olds & Schmitt 2006 ; Shindo *et al.* 2007). Cette diversité d'habitats s'observe même à une échelle géographique de l'ordre de quelques kilomètres (Figure 2). Son régime de reproduction majoritairement autogame a longtemps laissé penser que les populations naturelles étaient majoritairement monomorphes. Pourtant, des populations génétiquement et phénotypiquement polymorphes ont été mises en évidence (Le Corre 2005). De plus, le taux d'allogamie récemment calculé au sein de nombreuses populations naturelles d'*A. thaliana* est effectivement en moyenne de 2% (comme initialement décrit dans les années 1980 à partir de quelques populations, Abbott & Gomes 1989) mais peut atteindre jusqu'à 20 % dans certaines populations (Bomblies *et al.* 2010; Platt *et al.* 2010). Il est donc possible, en utilisant le modèle *A. thaliana*, d'étudier les patrons d'évolution des traits phénotypiques à différentes échelles géographiques, allant de l'aire de répartition de l'espèce à une échelle intra-populationnelle. L'importante quantité de ressources génétiques naturelles publiquement disponibles chez cette espèce, ainsi que la diversité des habitats où l'on peut la trouver font donc d'*A. thaliana* un modèle biologique de choix en génomique écologique évolutive.

B. Etat de l'art

Les avancées génomiques réalisées au cours de la dernière décennie chez *A. thaliana* en ont fait une espèce pionnière dans l'identification des bases génétiques associées à la variation phénotypique naturelle (Box 1). En effet, les efforts conjoints liés à la caractérisation de la taille moyenne du déséquilibre de liaison existant chez cette espèce et au développement

Box 2 : Etudes de GWA mapping chez *A.thaliana*

Publications	Traits phénotypiques	Gènes validés fonctionnellement
Todesco *et al.* 2010 Nature	Trade-off entre immunité et croissance végétative	Oui
Nemri *et al.* 2010 PNAS	Résistance au mildiou	Oui
Baxter *et al.* 2010 PLoS Genetics	Accumulation de sodium dans les feuilles	Oui
Li *et al.* 2010 PNAS	Date de floraison dans des climats simulés	Non
Brachi *et al.* 2010 PLoS Genetics	Date de floraison dans conditions écologiquement réalistes	Non
Chao *et al.* 2012 PLoS Genetics	Accumulation de cadmium dans les feuilles	Oui
Pagny *et al.* 2012 New phytologist	Résistance au virus *plumpox*	Non
Huard-chauveau *et al.* 2013 PLoS Genetics	Résistance quatitative à la bactérie pathogène Xanthomonas campestris	Oui
Rosas *et al.* 2013 PNAS	Architecture des racines	Oui

d'un grand nombre de marqueurs moléculaires ont rendu possible l'utilisation du GWA mapping chez cette espèce. Par rapport aux méthodes traditionnelles d'analyses de QTL (Quantitative Trait Loci) mapping nécessitant une étape de croisement entre deux lignées génétiques, la méthode du GWA mapping a l'avantage d'effectuer une cartographie fine des régions génomiques associées à la variation naturelle d'un trait en tirant bénéfice d'un déséquilibre de liaison de 10kb (en moyenne le long du génome) issu d'évènements de recombinaison se produisant dans les populations naturelles d'*A. thaliana* depuis des centaines de milliers d'années (Kim *et al.* 2007). En 2010, une étude fondatrice de GWA mapping basée sur 192 accessions naturelles génotypées pour 214k SNPs (Single Nucleotide Polymorphisms ; soit en moyenne un marqueur SNP tous les 500bp) et portant sur 107 phénotypes a confirmé la puissance de cette méthode pour identifier finement et rapidement des régions génomiques associées à la variation naturelle (Atwell *et al.* 2010). En effet, les régions génomiques significativement associées à la variation naturelle de la date de floraison et de la résistance aux pathogènes présentent un enrichissement significatif en gènes déjà connus pour intervenir dans ces fonctions. Cette validation de la puissance du GWA mapping chez *A. thaliana*, entraina une utilisation massive de cette méthode chez cette espèce afin d'identifier les bases génétiques de traits divers et variés (Box 2). Pourtant, malgré ces avancées récentes, peu d'études de GWA mapping ont été replacées dans un contexte écologiquement réaliste (voir Brachi *et al.* 2010 pour une exception). L'environnement très contrôlé qu'est le laboratoire doit cependant difficilement refléter les multiples et complexes signaux rencontrés par les plantes dans leurs habitats d'origine. Effectuer des études de GWA mapping à partir d'un phénotype exprimé en conditions contrôlées de laboratoire pourrait donc ne pas s'avérer optimal quant à l'identification des bases génétiques associées à la variation phénotypique exprimée dans les populations naturelles. Certains auteurs plaident

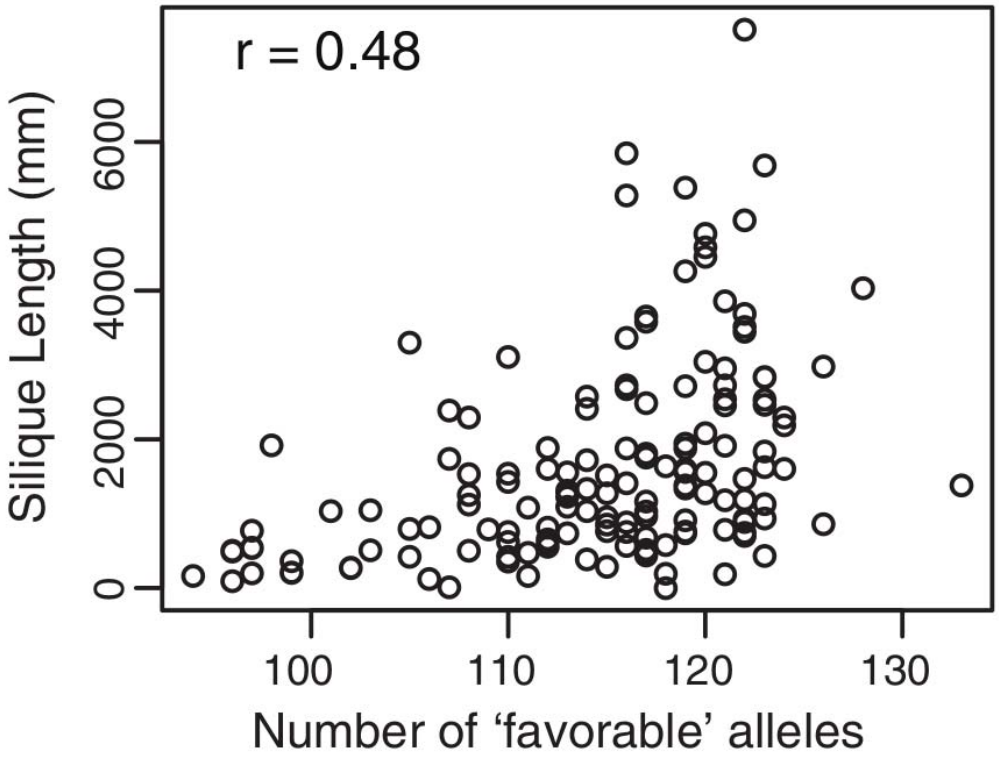


Figure 3 : Relation entre le nombre d'allèles favorables et la production de graines estimée ici par la longueur totale des fruits produits. D'après Hancock *et al.* 2011

donc pour le développement d'une approche de génomique écologique (Ungerer *et al.* 2008 ; Bergelson & Roux 2010) afin de replacer les études d'association phénotype – génétique dans un contexte écologique proche de ceux rencontrées par l'espèce étudiée.

Deux études de génomique écologique évolutive ont récemment été publiées chez l'espèce modèle *A. thaliana* (Hancock *et al.* 2011; Fournier-Level *et al.* 2011). Bien que ces deux études portent sur l'identification des bases génétiques de l'adaptation au climat chez *A. thaliana*, des résultats contradictoires en ressortent. Dans la première étude (Hancock *et al.* 2011), les auteurs ont identifié les polymorphismes génétiques associés au climat à partir de 948 d'accessions européennes génotypées pour environ 214 000 marqueurs SNPs. Les régions génomiques associées aux 13 variables climatiques utilisées dans cette étude montrent des enrichissements significatifs en mutations non-synonymes. Ces régions montrent également un enrichissement en gènes associés à des processus biologiques en relation avec une adaptation au climat. A partir de ces résultats, les auteurs ont par ensuite énoncé la prédiction suivante : si *A. thaliana* est adapté au climat, nous devrions être capable de prédire dans une localité donnée (et donc un climat donné) les accessions qui devraient avoir la fitness la plus importante. En se basant sur les SNPs les plus associés aux 13 facteurs climatiques, les auteurs ont tout d'abord déterminé les allèles les plus avantageux attendus pour le climat à Lille. Puis, à partir d'une expérience mise en place sur le terrain expérimental de l'Université de Lille 1, ils ont montré que les accessions mondiales qui avaient le plus d'allèles avantageux prédits avaient effectivement eu une production de graines plus importante que les accessions mondiales qui n'avaient pas ces allèles avantageux prédits (*rho*=0.48 ; *P*=0.003 ; Figure 3), même si une part importante de la variation de la production de graines reste inexpliquée. Ces résultats semblent indiquer que les auteurs ont bien réussi à identifier des régions génomiques impliquées dans l'adaptation au climat mais que ces régions n'expliquent pas la totalité de la

valeur sélective des individus. Les auteurs ont également identifié des traces de balayage sélectif dans ces régions associées au climat, ce qui suggèrerait que l'adaptation au climat chez *A. thaliana* résulterait d'un processus de sélection à partir de nouvelles mutations.

La deuxième étude porte sur un jeu de 287 accessions européennes, elles aussi génotypées pour 214 000 marqueurs SNPs. Ces accessions ont été plantées dans 4 sites répartis dans 4 pays d'Europe aux climats très contrastés (Valence, Espagne ; Norwich, Angleterre ; Halle, Allemagne ; Oulu, Finlande). Les accessions ayant la meilleur survie et produisant le plus grand nombre de fruits ne sont pas les mêmes entre sites et proviennent de régions géographiquement proches de leur site d'étude, ce qui est cohérent avec un patron d'adaptation locale. Les régions génomiques associées à la survie et à la production de graines sont différentes entre sites, laissant suggérer aux auteurs que les allèles sous-jacents à l'adaptation au climat dans un site donné sont neutres dans les autres sites. Les auteurs n'ont pas trouvé de traces de balayage sélectif associées à ces régions génomiques, ce qui semble suggérer que l'adaptation locale au climat se serait effectuée à partir de variants génétiques déjà présents dans les populations naturelles. Ce dernier constat est donc en désaccord avec l'étude précédente.

Ces deux études suggèrent qu'il existe une adaptation à une large échelle géographique pour le climat chez *A. thaliana*, mais tirent des conclusions différentes quant à l'identité des bases génétiques associées à cette adaptation et sur les traces de sélection génomique qui leur sont associées. Plusieurs explications peuvent être avancées pour expliquer ces différences de conclusion :

1. Le matériel biologique utilisé dans ces deux études n'est pas identique. En effet, ces études sont basées sur des jeux différents d'accessions européennes. Il faut

aussi noter que ces deux études n'utilisent qu'une seule accession par point géographique et donc négligent la variation phénotypique pouvant potentiellement exister au sein d'une population naturelle d'*A. thaliana* (Le Corre 2005).

2. On peut également noter que ces études se basent sur des données phénotypiques acquises sur différentes années (2009 pour la première étude; 2006 et 2007 pour la deuxième étude). Etant donné que la production totale de graines pour une accession donnée d'*A. thaliana* peut fortement varier d'une année sur l'autre (Frenkel *et al.* 2008), il aurait été souhaitable de répéter ces expériences sur plusieurs années.
3. Les régions génomiques trouvées dans la première étude ne permettent de prédire qu'une partie de la production totale de graines sur le site de l'Université de Lille. Une adaptation à des facteurs écologiques variant à une échelle géographique plus petite comme le sol, les attaques de pathogènes ou la compétition interspécifique pourrait potentiellement exister.
4. Bien que ces études détectent des régions génomiques potentiellement impliquées dans l'adaptation au climat, on peut se demander quels sont les traits sous-jacents (physiologie, morphologie, phénologie...) à cette adaptation locale et s'ils sont identiques entre les différents sites testés.

C. Quelles améliorations proposer pour les futures études de génomique écologique évolutive chez *A. thaliana* ?

Nous l'avons constaté : l'étude de la génomique écologique évolutive chez *A. thaliana* en est encore à ses balbutiements. Afin d'optimiser l'identification des bases génétiques de l'adaptation chez *A. thaliana* et d'obtenir une meilleure compréhension des dynamiques

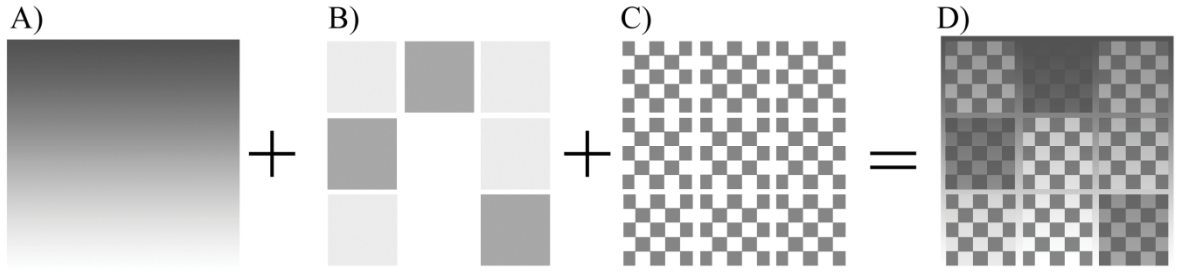
adaptatives présentes dans les populations naturelles, différentes améliorations détaillées dans les parties suivantes peuvent être proposées.

1) Processus neutres versus adaptatifs

La variation phénotypique spatiale observée pour un trait phénotypique au sein d'une espèce est la résultante de plusieurs forces évolutives. Une part de cette variation peut être la résultante d'une variation spatiale de l'optimum phénotypique due à une variation des conditions environnementales locales (Haldane 1948). Une autre part de cette variation peut également être la résultante d'autres processus tels que l'équilibre migration/dérive (Kimura 1955) ; mais également la sélection indirecte, c'est-à-dire lorsque la sélection agit indirectement sur un trait en sélectionnant directement un trait corrélé à celui-ci (Lande & Arnold 2009). L'identification des traits sous sélection dans un contexte multi-trait et plus précisément l'estimation de la part relative de la variation phénotypique due à des processus sélectifs apparaissent comme des prérequis indispensables à l'identification des bases génétiques de l'adaptation.

2) Echelle spatiale de la variation phénotypique adaptative : nécessité d'identifier les pressions de sélection

Il est important de déterminer l'échelle géographique à laquelle varie un trait phénotypique. Cette étape est un préliminaire à la détection de l'échelle géographique de la sélection et à l'identification des facteurs écologiques responsables de cette variation. En effet, si la variation observée pour un trait phénotypique varie très peu à une échelle fine mais varie fortement à une échelle très large, des facteurs écologiques variant à une échelle très large pourraient potentiellement être responsables de cette variation. Au contraire, si la variation observée pour un trait phénotypique présente autant de variation à une échelle géographique fine qu'a une échelle géographique large, des facteurs écologiques variant à une

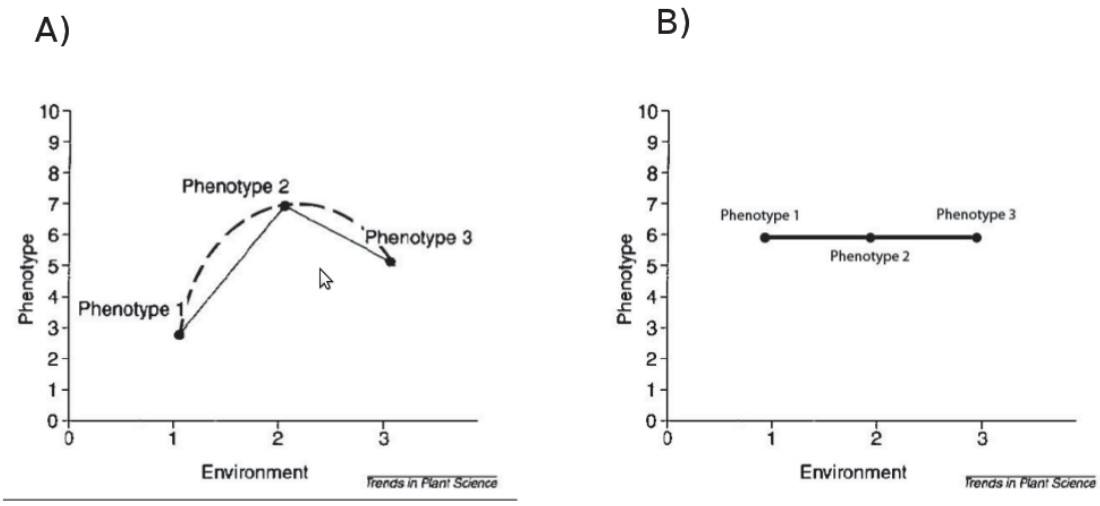


<u>Figure 4 :</u> Exemple théorique de variation de facteurs écologiques pouvant exercer une potentielle pression de sélection sur un trait phénotypique, ici la date de floraison. Les zones les plus foncées correspondent à des zones où la tardivité est sélectionnée. Les zones les plus claires correspondent à des zones où la précocité est sélectionnée. Variation de la sélection avec A) un gradient de variation (facteurs climatiques par exemple), B) une échelle grossière de variation (pratiques agricoles par exemple), C) une échelle fine de variation (présence de pathogènes, caractéristiques édaphiques par exemple). La figure D) présente la synthèse de toutes les variations théoriques de pression de sélection décrites précédemment.

échelle fine pourraient être potentiellement responsables de cette variation. Mais les deux scénarios présentés précédemment ne sont pas exclusifs. La variation phénotypique d'un trait peut être associée à la variation de différents facteurs écologiques variant à différentes échelles spatiales, correspondant ainsi à un recouvrement entre différents grains de l'environnement. En effet, il a été montré chez quelques espèces végétales, dont *A. thaliana*, que la date de floraison varie avec la latitude (Stinchcombe *et al.* 2004, Figure 4 A), ce qui suggère que la variation observée pour ce trait résulterait d'agents sélectifs agissant à une très large échelle géographique comme le climat. Mais il a aussi été démontré chez d'autres espèces végétales que la date de floraison répondrait à des facteurs environnementaux correspondant à un grain environnemental plus fin, tels que la richesse du sol (Sardansa *et al.* 2005) ou l'herbivorie (Poveda *et al.* 2003, Figure 4 B et C). La floraison répondrait donc potentiellement à des pressions de sélections variant à différentes échelles spatiales. La superposition de ces différents grains environnementaux peut amener à des paysages sélectifs extrêmement compliqués comme celui présenté dans la figure 4D. Cet exemple illustre bien la nécessité, afin de pouvoir estimer correctement la sélection agissant sur un trait, d'échantillonner un grand nombre de populations selon un échantillonnage hiérarchique ainsi que de caractériser un grand nombre de variables environnementales variant à différentes échelles.

3) Environnement de mesure phénotypique

Dans une étude cherchant à identifier les QTLs de la date de floraison dans une famille de lignées recombinantes consanguines (Recombinant Inbred Lines, RILs) d'*A. thaliana* en conditions contrôlées de serre et sur un terrain expérimental, Weinig *et al.* (2002) ont montré que les QTLs identifiés n'étaient pas les mêmes entre ces deux environnements. En effet, certains QTL à effets majeurs en serre n'ont pas été retrouvés sur le terrain ; et de nouveaux



(modifié d'après Sultan 2000)

(modifié d'après Sultan 2000)

Figure 5 : Normes de réaction A) d'un génotype plastique et B) d'un génotype fixé.

QTL à effet majeurs ont été détectés sur le terrain alors qu'ils n'étaient pas détectés en serre. Dans une autre étude de génomique écologique (Brachi *et al.* 2010), les auteurs se sont attachés à identifier les bases génétiques de la date de floraison sur un terrain expérimental de l'Université de Lille 1 en combinant GWA mapping et QTL mapping traditionnel. Les bases génétiques qu'ils ont identifiées sont là aussi différentes de celles identifiées dans d'autres études réalisées en serre avec le même matériel génétique. Il apparait donc important de mesurer aussi l'expression phénotypique d'un génotype dans un milieu proche de celui où il a été collecté. De plus, étant donnée l'hétérogénéité écologique (climatique, attaque de pathogènes…) pouvant exister entre années dans un même milieu, il apparait également nécessaire de reproduire les expériences plusieurs années de suite afin de proposer des scénarios évolutif fiables. Tenir compte de la variation génétique de la plasticité phénotypique exprimée entre ces différents environnements pourrait également s'avérer indispensable dans la détermination du potentiel adaptatif d'une espèce.

D. La plasticité phénotypique

1) Définition

Une confusion persiste dans la littérature autour du terme « plasticité phénotypique ». Celle-ci est induite par une mésentente sur la définition à donner à ce phénomène biologique et s'articule principalement autour de la question suivante : doit-on intégrer une notion adaptative dans la définition même de la plasticité phénotypique ? De ce fait, aucune définition précise n'existe (Debat & David 2001). Dans cette thèse, nous définirons la plasticité phénotypique comme la capacité d'un génotype à produire plusieurs phénotypes en fonction de l'environnement biotique ou abiotique auquel il est exposé (Agrawal 2001 ; Sultan 2000), en excluant toute notion adaptative dans sa définition. Cette définition nous permet donc de distinguer les génotypes plastiques pour un trait phénotypique (Figure 5 A),

des génotypes dit fixés pour lesquels le trait phénotypique ne change pas entre environnements (Figure 5 B).

2) Conditions favorisant théoriquement l'évolution d'une plasticité phénotypique

La plasticité phénotypique est favorisée par une forte hétérogénéité environnementale (Moran 1992 ; Sultan and Spencer 2002). Il est intéressant de noter que la plasticité phénotypique serait plus favorisée dans le cas d'une hétérogénéité temporelle que dans le cas d'une hétérogénéité spatiale (Moran 1992). En effet, contrairement à l'hétérogénéité temporelle, une hétérogénéité spatiale pourrait contenir des refuges, favorables aux génotypes fixés adaptés localement. La variation environnementale doit également être prédictible, être moins rapide que la réponse de l'organisme, afin de permettre au génotype plastique d'exprimer le phénotype adéquat dans l'environnement où le signal environnemental est perçu.

3) Coût et limites de la plasticité phénotypique

Malgré un avantage théorique certain, la plasticité phénotypique n'est pas toujours sélectionnée dans la nature, amenant à considérer dès les années 1960 la présence de coûts et de limites associés à la plasticité (Bradshaw 1965 ; Schlichting 1986). Dans une très bonne revue, DeWitt *et al.* (1998) listent de manière exhaustive différents coûts et limites de la plasticité. Nous ne parlerons dans cette partie que des coûts et limites les plus fréquemment rencontrés chez les plantes (van Kleunen & Fischer 2005) :

- Le maintien d'une machinerie de sensibilité et de réponse à l'environnement pourrait être coûteux énergétiquement, ce qui pourrait contre-sélectionner les génotypes plastiques face aux génotypes fixés. Ce type de coût pourrait avoir été mis en évidence

chez les plantes dans quelques études portant sur la plasticité des traits morphologiques en réponse à l'ombrage (Donohue *et al.* 2000 ; Steinger *et al.* 2003) et à la compétition (van Kleunen *et al.* 2000).

- Un autre coût fréquemment évoqué est le fait que les génotypes plastiques puissent avoir une instabilité développementale plus importante dans un environnement donné. Ce phénomène pourrait être notamment coûteux dans un environnement où s'opère une sélection stabilisante, puisque qu'une imprécision dans la production du phénotype adéquat pourrait entrainer une diminution de la fitness des génotypes plastique dans cet environnement. A ce jour, une seule étude portant sur la plasticité de traits morphologiques en réponse à différents apports en azote a mis en évidence l'existence d'un tel coût chez *A. thaliana* (Tonsor *et al.* 2013).
- La mise en place d'une réponse plastique dure un temps défini, ce qui peut limiter son évolution. En effet, lorsque l'environnement change, durant le temps que prend la réponse phénotypique, le génotype plastique n'est pas adapté à ce nouvel environnement. La plasticité phénotypique peut donc être défavorisée si l'environnement change plus rapidement que la réponse de l'organisme.
- L'évolution d'une plasticité phénotypique peut également être limitée si le signal environnemental perçu par la plante n'est pas relié à la pression de sélection. En effet, dans ce cas, le génotype plastique ne pourra pas exprimer le phénotype avantageux. Bien qu'il s'agisse d'une limite de la plasticité phénotypique et non d'un coût, une telle limite est également décrite sous le nom de coût écologique de la plasticité (Cipollini *et al.* 2003 ; van Kleunen & Fischer 2001).
- La plasticité phénotypique qu'exprime un organisme à un moment donné peut également limiter son potentiel à exprimer ultérieurement la même réponse plastique.

Box 3 :

Dans le but d'étudier des 'stratégies' phénotypiques plutôt que des traits individuels, nous avons mesuré plusieurs catégories de traits phénotypiques dans les différentes études présentées dans cette thèse.

Traits phénologiques :

Nous avons mesurés plusieurs traits phénologiques parcourant le cycle de vie d'une plante annuelle. Ces traits ont été choisis car ils sont cruciaux dans l'optimisation de la survie et de la reproduction des plantes annuelles (Rathcke & Lacey 1985). Nous avons donc mesuré: la date de germination, la date de montaison, la date de floraison, l'intervalle entre montaison et floraison, la durée de floraison, la durée de reproduction et le ratio durée de floraison sur durée de reproduction.

Traits d'accumulation de ressources :

Nous avons mesuré le diamètre de la rosette à floraison, indicateur de la taille de l'individu chez *A. thaliana* (Weinig *et al.* 2006).

Traits architecturaux :

Nous avons mesuré des traits architecturaux liés à la dispersion des graines (Wender *et al.* 2005) et à l'évitement de risques environnementaux comme la sécheresse (Shemesh & Novoplansky 2010). Dans cette thèse, je me suis principalement focalisé sur le nombre de branches primaires sur la tige principale, le nombre de branches basales, la hauteur depuis le sol jusqu'au premier fruit sur la tige principale et la hauteur maximale de la plante.

Estimation de la valeur sélective :

Nous avons estimé la survie et la production totale de graines, en estimant la longueur totale des fruits produits par un individu. *A. thaliana* étant une espèce avec un taux d'autogamie moyen de 98%, la composante mâle de la fitness a été négligée dans nos études.

Disposition des blocs expérimentaux sur le terrain extérieur de l'Université de Lille 1. Photographie prise le 1er mars 2011.

Cette limite appelée limite historique de la plasticité ('plasticity-history limit') a été découverte dans une étude portant sur la réponse de la plante *Abutilon theophrastii* (Weinig & Delph 2001) face à différentes environnements lumineux. Cette étude montre que l'induction d'une réponse plastique pour l'élongation des internodes à un stade précoce limite l'expression de cette même plasticité à un stade plus avancé. Ce phénomène pourrait en effet être particulièrement limitant chez les plantes, si les pressions de sélection varient durant leur cycle de vie.

3. Plan de thèse

Cette thèse fait partie intégrante d'un projet de génomique écologique évolutive chez *A. thaliana* au sein de l'équipe « Changements globaux : de la génomique écologique à l'écologie des communautés ». Durant cette thèse, je me suis attaché à :

- caractériser la variation phénotypique à différentes échelles spatiales, en mesurant de nombreux traits phénotypiques et leurs plasticités en conditions contrôlées de serre mais aussi sur un terrain expérimental de l'Université de Lille 1 (Box 3).
- identifier les traits sous sélection en utilisant trois approches : comparaison $F_{ST} - Q_{ST}$, relations phénotype – écologie et gradients de sélection génotypiques.
- identifier les agents sélectifs potentiellement responsables de ces variations phénotypiques adaptatives
- et identifier les bases génétiques associées à la variation naturelle par une approche de GWA mapping.

Cette thèse s'articule en trois parties :

Chapitre 1 - Etude de l'échelle géographique des variations phénologiques adaptatives et de leurs bases génétiques sous-jacentes chez *Arabidopsis thaliana*

Basée sur une expérience réalisée en conditions contrôlées de serre, ce premier chapitre vise à décrire l'échelle géographique de variation naturelle présente pour des traits phénologiques post-germination en utilisant 49 populations naturelles collectées dans 4 régions françaises. Afin de connaitre la nature adaptative de ces variations phénotypiques, la part de variation phénotypique due à des processus non sélectifs a été estimée. L'identification des facteurs écologiques potentiellement responsables de ces variations a également été réalisée, ceci afin de déterminer si les pressions de sélection agissant sur ces traits étaient les mêmes à différentes échelles géographiques et entre différentes régions géographiques. A partir d'un échantillonnage hiérarchique d'accessions naturelles génotypées pour 214 kSNPs, l'identification des bases génétiques associées à ces variations naturelles a également été réalisée afin de déterminer si les bases génétiques associées à la variation phénologique étaient identiques à différentes échelles géographiques.

Chapitre 2 - Réalisme écologique, contexte multi-traits et intégration de la plasticité dans l'étude de l'adaptation : une transition vers la génomique écologique évolutive

Ce second chapitre s'attèle à replacer dans un contexte écologiquement réaliste (i) la caractérisation de la variation naturelle d'une vingtaine de traits (morphologiques, phénologiques, architecturaux et reproducteurs) face à l'hétérogénéité environnementale temporelle et (ii) l'identification des traits sous sélection en utilisant trois approches

(comparaison F_{ST} – Q_{ST}, relations phénotype – écologie et gradients de sélection génotypiques). La date de germination pouvant grandement influencer la sélection opérant sur les traits post-germination (Donohue *et al.* 2005), les mesures des traits ont été effectuées suivant deux cohortes de germination au sein d'une même année. La sélection pouvant également être différente entre années (Siepielski *et al.* 2009), les mesures des traits ont également été effectuées durant deux années consécutives au sein de la même cohorte de germination. Je me suis également attaché à étudier la valeur adaptative de la plasticité phénotypique exprimée à deux échelles différentes de temps (entre cohortes et entre années). De nombreux modèles théoriques prédisant une contre sélection de la plasticité phénotypique si celle-ci est coûteuse, je me suis également attaché à identifier si un coût était associé à la plasticité phénotypique.

Chapitre 3 - Dynamique évolutive de populations naturelles d'*Arabidopsis thaliana* sur une courte échelle de temps

Les expériences décrites dans le chapitre 1 et le chapitre 2 ont majoritairement reposées sur 49 populations françaises qui ont toutes été récoltées au printemps 2009 et ne reflètent donc qu'une image instantanée. Les mêmes résultats auraient-ils été obtenus si les populations avaient été récoltées une année différente ? Il est difficile de répondre à cette question en absence de données empiriques sur la dynamique temporelle dans les populations naturelles d'*A. thaliana*. Pour pallier à ce manque, j'ai étudié l'évolution phénotypique sur une période d'une dizaine de générations dans deux populations naturelles décrites comme très polymorphes phénotypiquement dans le chapitre 1, ceci grâce à des études de résurrection.

CHAPITRE 1 – ETUDE DE L'ECHELLE GEOGRAPHIQUE DES VARIATIONS PHENOLOGIQUES ADAPTATIVES ET DE LEURS BASES GENETIQUES SOUS-JACENTES CHEZ *ARABIDOPSIS THALIANA*

II. CHAPITRE 1 – ETUDE DE L'ECHELLE GEOGRAPHIQUE DES VARIATIONS PHENOLOGIQUES ADAPTATIVES ET DE LEURS BASES GENETIQUES SOUS-JACENTES CHEZ *ARABIDOPSIS THALIANA* :

1. Introduction

Dans ce chapitre, nous avons voulu (i) caractériser l'échelle spatiale à laquelle varient six traits phénologiques recouvrant le cycle de vie d'*A. thaliana*, (ii) identifier la part de cette variation due à des processus non sélectifs, (iii) déterminer l'échelle spatiale de la variation adaptive en identifiant les agents sélectifs potentiellement responsables de cette sélection, et (iv) identifier les bases génétiques sous-jacentes de ces variations à différentes échelles géographiques par une approche de GWA mapping.

Dans ce but, 800 familles génétiques ont été constituées à partir d'un échantillonnage réalisé au printemps 2009 dans 49 patchs répartis dans 4 régions françaises climatiquement contrastées (Bretagne, Bourgogne, Languedoc, Nord) par Benjamin Brachi, Nathalie Faure et Fabrice Roux. Durant mon stage de M2, j'ai caractérisé en serre la variation naturelle de six traits phénologiques pour ce jeu de 800 familles, ainsi que pour deux jeux d'accessions naturelles (n = 394) issues d'échantillonnages mondiaux et français. La cohorte de germination étant connue pour affecter l'expression de traits phénologiques post-germination chez *A. thaliana* (Donohue *et al.* 2005), les deux principales cohortes de germinations observées chez *A. thaliana* (cohorte de printemps et cohorte d'automne) ont été simulées en serre afin d'estimer l'impact de la date de germination sur les patrons observés. J'ai également caractérisé ces 49 patchs français (i) génétiquement pour 135 marqueurs SNP et (ii)

écologiquement à l'aide de 42 variables écologiques dont 25 variables bioclimatiques, 14 variables édaphiques et 3 indices de compétition intra- et interspécifique.

L'identification des bases génétiques associées à la variation phénotypique a été réalisée à différentes échelles géographiques en utilisant les deux jeux d'accessions issus d'échantillonnages mondiaux et français et déjà caractérisés génétiquement pour 214 000 marqueurs SNP (Horton *et al.* 2012).

Cette étude a fait l'objet de la publication ci-jointe : B. BRACHI[§], **R. VILLOUTREIX[§],** N. FAURE, N. HAUTEKEETE,Y. PIQUOT, M. PAUWELS, D. ROBY, J. CUGUEN, J. BERGELSON and F. ROUX. 2013. Investigation of the geographical scale of adaptive phenological variation and its underlying genetics in *Arabidopsis thaliana*. *Molecular Ecology* 22:4222-4240. [§]**These authors contributed equally to this work**

MANUSCRIT: INVESTIGATION OF THE GEOGRAPHICAL SCALE OF ADAPTIVE PHENOLOGICAL VARIATION AND ITS UNDERLYING GENETICS IN *ARABIDOPSIS THALIANA*

Investigation of the geographical scale of adaptive phenological variation and its underlying genetics in *Arabidopsis thaliana*

Benjamin Brachi[1, 2,†], Romain Villoutreix[1,†], Nathalie Faure[1], Nina Hautekèete[1], Yves Piquot[1], Maxime Pauwels[1], Dominique Roby[3,4], Joël Cuguen[1], Joy Bergelson[2], Fabrice Roux[1,*]

[1]*Laboratoire Génétique et Evolution des Populations Végétales, UMR CNRS 8198, Université des Sciences et Technologies de Lille – Lille 1, F-59655 Villeneuve d'Ascq cedex France*

[2] *Department of Ecology and Evolution, University of Chicago, Chicago, Illinois 60637, USA.*

[3] *INRA, Laboratoire des Interactions Plantes-Microorganismes (LIPM), UMR441, F-31326*

Castanet-Tolosan, France.

[4] *CNRS, Laboratoire des Interactions Plantes-Microorganismes (LIPM), UMR2594, F-31326*

Castanet-Tolosan, France.

[†] These authors contributed equally to this work.

Keywords: *Arabidopsis thaliana*, flowering time, ecology, adaptation, selection

Corresponding author: Fabrice Roux

Laboratoire Génétique et Evolution des Populations Végétales, UMR CNRS 8198, Université des Sciences et Technologies de Lille – Lille 1, F-59655 Villeneuve d'Ascq cedex France

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

Phone number: +33 (0)3 20 43 40 33

Fax number: +33 (0)3 20 43 69 79

Running title: The scale of adaptation: *A. thaliana* phenology.

Summary

Despite the increasing number of genomic tools, identifying the genetics underlying adaptive complex traits remains challenging in the model species *Arabidopsis thaliana*. This is due, at least in part, to the lack of data on the geographical scale of adaptive phenotypic variation. The aims of this study were (i) to tease apart the historical roles of adaptive and non-selective processes in shaping phenological variation in *A. thaliana* in France and (ii) to gain insights into the spatial scale of adaptive variation by identifying the putative selective agents responsible for this selection. Forty-nine natural stands from four climatically contrasted French regions were characterized (1) phenologically for six traits; (2) genetically using 135 SNP markers; and (3) ecologically for 42 variables. Up to 63% of phenological variation could be explained by neutral genetic diversity. The remaining phenological variation displayed stronger associations with ecological variation within regions than among regions, suggesting the importance of local selective agents in shaping adaptive phenological variation. Although climatic conditions have often been suggested as the main selective agents acting on phenology in *A. thaliana*, both edaphic conditions and interspecific competition appear to be strong selective agents in some regions. In a first attempt to identify the genetics of phenological variation at different geographical scales, we phenotyped worldwide accessions and local polymorphic populations from the French RegMap in a GWA mapping study. The genomic regions associated with phenological variation depended upon the geographical scale considered, stressing the need to account for the scale of adaptive phenotypic variation when choosing accession panels for GWAS.

Introduction

Numerous phenotypic traits display extensive variation among natural populations. Many evolutionary forces may jointly shape this variation, including genetic drift, migration, demographic history and natural selection (Belotte *et al.* 2003; Kawecki & Ebert 2004; Olson-Manning *et al.* 2012). The extent to which this natural variation is adaptive is still an open question (Mitchell-Olds *et al.* 2007).

Characterization of how selection has shaped natural variation in complex traits requires that at least two key issues be addressed. The first challenge is to estimate the part played by adaptive processes in shaping phenotypic natural variation, and the geographical scale at which adaptation takes place. To date, much of the work on how selection shapes complex traits has focused on local adaptation. Two complementary approaches have been applied. Reciprocal transplants address how genotypes perform in their population of origin relative to more distant populations (Leimu & Fischer 2008; Hereford 2009). Fitness superiority of local individuals at each site demonstrates local adaptation (following the "local vs foreign" criterion, Kawecki & Ebert 2004). In the second approach, selection gradients are used to quantify the effects of trait variation on an estimator of fitness (Lande & Arnold 1983; Munguía-Rosas *et al.* 2011). Because it is time consuming to measure fitness, both approaches require that adaptation be studied in a restricted number of natural populations, making generalizations difficult (but see Belotte *et al.* 2003; Laine 2005; Becker *et al.* 2006). One indirect method is to study the relationships between ecological and phenotypic variations (Linhart & Grant 1996; Merilä *et al.* 2001). While correlational analyses such as these do not allow definitive conclusions, this approach facilitates the inclusion of more individuals and more populations and therefore enables study of adaptation at different

geographical scales (Conner 2010). Correlations with environmental variables, if they are driven by local adaptation, integrate multiple generations of selection and multiple components of fitness (Merilä *et al.* 2001).

The second challenge is to study the genetic architecture and identify the genetic basis of adaptation as a first step towards reconstructing the adaptive walk followed by a natural population towards the local optimum phenotype (Orr 2005; Hermisson & Pennings 2005; Kopp & Hermisson 2007; 2009). Genome-wide association (GWA) mapping provides a powerful tool to start tackling this question. The power of GWA mapping to identify the genetics underlying phenotypic variation observed at broad geographical scales has been demonstrated in multiple plant studies, but so have its limitations (Atwell *et al.* 2010; Brachi *et al.* 2010; Huang *et al.* 2010). First, strong confounding by population structure introduces false positives and false negatives (Brachi *et al.* 2010). Second, rare alleles, potentially important in local adaptation processes, are difficult to detect (Atwell *et al.* 2010). Third, because the same phenotypic value may be caused by different alleles of the same gene among natural populations, allelic heterogeneity may hinder the detection of genomic regions associated with phenotypic natural variation (Bergelson & Roux 2010). Traditional linkage mapping based on crosses among populations may solve these issues but has the disadvantage of coarse resolution (Bergelson & Roux 2010). An alternative approach would be to define regional or local mapping panels that match the geographical scale at which adaptive phenotypic variation is observed. As a consequence, confounding by population structure could be greatly reduced and causal genes with rare alleles or allelic heterogeneity at the species-wide scale could become detectable (Bergelson & Roux 2010; Brachi *et al.* 2011; Horton *et al.* 2012).

Phenological traits are excellent candidates for studying local adaptation because they display extensive variation in many plant species (Rathcke & Lacey 1985) and have often been found to be adaptive. In a meta-analysis of selection gradients, Munguía-Rosas (2011) found the strength of selection on phenology to vary with latitude, as expected if climatic variation acts as a selective agent. Supporting evidence is provided by other studies that similarly reveal significant latitudinal gradients (Van Dijk *et al.* 1997; Stinchcombe *et al.* 2004; Wagmann *et al.* 2012) or a response of phenological traits to global climate change (Franks *et al.* 2007; Inouye 2008). However, climatic factors may not be the only selective agents acting on phenology. Many studies have found evidence for local selection by pollinators or herbivores (Parachnowitsch & Caruso 2008; Sandring & Ågren 2009). Taken together, there is overwhelming evidence that phenological traits are adaptive and that relevant selective agents can act globally or locally.

Arabidopsis thaliana, the flagship species of plant genomics, is a widely distributed annual selfing species found in diverse habitats (Mitchell-Olds & Schmitt 2006) that displays tremendous variation in phenological traits (Donohue 2005; Atwell *et al.* 2010; Brachi *et al.* 2010). Despite its emerging status as a model species in evolutionary ecology (Fournier-Level *et al.* 2011; Hancock *et al.* 2011; Gaut 2012), little is known about the geographic scale of phenological adaptation, i.e. the selective pressures acting on its phenology, in natural populations (Méndez-Vigo *et al.* 2011; Montesinos-Navarro *et al.* 2011). The relationship between flowering time and latitude (Stinchcombe *et al.* 2004; Brachi *et al.* 2010), and between flowering time and several climatic variables (e.g., the number of consecutive frost-free days, maximum temperature in the warmest month and photosynthetically active radiation, Hancock *et al.* 2011), suggests that climate may impose selection shaping flowering time variation. A large proportion of phenological variation, however, remains unexplained at

the continental scale, suggesting that phenological variation at smaller geographical scales results from either non-selective processes or local selective agents.

Here, we investigated the geographical scale of adaptive phenological variation and its underlying genetics in *A. thaliana*. In the first part of the study, (i) we estimated the portions of natural phenological variation in *A. thaliana* that could be the result of non-selective or adaptive processes and (ii) we investigated the geographical scale of adaptive variation by identifying the putative selective agents responsible for this selection. To achieve these goals, we used a hierarchical sampling design to collect 800 individuals from 49 natural stands located in four climatically contrasted regions of France. We characterized the selfed progeny of those individuals: phenologically for six traits spanning the annual plant life cycle, genetically using 135 neutral SNPs, and ecologically for 42 variables. We found that neutral genetic diversity could explain up to 63% of phenotypic variation for some phenological traits. The remaining variation was associated with many ecological factors, including edaphic variation and competition. Relationships with ecological variables were stronger within regions than considering all regions together, suggesting a prevalence of local adaptation in shaping the natural variation of phenological traits in our study. We made all plant material used in this study publicly available in order to allow follow up studies (http://publiclines.versailles.inra.fr/).

In the second part of the study, we aimed at testing whether GWA mapping studies were more successful in identifying the genetic basis of phenological variation with a mapping panel of accessions selected to match the geographical scale of adaptive variation observed within the 49 stands. To do so, we performed a GWA study using samples of *A. thaliana* collected from different geographical scales using both worldwide accessions and accessions from the French RegMap panel (Horton *et al.* 2012). Consistent with the scale of

adaptive phenological variation observed within the 49 stands, GWA mapping revealed strong signals of association at the population and regional scales that were often located in different genomic regions than those detected at the worldwide scale.

Materials and Methods

Plant material

Forty-nine stands of *A. thaliana* were collected from early March to late April 2009 in 42 locations from four regions of France that have contrasting climates (Brittany: oceanic, 11 stands; Burgundy: continental, 11 stands; Languedoc: Mediterranean, 16 stands; North of France: semi-oceanic, 11 stands; Table S1). We defined stands as a single patch of plants growing in relatively homogeneous ecological conditions. The average within-region distance among stands was 33.1 km (SD = 19.6 km). The pairwise minimal distances among regions ranged from 292 km (Burgundy - Languedoc) to 758 km (Brittany – Languedoc).

In each stand, between 10 and 30 plants were collected randomly and brought back to a cold frame greenhouse (no additional light or heating). Seeds were collected from individual plants to constitute seed families. To reduce maternal effects, families were grown for one generation from June to December 2009 under controlled greenhouse conditions (16 h photoperiod, 20°C) at the University of Lille 1. For the phenotyping experiment, 20 families were randomly chosen from each stand when possible. If less than 20 families were available in a given stand, all families were phenotyped (Table S1). In all, each French region was represented by 200 families. Seeds of the 800 families will be publicly available from the Centre de Ressources Biologiques (CRB, INRA Versailles, France, http://publiclines.versailles.inra.fr/).

A set of 184 worldwide accessions and a set of 210 French accessions corresponding to the French RegMap (Horton *et al.* 2012) were also included in the experiment (Table S1). In order to (i) have enough power to run GWA mapping analyses at small geographical scales and (ii) use accessions collected in the same geographical region than the 49 stands, French accessions were over-represented. For the purpose of our study, we excluded from our statistical analyses the natural accessions that were likely to be contaminants, *i.e.* accessions for which geographical origin is suspicious (Anastasio *et al.* 2011), leaving us with a set of 352 accessions (Table S1). From these 352 accessions, we designed six sets of accessions corresponding to different geographical scales: WORLD (n = 167), EUROPE (n = 143), FRANCE (n = 203), BURGUNDY (regional scale, n = 121), and MIB and TOU (two local populations in Burgundy; n = 52 and n = 69, respectively). The WORLD, EUROPE and FRANCE sets shared 18 French accessions that are representative of the French diversity (Atwell *et al.* 2010).

Phenological characterization

A greenhouse experiment of 4,928 plants (French families, worldwide accessions and accessions from the French Regmap) was set up at the University of Lille (North, France) in January 2010 using a split-plot design arranged as a Randomized Complete Block Design (RCBD) with two seasonal germination cohorts nested within two experimental blocks. The two germination cohorts were grown to mimic the two germination flushes found in natural stands, *i.e.* fall and spring (Donohue 2005; Picó 2012). Each block was represented by 19 arrays of 66 individual wells (Ø4 cm, vol. ~38 cm3) (TEKU, JP 3050/66) filled with damp standard culture soil (Huminsubstrat N3, Neuhaus). Each block corresponded to 1,232 plants with one replicate per family (n = 800), one replicate per worldwide accession (n = 184), one

replicate per French accession (n = 210) and a control worldwide accession Bg-2 placed in the same two positions within each array (n = 38 = 19 trays * 2 replicates). Those control plants allowed correcting for micro-environmental variation within blocks. At least 5 seeds were placed in each well. To promote germination, seeds were stratified for 7 days at 4°C. After the 7 day stratification treatment, plants were grown at 20°C under natural light supplemented by artificial light to provide a 16-hr photoperiod. Germination date was monitored daily in each well. *A. thaliana* seeds that had not germinated 13 days after sowing were replaced by extra seedlings of the same family or accession from other blocks. Seedlings were thinned to one per well 14 days after the stratification treatment, keeping the first germinated seedling. Plants in the spring germination cohort underwent the same greenhouse conditions. A winter treatment was simulated for plants in the fall germination cohort. Twenty-two days after the stratification treatment, plants in the fall germination cohort were grown in winter conditions (4°C and 12 hr photoperiod) for 3 weeks and then moved back to greenhouse conditions (20°C and 16 hr photoperiod). During the whole growing period, arrays within blocks were rotated every day to minimize potential effects of uneven lighting across the growth room. Plants were monitored every 2 or 3 days for bolting date (inflorescence distinguishable from the leaves at a size < 5mm), flowering date (appearance of the first open flower), date of senescence of the last flower on the main stem and date of maturation of the last fruit on the main stem. A period of 3 weeks, corresponding to the length of the winter treatment, was subtracted for plants in the fall germination cohort. We measured six phenological traits spanning the life cycle of *A. thaliana*, as previously described in Brachi *et al.* (2012). Bolting time (BT) was measured as the number of days between germination and bolting. The flowering interval (INT) was measured as the difference between bolting and flowering dates. Flowering time (FT) was measured as the number of days between germination and

flowering. Flowering period (FP) was measured as the number of days between the onset of flowering and the senescence of the last flower. The reproductive period (RP) was measured as the number of days between the onset of flowering and maturation of the last fruit on the main stem. The flowering-to-reproductive ratio (FRR) was calculated as the ratio between flowering (FP) and reproductive period ratio (RP). FRR may indicate a trade-off between seed number and seed quality as plants that spend a smaller fraction of their reproductive period engaged in flowering may spend relatively more time filling and maturing seeds (Brachi *et al.* 2012). Plants that had not bolted 100 days after sowing (*i.e.* 3.3 %) were assigned a bolting date value of 100. INT, FT, FP, RP and FRR were therefore not available for these plants.

DNA extraction and genotyping of SNP markers

After one generation of multiplication, seeds from the 800 French families were sown in March 2010 in arrays of 66 individual wells (Ø4 cm, vol. ~38 cm3) (TEKU, JP 3050/66) filled with damp standard culture soil (Huminsubstrat N3, Neuhaus). After a 7-day stratification treatment at 4°C, plants were grown at 20°C and under natural light supplemented by artificial light to provide a 16-hr photoperiod. Seedlings were thinned to one or two per well 14 days after the stratification treatment. Four weeks after the stratification treatment, plants were cut and oven-dried for 2 days at 65°C. We used a NucleoSpin_96 Plant Kit (Macherey-Nagel) to extract and purify total DNA from 10 mg dried leaf tissue. All DNA samples were adjusted to 5ng/µl. A total of 765 French families were genotyped for a set of 149 SNPs spread across the genome (Clark *et al.* 2007), already used to describe the scale of population structure in *A. thaliana* (Platt, Horton, *et al.* 2010a) and genetic variation in natural stands of *A. thaliana* in Germany (Bomblies *et al.* 2010). The remaining 35 French families that were not genotyped correspond to those for which plant tissue was not available. The

genotyping of 149 SNPs was performed by University of Chicago DNA sequencing facility (Chicago, IL) using the Sequenom MassArray system.

Following (Platt, Horton, *et al.* 2010a), seven French families were removed due to an excess of missing genotype calls (>50 of 149). Fourteen SNP assays were also removed due to an excess of missing genotypes or heterozygous calls (>25% of families).

Ecological characterization of stands

The wide range of ecological conditions sampled in each region can be summarized in three broad habitat types: grassland (n = 23), hoed land (n = 16) and meadow (n = 10; Table S1). Each stand was characterized for climatic and edaphic variables, as well as for plant-plant interactions (Table 1). To obtain biologically meaningful climatic variables (Hijmans *et al.* 2005), we chose 25 climatic variables with a grid resolution smaller than the average distance among stands within regions (*i.e.* 33.1 km, SD = 19.6 km). Nineteen bioclimatic variables were obtained from the Worldclim database (www.worldclim.org) and extracted using DIVAgis software (www.diva-gis.org). These 19 bioclimatic variables were derived from monthly temperature and rainfall values, based on averages calculated for the 1960-90 period with a spatial resolution of 1 km². Six additional climatic variables corresponding to relative humidity for each of the four seasons, altitude and aridity were obtained as described in Hancock *et al.* (2011).

A sample of the upper soil layer was collected in each stand. Soil samples were transferred to the greenhouse and air-dried. Soil samples were then stored in the laboratory at room temperature. Each stand was characterized for 14 edaphic factors (Table 1). Mean pH per stand was calculated using two soil subsamples. Maximal water holding capacity of two

samples of 200 cm^3 per stand was measured as the amount of water held in soil after excess water had been drained away. The twelve other soil properties (content of total N and organic C, C/N ratio, content of organic matter, concentrations of P_2O_5, K, Ca, Mg, Mn, Al, Na and Fe) were assessed at INRA Arras (France, www.lille.inra.fr/las).

To estimate intra- and inter-specific competition intensities, a 50 x 50 cm quadrat divided into 25 smaller squares (10 x 10 cm) was established in a representative area of each stand. Intraspecific competition was calculated as the density of intraspecific competitors based on the presence/absence of *A. thaliana* in each 10 x 10 cm square. Two interspecific competition indices were estimated as the presence/absence in each 10 x 10 cm square of either grasses or other herbs.

Data analysis

Phenological variation and spatial scale

Based on the 49 stands, we studied phenological variation in France by using the following statistical model according to our split-plot design:

$$\begin{aligned} Y_{ijklmc} = {} & \mu_{trait} + block_i + cohort_j + block_i \times cohort_j + region_k + stand_l(region_k) + \\ & family_m\left(stands_l\left(region_k\right)\right) + cohort_j \times region_k + cohort_j \times stands_l\left(region_k\right) + \\ & cohort_j \times family_m\left(stands_l\left(region_k\right)\right) + covBg2_c + \varepsilon_{ijklmc} \end{aligned} \qquad (1)$$

In this model, 'Y' is one of the six phenological traits, 'μ' is the overall mean; 'block' accounts for differences in micro-environment among the two experimental blocks; 'cohort' corresponds to the two germination cohorts grown; 'region', 'stand' and 'family' measure the effect of three spatial scales in France, *i.e.* regions, stands within regions and families within

stands; interaction terms involving the 'cohort' factor account for genetic variation in reaction norms among regions, among stands within regions and among families within stands; $cov_{Bg\text{-}2}$ is a covariate accounting for array effects within blocks (phenotypic mean of the two Bg-2 replicates per array was used as a covariate); and 'ε' is the residual term. All factors were treated as fixed effects, except for 'family' that was treated as a random effect. For calculating *F*-values, terms were tested over their appropriate denominators. Given the split-plot design used in this study, the variance associated with 'Block x Cohort' was for example used as the error term for testing the 'Block' and 'Cohort' effects. Raw data were Box-Cox transformed to satisfy the normality and equal variance assumptions of linear regression. Model fitting was conducted using the PROC MIXED procedure in SAS 9.3 (SAS Institute Inc., Cary, North Carolina, USA).

Within each cohort, the model described above was used excluding the terms involving '$cohort_j$'. The contribution of each factor to the total phenotypic variance was estimated using variance component analysis, treating all factors as random effects and excluding $cov_{Bg\text{-}2}$. Model fitting was conducted using the PROC VAR COMP procedure in SAS 9.3 (REML method; SAS Institute Inc., Cary, North Carolina, USA).

Within each cohort, a within-stand coefficient of genetic variation CV was calculated for each phenological trait based on variance components estimated by REML.

Phenological variation scored on the worldwide accessions and the natural accessions from the French Regmap was analyzed with the following model treating all the factors as fixed effects, except for 'accession' that was treated as a random effect (PROC MIXED procedure in SAS 9.3):

$$Y_{ijnc} = \mu_{trait} + block_i + cohort_j + block_i \times cohort_j + accession_n + cohort_j \times accessions_n + covBg2_c + \varepsilon_{ijnc} \quad (2)$$

For calculating *F*-values, terms were tested over their appropriate denominators. Raw data were Box-Cox transformed to satisfy the normality and equal variance assumptions of linear regression.

For each cohort, Best Linear Unbiased Predictions (BLUPs) were obtained for each genotype (*i.e.* natural accessions and families) using the PROC MIXED procedure in SAS 9.3 (SAS Institute Inc., Cary, North Carolina, USA):

$$Y_{igc} = \mu_{trait} + block_i + genotype_g + covBg2_c + \varepsilon_{igc} \quad (3)$$

Genetic diversity and spatial scale of genetic variation

Six genetic diversity parameters were estimated for each of the 49 stands. Mean gene diversity (H_S), percentage of polymorphic loci (*PL*), mean number of observed alleles per locus (n_a) and mean allelic richness per locus (R_S) were estimated using FSTAT v.2.9.3 (Goudet 1995). Number of haplogroups (*HG*) and number of private haplogroups (*PHG*) were estimated as described in Platt *et al.* (2010a), with a 0.01 per-site genotyping error rate and a 0.05 *p*-value threshold to exclude an additional French family from a haplogroup. Only the French families with a level of heterozygosity below 1% (n = 710) were considered for the estimation of *HG* and *PHG*. A Multivariate Analysis of Variance (MANOVA) was used to test the effect of either French region or habitat type on the six genetic diversity parameters simultaneously (PROC MANOVA in SAS 9.3, SAS Institute Inc., Cary, North Carolina, USA). Raw data for *HG* and *PHG* were log transformed to satisfy the normality and equal variance assumptions of MANOVA.

Spatial partitioning of the total genetic variation was estimated for all loci using the *R* package 'hierfstat' (Goudet 1995; de Meeûs & Goudet 2007). Hierarchical *F*-statistics and variance components were calculated for the 'region', 'stands' and 'individual' spatial levels. Ten thousand bootstraps were performed to estimate confidence intervals for the *F*-statistics. Within region, an among population *F*-statistic (Weir & Cockerham 1984) was calculated by averaging over the 135 SNPs F_{ST} values obtained using the *R* package 'pegas' (Paradis 2010).

Spatial scale of ecological variation

For each ecological factor, the contribution of regions to the total ecological variance was estimated by variance component analysis (PROC VARCOMP procedure in SAS 9.3), according to the following statistical model:

$$Y_i = \text{region}_i + \varepsilon_i \qquad (4)$$

where Y stands for the different ecological variables and 'region' corresponds to the four French regions. Raw data for the three competition indices were arc-sin transformed to satisfy the normality and equal variance assumptions of linear regression. Relationships among the 42 ecological factors were assessed with both Pearson and Spearman correlation coefficients.

Untangling the portions of phenotypic variation resulting from adaptive and non-selective processes.

To estimate the portion of phenological variation that could be explained by neutral genetic variation, a principal component analysis (PCA) was first performed on the 135 SNPs, recoded as binary haploid SNP genotypes (0 and 1 corresponding to the common and rare

allele at each SNP, respectively). Principal components (PCs) with eigenvalues ≥1 were then included in the following linear model:

$$Y_i = \mu_{trait} + PC1_i + ... + PCN_i + \varepsilon_i \quad (5)$$

where Y is a vector of phenological BLUPs, μ_{trait} is the overall phenological trait mean, PC1…PCN are the genetic principal components included in the model, and ε_i are the residuals of the model. The residuals were considered to represent the phenological variation minus the effect of non-selective processes. The adjusted r-square obtained for each trait can be seen as an upper limit for the amount of phenotypic variation explained by non-selective processes.

Estimating the geographical scale of adaptive variation through the identification of the putative selective agents on phenology.

Partial least square regression (PLSR) was used to identify the ecological factors potentially acting as selective pressures on phenological traits in France. PLSR identifies combinations of ecological variables that, unlike PCA, maximize the variation explained in a response variable (Geladi & Kowalski 1986; Carrascal *et al.* 2009), such as seed dormancy (Wagmann *et al.* 2012) or our phenological traits in this study. PLSR was chosen for two reasons. First, the number of explanatory variables (*i.e.* ecological factors) is not limited by the number of observations (*i.e.* stands). Second, PLSR can be carried out on non-independent explanatory variables.

Using the *R* package 'pls' (Mevik & Wehrens 2007), PLSR was performed twice. In the first analysis, we investigated the relationships between standardized ecological factors and phenological median per stand calculated from standardized BLUPs. In the second analysis, we investigated the relationship between standardized ecological factors and phenological median per stand calculated from standardized residuals obtained from equation 5. Standardization of BLUPs and residuals allowed us to compare regression coefficients obtained from PLSR before or after accounting for non-selective processes, respectively. In both analyses, leave-one-out cross-validation was used to determine the optimal number of components to be included in the model. Significance of regression coefficients was tested by approximate *t*-tests based on jackknife variance estimates.

GWA mapping at different geographical scales

All 352 natural accessions corresponding to the WORLD, EUROPE, FRANCE, BURGUNDY, MIB and TOU sets have been genotyped for 214,051 SNPs evenly spaced across the genome (Horton *et al.* 2012, http://bergelson.uchicago.edu/regmap-data/). In order to fine-map genomic regions associated with natural phenological variation in each set, we first ran a Wilcoxon rank-sum test on the association between phenotypes and genotypes for each marker (Atwell *et al.* 2010). GWA mapping was then run using a mixed-model approach implemented in the software EMMAX (Efficient Mixed-Model Association eXpedited, Kang *et al.* 2010). To control for population structure, this model includes a genetic kinship matrix *K* accounting for genome-wide patterns of relatedness among the accessions (i.e. identity-by-state). These analyses were based on BLUPs obtained by the statistical model described in equation (3). Because of bias due to rare alleles, we only considered SNPs with Minor Allele

Relative Frequency (MARF) > 10% (Kang *et al.* 2010; Brachi *et al.* 2010) at the geographic scale considered.

Candidate genes close to highly associated SNPs were identified among a list of 282 *a priori* flowering time candidate genes described in Brachi *et al.* (2010). This list was enriched with flowering time candidate genes from the following website (Max Planck Institute for plant breeding research, http://www.mpipz.mpg.de/14637/Arabidopsis_flowering_genes), resulting in a list of 328 candidate genes in total (Table S2).

Results

Natural diversity of *A. thaliana* in France: phenological, genetic and ecological characterization of 49 stands

Phenological variation: comparison with a set of worldwide natural accessions

The distribution of bolting time (BT) at large geographical scales, i.e. WORLD and EUROPE scales, was bimodal with an apparent excess of early and late bolting accessions (Figure 1). Apparent excess of late bolting accessions in the fall cohort might result, in part, to an unsatisfied vernalization requirement in Scandinavian accessions (Shindo *et al.* 2005). Interestingly, in the spring cohort, the natural variation observed for BT in the local population TOU (n = 69) almost covered the range of natural variation for BT at the worldwide scale. The distributions of BT in the four French regions also differed from its distribution in worldwide accessions for both germination cohorts (Figure 1). Noteworthy was the excess of intermediate values of BT in the Languedoc region compared to worldwide accessions. Trends similar to those for BT were observed for flowering time (FT, Figure S1).

Distributions of the flowering interval (INT) showed very small differences among geographic scales, with within-region (and even within-population) diversity spanning worldwide diversity (Figure S1). Flowering and reproductive periods (FP and RP) displayed patterns similar to INT except for the Languedoc region. In the spring cohort, Languedoc families exhibited an excess of short flowering and reproductive periods. In contrast, in the fall cohort, an excess of long flowering periods was observed for Languedoc (Figures 1 and S1). In both cohorts, Languedoc families exhibited extensive flowering-to-reproductive ratio (FRR) natural variation in comparison to worldwide accessions and to the other three French regions (Figure 1).

The interaction between the 352 natural accessions and germination cohort was highly significant for BT and FT, and non-significant for the other traits (Table 2). In the French populations, the interaction between germination cohort and either region or stand within region was highly significant for all traits (except for FT and FRR at the regional scale; Table 3). Contrasting responses to winter treatment among phenological traits can be observed at the scale of stands. For example, while the winter treatment decreased BT in the BAU and BRI stands from the North, it increased FP in BAU and decreased FP in BRI (Figure 2). The interaction between germination cohort and the families was only significant for FP, suggesting similar responses to winter treatment among all families within a specific stand for the other five traits (Table 3).

In the four French regions, the partitioning of phenological variation was similar in the two germination cohorts (Tables S3 and S4). Variation in BT was partitioned among regions (Fall: 26.3%, Spring: 28.7%), among stands within regions (Fall: 37.1%, Spring: 37.6%) and among families within stands (Fall: 19.0%, Spring, 19.0%; Figure 2). Extensive within-region variation was observed in Burgundy and Languedoc, whereas Brittany and North regions

were less variable, displaying mainly early bolting families (Figure 2). Results similar to BT were found for FT (Tables S3 and S4). For INT, FP, RP and FRR, a large fraction of phenotypic variation (from 50.3 % to 86.7 %) remained unexplained, suggesting high levels of phenotypic plasticity in those traits to uncontrolled micro-environmental variation in the greenhouse. For these phenological traits, the remaining phenotypic variation was mostly observed at local scales, *i.e.* within stands and families. Stands from Burgundy and Languedoc were significantly more diverse than stands from Brittany and North for BT, INT and FT in the fall cohort and for BT, FT and RP in the spring cohort (Figure 2, Table S5).

Neutral genetic variation

A total of 758 French families were successfully genotyped for this set of 135 SNPs (Table S1). Extensive variation was observed across the 49 French stands for mean gene diversity (0 – 0.332) and percentage of polymorphic loci (0% – 88.2%) (Table S6). A significant 'region' effect was detected on the six genetic diversity parameters tested simultaneously (Wilks' Lambda F value = 2.26, P = 0.0051). No significant 'region' effect was detected for mean gene diversity, the percentage of polymorphic loci, the mean number of observed alleles per locus, the mean allelic richness per locus and the number of haplogroups (Table S6). A significant 'region' effect was detected for the number of private haplogroups per stand (F = 5.50, P = 0.0026), with stands from Burgundy having on average more private haplogroups than stands from Brittany, Languedoc and North after a Tukey's Studentized Range (HSD) test (Table S6). No significant 'habitat type' effect was detected on the six genetic diversity parameters tested simultaneously (Wilks' Lambda F value = 0.74, P = 0.71). No genetic diversity parameter was significantly associated with either altitude or the number of families genotyped per stand (data not shown). Noteworthy was the significant,

positive correlation observed between the coefficient of genetic variation in BT (or FT) and mean gene diversity (fall cohort: BT Spearman's rho = 0.506 P = 0.0003, FT Spearman's rho = 0.499 P = 0.0005; spring cohort: BT Spearman's rho = 0.445 P = 0.0017, FT Spearman's rho = 0.411 P = 0.0056).

A hierarchical analysis of molecular variation (AMOVA) revealed that genetic variation was partitioned among regions (9.1 %), among stands within regions (55.4%) and among families within stands (33.7%). A strong stand subdivision was observed within each region, with fixation index F_{ST} (from 0.48 in Burgundy to 0.61 for Brittany) similar to values reported in a previous study on other French stands (Le Corre 2005).

Ecological variation

Climate variables appeared more strongly correlated with each other than with edaphic variables and competition indices (Figure S2). Both edaphic variables and competition indices were only weakly inter-correlated (Figure S2). The variance within bioclimatic variables appeared to be mainly partitioned among regions, the 'region' effect explaining on average 90.3% of the bioclimatic variance in France (Figure S3). In contrast, the variance of edaphic variables and competition indices was mostly partitioned among stands within regions, the 'region' effect explaining on average only 22.5% and 5.2% of the edaphic and competition variance in France, respectively (Figure S3).

Investigation of the geographical scale of adaptive phenological variation

The percentage of phenological variation explained by genetic Principal Components with an eigenvalue ≥ 1 (n = 31) ranged from ~52% (BT and FT) to 8% (RP) in the fall cohort

(mean = 27.2%), and from ~63% (BT and FT) to 10% (INT) in the spring cohort (mean = 35.6%).

Five main features characterized the relationships between phenological and ecological variation. First, non-selective processes may have generated spurious relationships between ecological factors and phenological variation. At the scale of France, all significant phenology - ecology relationships disappeared when neutral genetic similarities among families were accounted for (Figure 3). In contrast, within regions, correcting for neutral genetic similarities revealed new significant phenology - ecology relationships (Figure 3). Second, despite the small number of stands sampled within each region, the percentage of phenological variation explained by PLSR components was generally higher at the within-region than at the broader scale of the France (Figure 3). Third, at the within-region scale, edaphic factors and competition indices showed relationships with phenological traits that were as strong as those observed for climatic factors. For example, significant phenology – ecology relationships was detected between the density of grasses and BT (spring cohort), FT and FRR (fall cohort) in Burgundy (Figure 3). Fourth, phenological traits generally displayed significant relationships with a particular ecological variable in only one or two of the four regions. For example, in the spring cohort, FT variation was significantly associated with climate variation in Burgundy and North, but not in Brittany and Languedoc (Figure 3). Finally, the phenology – ecology relationships depended on the germination cohort season. In the North, significant relationships were detected between FP and climate variation only in the fall cohort. FP and RP displayed a significant relationship with climate only in the spring cohort in Languedoc.

Identification of the genomic regions associated with phenological variation from worldwide scale to local population scale

For BT and FT, an excess of low p-values due to confounding by population structure was found at the worldwide and European scale (Figure S4). This excess of low *p*-values decreased at the French and regional (*i.e.* Burgundy) scale, and was almost eliminated at the local scales. For INT, FP and FRR, the effect of population structure was mainly observed at the French and regional scales, and was almost undetectable at the local scales (Figure S4). No excess of low *p*-values due to confounding by population structure was detected for RP at any geographical scale. For each trait, cohort and geographical scale, the excess of low *p*-values detected from the Wilcoxon rank-sum analyses was eliminated from a mixed-model approach that takes genetic similarity among accessions into account (Figure S4).

As illustrated by BT in the spring cohort, the identity of the genomic regions associated with natural variation depends on the geographical scale considered (Figure 4 and Table S7). While a neat peak of association was detected in the vicinity of *DOG1* (*DELAY OF GERMINATION 1*) on chromosome 5 at the worldwide and European scales, GWA mapping revealed a unique and strong peak of association centered on the flowering time gene *FRI* (*FRIGIDA*) on chromosome 4 at the French, regional and local scales (Figures 4 and S5). In the TOU local population, the highest associated SNP is located within *FRI* (Figure S5). In the fall cohort, association peaks centered on *FRI* were also found at the French and smaller geographical scales. In contrast, six and five new association peaks were detected at the Bonferroni threshold at the worldwide and European scales, respectively. Two of these new peaks correspond to the circadian gene *CIRCADIAN CLOCK ASSOCIATED 1* (*CCA1*) on chromosome 2 and *DWARF IN LIGHT 2* (*DFL2*) on chromosome 4. No flowering time candidate gene was found in a ~30 kb genomic region of chromosome 5 strongly associated

with BT in the world and in Europe (i.e. region 6,623 kb – 6,652 kb on chromosome 5, Figure S5). GWA mapping also revealed an association peak after the winter treatment that is only present in the local TOU population. This association peak located 120 kb upstream of the strongest association peak detected at the worldwide and European scales in the fall cohort contains no flowering time candidate genes (Figure S5).

For FT, GWA mapping revealed the same genomic regions as for BT at the French and smaller geographical scales (Figure S4). In contrast, no shared association peak was detected between BT and FT at the worldwide and European scales (Figure S4). Two and nine association peaks were detected for FT in the fall cohort at the Bonferroni threshold at the worldwide and European scales, respectively. Only one association peak with no flowering time candidate gene is shared between these two scales. Two association peaks for FT in the fall cohort at the European scale colocalizes with *SQUAMOSA PROMOTER BINDING PROTEIN-LIKE 4* (*SPL4*) and *SPA-RELATED 4* (*SPA4*). In the fall cohort, *SPL4* and *SPA4* were also found in the vicinity of strong association peaks detected for INT at both the worldwide and European scales (Figure S4). The strongest SNP ($P = 4.02 \times 10^{-10}$) for INT in the fall cohort resides in the gene *At1g52880*, a homolog of the petunia gene *NAM* that is involved in the development of the shoot. No association peak was detected at the Bonferoni threshold for INT either in the fall cohort at the French (or smaller geographical) scales or in the spring cohort.

Few association peaks have been detected for post-flowering traits. No association peak was detected for RP, whereas only three association peaks were detected at the Bonferoni threshold for FP across the two cohorts and the six geographical scales (Figure S4). For FRR, GWA mapping revealed a unique peak of association in the region centered on *FRI* at the French, regional and local scales in the spring cohort, but only at the French and

regional scales in the fall cohort (Figure S4). No association peak was detected for FRR at the worldwide and European scales.

A list of flowering time candidate genes in the vicinity of 20 kb of the 100 most associated SNPs for each 'trait x cohort x geographical scale' combination is available in Table S7.

Discussion

Extensive natural phenological variation at the local scale

Our phenological characterization of 49 French stands, worldwide accessions and accessions from the French Regmap revealed a relatively continuous distribution of bolting and flowering time, with an excess of intermediate values from populations collected in the South-East of France. This stands in contrast to the distribution observed in current worldwide collections that show a bimodal distribution encompassing two major types in controlled growth conditions, i.e. summer and winter annuals (Shindo *et al.* 2007). This phenological discrepancy suggests that thorough characterization of new populations should improve our picture of the natural variation observed in *A. thaliana*.

Our study also revealed highly polymorphic local populations at the phenological level, with some populations (like TOU) almost spanning the range of natural variation observed at the worldwide scale. For this reason, single accessions cannot reliably be used to characterize the phenotypes or genotypes present at ecologically distinct sites. At least three hypotheses can explain the high phenological polymorphism observed in natural populations of *A. thaliana* within France. First, non-selective processes may have played a major role in shaping the phenological variation at the within-population scale. This hypothesis is

consistent with the patterns observed for BT and FT, which showed significant, positive relationships between phenological diversity and neutral genetic diversity. Second, phenological variation at the within-population scale may reflect adaptation to a fine-grained environmental variation, as previously observed in other plant species (inter-specific competition: Turkington & Harper 1979a; b; herbivory: Schemske 1984). Third, the strength and direction of selection in natural populations may vary considerably over time (Siepielski *et al.* 2009). Although temporal dynamics of phenotypic selection may lead to the selection of plastic genotypes (Stomp *et al.* 2008), the cost of plasticity may promote the coexistence of specialist genotypes in the same population (Dewitt *et al.* 1998).

Geographical scale of adaptive phenological variation

The need to account for effects of non-selective processes is increasingly considered in the study of phenology – ecology relationships (Keller & Taylor 2008; Keller *et al.* 2009; Méndez-Vigo *et al.* 2011; Chun *et al.* 2011; Hancock *et al.* 2011; Kooyers & Olsen 2012; Lee & Mitchell-Olds 2012). In this study, phenology – ecology relationships were strongly affected by neutral genetic diversity at the scale of France, suggesting that phenology – ecology relationships could be spurious at this geographical scale. Interestingly, controlling for neutral genetic diversity revealed new phenology – ecology relationships at the regional scale. This is a similar phenomenon to the GWA peaks that become evident only after population structure is controlled in mixed model association mapping (Kang *et al.* 2008). Here, correction for non-selective processes may reduce the phenotypic variance within populations or within types of environment, revealing otherwise non-significant relationships.

Similar to studies on the identification of genomic regions associated with natural phenotypic variation by GWA mapping, we should acknowledge that our approach may

produce false negatives after controlling for the effects of non-selective processes (Brachi *et al.* 2010), i.e. true phenology – ecology associations that overlap with neutral genetic diversity. Significant phenology – ecology relationships detected without controlling for the effects of non-selective processes can be validated in experiments in controlled conditions (see below).

The ecological characterization of 49 French stands suggests an overlap of environmental grains, with variation for edaphic and intra- and inter-specific competition factors observed at a finer scale than climate variables. This is in agreement with the ecological characterization of natural populations of *A. thaliana* in the Iberian Peninsula (Montesinos-Navarro *et al.* 2011). As in other plant species (Bischoff *et al.* 2006; Becker *et al.* 2008; Manel *et al.* 2010), the selective agents shaping the adaptive population differentiation for phenological traits appear to act at multiple scales in *A. thaliana*, but seem to be stronger at the within-region level.

The apparent lack of relationship between some traits and ecological factors may originate from (i) a lack of statistical power; (ii) reduced phenological or ecological variation in some within-region samples; (iii) uncharacterized ecological factors acting on phenology, especially biotic agents in natural populations like herbivores (Lennartsson *et al.* 1997), pathogen attacks (Roux *et al.* 2010), pre-dispersal seed predators (Elzinga *et al.* 2007) or pollinators (Hoffmann *et al.* 2003); (iv) the absence of relationships between phenological variation and fitness in some locations; (v) fine-scale environmental differentiation within populations and/or (vi) selection for phenotypic plasticity in natural habitats (Pérez de la Vega 1996); and (vii) inaccurate estimates of phenological variation in our greenhouse conditions. It is important to note that our estimation of natural variation in phenology was done under controlled standard greenhouse conditions. Although phenological traits in *A. thaliana* often

display strong genotype by environment interactions when tested across greenhouse and common garden conditions (Bergelson & Roux 2010; but see Méndez-Vigo *et al.* 2012), we believe our measures are meaningful as they reflect genetically fixed variation among families.

In a recent study based on an interconnected mapping population of 117 Recombinant Inbred Lines (RILs), the influence of reproductive timing on fitness at the phenotypic level was found to greatly differ among four field sites across the native European range of *A. thaliana* (Fournier-Level *et al.* 2013). It would be worth measuring phenological traits on the 49 French stands of this study in more ecologically realistic conditions to check whether an equivalent amount of phenological variation explained by non-selective processes is consistent with the results of this study, and whether adaptive phenological variation estimated in field settings is associated with others ecological factors.

Diversity of selective agents acting on phenology among French accessions

The putative selective agents acting on phenological traits were clearly dependent on the region, the trait considered, and the germination cohort; the latter confirming the importance of germination cohort in determining the environmental conditions experienced by plants after germination in *A. thaliana* (Donohue 2002; Donohue *et al.* 2005). The geographical heterogeneity in the relationships between phenological traits and ecological variables is also consistent with results from a recent study suggesting the selection acts on different traits and loci in different locations across Europe (Fournier-Level *et al.* 2013).

Although rarely mentioned in *A. thaliana*, relationships with edaphic factors and plant-plant interactions were as strong as correlations with climatic factors, which were previously

suggested to be selective agents on bolting time in *A. thaliana* (Stinchcombe *et al.* 2004). Naturally occurring variation in edaphic factors, as well as long-term application of different fertilizer treatments, are known to act as selective pressures on the phenology of plants (Snaydon & Davies 1982; Rajakaruna & Bohm 1999; Kittelson & Maron 2001; Antonovics 2006), but have seldom been tested in *A. thaliana*.

Significant phenotype – ecology correlations, however, are only suggestive of the role of ecological factors in selecting on phenology. Experiments in controlled conditions are clearly needed for validation. In a previous study, both estimates of seed production and experimental evolution were used to estimate the adaptive values of phenological traits in two stressful environments, i.e. water stress and interspecific competition with grasses (Brachi *et al.* 2012). The relationships of RP and FRR with aridity and interspecific competition with grasses that were detected in this study are congruent with the predictions of our previous study. This congruence suggests that aridity and interspecific competition with grasses are true selective agents acting on phenology in natural populations in France.

Complementarity of GWA mapping studies at different geographical scales

The geographical scale of adaptive phenological variation observed in the 49 stands led us to test whether the identification of the genomic regions associated with phenological variation differed across geographical scales. In a first attempt, combining both worldwide accessions and accessions from the French Regmap indicates that the genomic regions associated with phenological variation appear to depend on the geographical scale considered. Strong association peaks were detected at the continental and local geographical scales. The variants underlying those peaks may reflect adaptation to coarse-grained and fine-grained ecological variation, respectively.

Performing GWA using sets of accessions spanning different geographical scales also proved to be promising in resolving major limitations of GWA mapping with regard to population structure, rare alleles and allelic heterogeneity. First, as expected from the pattern of isolation by distance observed across the species range of *A. thaliana* (Platt, Horton, *et al.* 2010a), confounding by population structure was greatly reduced at small geographical scales. This was especially true for phenotypic traits like bolting time or flowering time whose natural variation overlaps with population structure at the worldwide scale (Zhao *et al.* 2007). While dual linkage-association mapping has been shown to reduce the rate of false positives and false negatives in GWA studies (Brachi *et al.* 2010), running GWA mapping in regional or local panels may also circumvent confounding by population structure.

Second, association studies often lack power to detect rare variants. However, rare variants at the worldwide scale may be common in local populations, making them easier to detect in association studies. For example, using a regional collection of wild *A. thaliana* genotypes in the Iberian Peninsula, Sánchez-Bermejo *et al.* (2012) recently identified a novel *cis*-regulatory *FLC* polymorphism located only in the North-East of Spain that is associated with an increase in vernalization sensitivity. In our study, bolting time (BT) in the fall cohort revealed a significant association peak located on chromosome 5 that was only detected in the TOU population, suggesting that variants may be so rare that they are only present in one local population.

Third, because the flowering time gene *FRI* is a classic example of allelic heterogeneity in *A. thaliana* (Atwell *et al.* 2010), we expected to detect this gene when running GWA mapping at geographical scales smaller than the worldwide scale. Despite the description of 13 *FRI* non-functional alleles in France (Le Corre *et al.* 2002; Le Corre 2005; Shindo *et al.* 2005), *FRI* was detected using the French mapping panel and local French

populations. Two hypotheses can explain our ability to map *FRI* in the French mapping panel despite the presence of numerous *FRI* non-functional alleles. First, one of the non-functional allele may be much more prevalent than the other non-functional alleles. Second, one polymorphism located in *FRI* may be shared by several non-functional alleles (Platt, Vilhjálmsson, *et al.* 2010b).

In this study, we used all the worldwide and French accessions for which both seeds and 214k SNPs data were available at the beginning of the experiment. The on-going genomic characterization of the 49 stands will soon reveal whether (i) we obtain similar biological conclusions as with the French Regmap panel and (ii) the genomic regions associated with phenological variation differ across the four French regions, and overlap with imprints of selection.

Conclusion

While next-generation sequencing (NGS) technologies will facilitate the identification of causal polymorphisms underlying natural variation of complex traits (Brachi *et al.* 2011), it should not be forgotten that the genetics of adaptation may largely depend on the environmental grain both at the spatial and temporal scales (Kopp & Hermisson 2007; Roux & Reboud 2007; Roux *et al.* 2008; Kopp & Hermisson 2009). By investigating putative selective pressures acting on phenological traits, we improved our understanding of the geographical scale of adaptive variation in *A. thaliana.* Our results suggest that phenological variation at small geographical scales might be adaptive and that different phenological traits could be under selection in different regions. While GWA mapping is powerful in detecting common genes underlying natural variation at a worldwide scale, it suffers from limitations like confounding by population structure, rare alleles and allelic heterogeneity. In our study,

the geographical scale of adaptive variation suggested for phenological traits and the scale at which associations were detected are consistent. Mapping in regional panels of accessions or even in local populations may resolve limitations of GWA studies. Overall our study suggests that mapping panels that span geographical areas over which phenotypic variation appears to be adaptive have the potential to unravel the genetics underlying adaptive complex trait and allow reconstruction of the adaptive walks that natural populations follow towards local phenotypic optima.

Acknowledgments

We are grateful to Eric Imbert and Valérie Le Corre for their assistance in locating Arabidopsis stands. Special thanks are given to Adeline Courseaux, Cédric Glorieux and Cécile Godé for their assistance during the greenhouse experiment and for DNA extraction protocols. We also thank Alexander Platt for sharing his script for estimating the number of haplogroups per stand. This study was supported by a Ph.D. fellowship from the French Ministry of Higher Education and Research to B.B. and R.V., an ANR BLANC grant for the QUANTIREX project (NT09_473214), and by a Dropkin fellowship and an NIH grant to J. Bergelson.

References

Adler D (2005) vioplot: Violin plot. *R package version 0.2, URL http://CRAN. R-project. org/package= vioplot*.

Anastasio AE, Platt A, Horton M *et al.* (2011) Source verification of mis-identified *Arabidopsis thaliana* accessions. *The Plant Journal*, **67**, 554–566.

Antonovics J (2006) Evolution in closely adjacent plant populations X: long-term persistence of prereproductive isolation at a mine boundary. *Heredity*, **97**, 33–37.

Atwell S, Huang YS, Vilhjálmsson BJ *et al.* (2010) Genome-wide association study of 107 phenotypes in *Arabidopsis thaliana* inbred lines. *Nature*, **465**, 627–631.

Becker U, Colling G, Dostal P, Jakobsson A, Matthies D (2006) Local adaptation in the monocarpic perenial *Carlin vulgaris* at different spatial scales across Europe. *Oecologia*, **150**, 506–518.

Becker U, Dostal P, Jorritsma-Wienk LD, Matthies D (2008) The spatial scale of adaptive population differentiation in a wide spread, well-dispersed plant species. *Oikos*, **117**, 1865–1873.

Belotte D, Curien JB, Maclean RC, Bell G (2003) An experimental test of local adaptation in soil bacteria. *Evolution*, **57**, 27–36.

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

Bischoff A, Crémieux L, Smilauerova M *et al.* (2006) Detecting local adaptation in widespread grassland species- the importance of the scale and local plant community. *Journal of Ecology*, **94**.

Bomblies K, Yant L, Laitinen RA *et al.* (2010) Local-scale patterns of genetic variability, outcrossing, and spatial structure in natural stands of *Arabidopsis thaliana*. *PLoS genetics*, **6**, e1000890.

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

Brachi B, Faure N, Horton M *et al.* (2010) Linkage and Association Mapping of *Arabidopsis thaliana* Flowering Time in Nature. *PLoS genetics*, **6**, e1000940.

Brachi B, Morris GP, Borevitz JO (2011) Genome-wide association studies in plants: the missing heritability is in the field. *Genome biology*, **12**, 232.

Carrascal LM, Galvan I, Gordo O (2009) Partial least squares regressions as an alternative to current regression methods used in ecology. *Oikos*, **118**, 618–690.

Chun YJ, Le Corre V, Bretagnolle F (2011) Adaptive divergence for a fitness-related trait among invasive *Ambrosia artemisiifolia* populations in France. *Molecular Ecology*, **20**, 1378–1388.

Clark RM, Schweikert G, Toomajian C *et al.* (2007) Common Sequence Polymorphisms Shaping Genetic Diversity in *Arabidopsis thaliana*. *Science*, **317**, 338–342.

Conner JK (2010) Natural Selection in Plants 151 Years after The Origin: Introduction. *International Journal of Plant Sciences*, **171**, 927–929.

de Meeûs T, Goudet J (2007) A step-by-step tutorial to use HierFstat to analyse populations hierarchically structured at multiple levels. *Infection, Genetics and Evolution*, **7**, 731–735.

Dewitt T, Sih A, Wilson D (1998) Costs and limits of phenotypic plasticity. *Trends in Ecology & Evolution*, **13**, 77–158.

Donohue K (2002) Germination timing influences natural selection on life-history characters

in *Arabidopsis thaliana*. *Ecology*, **83**, 1006–1016.
Donohue K (2005) Seeds and seasons: interpreting germination timing in the field. *Seed Science Research*, **15**, 175–187.
Donohue K, Dorn LA, Griffith C *et al.* (2005) The evolutionary ecology of seed germination of *Arabidopsis thaliana*: variable natural selection on germination timing. *Evolution*, **59**, 758–770.
Elzinga JA, Atlan A, Biere A *et al.* (2007) Time after time: flowering phenology and biotic interactions. *Trends in Ecology & Evolution*, **22**, 432–439.
Fournier-Level A, Korte A, Cooper MD *et al.* (2011) A Map of Local Adaptation in *Arabidopsis thaliana*. *Science*, **334**, 86–89.
Fournier-Level A, Wilczek AM, Cooper MD *et al.* (2013) Paths to selection on life history loci in different natural environments across the native range of *Arabidopsis thaliana*. *Molecular Ecology*.
Franks SJ, Sim S, Weis AE (2007) Rapid evolution of flowering time by an annual plant in response to a climate fluctuation. *Proceedings of the National Academy of Sciences*, **104**, 1278–1282.
Gaut B (2012) *Arabidopsis thaliana* as a model for the genetics of local adaptation. *Nature Genetics*, **44**, 115–121.
Geladi P, Kowalski BR (1986) Partial least-squares regression: a tutorial. *Analytica Chimica Acta*, **185**.
Goudet J (1995) FSTAT (version 1.2): a computer program to calculate F-statistics. *Journal of Heredity*, **86**, 485–971.
Hancock AM, Brachi B, Faure N *et al.* (2011) Adaptation to Climate Across the *Arabidopsis thaliana* Genome. *Science*, **334**, 83–86.
Hereford J (2009) A Quantitative Survey of Local Adaptation and Fitness Trade-Offs. *The American Naturalist*, **173**, 579–588.
Hermisson J, Pennings PS (2005) Soft sweeps: molecular population genetics of adaptation from standing genetic variation. *Genetics*, **169**, 2335–2352.
Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, **25**, 1965–1978.
Hintze JL, Nelson RD (1998) Violin Plots: A Box Plot-Density Trace Synergism. *The American Statistician*, **52**, 181–184.
Hoffmann MH, Bremer M, Schneider K *et al.* (2003) Flower visitors in a natural population of *Arabidopsis thaliana*. *Plant Biology*, **5**, 491–494.
Horton MW, Hancock AM, Huang YS *et al.* (2012) Genome-wide patterns of genetic variation in worldwide *Arabidopsis thaliana* accessions from the RegMap panel. *Nature Genetics*, **44**, 212–216.
Huang X, Wei X, Sang T *et al.* (2010) Genome-wide association studies of 14 agronomic traits in rice landraces. *Nature Genetics*, **42**, 961–967.
Inouye DW (2008) Effects of climate change on phenology, frost damage, and floral abundance of montane wildflowers. *Ecology*, **89**, 353–362.
Kang HM, Zaitlen NA, Wade CM *et al.* (2008) Efficient control of population structure in model organism association mapping. *Genetics*, **178**, 1709–1723.
Kang H, Sul J, Service S *et al.* (2010) Variance component model to account for sample structure in genome-wide association studies. *Nature Genetics*, **42**, 348–402.
Kawecki TJ, Ebert D (2004) Conceptual issues in local adaptation. *Ecology Letters*, **7**, 1225–1241.

Keller S, Taylor D (2008) History, chance and adaptation during biological invasion: separating stochastic phenotypic evolution from response to selection. *Ecology Letters*, **11**, 852–918.
Keller SR, Sowell DR, Neiman M, Wolfe LM, Taylor DR (2009) Adaptation and colonization history affect the evolution of clines in two introduced species. *New Phytologist*, **183**, 678–690.
Kittelson PM, Maron JL (2001) Fine-scale genetically based differentiation of life-history traits in perennial shrub *Lupinus arboreus*. *Evolution*, **55**, 2429–2438.
Kooyers NJ, Olsen KM (2012) Rapid evolution of an adaptive cyanogenesis cline in introduced North American white clover (*Trifolium repens L.*). *Molecular Ecology*, **21**, 2455–2468.
Kopp M, Hermisson J (2007) Adaptation of a Quantitative Trait to a Moving Optimum. *Genetics*, **176**, 715–719.
Kopp M, Hermisson J (2009) The Genetic Basis of Phenotypic Adaptation I: Fixation of Beneficial Mutations in the Moving Optimum Model. *Genetics*, **182**, 233–249.
Laine AL (2005) Spatial scale of local adaptation in a plant-pathogen metapopulation. *Journal of Evolutionary Biology*, **18**, 930–938.
Lande R, Arnold SJ (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. *Molecular Ecology*, **14**, 4181–4192.
Le Corre V, Roux F, Reboud X (2002) DNA polymorphism at the FRIGIDA gene in *Arabidopsis thaliana*: extensive nonsynonymous variation is consistent with local selection for flowering time. *Molecular Biology and Evolution*, **19**, 1261–1271.
Lee CR, Mitchell-Olds T (2012) Environmental Adaptation Contributes to Gene Polymorphism across the *Arabidopsis thaliana* Genome. *Molecular Biology and Evolution*, **29**, 3721–3728.
Leimu R, Fischer M (2008) A Meta-Analysis of Local Adaptation in Plants (A Buckling, Ed,). *PloS one*, **3**, e4010.
Lennartsson T, Tuomi J, Nilsson P (1997) Evidence for an evolutionary history of overcompensation in the grassland biennial *Gentianella campestris* (Gentianaceae). *The American Naturalist*, **149**, 1147–1155.
Linhart YB, Grant MC (1996) Evolutionary significance of local genetic differentiation in plants. *Annual Review of Ecology and Systematics*, 237–277.
Manel S, Poncet B, Legendre P, Gugerli F, Holderegger R (2010) Common factors drive adaptive genetic variation at different spatial scales in *Arabis alpina*. *Molecular Ecology*, **19**, 3824–3859.
Merilä J, Sheldon BC, Kruuk LE (2001) Explaining stasis: microevolutionary studies in natural populations. *Genetica*, **112-113**, 199–222.
Mevik BH, Wehrens R (2007) The pls package: Principal component and partial least squares regression in R. *Journal of Statistical Software*, **18**, 1–25.
Méndez-Vigo B, Gomaa NH, Alonso-Blanco C, Xavier Picò F (2012) Among- and within-population variation in flowering time of Iberian *Arabidopsis thaliana* estimated in field and glasshouse conditions. *New Phytologist*, **197**, 1332–1343.
Méndez-Vigo B, Picó FX, Ramiro M, Martínez-Zapater J, Alonso-Blanco C (2011) Altitudinal and climatic adaptation is mediated by flowering traits and *FRI*, *FLC*, and *PHYC* genes in Arabidopsis. *Plant Physiology*, **157**, 1942–1997.

Mitchell-Olds T, Schmitt J (2006) Genetic mechanisms and evolutionary significance of natural variation in Arabidopsis. *Nature*, **441**, 947–952.
Mitchell-Olds T, Willis JH, Goldstein DB (2007) Which evolutionary processes influence natural genetic variation for phenotypic traits? *Nature Reviews Genetics*, **8**, 845–856.
Montesinos-Navarro A, Wig J, Pico F, Tonsor S (2011) *Arabidopsis thaliana* populations show clinal variation in a climatic gradient associated with altitude. *The New phytologist*, **189**, 282–376.
Munguía-Rosas MA, Ollerton J, Parra-Tabla V, De-Nova JA (2011) Meta-analysis of phenotypic selection on flowering phenology suggests that early flowering plants are favoured. *Ecology Letters*, **14**, 511–521.
Olson-Manning CF, Wagner MR, Mitchell-Olds T (2012) Adaptive evolution: evaluating empirical support for theoretical predictions. *Nature Reviews Genetics*, **13**, 867–877.
Orr HA (2005) The genetic theory of adaptation: a brief history. *Nature Reviews Genetics*, **6**, 119–127.
Parachnowitsch AL, Caruso CM (2008) Predispersal seed herbivores, not pollinators, exert selection on floral traits via female fitness. *Ecology*, **89**, 1802–1810.
Paradis E (2010) pegas: an R package for population genetics with an integrated-modular approach. *Bioinformatics (Oxford, England)*, **26**, 419–439.
Pérez de la Vega M (1996) Plant genetic adaptedness to climatic and edaphic environment. *Euphytica*, **92**, 27–38.
Picó FX (2012) Demographic fate of *Arabidopsis thaliana* cohorts of autumn-and spring-germinated plants along an altitudinal gradient. *Journal of Ecology*, **100**, 1009–1018.
Platt A, Horton M, Huang YS *et al.* (2010a) The Scale of Population Structure in *Arabidopsis thaliana* (J Novembre, Ed,). *PLoS genetics*, **6**, e10000843.
Platt A, Vilhjálmsson B, Nordborg M (2010b) Conditions Under Which Genome-Wide Association Studies Will be Positively Misleading. *Genetics*, **186**, 1045–1052.
Rajakaruna N, Bohm BA (1999) The edaphic factors and patterns of variation in *Lasthenia californica* (Asteraceae). *American Journal of Botany*, **86**, 1576–1596.
Rathcke B, Lacey EP (1985) Phenological patterns of terrestrial plants. *Annual Review of Ecology and Systematics*, **16**, 179–214.
Roux F, Reboud X (2007) Herbicide resistance dynamics in a spatially heterogeneous environment. *Crop Protection*, **26**, 335–341.
Roux F, Gao L, Bergelson J (2010) Impact of initial pathogen density on resistance and tolerance in a polymorphic disease resistance gene system in *Arabidopsis thaliana*. *Genetics*, **185**, 283–291.
Roux F, Paris M, Reboud X (2008) Delaying weed adaptation to herbicide by environmental heterogeneity: a simulation approach. *Pest Manag Sci*, **64**, 16–29.
Sandring S, Ågren J (2009) Pollinator-mediated selection on floral display and flowering time in the perennial herb *Arabidopsis lyrata*. *Evolution*, **63**, 1292–1300.
Sánchez-Bermejo E, Méndez-Vigo B, Picó FX, Martinez-Zapater JM, Alonso-Blanco C (2012) Novel natural alleles at FLC and LVR loci account for enhanced vernalization responses in *Arabidopsis thaliana*. *Plant, Cell & Environment*, **35**, 1672–1684.
Schemske DW (1984) Population structure and local selection in *Impatiens pallida* (Balsaminaceae), a selfing annual. *Evolution*, 817–1649.
Shindo C, Aranzana MJ, Lister C *et al.* (2005) Role of *FRIGIDA* and *FLOWERING LOCUS C* in determining variation in flowering time of Arabidopsis. *Plant Physiology*, **138**, 1163–1173.

Shindo C, Bernasconi G, Hardtke CS (2007) Natural Genetic Variation in Arabidopsis: Tools, Traits and Prospects for Evolutionary Ecology. *Annals of Botany*, **99**, 1043–1054.
Siepielski AM, DiBattista JD, Carlson SM (2009) It's about time: the temporal dynamics of phenotypic selection in the wild. *Ecology Letters*, **12**, 1261–1337.
Snaydon RW, Davies TM (1982) Rapid divergence of plant populations in response to recent changes in soil conditions. *Evolution*, 289–297.
Stinchcombe JR, Weinig C, Ungerer M *et al.* (2004) A latitudinal cline in flowering time in *Arabidopsis thaliana* modulated by the flowering time gene *FRIGIDA*. *Proceedings of the National Academy of Sciences, USA*, **101**, 4712–4717.
Stomp M, van Dijk M, van Overzee HT *et al.* (2008) The timescale of phenotypic plasticity and its impact on competition in fluctuating environments. *The American Naturalist*, **172**, 169–254.
Turkington R, Harper JL (1979a) The Growth, Distribution and Neighbour Relationships of *Trifolium Repens* in a Permanent Pasture: II. Inter-and Intra-Specific Contact. *Journal of Ecology*, **67**, 219–230.
Turkington R, Harper JL (1979b) The Growth, Distribution and Neighbour Relationships of *Trifolium Repens* in a Permanent Pasture: IV. Fine-Scale Biotic Differentiation. *Journal of Ecology*, **67**, 245–254.
Van Dijk H, Boudry P, McCombie H, Vernet P (1997) Flowering time in wild beet (*Beta vulgariss* sp. maritima) along a latitudinal cline. *Acta Oecologica*, **18**, 47–60.
Wagmann K, Hautekèete NC, Piquot Y *et al.* (2012) Seed dormancy distribution: explanatory ecological factors. *Annals of Botany*.
Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. *Evolution*, 1358–2728.
Zhao K, Nordborg M, Marjoram P (2007) Genome-wide association mapping using mixed-models: application toGAW15Problem 3. *BMC Proceedings*, **1 Suppl 1**, S164.

Data Accessibility

Phenotypic data, ecological variables and genotypes are available in the Dryad database: doi:10.5061/dryad.07s25

Figure Legends

Figure 1. Violin plots of natural variation for bolting time (BT), flowering period (FP) and flowering-to-reproductive ratio (FRR) for all accessions (Worldwide, French Regmap and new set of 49 French populations). BT and FP are expressed in days. 'F' and 'S' denote Fall and Spring germination cohorts, respectively. Violin plots are based on BLUPs calculated for each of the 352 natural accessions or 800 French families. Violin plots are a combination of a boxplot (the white dot represents the median, the black solid box around it the range from the first to the third quartile and the thin black lines above and below the box extent to the 1.5 times the inner quartile ranges) and a rotated kernel density plots (Hintze & Nelson 1998; Adler 2005).

Figure 2. Box-and-whiskers plots of diversity in BT and FP variation (both expressed in days) for each stand studied in each French region. Code names for each stand are given in Table S1 and 'F' and 'S' denote Fall and Spring germination cohorts, respectively. On boxplots, the horizontal bold black line represents the median, the black hollow box around it the range from the first to the third quartile. The dashed black lines above and below the box extent to the 1.5 time the inner quartile ranges. The data points outside this interval are plotted as black circles.

Figure 3. Identification of the putative selective agents acting on phenology in *A. thaliana* in France. A. Results from the PLSR, without control for neutral genetic variation. B. Results of the PLSR obtained while accounting for neutral genetic variation. Each column corresponds to a combination of scale/region, germination cohort and phenological trait. The first 42 lines

correspond to the ecological variables used in this study to characterize the 49 stands, separated in three categories: climate, soil and competition (Comp.). For each line and column, the colored squares indicate significant regression coefficients (p-value < 0.05) and the colors represent the strength of the regression coefficient estimated between phenological variation and ecological variation. The last three lines correspond to the number of PLSR components retained after cross-validation ('axis'), the percentage of ecological variation explained by PLSR components ('Var X') and the percentage of phenological variation explained by PLSR components ('Var Y'). Because analyses were performed on standardized data, regression coefficients are comparable between panels A and B. See Table 1 for description of the ecological variables.

Figure 4. Manhattan plots of the genome-wide association mapping (EMMAX) results for bolting time in each cohort and at different geographical scales. The x-axis indicates the position along each chromosome. The five chromosomes are presented in a row along the x-axis in different degrees of grey. The y-axis indicates the $-\log^{10}$ p-values using the EMMAX method. MARF >10%. The dashed line denotes the Bonferroni threshold. Vertical dotted lines indicate the physical position of the flowering time genes *CCA1*, *FRI*, *DFL2* and *DOG1*.

Table Legends

Table 1. Summary of 42 ecological variables gathered for this study.

Table 2. Phenological variation among worldwide accessions and natural accessions from the French Regmap. Model random terms were tested with likelihood ratio tests of models with and without these effects. Random effects are in italic. Bold *p*-values indicate significant effect after Bonferroni correction. Because the variance associated with the model term 'Block*Cohort' is the correct error term for testing the 'Block' and 'Cohort' effects, *F* value and *p*-value were not reported for the 'Block*Cohort' term.

Table 3. Phenological variation of the 49 French stands. Model random terms were tested with likelihood ratio tests of models with and without these effects. Random effects are in italic. Bold *p*-values indicate significant effect after Bonferroni correction. Because the variance associated with the model term 'Block*Cohort' is the correct error term for testing the 'Block' and 'Cohort' effects, *F* value and *p*-value were not reported for the 'Block*Cohort' term.

Figure 1

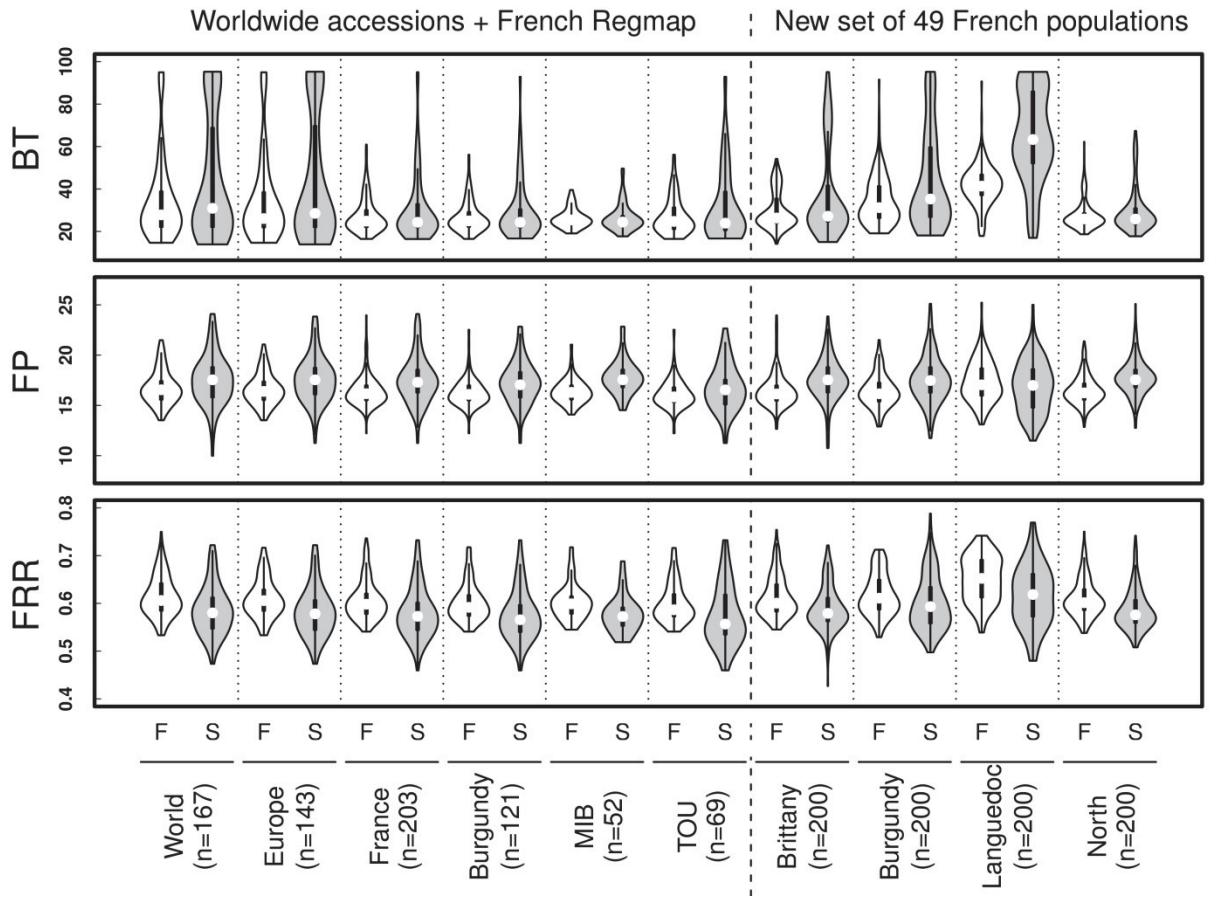
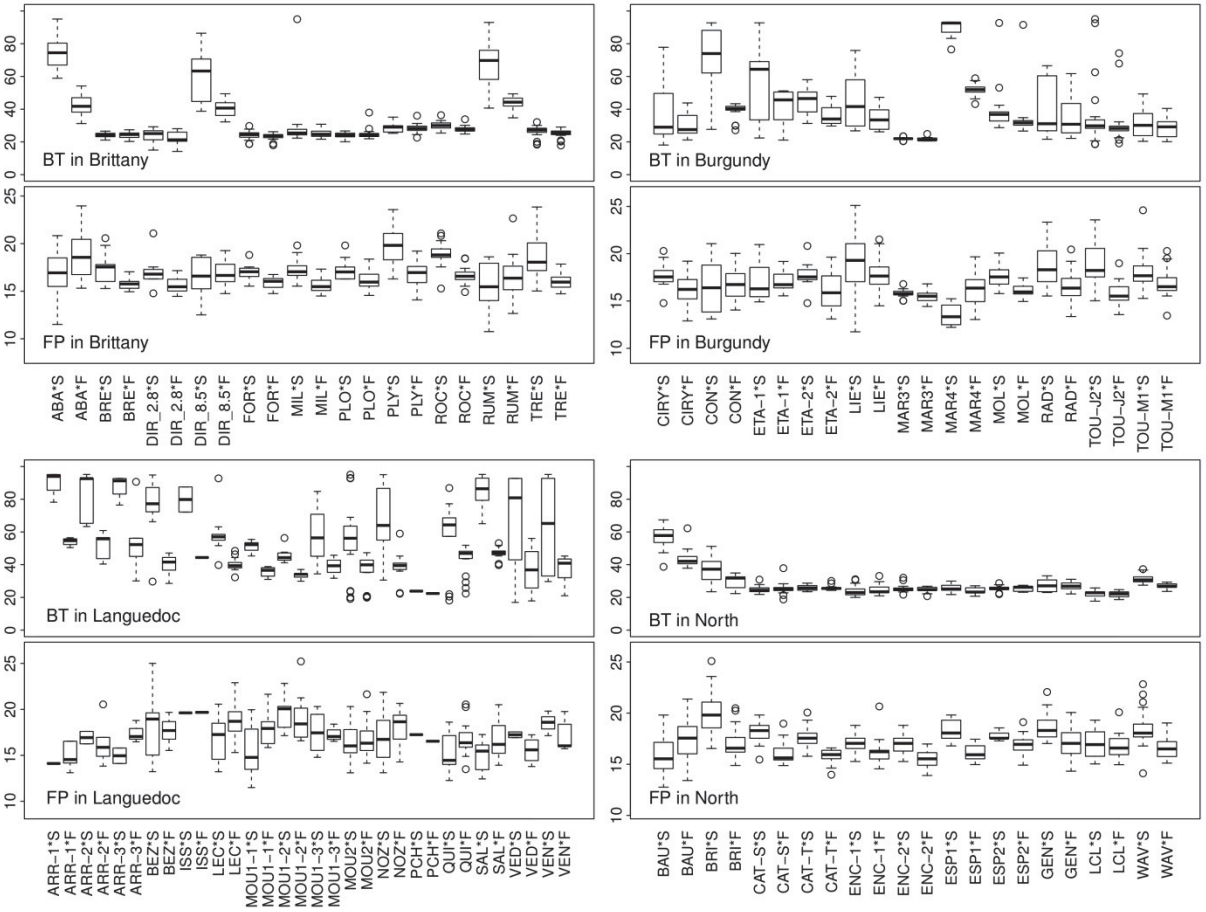
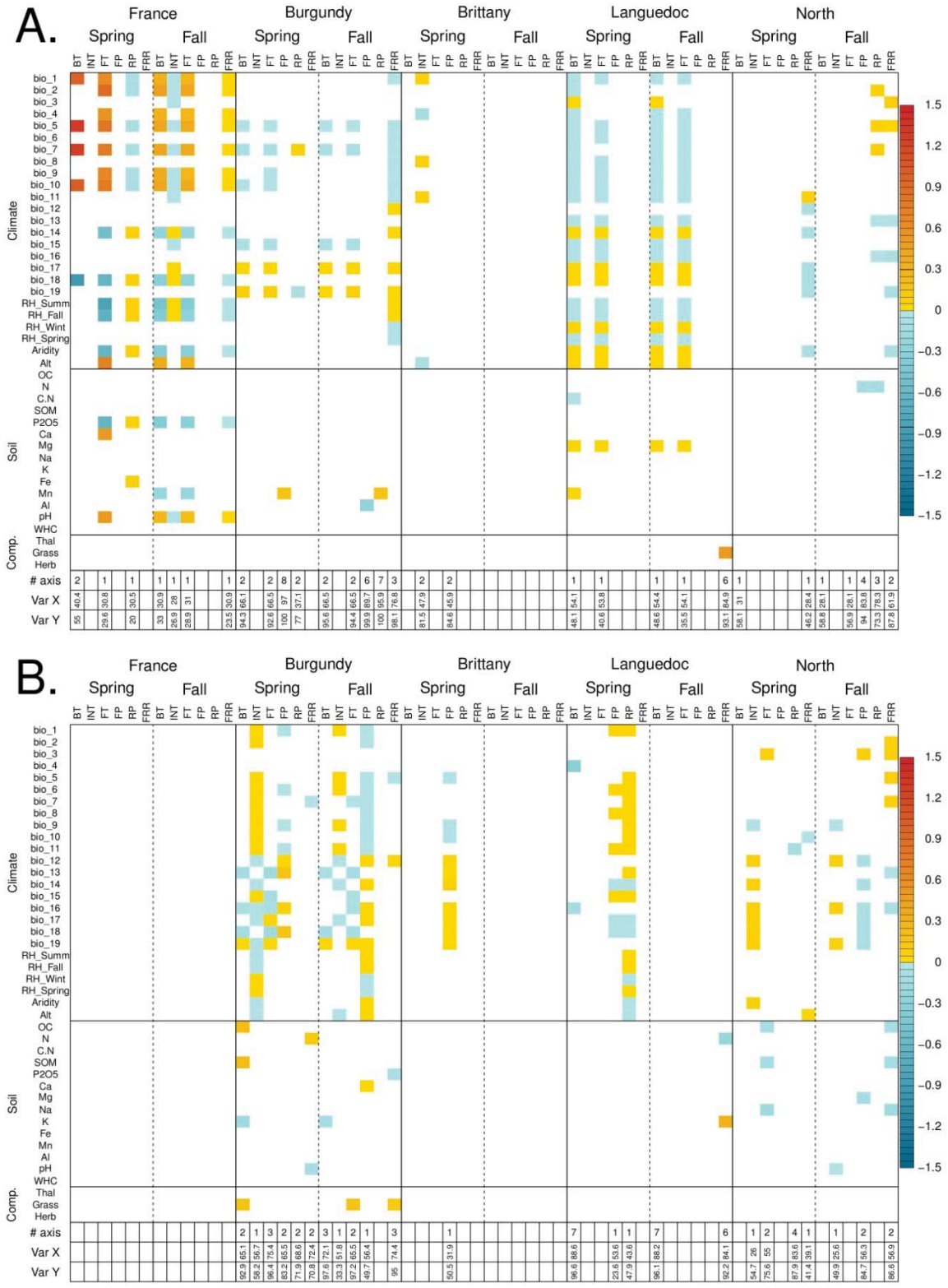


Figure 2



26 **Figure 3**

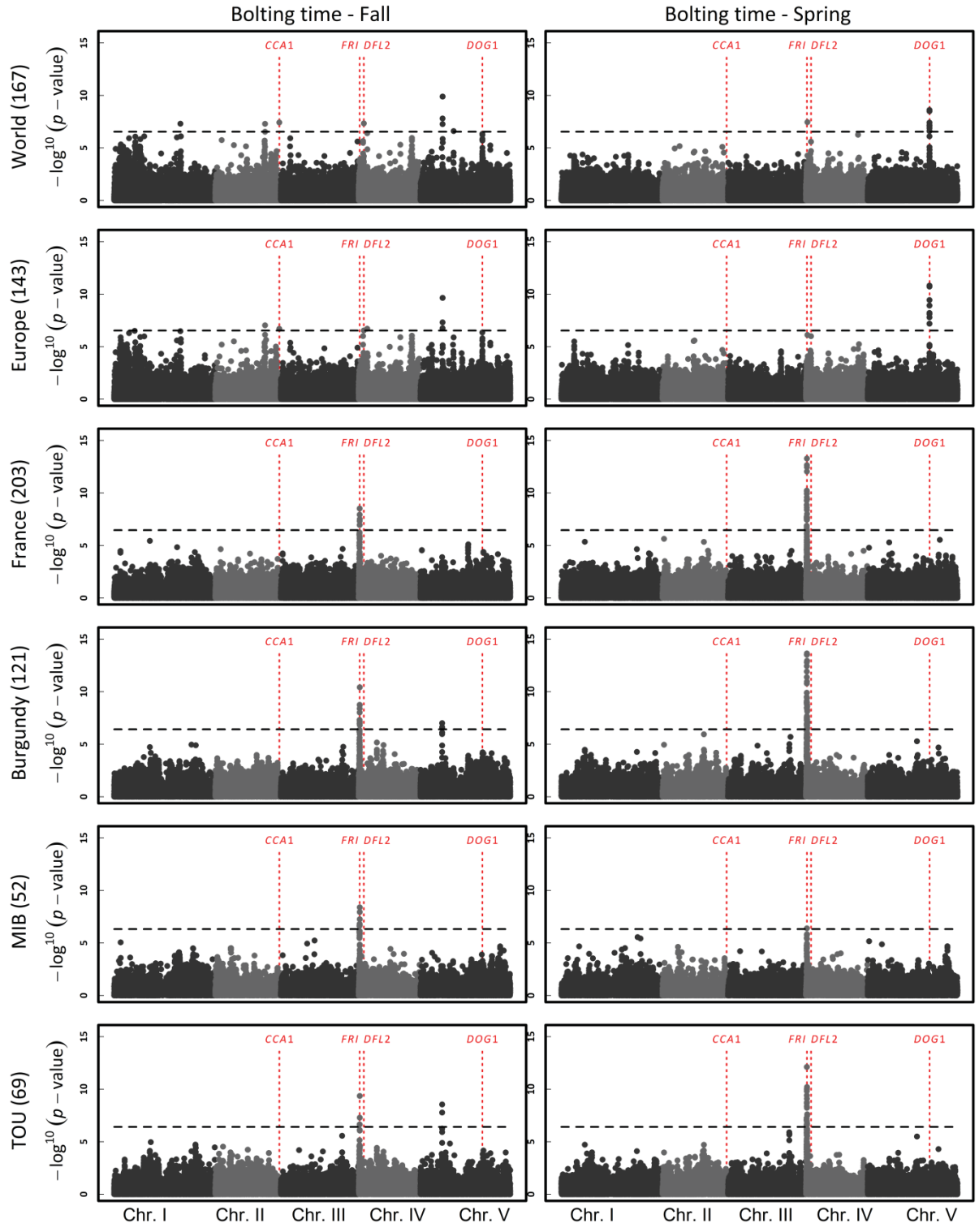
27



28

29 **Figure 4**

30



31

Table 1

Variable	Description	Source or method	Grid resolution
Alt	Altitude (m)	-	-
RH Wint	relative humidity in Winter (%)	NCAR/NCEP	5 arc-Minutes (~9 km)
RH Spring	relative humidity in Spring (%)	NCAR/NCEP	5 arc-Minutes (~9 km)
RH Summer	relative humidity in Summer (%)	NCAR/NCEP	5 arc-Minutes (~9 km)
RH Fall	relative humidity in Fall (%)	NCAR/NCEP	5 arc-Minutes (~9 km)
Aridity	aridity (mm/day)	FAO GeoNetwork	0.17°x 0.17° grid (~20 km)
Bio 1	annual mean temperature (°C x 10)	Wordclim	30 arc-seconds (~1 km)
Bio 2	mean diurnal range (mean of monthly temperature range (max temp - min temp))	Wordclim	30 arc-seconds (~1 km)
Bio 3	isothermality (Bio 2 / Bio 7)(*100)	Wordclim	30 arc-seconds (~1 km)
Bio 4	temperature seasonality (standard deviation * 100)	Wordclim	30 arc-seconds (~1 km)
Bio 5	maximum temperature of warmest month (°C x 10)	Wordclim	30 arc-seconds (~1 km)
Bio 6	min temperature of coldest month (°C x 10)	Wordclim	30 arc-seconds (~1 km)
Bio 7	temperature annual range (Bio 5 - Bio 6)	Wordclim	30 arc-seconds (~1 km)
Bio 8	mean temperature of wettest quarter (°C x 10)	Wordclim	30 arc-seconds (~1 km)
Bio 9	mean temperature of driest quarter (°C x 10)	Wordclim	30 arc-seconds (~1 km)
Bio 10	mean temperature of warmest quarter (°C x 10)	Wordclim	30 arc-seconds (~1 km)
Bio 11	mean temperature of coldest quarter (°C x 10)	Wordclim	30 arc-seconds (~1 km)
Bio 12	annual precipitation (mm)	Wordclim	30 arc-seconds (~1 km)
Bio 13	precipitation of wettest month (mm)	Wordclim	30 arc-seconds (~1 km)
Bio 14	precipitation of driest month (mm)	Wordclim	30 arc-seconds (~1 km)
Bio 15	precipitation seasonality (coefficient of variation)	Wordclim	30 arc-seconds (~1 km)
Bio 16	precipitation of wettest quarter (mm)	Wordclim	30 arc-seconds (~1 km)
Bio 17	precipitation of driest quarter (mm)	Wordclim	30 arc-seconds (~1 km)
Bio 18	precipitation of warmest quarter (mm)	Wordclim	30 arc-seconds (~1 km)
Bio 19	precipitation of coldest quarter (mm)	Wordclim	30 arc-seconds (~1 km)

Table 1. continued.

Variable	Description	Source or method	Grid resolution
OC	organic carbon ($g.kg^{-1}$)	NF ISO 10694 and NF ISO 13878 standards	-
N	total nitrogen ($g.kg^{-1}$)	NF ISO 10694 and NF ISO 13878 standards	-
C/N	carbon/nitrogen ratio	NF ISO 10694 and NF ISO 13878 standards	-
SOM	soil organic matter ($g.kg^{-1}$)	NF ISO 10694 and NF ISO 13878 standards	-
P2O5	phosphorus (P_2O_5) ($g.kg^{-1}$)	Olsen method (ISO 11263 standard)	-
Ca	exchangeable calcium ($cmol+.kg^{-1}$)	Cobaltiexamine method (ICP-AES, INRA method)	-
Mg	exchangeable magnesium ($cmol+.kg^{-1}$)	Cobaltiexamine method (ICP-AES, INRA method)	-
Na	exchangeable sodium ($cmol+.kg^{-1}$)	Cobaltiexamine method (ICP-AES, INRA method)	-
K	exchangeable potassium ($cmol+.kg^{-1}$)	Cobaltiexamine method (ICP-AES, INRA method)	-
Fe	exchangeable potassium ($cmol+.kg^{-1}$)	Cobaltiexamine method (ICP-AES, INRA method)	-
Mn	exchangeable iron ($cmol+.kg^{-1}$)	Cobaltiexamine method (ICP-AES, INRA method)	-
Al	exchangeable aluminium ($cmol+.kg^{-1}$)	Cobaltiexamine method (ICP-AES, INRA method)	-
WHC	soil water holding capacity ($ml.g^{-1}$)	Granier & Tardieu (1999)	-
pH	pH	NF ISO 10390 standard	-
Herb	interspecific competition with herbs which are not grasses	50 x 50 cm quadrat	-
Grass	interspecific competition with grasses	50 x 50 cm quadrat	-
Thal	intraspecific competition	50 x 50 cm quadrat	-

Table 2.

Traits	BT		INT		FT		FP		RP		FRR	
Model terms	***F* or LRT**	***P***	***F* or LRT**	***P***	***F* or LRT**	***P***	***F* or LRT**	***P***	***F* or LRT**	***P***	***F* or LRT**	***P***
Block	1.38	0.4477	2.51	0.1133	0.94	0.5095	3.00	0.0836	0.00	0.9985	1.02	0.4955
Cohort	3.49	0.0622	1.23	0.2685	9.44	**0.0022**	0.39	0.5301	9.68	**0.0020**	0.41	0.5247
Accession	739.20	**0.0001**	69.10	**0.0001**	566.00	**0.0001**	35.50	**0.0001**	34.20	**0.0001**	65.70	**0.0001**
*Cohort*Accession*	12.30	**0.0005**	0.00	1.0000	23.60	**0.0001**	2.20	0.1380	0.60	0.4386	0.00	1.0000
Control Bg-2	8.68	**0.0002**	2.33	0.0976	14.42	**0.0001**	0.74	0.4757	5.74	**0.0033**	0.72	0.4852

Table 3.

Traits	BT		INT		FT		FP		RP		FRR	
Model terms	***F* or LRT**	***P***	***F* or LRT**	***P***	***F* or LRT**	***P***	***F* or LRT**	***P***	***F* or LRT**	***P***	***F* or LRT**	***P***
Block	2.54	0.3529	0.99	0.3189	9.73	0.1975	0.01	0.9325	0.15	0.7637	0.69	0.5583
Cohort	0.15	0.6952	0.32	0.5727	3.19	0.0743	0.73	0.3938	14.24	**0.0002**	0.63	0.4307
Region	158.12	**0.0001**	6.90	**0.0001**	116.07	**0.0001**	3.49	0.0153	9.31	**0.0001**	14.89	**0.0001**
Stand(Region)	26.33	**0.0001**	4.01	**0.0001**	26.30	**0.0001**	6.51	**0.0001**	5.16	**0.0001**	7.52	**0.0001**
Family(Stand(Region))	397.50	**0.0001**	11.60	**0.0007**	384.90	**0.0001**	0.00	1.0000	17.80	**0.0001**	1.23	0.1213
Cohort*Region	9.49	**0.0001**	4.09	**0.0067**	2.72	0.0432	10.93	**0.0001**	14.85	**0.0001**	0.74	0.5259
Cohort*Stand(Region)	3.28	**0.0001**	2.35	**0.0001**	2.87	**0.0001**	4.59	**0.0001**	4.14	**0.0001**	1.91	**0.0006**
Cohort(Family(Stand(Region)))*	0.00	1.0000	1.70	0.1923	0.00	1.0000	17.70	**0.0001**	0.00	1.0000	1.13	1.0000
Control Bg-2	5.14	**0.0060**	0.73	0.4839	6.99	**0.0010**	1.13	0.3225	8.64	**0.0002**	2.02	0.1336

SUPPORTING INFORMATION: INVESTIGATION OF THE GEOGRAPHICAL SCALE OF ADAPTIVE PHENOLOGICAL VARIATION AND ITS UNDERLYING GENETICS IN *ARABIDOPSIS THALIANA*

Figure S1. Violin plots of natural variation for the interval between bolting and flowering (INT), flowering time (FT) and reproductive period (RP). White circles indicate the median value and height corresponds to the interquartile range, with the width representing the probability density of the data at different values. INT, FT and RP are expressed in days. 'F' and 'S' denote Fall and Spring germination cohorts, respectively. Violin plots are based on BLUPs calculated for each of the 352 natural accessions or 800 families.

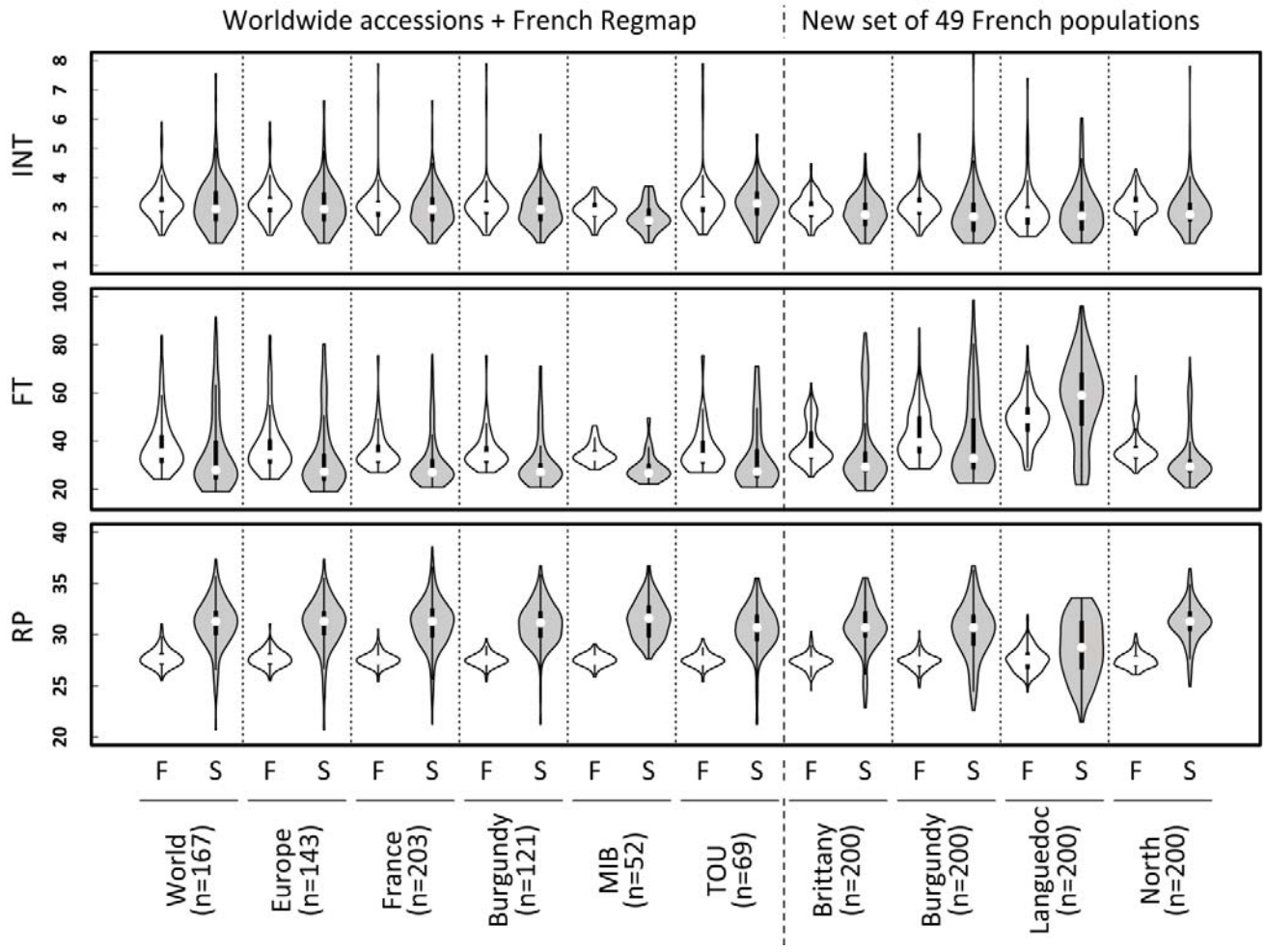
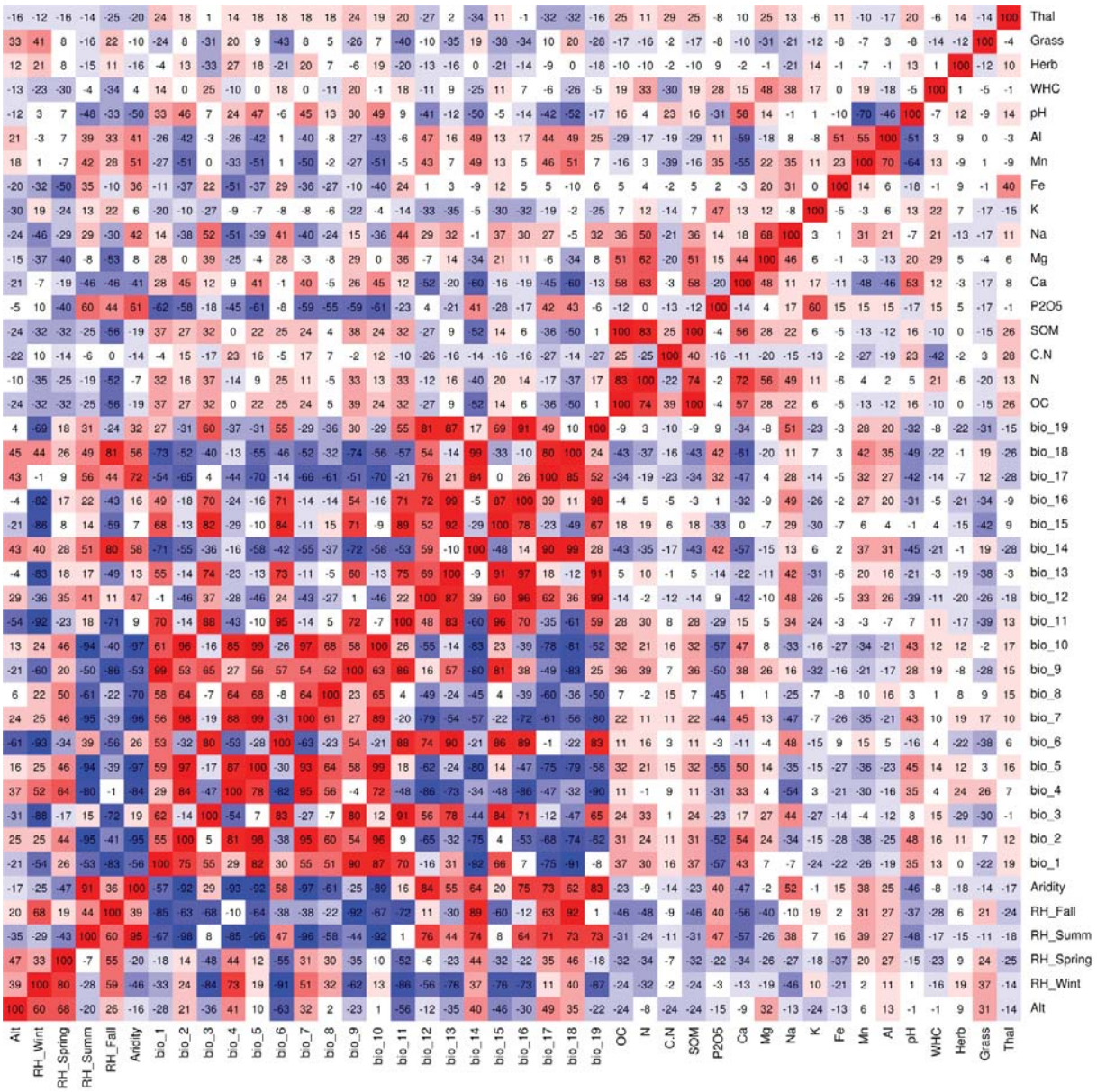


Figure S2. Pairwise matrix of similarity between ecological variables. Spearman and Pearson correlations (multiplied by 100) are given above and below the diagonal, respectively.



Climatic variables: **Alt**, Altitude (m); **RH Wint**, **RH Spring**, **RH Summ** and **RH Fall** stand for relative humidity (%) in Winter, Spring, Summer and Fall, respectively; **Aridity**, Aridity; **Bio 1**, annual mean temperature (°C x 10); **Bio 2**, mean diurnal range (mean of monthly temperature range (max temp - min temp)); **Bio 3**, isothermality (Bio 2/Bio 7) (* 100); **Bio 4**, temperature seasonality (standard deviation *100); **Bio 5**, max temperature of warmest month (°C x 10); **Bio 6**, min temperature of coldest month (°C x 10); **Bio**

7, temperature annual range (Bio 5 - Bio 6); **Bio 8**, mean temperature of wettest quarter (°C x 10); **Bio 9**, mean temperature of driest quarter (°C x 10); **Bio 10**, mean temperature of warmest quarter (°C x 10); **Bio 11**, mean temperature of coldest quarter (°C x 10); **Bio 12**, annual precipitation (mm); **Bio 13**, precipitation of wettest month (mm); **Bio 14**, precipitation of driest month (mm); **Bio 15**, precipitation seasonality (coefficient of variation); **Bio 16**, precipitation of wettest quarter (mm); **Bio 17**, precipitation of driest quarter (mm); **Bio 18**, precipitation of warmest quarter (mm); **Bio 19**, precipitation of coldest quarter (mm).

Edaphic variables: OC, organic carbon ($g.kg^{-1}$); **N**, total nitrogen ($g.kg^{-1}$); **C/N**, carbon/nitrogen ratio; **SOM**, soil organic matter ($g.kg^{-1}$); **P2O5**, phosphorus (P_2O_5) ($g.kg^{-1}$); **Ca**, exchangeable calcium ($cmol+.kg^{-1}$); **Mg**, exchangeable magnesium ($cmol+.kg^{-1}$); **Na**, exchangeable sodium ($cmol+.kg^{-1}$); **K**, exchangeable potassium ($cmol+.kg^{-1}$); **Fe**, exchangeable iron ($cmol+.kg^{-1}$); **Mn**, exchangeable manganese ($cmol+.kg^{-1}$); **Al**, exchangeable aluminium ($cmol+.kg^{-1}$); **WHC**, soil water holding capacity ($ml.g^{-1}$); **pH**.

Competition variables: Herb, interspecific competition with herbs which are not grasses; **Grass**, interspecific competition with grasses; **Thal**, intraspecific competition.

Figure S3. Box-and-whiskers plots of ecological variation in French stands. In brackets, significance of the 'Region' effect after Bonferroni correction (***: $P < 0.001$, **: $P < 0.01$, *: $P < 0.05$, ns: $P > 0.05$) and percentage of ecological variance explained by the 'Region' effect. See legends of Figure S2 for description of the variables.

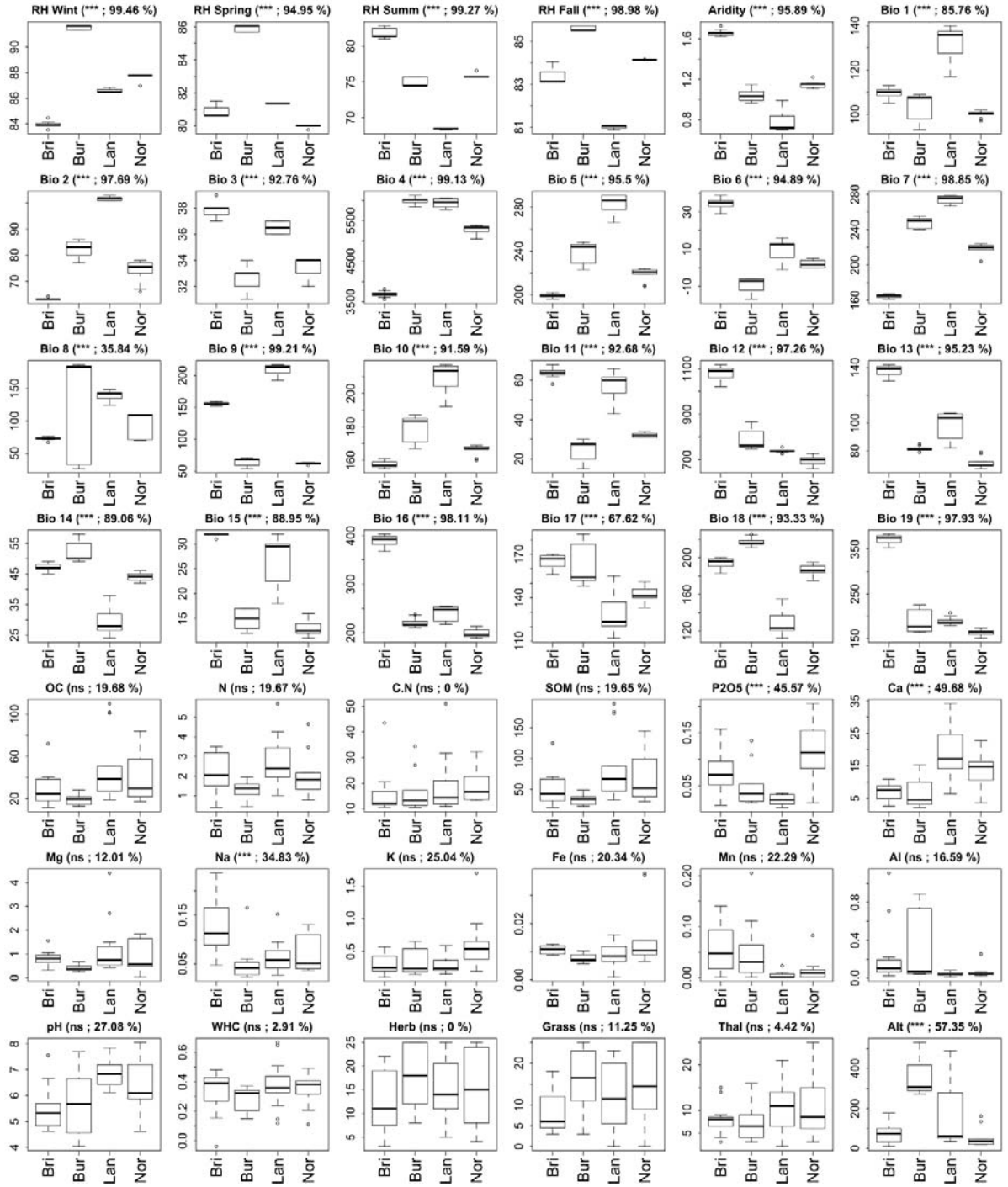
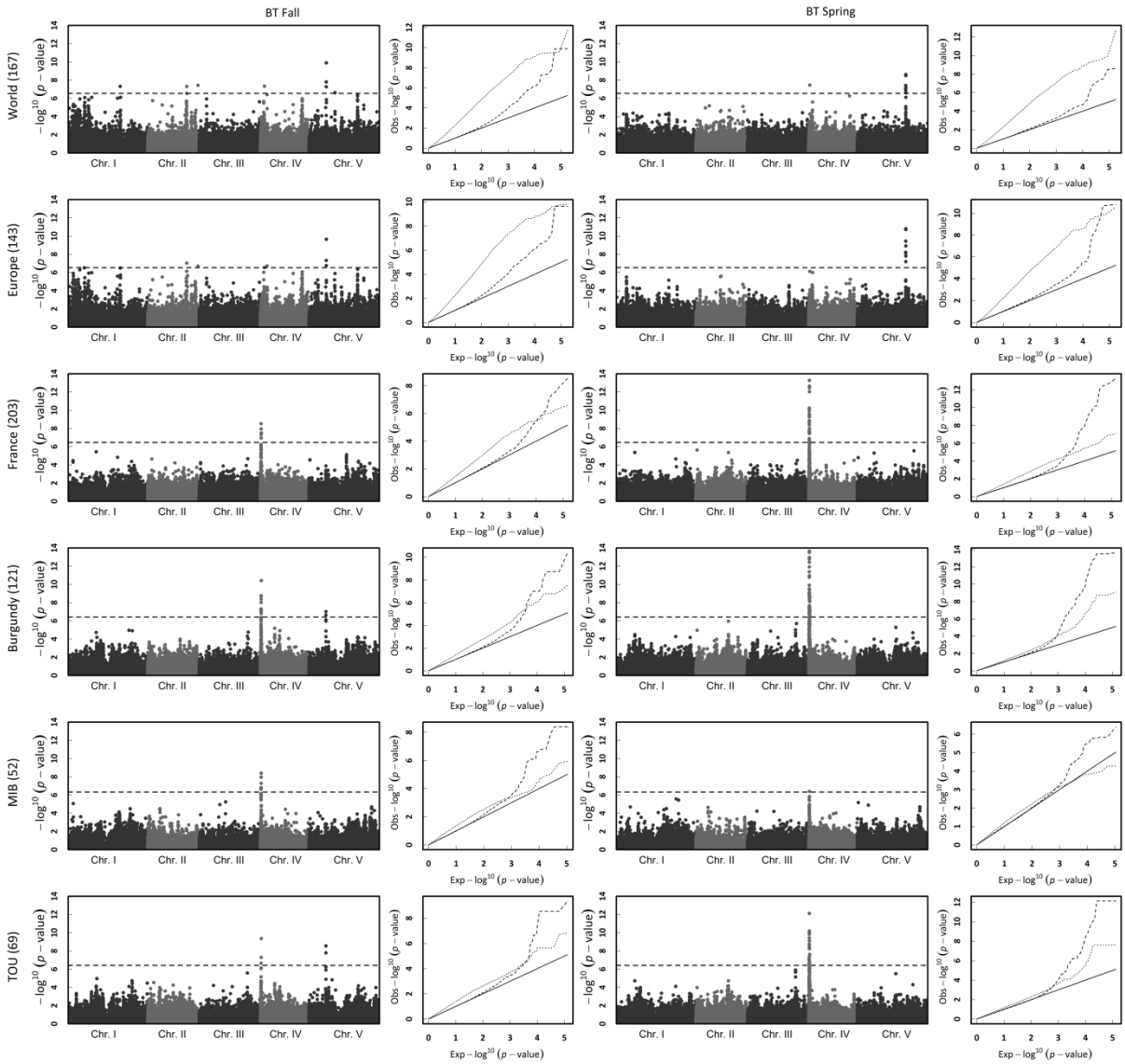


Figure S4. Manhattan plots of the genome-wide association mapping (EMMAX) results and quantile-quantile plot of p-values (negative logarithm, Wilcoxon and EMMAX) for BT, INT, FT, FP, RP and FRR in each cohort and at different geographic scales.



Manhattan plots: The *x*-axis indicates the position along each chromosome. The five chromosomes are presented in a row along the *x*-axis in different shades of grey. The *y*-axis indicates the $-\log^{10} p$-values using the EMMAX method. The dashed line denotes the Bonferroni threshold. MARF >10%.

Quantile-quantile plots: The different curves of the Q-Q plots correspond to different analyses of GWA mapping. Solid line: expected; dotted line: Wilcoxon with minor allele relative frequency (MARF) > 10%; dotted line: EMMAX with MARF > 10%.

Figure S4. (continued)

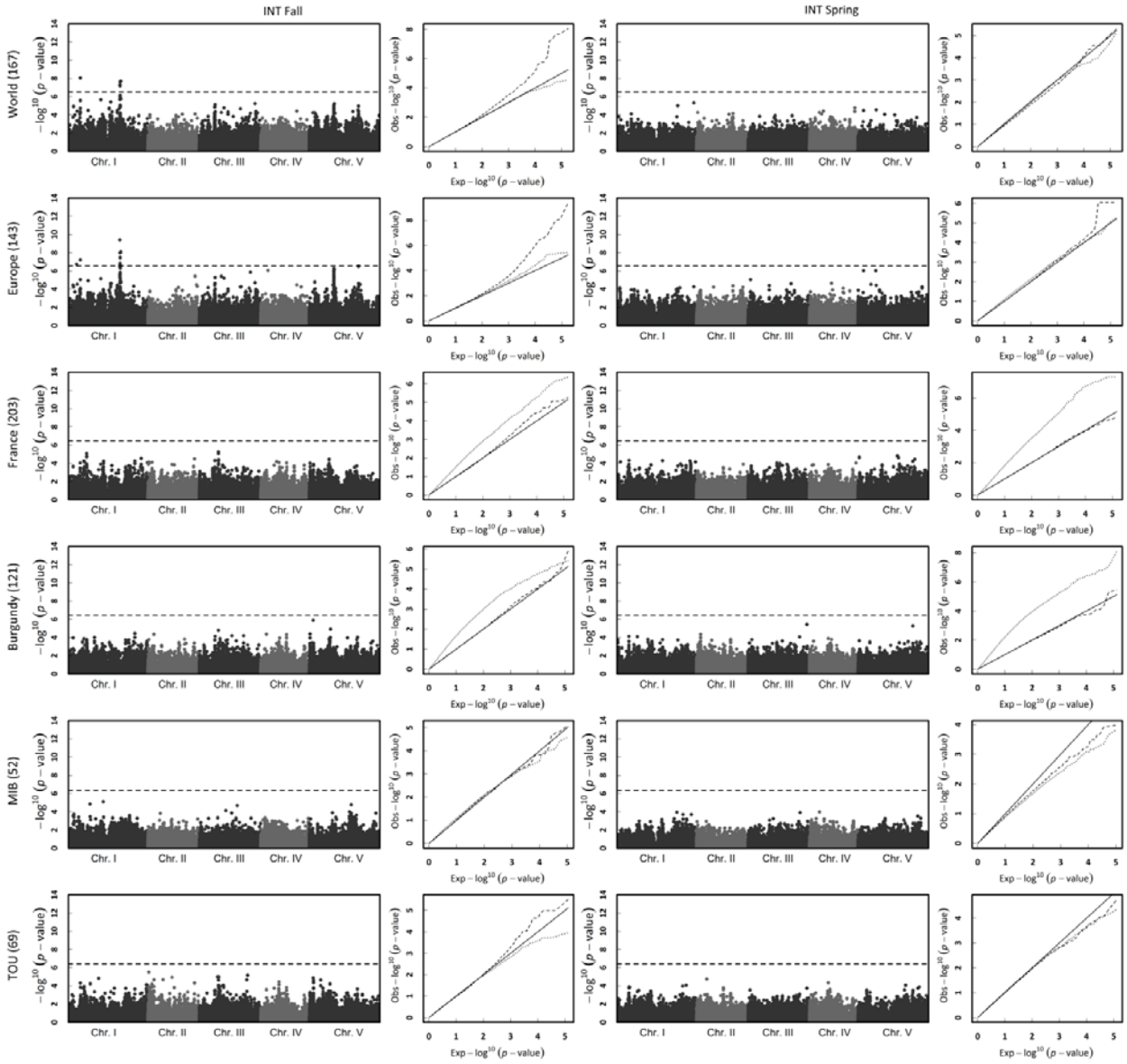


Figure S4. (continued)

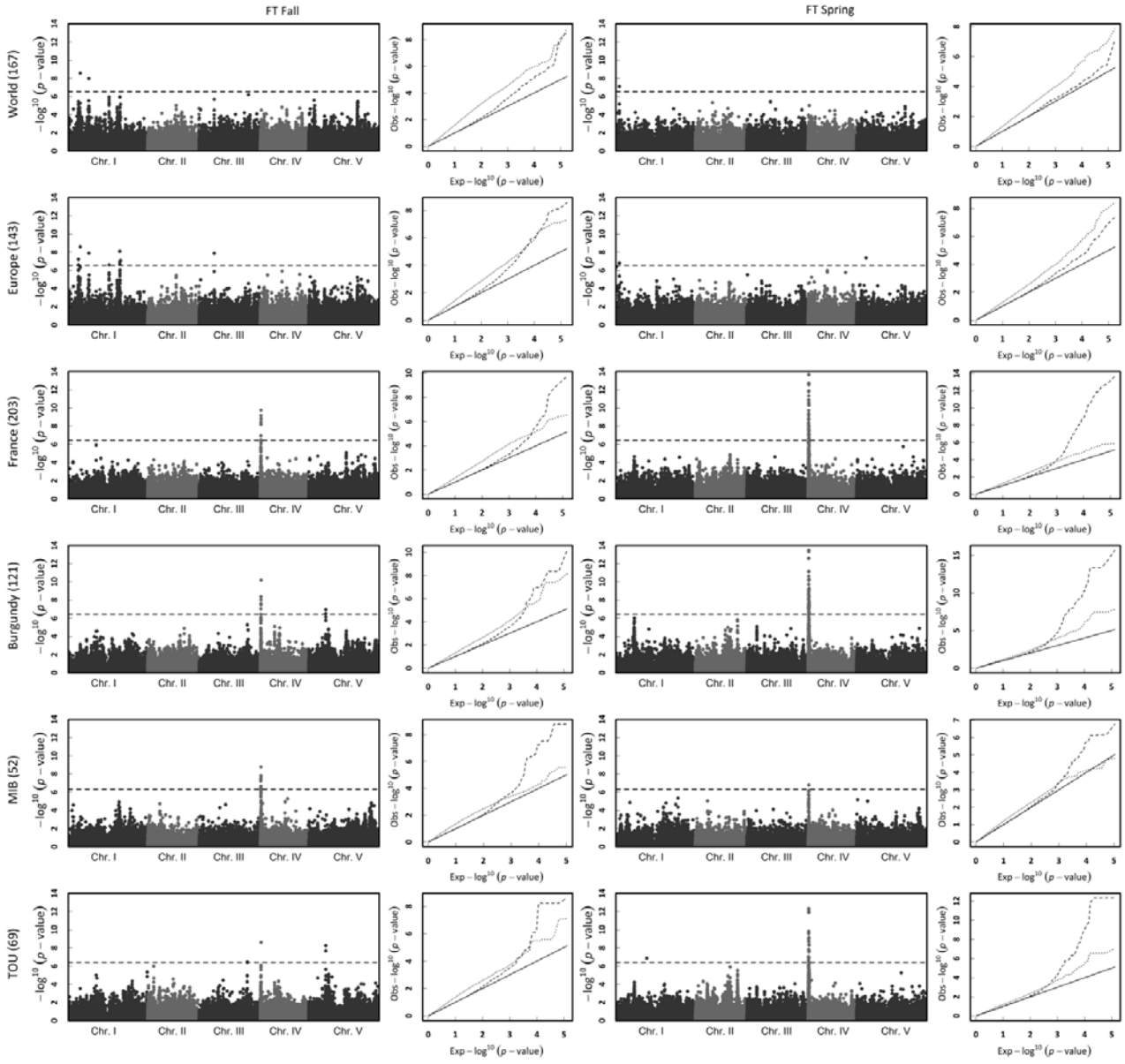


Figure S4. (continued)

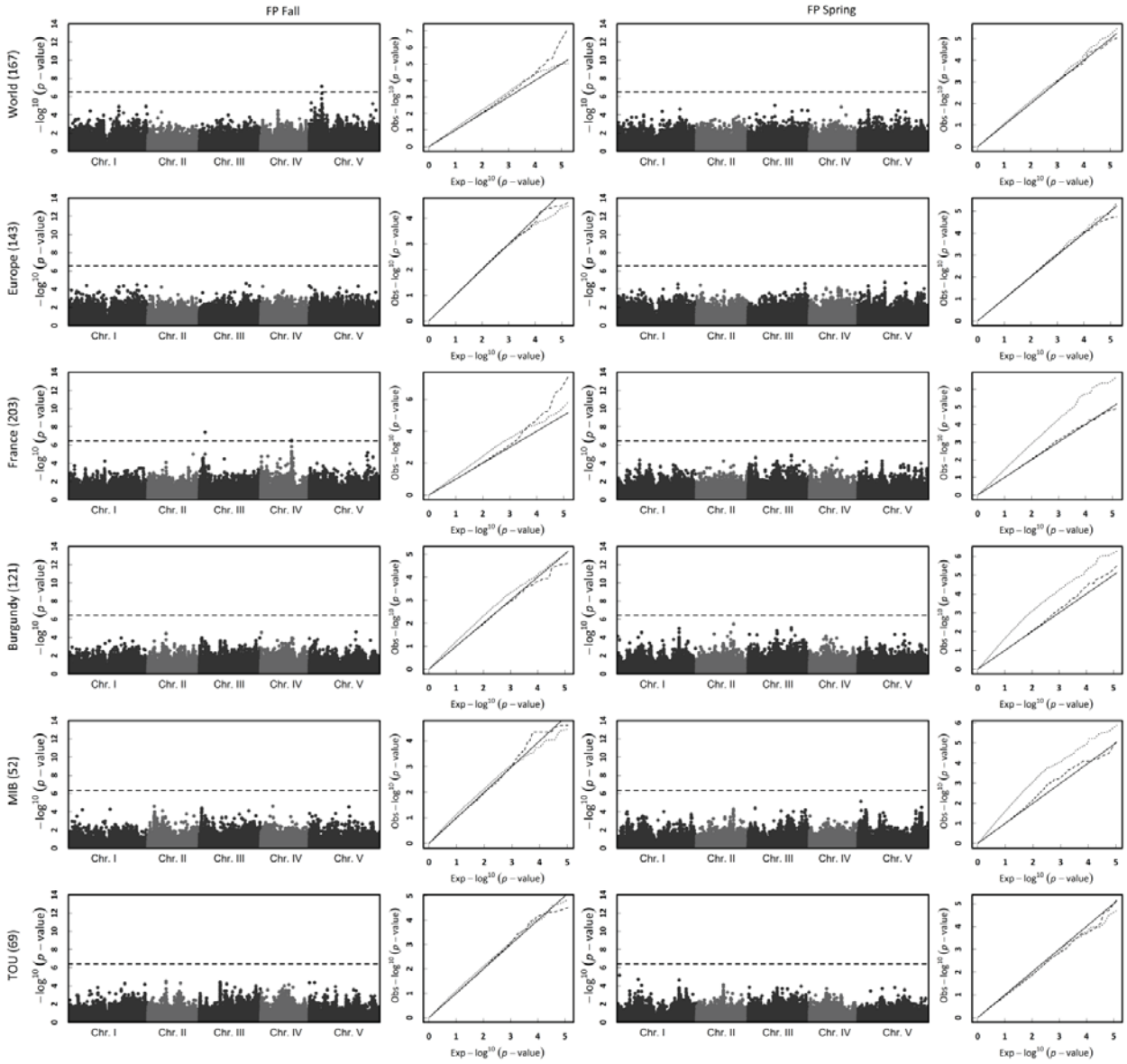


Figure S4. (continued)

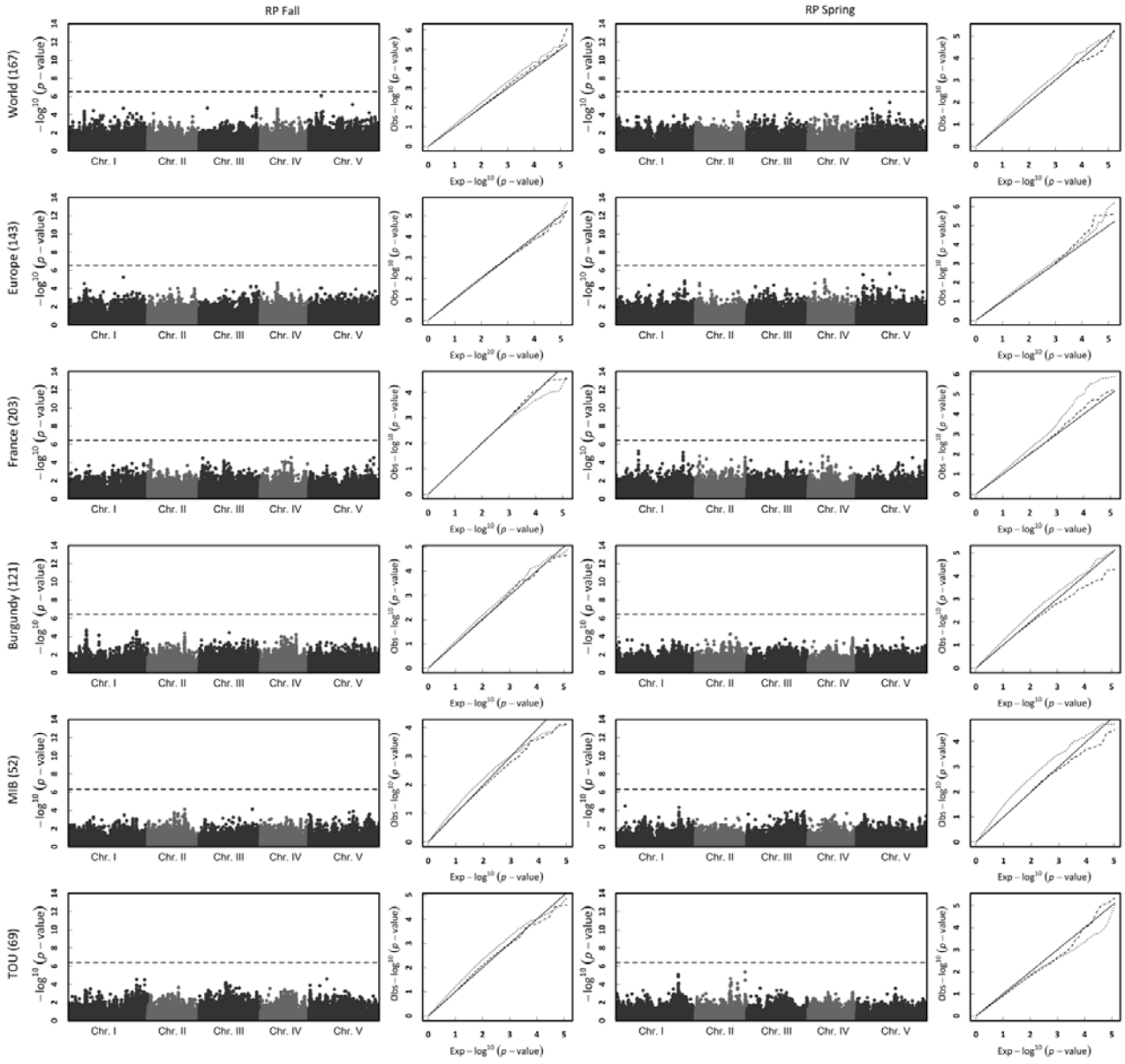


Figure S4. (continued)

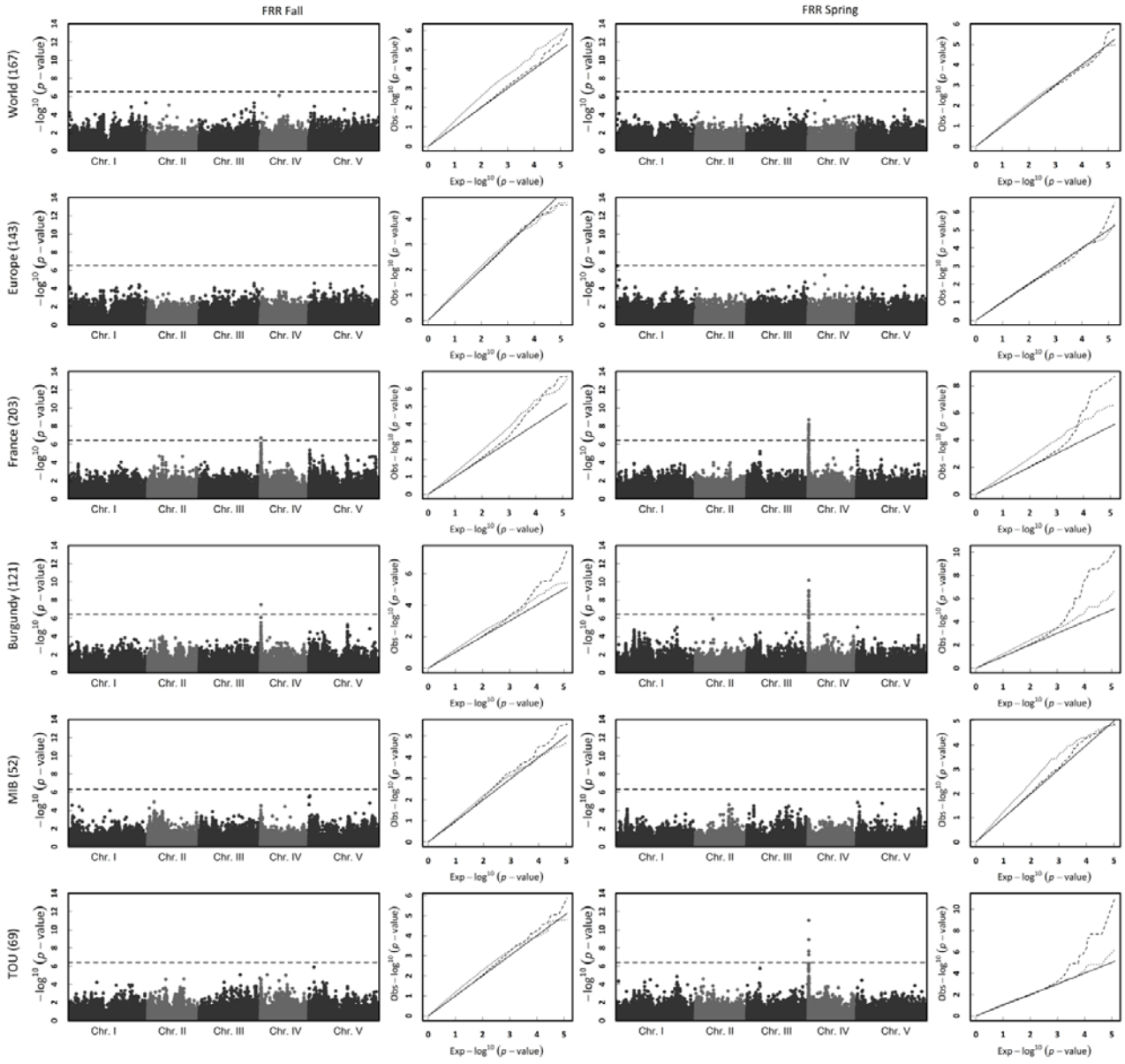


Figure S5. Close up of the association peaks at the beginning of chromosome 5 (left) and around *FRIGIDA* (right) at three different geographical scales for bolting time. Moccasin colored area indicates the genomic position of *FRIGIDA*. Method = EMMAX, MARF > 10%.

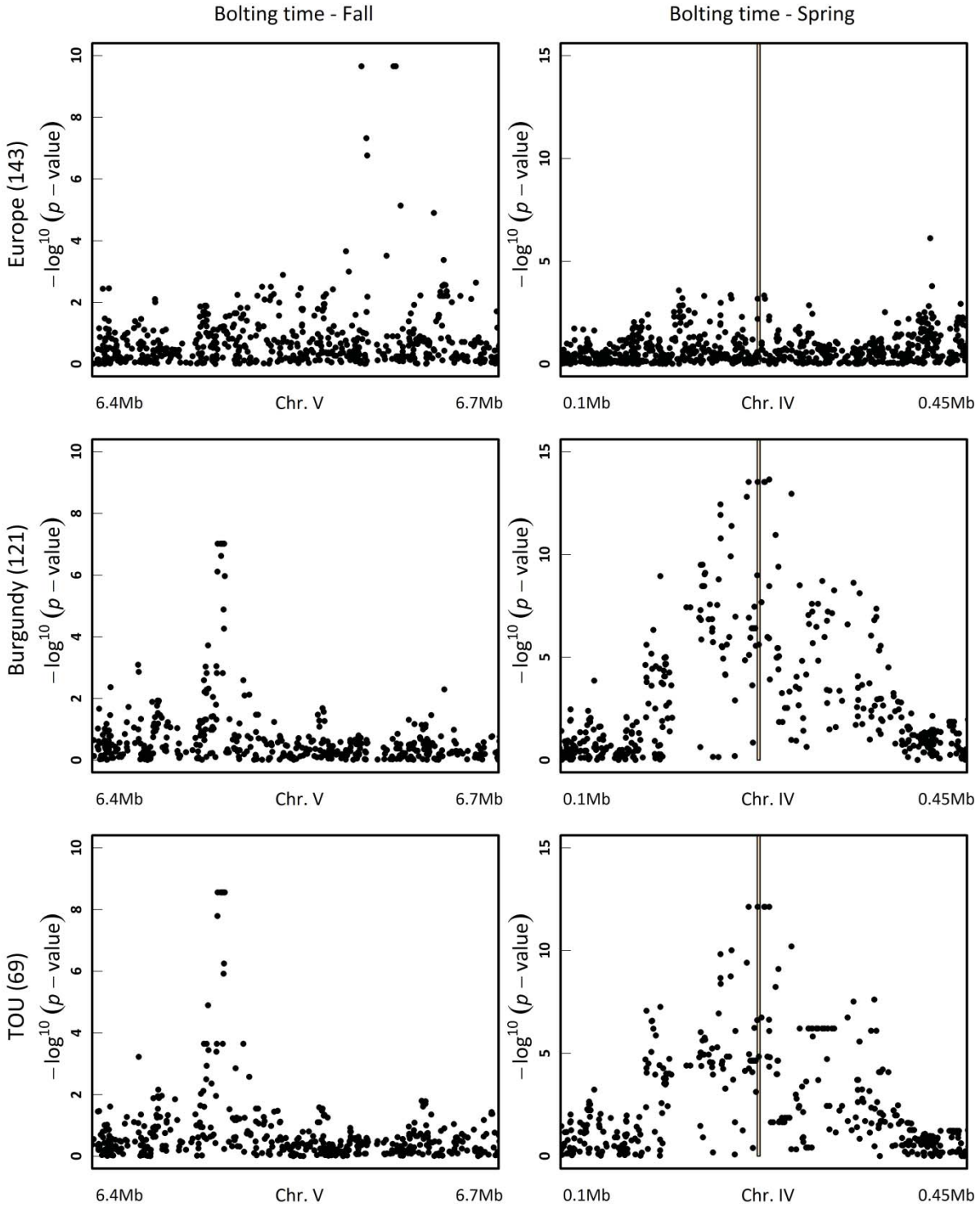


Table S1: Plant material.

A. French stands

Stand	Region	Town	Latitude	Longitude	Habitat	Sub-habitat	Families (N)	49 SNPs[a]
CIRY	Burgundy	Ciry-le-Noble	46°36.257'N	4°17.971'E	grassland	railway station	21	7
CON	Burgundy	Conforgien	47°14.011'N	4°11.893'E	grassland	river bank	20	0
ETA-1	Burgundy	Etang sur Arroux	46°52.060'N	4°11.078'E	grassland	railway station	10	0
ETA-2	Burgundy	Etang sur Arroux	46°52.060'N	4°11.078'E	hoed land	railway station	10	
LIE	Burgundy	Liernais	47°12.879'N	4°16.036'E	grassland	parking lot	20	8
MAR-3	Burgundy	Marigny l'Eglise	47°21.333'N	3°56.183'E	grassland	roadside	20	0
MAR-4	Burgundy	Marigny l'Eglise	47°21.356'N	3°56.024'E	meadow	perennial meadow	18	8
MOL	Burgundy	Saint- Léger sous Beuvray	46°54.732'N	4°06.425'E	meadow	meadows-footpath	20	9
RAD	Burgundy	Sainte-Radegonde	46°41.010'N	4°04.982'E	meadow	perennial meadow	23	2
TOU-J2	Burgundy	Toulon sur Arroux	46°39.060'N	4°06.466'E	hoed land	perennial meadow	21	1
TOU-M1	Burgundy	Toulon sur Arroux	46°39.071'N	4°06.758'E	hoed land	meadows-footpath	17	7
ABA	Brittany	Abaty	48° 16.599'N	4° 11.367'W	grassland	courtyard	20	9
BRE	Brittany	Brendaouez	48° 36.909'N	4° 25.129'W	grassland	roadside, along a stone wall	22	2
DIR-2.8	Brittany	Dirinon	48° 23.571'N	4° 17.134'W	grassland	railway station	11	1
DIR-8.5	Brittany	Dirinon	48° 23.571'N	4° 17.134'W	grassland	railway station	11	0
FOR	Brittany	La Forest Landerneau	48° 25.560'N	4° 18.413'W	meadow	parking lot	20	9
MIL	Brittany	Milizac	48° 28.176'N	4° 33.526'W	meadow	roadside	20	6

Stand	Region	Town	Latitude	Longitude	Habitat	Sub-habitat	Families (N)	49 SNPs[a]
PLO	Brittany	Plougastel-Daoulas	48° 21.839'N	4° 22.348'W	hoed land	garden	20	0
PLY	Brittany	Ploudiry	48°27 .502'N	4° 8.383'W	hoed land	perennial meadow	14	4
ROC	Brittany	La Roche Maurice	48° 28.450'N	4° 12.269'W	grassland	railway station	20	0
RUM	Brittany	Rumiqueal	48° 38.989' N	4° 21.607'W	hoed land	roadside	23	3
TRE	Brittany	Treflaouenan	48° 37.868'N	4° 5.047'W	hoed land	cultivated field	19	9
ARR-1	Languedoc	Arrigas	43° 59.254'N	3° 28.813' E	meadow	river bank	7	
ARR-2	Languedoc	Arrigas	43° 59.254'N	3° 28.813' E	meadow	river bank	7	
ARR-3	Languedoc	Arrigas	43° 59.254'N	3° 28.813' E	meadow	river bank	7	
BEZ	Languedoc	Bez et Esparon	43° 58.551'N	3° 31.895' E	grassland	roadside	21	1
ISS	Languedoc	Issensac	43° 50.296'N	3° 42.036'E	grassland	clearing in forest (*Quercus ilex*)	6	
LEC	Languedoc	Lecques	43° 50.280'N	4° 04.259' E	grassland	waste ground	20	0
MOU1-1	Languedoc	Moussac 1	43° 57.694'N	4° 14.900'E	grassland	river bank -fluvial deposit	7	
MOU1-2	Languedoc	Moussac 1	43° 57.694'N	4° 14.900'E	grassland	river bank -fluvial deposit	9	
MOU1-3	Languedoc	Moussac 1	43° 57.694'N	4° 14.900'E	meadow	river bank -fluvial deposit	4	
MOU2	Languedoc	Moussac 2	43° 58.064'N	4° 14.265'E	hoed land	cultivated field	22	2
NOZ	Languedoc	Nozières-Brignon	43° 58.548'N	4° 12.474'E	grassland	railway station	20	8
PCH	Languedoc	Puechredon	43° 57.500'N	4° 02.603' E	grassland	garrigue	10	
QUI	Languedoc	Quissac	43° 54.600' N	4° 00.285' E	grassland	waste ground (railway station)	20	0
SAL	Languedoc	Salinelles	43° 47.646'N	4° 04.462' E	grassland	rocky hillside	20	0

Stand	Region	Town	Latitude	Longitude	Habitat	Sub-habitat	Families (N)	49 SNPs[a]
VED	Languedoc	St Jean de Vedas	43° 33.686'N	3° 49.268' E	grassland	garrigue	8	
VEN	Languedoc	Vendargue	43° 39.259'N	3° 57.657' E	grassland	railway	12	2
BAU	North	Bauvin	50° 30.358'N	2° 54.095'E	grassland	railway	20	0
BRI	North	Brillon	50° 25.947'N	3° 19.719'E	hoed land	flower bed	21	1
CAT-S	North	Mont des Cats	50° 47.010'N	2° 40.101'E	hoed land	perennial meadow	20	9
CAT-T	North	Mont des Cats	50° 46.904'N	2° 39.961'E	hoed land	roadside	20	8
ENC-1	North	Mont de l'Enclus	50° 45.195'N	3° 29.501'E	hoed land	roadside	21	1
ENC-2	North	Mont de l'Enclus	50° 45.400'N	3° 28.480'E	hoed land	roadside	21	1
ESP1	North	Esplechin	50° 34.569'N	3° 18.601'E	meadow	perennial meadow	11	1
ESP2	North	Esplechin	50° 34.569'N	3° 18.601'E	hoed land	meadow, old gravel deposit	11	1
GEN	North	Genech	50° 32.217'N	3° 13.283'E	hoed land	cultivated field	15	5
LCL	North	Lecelles	50° 28.021'N	3° 24.222'E	grassland	roadside	20	0
WAV	North	Wavrin	50° 33.366'N	2° 55.251'E	hoed land	garden	20	9

[a] Number of families genotyped for 149 SNPs

B. Worldwide accessions and French accessions from the French RegMap panel (http://bergelson.uchicago.edu/).

accession ID[a]	Geographic group[b]
1	
2	FRANCE
4	
5	FRANCE
6	FRANCE
7	FRANCE
8	FRANCE
9	
12	FRANCE
23	FRANCE
32	FRANCE
45	FRANCE
48	FRANCE
51	FRANCE
60	FRANCE
62	FRANCE
66	FRANCE
69	FRANCE
74	FRANCE
77	
79	FRANCE
80	FRANCE
81	FRANCE
82	

accession ID[a]	Geographic group[b]
159	FRANCE
160	FRANCE,BURGUNDY,MIB
162	FRANCE,BURGUNDY,MIB
163	FRANCE,BURGUNDY,MIB
165	FRANCE,BURGUNDY,MIB
166	FRANCE,BURGUNDY,MIB
167	FRANCE,BURGUNDY,MIB
168	FRANCE,BURGUNDY,MIB
169	FRANCE,BURGUNDY,MIB
170	FRANCE,BURGUNDY,MIB
171	FRANCE,BURGUNDY,MIB
173	FRANCE,BURGUNDY,MIB
174	FRANCE,BURGUNDY,MIB
175	FRANCE,BURGUNDY,MIB
178	FRANCE,BURGUNDY,MIB
179	FRANCE,BURGUNDY,MIB
183	FRANCE,BURGUNDY,MIB
184	FRANCE,BURGUNDY,MIB
185	FRANCE,BURGUNDY,MIB
186	FRANCE,BURGUNDY,MIB
187	FRANCE,BURGUNDY,MIB
188	FRANCE,BURGUNDY,MIB
190	FRANCE,BURGUNDY,MIB
191	FRANCE,BURGUN

83	FRANCE
84	FRANCE
85	FRANCE
86	FRANCE
87	FRANCE
88	FRANCE
89	FRANCE
91	
94	FRANCE
96	FRANCE
104	FRANCE
106	FRANCE
116	FRANCE
121	FRANCE
122	FRANCE
123	FRANCE
126	FRANCE
137	
146	FRANCE
148	FRANCE
149	FRANCE
151	FRANCE
153	FRANCE
156	
157	FRANCE

	DY,MIB
194	FRANCE,BURGUNDY,MIB
196	FRANCE,BURGUNDY,MIB
198	FRANCE,BURGUNDY,MIB
200	FRANCE,BURGUNDY,MIB
201	FRANCE,BURGUNDY,MIB
202	FRANCE,BURGUNDY,MIB
203	FRANCE,BURGUNDY,MIB
204	FRANCE,BURGUNDY,MIB
205	FRANCE,BURGUNDY,MIB
206	FRANCE,BURGUNDY,MIB
207	FRANCE,BURGUNDY,MIB
208	FRANCE,BURGUNDY,MIB
210	FRANCE,BURGUNDY,MIB
212	FRANCE,BURGUNDY,MIB
213	FRANCE,BURGUNDY,MIB
214	FRANCE,BURGUNDY,MIB
215	FRANCE,BURGUNDY,MIB
216	FRANCE,BURGUNDY,MIB
217	FRANCE,BURGUNDY,MIB
219	FRANCE,BURGUNDY,MIB
222	FRANCE,BURGUNDY,MIB
223	FRANCE,BURGUNDY,MIB
224	FRANCE,BURGUNDY,MIB
225	FRANCE,BURGUNDY,MIB
227	FRANCE,BURGUNDY,MIB

accession ID[a]	Geographic group[b]
228	FRANCE,BURGUNDY,MIB
229	FRANCE,BURGUNDY,MIB
230	FRANCE,BURGUNDY,MIB
231	FRANCE,BURGUNDY,MIB
236	FRANCE
237	FRANCE
242	FRANCE
244	FRANCE
252	FRANCE
257	FRANCE
258	FRANCE
259	FRANCE
261	FRANCE
262	FRANCE
263	FRANCE
264	FRANCE
266	
267	FRANCE
268	FRANCE
269	FRANCE
273	FRANCE,BURGUNDY,TOU
275	FRANCE,BURGUNDY,TOU
27	FRANCE,BURGUN

accession ID[a]	Geographic group[b]
323	FRANCE,BURGUNDY,TOU
326	FRANCE,BURGUNDY,TOU
327	FRANCE,BURGUNDY,TOU
328	FRANCE,BURGUNDY,TOU
329	FRANCE,BURGUNDY,TOU
331	FRANCE,BURGUNDY,TOU
332	FRANCE,BURGUNDY,TOU
334	
335	FRANCE,BURGUNDY,TOU
337	
338	FRANCE,BURGUNDY,TOU
339	FRANCE,BURGUNDY,TOU
340	FRANCE,BURGUNDY,TOU
341	FRANCE,BURGUNDY,TOU
343	FRANCE,BURGUNDY,TOU
346	FRANCE,BURGUNDY,TOU
347	FRANCE,BURGUNDY,TOU
348	FRANCE,BURGUNDY,TOU
349	FRANCE,BURGUNDY,TOU
355	FRANCE,BURGUNDY,TOU
357	FRANCE,BURGUNDY,TOU
359	FRANCE,BURGUNDY,TOU
360	FRANCE,BURGU

7	DY,TOU
278	FRANCE,BURGUNDY,TOU
280	FRANCE,BURGUNDY,TOU
281	FRANCE,BURGUNDY,TOU
282	FRANCE,BURGUNDY,TOU
283	FRANCE,BURGUNDY,TOU
287	FRANCE,BURGUNDY,TOU
288	FRANCE,BURGUNDY,TOU
290	FRANCE,BURGUNDY,TOU
291	FRANCE,BURGUNDY,TOU
292	FRANCE,BURGUNDY,TOU
293	FRANCE,BURGUNDY,TOU
295	FRANCE,BURGUNDY,TOU
296	FRANCE,BURGUNDY,TOU
297	FRANCE,BURGUNDY,TOU
298	FRANCE,BURGUNDY,TOU
302	FRANCE,BURGUNDY,TOU
306	FRANCE,BURGUNDY,TOU
309	FRANCE,BURGUNDY,TOU
310	FRANCE,BURGUNDY,TOU
311	FRANCE,BURGUNDY,TOU
314	FRANCE,BURGUNDY,TOU
316	FRANCE,BURGUNDY,TOU
318	FRANCE,BURGUNDY,TOU
	NDY,TOU
361	FRANCE,BURGUNDY,TOU
363	FRANCE,BURGUNDY,TOU
364	FRANCE,BURGUNDY,TOU
366	FRANCE,BURGUNDY,TOU
367	FRANCE,BURGUNDY,TOU
368	FRANCE,BURGUNDY,TOU
369	
371	FRANCE,BURGUNDY,TOU
372	FRANCE,BURGUNDY,TOU
374	FRANCE,BURGUNDY,TOU
375	FRANCE,BURGUNDY,TOU
377	FRANCE,BURGUNDY,TOU
378	FRANCE,BURGUNDY,TOU
379	FRANCE,BURGUNDY,TOU
380	FRANCE,BURGUNDY,TOU
383	FRANCE,BURGUNDY,TOU
385	FRANCE,BURGUNDY,TOU
386	FRANCE,BURGUNDY,TOU
387	FRANCE,BURGUNDY,TOU
388	FRANCE,BURGUNDY,TOU
389	FRANCE,BURGUNDY,TOU
390	FRANCE
391	FRANCE

320	FRANCE,BURGUNDY,TOU
321	FRANCE,BURGUNDY,TOU
322	

accession ID[a]	Geographic group[b]
392	FRANCE
393	FRANCE
394	
395	FRANCE
396	FRANCE
397	FRANCE
5719	WORLD,EUROPE
5837	WORLD
6008	WORLD,EUROPE
6009	WORLD,EUROPE
6016	WORLD,EUROPE
6039	WORLD,EUROPE
6040	WORLD,EUROPE
6042	
6043	WORLD,EUROPE
6046	WORLD,EUROPE
6064	WORLD,EUROPE
6074	WORLD,EUROPE
6243	WORLD,EUROPE
6709	WORLD
6730	WORLD,EUROPE
6897	WORLD,EUROPE,FRANCE
6893	WORLD,EUROPE
6936	WORLD,EUROPE,FRANCE
6937	WORLD,EUROPE
6938	WORLD,EUROPE
6939	WORLD
6940	WORLD,EUROPE
6942	WORLD,EUROPE
6943	WORLD,EUROPE
6944	WORLD,EUROPE
6945	WORLD,EUROPE
6946	WORLD
6951	WORLD,EUROPE
6956	
6957	WORLD,EUROPE
6958	WORLD,EUROPE,FRANCE
6959	WORLD,EUROPE,FRANCE
6960	WORLD,EUROPE,FRANCE
6961	WORLD,EUROPE
6962	WORLD,EUROPE
696	WORLD,EUROPE

8	E
689 9	WORLD,EUROPE
690 0	WORLD,EUROPE
690 1	WORLD,EUROPE
690 3	WORLD,EUROPE
690 4	WORLD
690 6	WORLD,EUROPE
690 7	WORLD,EUROPE
690 9	
691 0	
691 1	WORLD
691 3	WORLD,EUROPE
691 4	WORLD,EUROPE
691 5	WORLD,EUROPE
691 6	WORLD,EUROPE
691 7	WORLD,EUROPE
691 8	WORLD,EUROPE
691 9	WORLD,EUROPE
692 0	WORLD
692 1	WORLD
692 2	WORLD,EUROPE
692 3	WORLD,EUROPE
692 4	WORLD,EUROPE
692 6	

3	
696 4	
696 5	WORLD,EUROPE
696 6	WORLD,EUROPE
696 7	WORLD,EUROPE
696 8	WORLD,EUROPE
696 9	WORLD,EUROPE
697 0	WORLD,EUROPE
697 1	WORLD,EUROPE
697 2	
697 3	WORLD,EUROPE
697 4	WORLD,EUROPE
697 5	WORLD,EUROPE
697 6	WORLD
697 7	WORLD
697 9	WORLD,EUROPE
698 0	WORLD,EUROPE
698 1	
698 2	WORLD,EUROPE
698 3	WORLD
698 4	WORLD,EUROPE
698 5	WORLD,EUROPE
698 8	WORLD,EUROPE
699 2	WORLD,EUROPE

6927	WORLD
6928	WORLD
6929	WORLD
6930	WORLD,EUROPE
6931	WORLD,EUROPE
6932	

accession ID[a]	**Geographic group[b]**
7092	WORLD,EUROPE,FRANCE
7098	WORLD,EUROPE,FRANCE
7296	WORLD,EUROPE
7307	WORLD,EUROPE,FRANCE
7320	
7323	WORLD
7394	
7430	WORLD,EUROPE,FRANCE
7438	WORLD,EUROPE
7461	WORLD
7506	WORLD,EUROPE,FRANCE
7507	WORLD,EUROPE,FRANCE
7508	WORLD,EUROPE,FRANCE
7514	WORLD
7515	WORLD
7516	WORLD,EUROPE
7517	WORLD,EUROPE

6994	WORLD,EUROPE,FRANCE
7015	WORLD,EUROPE
7033	WORLD
7071	WORLD,EUROPE,FRANCE
7075	WORLD,EUROPE,FRANCE
7081	WORLD,EUROPE

accession ID[a]	**Geographic group[b]**
8271	WORLD,EUROPE
8274	WORLD,EUROPE
8275	WORLD,EUROPE,FRANCE
8283	WORLD,EUROPE
8284	WORLD,EUROPE
8285	WORLD,EUROPE
8290	
8296	WORLD,EUROPE
8297	WORLD,EUROPE
8300	
8304	WORLD,EUROPE
8306	WORLD,EUROPE
8310	WORLD,EUROPE
8311	WORLD,EUROPE
8312	WORLD,EUROPE
8313	
8314	WORLD,EUROP

17	
7518	WORLD,EUROPE
7519	WORLD,EUROPE
7520	WORLD,EUROPE
7521	WORLD,EUROPE
7522	
7523	WORLD
7524	WORLD
7525	WORLD
7526	WORLD
8213	WORLD,EUROPE
8214	WORLD,EUROPE, FRANCE
8215	WORLD,EUROPE
8222	WORLD,EUROPE
8230	WORLD,EUROPE
8231	WORLD,EUROPE
8233	WORLD
8235	WORLD,EUROPE
8236	WORLD,EUROPE
8238	
8239	WORLD,EUROPE
8240	WORLD,EUROPE
8241	
8242	WORLD,EUROPE

	E
8323	WORLD,EUROPE
8325	WORLD,EUROPE
8326	
8329	WORLD,EUROPE,FRANCE
8334	WORLD,EUROPE
8335	WORLD,EUROPE
8337	WORLD,EUROPE
8343	
8348	WORLD,EUROPE
8351	WORLD,EUROPE
8353	WORLD,EUROPE
8354	WORLD,EUROPE
8357	WORLD,EUROPE
8365	WORLD,EUROPE
8366	WORLD,EUROPE
8369	WORLD,EUROPE
8374	
8376	WORLD,EUROPE
8378	WORLD,EUROPE
8386	
8387	
8388	WORLD,EUROPE
8389	

8243	WORLD,EUROPE
8245	
8247	WORLD,EUROPE
8249	WORLD,EUROPE
8256	
8258	WORLD,EUROPE
8259	WORLD,EUROPE
8266	WORLD,EUROPE
8270	WORLD,EUROPE
8395	
8411	WORLD,EUROPE
8412	WORLD,EUROPE
8420	WORLD,EUROPE
8422	WORLD,EUROPE
8423	
8426	WORLD,EUROPE
8430	WORLD,EUROPE
9057	WORLD,EUROPE
9058	
100000	WORLD,EUROPE

[a] Accession identities have been retrieved from Horton *et al.* (2012).

[b] Six different geographical scales. Grey cases correspond to natural accessions that were likely to be contaminants, *i.e.* accessions for which geographical origin is suspicious (Anastasio *et al.* 2011).

Table S2. List of the 328 *a priori* candidate genes for flowering time.

Name	LocusTag	Start	End
LHY	AT1G01060	33379	37840
GA2ox6	AT1G02400	486801	489577
CRY2	AT1G04400	1185550	1188517
CKL13	AT1G04440	1202255	1205803
STO	AT1G06040	1828413	1829890
NFYC3	AT1G08970	2882524	2884337
PIF3	AT1G09530	3076582	3079539
PHYA	AT1G09570	3095256	3100357
HYL1	AT1G09700	3137767	3140353
AT1G10588	AT1G10588	3501145	3501904
DDF1	AT1G12610	4289944	4291017
LWD1	AT1G12910	4394897	4396291
RAV1	AT1G13260	4542168	4543742
UBC1	AT1G14400	4927011	4928690
GAI	AT1G14920	5149226	5151354
GA4	AT1G15550	5344478	5346166
Cstf77	AT1G17760	6110107	6116617
miR159b	AT1G18075	6220648	6220833
ATARP4	AT1G18450	6348107	6351975
AT1G22690	AT1G22690	8027298	8028125
GI	AT1G22770	8061844	8067716
SEPALLATA3	AT1G24260	8593642	8596098
PFT1	AT1G25540	8969065	8974660
TEM1	AT1G25560	8981677	8983041
CAL	AT1G26310	9100153	9103603
ATGA2OX2	AT1G30040	10537648	10539831
UFO	AT1G30950	11036180	11037508
AT1G30960	AT1G30960	11037612	11040033
SUF4	AT1G30970	11040281	11043751
FRL2	AT1G31814	11412608	11414502
VIP1	AT1G43700	16484231	16486241
ATGA20OX5	AT1G44090	16760677	16762486
CH1	AT1G44446	16848359	16851224
GA2ox4	AT1G47990	17698655	17700834
GCR1	AT1G48270	17827953	17830420
RTV1	AT1G49480	18314178	18316645
ATGA2OX7	AT1G50960	18889549	18891719
AT1G52800	AT1G52800	19664044	19665362
SPA4	AT1G53090	19783352	19786902
SPL4	AT1G53160	19806419	19807608
NFYC9	AT1G54830	20451083	20452671
AT1G55080	AT1G55080	20553011	20554254
HUB2	AT1G55250	20607214	20612378
NFYC2	AT1G56170	21024764	21025883
ORTH2	AT1G57820	21414170	21417946
ARR3	AT1G59940	22065617	22066973
ATGA20OX4	AT1G60980	22452573	22454140
VIP5	AT1G61040	22483207	22485969
ATSCO1	AT1G62750	23233434	23236447

Name	LocusTag	Start	End
LDL1	AT1G62830	23264490	23267202
DDF2	AT1G63030	23367407	23368410
FT	AT1G65480	24331428	24333934
RGL1	AT1G66350	24748195	24750043
FKF1	AT1G68050	25508676	25510889
TEM2	AT1G68840	25880327	25881736
AP1	AT1G69120	25982330	25986313
CDF5	AT1G69570	26161528	26163496
AT1G69935	AT1G69935	26341801	26343126
MMP	AT1G70170	26423874	26425356
AGL12	AT1G71692	26952645	26955127
Cstf64	AT1G71800	26999452	27002342
AT1G72050	AT1G72050	27115024	27117470
HAP2C	AT1G72830	27405456	27407969
miR159a	AT1G73687	27713234	27713415
MIF1	AT1G74660	28047576	28048127
AT1G74670	AT1G74670	28053286	28054149
ASHH1	AT1G76710	28789729	28792676
FLM	AT1G77080	28955637	28960096
EFS	AT1G77300	29039922	29048810
ATGA2OX1	AT1G78440	29511599	29513051
NUA	AT1G79280	29819176	29832978
GA2	AT1G79460	29890392	29894587
ELF7	AT1G79730	30000538	30004005
ATGA3OX4	AT1G80330	30198061	30199537
GA4H	AT1G80340	30200693	30202200
MOS3	AT1G80680	30324008	30328769
RGA1	AT2G01570	255249	257550
CAND1	AT2G02560	689788	697596
UBC2	AT2G02760	773684	775371
PKS1	AT2G02950	854947	856538
SEPALLATA4	AT2G03710	1129268	1131838
CR88	AT2G04030	1281841	1286104
ELF8	AT2G06210	2428903	2436687
AT2G14900	AT2G14900	6404175	6405330
PHYB	AT2G18790	8139881	8144430
LKP2	AT2G18915	8194572	8197483
miR156g	AT2G19425	8412516	8412618
FVE	AT2G19520	8455936	8459525
PIL5	AT2G20180	8704024	8706892
FIO1	AT2G21070	9040863	9043528
ATGRP7	AT2G21660	9265249	9266393
SVP	AT2G22540	9579874	9583893
AGL17	AT2G22630	9618372	9621957
CLF	AT2G23380	9955553	9960359
COL3	AT2G24790	10566898	10568145
miR156a	AT2G25095	10676472	10676553
ELF3	AT2G25930	11059035	11063324
HAP2C	AT2G26710	11380492	11383612

Table S2. (continued)

Name	LocusTag	Start	End
ATC	AT2G27550	11773251	11774681
PNF	AT2G27990	11921433	11924698
miR172a	AT2G28056	11941661	11943604
SYD	AT2G28290	12056213	12073083
TOE1	AT2G28550	12225951	12228543
AT2G30810	AT2G30810	13127826	13128666
ATX1	AT2G31650	13455272	13462181
COP1	AT2G32950	13977933	13983535
SPL3	AT2G33810	14305001	14306072
FES1	AT2G33835	14311787	14314700
ATGA2OX3	AT2G34555	14556988	14558697
NFYA4	AT2G34720	14649767	14651627
FHY1	AT2G37678	15801465	15802793
NFYB1	AT2G38880	16238476	16240834
SNZ	AT2G39250	16388886	16391073
AT2G39540	AT2G39540	16500866	16501241
HOS1	AT2G39810	16612800	16618057
ELF4	AT2G40080	16734294	16734912
SPL9	AT2G42200	17587407	17589630
SHP2	AT2G42830	17820255	17824013
PIF4	AT2G43010	17886427	17889050
FPA	AT2G43410	18025247	18031243
CKB4	AT2G44680	18426546	18428357
HUB1	AT2G44950	18542213	18548591
SAP18	AT2G45640	18799610	18801383
AGL6	AT2G45650	18804350	18806522
SOC1	AT2G45660	18807538	18811047
miR159c	AT2G46255	18994632	18994856
SPA1	AT2G46340	19022173	19027413
APRR9	AT2G46790	19232649	19235087
CCA1	AT2G46830	19245672	19248914
AT2G47310	AT2G47310	19423684	19427277
RFI2	AT2G47700	19552314	19554584
MBD9	AT3G01460	173316	182454
SEPALLATA2	AT3G02310	464279	467074
COL2	AT3G02380	487236	488693
GASA5	AT3G02885	638021	639055
ATVGT1	AT3G03090	700456	704769
AT3G04510	AT3G04510	1215636	1216958
FLK	AT3G04610	1250553	1254873
WNK1	AT3G04910	1354635	1358219
HST	AT3G05040	1401271	1408197
ATGID1A	AT3G05120	1430471	1432778
ATHAP2B	AT3G05690	1676546	1678932
AT3G06910	AT3G06910	2178630	2181197
AtPRMT4b	AT3G06930	2185143	2189387
COL9	AT3G07650	2441657	2444532
AT3G10185	AT3G10185	3145579	3146199
FLD	AT3G10390	3229293	3232345

Name	LocusTag	Start	End
miR172c	AT3G11435	3599776	3599908
ATMYB65	AT3G11440	3602093	3605104
SPY	AT3G11540	3631887	3637955
PIE1	AT3G12810	4065042	4074078
LDL2	AT3G13682	4479193	4481509
SPL5	AT3G15270	5140365	5141348
SPA3	AT3G15354	5169095	5172837
VRN1	AT3G18990	6548869	6551853
HAF2	AT3G19040	6567157	6575282
ATFYPP3	AT3G19980	6961736	6965108
DDL	AT3G20550	7174464	7177942
FIE	AT3G20740	7248809	7252452
AT3G21320	AT3G21320	7499053	7501841
TIC	AT3G22380	7912905	7919510
VRN5	AT3G24440	8876027	8878171
TEL1	AT3G26120	9546398	9549186
LWD2	AT3G26640	9793220	9794457
FUS3	AT3G26790	9853828	9855989
BR6OX2	AT3G30180	11810737	11813765
ATARP6	AT3G33520	14093656	14095549
PCL1	AT3G46640	17183090	17185218
CDF3	AT3G47500	17504000	17506058
REF6	AT3G48430	17935609	17940746
NFYC1	AT3G48590	18008657	18009982
UBP26	AT3G49600	18380549	18387128
AMP1	AT3G54720	20254725	20257848
SMZ	AT3G54990	20373718	20376522
AGL16	AT3G57230	21177423	21180932
AT3G57300	AT3G57300	21199488	21207885
AGL18	AT3G57390	21233701	21235911
SPL15	AT3G57920	21444321	21446035
GIS	AT3G58070	21506613	21507654
SHP1	AT3G58780	21738460	21741907
PIL6	AT3G59060	21827978	21830507
CKB3	AT3G60250	22270337	22272113
AGL13	AT3G61120	22618259	22620491
PIL2	AT3G62090	22988547	22990709
ATGID1B	AT3G63010	23289425	23291486
CRP	AT4G00450	202416	211003
FRI	AT4G00650	269026	271503
ETC3	AT4G01060	460472	461085
EZA1	AT4G02020	886600	891955
LD	AT4G02560	1123490	1128421
GA1	AT4G02780	1237767	1244813
DFL2	AT4G03400	1497536	1499865
PDF2	AT4G04890	2476489	2482345
CRY1	AT4G08920	5724103	5727253
GASA2	AT4G09610	6074770	6075645
SPA2	AT4G11110	6771605	6777225

Table S2. (continued)

Name	LocusTag	Start	End
AGL14	AT4G11880	7143115	7147222
pEARLI1	AT4G12480	7406105	7406937
COP9	AT4G14110	8132886	8134920
NFYB3	AT4G14540	8344616	8345218
ELIP2	AT4G14690	8418283	8419263
FAR1	AT4G15090	8614067	8618145
AT4G15180	AT4G15180	8651406	8662587
ESD4	AT4G15880	9012645	9016116
PHYD	AT4G16250	9195602	9199486
FCA	AT4G16280	9206597	9214825
HAT4	AT4G16780	9449114	9450743
AT4G16810	AT4G16810	9459870	9462253
VRN2	AT4G16845	9476143	9479878
CKB2	AT4G17640	9825197	9827272
PHYE	AT4G18130	10042137	10046082
TSF	AT4G20370	11000771	11002996
Jmj4	AT4G20400	11008666	11013860
ATGA2OX8	AT4G21200	11302685	11306601
ATGA3OX3	AT4G21690	11527229	11529060
EBS	AT4G22140	11727726	11730509
AGL19	AT4G22950	12023915	12027421
AT4G23340	AT4G23340	12195453	12196793
SLY1	AT4G24210	12563553	12564482
AGL24	AT4G24540	12670965	12674072
PGI1	AT4G24620	12708752	12712825
GA5	AT4G25420	12990884	12992458
FWA	AT4G25530	13038360	13042443
TOR1	AT4G27060	13581401	13585155
CIP7	AT4G27430	13718679	13723324
ATHXK1	AT4G29130	14352037	14355103
VIP3	AT4G29830	14597661	14599300
AT4G30200	AT4G30200	14786633	14790503
miR156b	AT4G30972	15074945	15075024
ATPRMT5	AT4G31120	15132011	15136639
FLP1	AT4G31380	15229785	15230718
CYP83B1	AT4G31500	15273471	15275310
miR156c	AT4G31877	15413319	15415873
KNAT5	AT4G32040	15494065	15496356
ATH1	AT4G32980	15914722	15918044
AT4G33280	AT4G33280	16047354	16049355
CIB1	AT4G34530	16498391	16500174
FD	AT4G35900	17004595	17006287
AP2	AT4G36920	17400847	17403332
HLS1	AT4G37580	17658612	17660878
AGL21	AT4G37940	17835695	17838621
BRI1	AT4G39400	18324661	18328826
PNY	AT5G02030	395634	399041
FHL	AT5G02200	437460	438894
PRR7	AT5G02810	637897	641977

Name	LocusTag	Start	End
LCL1	AT5G02840	648704	651972
ATHB51	AT5G03790	1004983	1006373
TFL1	AT5G03840	1024641	1025812
ELF6	AT5G04240	1169544	1174878
miR172b	AT5G04275	1188211	1188299
CPD	AT5G05690	1702688	1706787
MYB33	AT5G06100	1837907	1840727
YAP169	AT5G07200	2243553	2245339
CHE	AT5G08330	2680744	2681813
FLC	AT5G10140	3173497	3179448
AT5G10625	AT5G10625	3358787	3359781
miR156d	AT5G10945	3456647	3456732
HY5	AT5G11260	3593380	3594992
EMF1	AT5G11530	3695862	3701548
miR156e	AT5G11977	3867213	3867308
HAP2A	AT5G12840	4050691	4053606
FY	AT5G13480	4326528	4331699
AGL15	AT5G13790	4449014	4450843
AT5G14920	AT5G14920	4826479	4827980
GASA4	AT5G15230	4944900	4946216
CO	AT5G15840	5171182	5172758
COL1	AT5G15850	5176091	5177897
KIN1	AT5G15960	5209898	5210727
KIN2	AT5G15970	5211911	5212665
FRL1	AT5G16320	5344502	5346019
TFL2	AT5G17690	5827171	5829682
ASP2	AT5G19550	6598017	6601819
NPH4	AT5G20730	7016445	7022113
HUA2	AT5G23150	7785835	7792489
APRR5	AT5G24470	8355951	8358873
FPF1	AT5G24860	8541778	8542449
TNY	AT5G25810	8986771	8987787
GA3	AT5G25900	9036018	9038406
miR156f	AT5G26147	9136126	9136215
AT5G27230	AT5G27230	9584092	9588049
ATGID1C	AT5G27320	9629087	9631210
AT5G28450	AT5G28450	10372938	10374190
LSH1	AT5G28490	10454393	10455196
PHYC	AT5G35840	14007826	14011764
SEF	AT5G37055	14641551	14642440
CIR1	AT5G37260	14751344	14753088
TCH2	AT5G37770	14998854	14999619
PMI15	AT5G38150	15223116	15224947
CDF2	AT5G39660	15878699	15881044
XPB2	AT5G41360	16544340	16549280
CIP1	AT5G41790	16727530	16732847
CUL4	AT5G46210	18731418	18736810
AT5G46910	AT5G46910	19047780	19050880
LBA1	AT5G47010	19072009	19079334

Table S2. (continued)

Name	LocusTag	Start	End
CKB1	AT5G47080	19124612	19126611
NFYB2	AT5G47640	19309227	19310272
PAT1	AT5G48150	19522255	19524698
AtPRMT4a	AT5G49020	19871251	19874920
EMF2	AT5G51230	20823736	20829564
AT5G51310	AT5G51310	20852854	20854718
GA20ox2	AT5G51810	21055188	21056808
PGM	AT5G51820	21063368	21068057
DFL1	AT5G54510	22131093	22133678
miR156h	AT5G55835	22597012	22597117
ZTL	AT5G57360	23241427	23244590
VIN3	AT5G57380	23246395	23249504
COL5	AT5G57660	23355464	23356989
MSI1	AT5G58230	23556012	23558245
SRR1	AT5G59560	24000556	24001957
AT5G59570	AT5G59570	24003888	24005512
VIP2	AT5G59710	24057407	24061918
AT5G59845	AT5G59845	24111324	24112020
APRR3	AT5G60100	24197999	24201364
TOE2	AT5G60120	24207786	24211724
AGL8	AT5G60910	24502482	24506143
VIP4	AT5G61150	24603656	24607725
TOC1	AT5G61380	24675064	24678550
LFY	AT5G61850	24844295	24846933
AT5G62040	AT5G62040	24922810	24923709
CDF1	AT5G62430	25069093	25070934
ELF5	AT5G62640	25149433	25152541
LIP1	AT5G64813	25910279	25912896
MAF2	AT5G65050	25982254	25986326
MAF3	AT5G65060	25987429	25991315
MAF4	AT5G65070	25992260	25996134
MAF5	AT5G65080	25997504	26002465
AT5G65540	AT5G65540	26195689	26198322
ICU2	AT5G67100	26776994	26785104

Table S3. Structure of phenological variation in France in the Fall cohort.

raits	Model terms	*F* or LRT	*P*	Var Comp (%)
		18.	**0.0**	
T	Block	56	**001**	0.6
		11	**0.0**	
	Region	3.21	**001**	26.3
		22.	**0.0**	
	Stand(Region)	38	**001**	37.1
	Family(Stand(	19	**0.0**	
	Region))	1.10	**001**	19.0
		1.3	0.2	
	Control Bg-2	6	441	-
	Error			17.0
		0.2	0.6	
NT	Block	7	057	0.0
		12.	**0.0**	
	Region	69	**001**	5.4
		2.9	**0.0**	
	Stand(Region)	7	**001**	6.9
	Family(Stand(	7.4	**0.0**	
	Region))	0	**065**	10.5
		1.0	0.3	
	Control Bg-2	3	109	-
	Error			77.3
		20.	**0.0**	
T	Block	88	**001**	0.7
		95.	**0.0**	
	Region	11	**001**	23.5
		22.	**0.0**	
	Stand(Region)	47	**001**	38.2
	Family(Stand(	16	**0.0**	
	Region))	3.30	**001**	18.8
		0.8	0.3	
	Control Bg-2	6	544	-
	Error			18.9
		4.6	**0.0**	
P	Block	2	**320**	0.7
		5.6	**0.0**	
	Region	6	**008**	2.2
		4.6	**0.0**	
	Stand(Region)	0	**001**	13.1
	Family(Stand(	6.3	**0.0**	
	Region))	0	**121**	8.8
	Control Bg-2	1.8	0.1	-

		9	700	
	Error			75.2
		14.	**0.0**	
P	Block	36	**002**	1.6
		2.9	**0.0**	
	Region	8	**308**	0.3
		3.6	**0.0**	
	Stand(Region)	4	**001**	9.8
	Family(Stand(	0.3	0.5	
	Region))	0	839	1.6
		16.	**0.0**	
	Control Bg-2	22	**001**	-
	Error			86.7
		0.0	0.8	
RR	Block	2	932	0.0
		15.	**0.0**	
	Region	48	**001**	8.4
		5.7	**0.0**	
	Stand(Region)	6	**001**	14.4
	Family(Stand(	0.2	0.6	
	Region))	0	547	0.0
		0.1	0.7	
	Control Bg-2	0	489	-
	Error			77.2

Model random terms were tested with likelihood ratio tests of models with and without these effects. Random effects are in italic. Significant effects are highlighted in bold.

'VarComp (%)' Percentage of variance explained by the corresponding term of the model. Because 'Control Bg-2' is a covariate, this term was not included in the variance component analysis.

Table S4. Structure of phenological variation in France in the Spring cohort.

raits	Model terms	*F* or LRT	*P*	Var Comp (%)
		1.1	0.2	
T	Block	4	864	0.0
		13	**0.0**	
	Region	4.26	**001**	28.7
		22.	**0.0**	
	Stand(Region)	26	**001**	37.6
	Family(Stand(	20	**0.0**	
	Region))	9.40	**001**	19.0
		5.7	**0.0**	
	Control Bg-2	3	**168**	-
	Error			14.8
		0.7	0.3	
NT	Block	4	912	0.0
		1.8	0.1	
	Region	6	350	0.0
		3.8	**0.0**	
	Stand(Region)	4	**001**	10.5
	Family(Stand(	10.	**0.0**	
	Region))	90	**010**	14.4
		0.0	0.9	
	Control Bg-2	0	661	-
	Error			75.1
		3.8	**0.0**	
T	Block	8	**493**	0.0
		73.	**0.0**	
	Region	24	**001**	22.1
		20.	**0.0**	
	Stand(Region)	70	**001**	43.7
	Family(Stand(	23	**0.0**	
	Region))	0.40	**001**	21.7
		6.7	**0.0**	
	Control Bg-2	1	**098**	-
	Error			12.6
		6.0	**0.0**	
P	Block	4	**143**	0.6
		8.7	**0.0**	
	Region	2	**001**	1.6
		6.1	**0.0**	
	Stand(Region)	4	**001**	25.1
	Family(Stand(	26.	**0.0**	
	Region))	00	**001**	20.0
	Control Bg-2	0.0	0.7	-

		9	630	
	Error			52.7
		2.8	0.0	
P	Block	4	928	0.1
		17.	**0.0**	
	Region	62	**001**	9.5
		5.8	**0.0**	
	Stand(Region)	0	**001**	22.2
	Family(Stand(	18.	**0.0**	
Region))		50	**001**	17.9
		0.4	0.4	
	Control Bg-2	9	838	-
	Error			50.3
		3.2	0.0	
RR	Block	8	706	0.9
		4.5	**0.0**	
	Region	5	**036**	1.5
		4.8	**0.0**	
	Stand(Region)	6	**001**	17.5
	Family(Stand(	9.4	**0.0**	
Region))		0	**022**	14.3
		3.0	0.0	
	Control Bg-2	3	820	-
	Error			65.7

Random terms were tested with likelihood ratio tests of models with and without these effects. Random effects are in italics. Significant effects are highlighted in bold.

'VarComp (%)' Percentage of variance explained by the corresponding term of the model. Because 'Control Bg-2' is a covariate, this term was not included in the variance component analysis.

Table S5. Genetic coefficient of phenological variation for BT, INT, FT, FP, RP and FRR for each region and cohort.

ohort	rait	B rittany	B urgundy	La nguedoc	orth	*P*
all	T	0 .062	0. 231	0.1 43 [b]	.091	**0.0 020**
	NT	0 .165	0. 215	0.5 89	.095	**0.0 006**
	T	0 .050	0. 186	0.1 52 [b]	.054 [c]	**0.0 009**
	P	0 .086	0. 133	0.1 50	.077	0.3 706
	P	0 .035	0. 039	0.0 80	.048	0.4 668
	RR	0 .048	0. 091	0.0 85	.028	0.1 734
pring	T	0 .124	0. 296	0.1 98 [b]	.074	**0.0 096**
	NT	0 .287	0. 549	0.2 86	.178	0.1 956
	T	0 .072	0. 265	0.2 16	.059	**0.0 004**
	P	0 .136	0. 185	0.1 53	.062	0.1 859
	P	0 .037	0. 120 [b]	0.1 48	.032	**0.0 037**
	RR	0 .089	0. 135	0.0 81	.038	0.1 853

Bold *p*-values indicate a significant 'region' effect after Bonferroni correction. When a 'region' effect is significant, different letters indicate different coefficients of variation among regions based on a Tukey's test of multiple comparisons of means ($P = 0.05$).

Table S6. Mean values for genetic diversity parameters of the 49 French stands.

Region	Stands	Habitat	Alt[a]	H_S[b]	PL[c]	n_a[d]	R_S[e]	HG[f]	PHG[g]
Brittany	**ABA**	grassland	95	0.008	1.48	1.01	1.01	1	1
	BRE	grassland	41	0.264	83.70	1.84	1.26	5	2
	DIR-2.8	grassland	105	0.230	55.56	1.56	1.22	5	5
	DIR-8.5	grassland	105	0.138	60.00	1.60	1.13	1	1
	FOR	meadow	9	0.184	75.56	1.76	1.18	4	2
	MIL	meadow	75	0.092	57.78	1.58	1.09	3	2
	PLO	hoed land	86	0.215	56.30	1.56	1.21	4	2
	PLY	hoed land	177	0.177	75.56	1.76	1.17	3	1
	ROC	grassland	22	0.005	1.48	1.01	1.00	1	0
	RUM	hoed land	14	0.114	33.33	1.33	1.11	2	1
	TRE	hoed land	42	0.233	85.19	1.85	1.23	5	3
			mean	0.15	53.27	1.53	1.15	3.09	1.82
			SD	0.09	29.63	0.30	0.09	1.64	1.33
Burgundy	**CIRY**	grassland	271	0.268	77.04	1.77	1.26	7	6
	CON	grassland	508	0.081	41.48	1.41	1.08	2	2
	ETA-1	grassland	300	0.240	56.30	1.56	1.23	5	5
	ETA-2	hoed land	300	0.187	59.26	1.59	1.17	3	3
	LIE	grassland	530	0.308	81.48	1.81	1.30	7	6
	MAR-3	grassland	439	0.003	0.74	1.01	1.00	1	1
	MAR-4	meadow	416	0.005	0.74	1.01	1.00	1	1
	MOL	meadow	344	0.241	64.44	1.64	1.24	10	9
	RAD	meadow	289	0.314	87.41	1.87	1.31	16	16
	TOU-J2	hoed land	279	0.151	56.30	1.56	1.15	4	4
	TOU-M1	hoed land	312	0.332	88.15	1.88	1.32	11	11
			mean	0.19	55.76	1.56	1.19	6.09	5.82
			SD	0.12	30.84	0.31	0.12	4.72	4.62
Languedoc	**ARR-1**	meadow	489	0.007	2.22	1.02	1.01	1	0
	ARR-2	meadow	489	0.192	34.07	1.34	1.18	2	1
	ARR-3	meadow	489	0.124	37.04	1.37	1.11	2	1
	BEZ	grassland	357	0.038	36.30	1.36	1.04	1	1
	ISS	grassland	199	0.316	48.15	1.48	1.26	2	1
	LEC	grassland	43	0.037	34.07	1.34	1.04	2	2
	MOU1-1	grassland	61	0.183	32.59	1.33	1.17	2	0
	MOU1-2	grassland	61	0.077	34.07	1.34	1.07	2	0
	MOU1-3	meadow	61	0.363	64.44	1.64	1.31	4	3
	MOU2	hoed land	63	0.306	81.48	1.81	1.30	12	10
	NOZ	grassland	78	0.185	55.56	1.56	1.18	3	1
	PCH	grassland	59	0.267	45.19	1.45	1.23	3	1
	QUI	grassland	88	0.029	28.15	1.28	1.03	2	0
	SAL	grassland	34	0.043	38.52	1.39	1.04	2	0
	VED	grassland	39	0.350	70.37	1.70	1.31	4	3
	VEN	grassland	42	0.098	18.52	1.19	1.09	2	1
			mean	0.163	41.296	1.413	1.147	2.875	1.563
			SD	0.125	19.601	0.196	0.109	2.579	2.449
North	**BAU**	grassland	26	0.046	11.11	1.11	1.05	4	4
	BRI	hoed land	19	0.192	66.67	1.67	1.19	3	3
	CAT-S	hoed land	161	0.110	60.00	1.60	1.11	2	0
	CAT-T	hoed land	135	0.027	22.96	1.23	1.03	1	0
	ENC-1	hoed land	74	0.287	72.59	1.73	1.28	7	7
	ENC-2	hoed land	39	0.197	57.78	1.58	1.19	3	3
	ESP1	meadow	47	0.042	43.70	1.44	1.04	1	0
	ESP2	hoed land	47	0.153	66.67	1.67	1.15	3	1
	GEN	hoed land	34	0.207	64.44	1.64	1.20	5	5
	LCL	grassland	17	0.288	64.44	1.64	1.28	3	1
	WAV	hoed land	19	0.000	0.00	1.00	1.00	1	1
			mean	0.14	48.22	1.48	1.14	3.00	2.27
			SD	0.10	25.28	0.25	0.10	1.84	2.33

[a] Alt: Altitude (m), [b] H_S: Mean gene diversity, [c] *PL:* percentage of polymorphic loci, [d] n_a: mean number of observed alleles per locus, [e] R_S: mean allelic richness per locus, [f] *HG*: number of haplogroups, [g] *PHG*: number of private haplogroups. SD: standard deviation.

Table S7. List of flowering time candidate genes close to the 100 most associated SNPs for each 'trait x cohort x geographical scale' combination.

[a] SNP: number of top SNPs close to the candidate gene.

[b] Top rank: the highest rank among the top SNPs close to the candidate gene.

CHAPITRE 1 – CONCLUSIONS ET PERSPECTIVES

2. Conclusions et perspectives

Un des résultats marquants de cette étude concerne l'étendue de la variation phénologique observée au sein de certaines populations françaises. En effet, certaines populations présentent presque autant de variation phénologique qu'un panel d'accessions naturelles mondiales. Le cas des populations MIB et TOU illustre bien ce phénomène et soulève de nombreuses questions. La nature adaptative d'une telle variation peut effectivement être interrogée. L'hypothèse neutre H0 serait que cette variation résulte de processus non-sélectifs. Etant donné la relation positive observée entre diversité phénologique et diversité génétique neutre entre les 49 patchs français, cette hypothèse pourrait être favorisée. Cependant, une large proportion de la variation phénologique semble rester malgré tout inexpliquée par des processus non-sélectifs. L'hypothèse alternative serait alors que cette variation observée à l'échelle d'une population résulte d'un phénomène d'adaptation micro-locale à des facteurs écologiques variant à une échelle fine comme le sol, la compétition avec d'autres espèces végétales ou encore l'attaque d'agents pathogènes. Le peu de connaissances disponibles sur le fonctionnement et la dynamique de populations naturelles d'*A. thaliana* nous empêche de privilégier l'une de ces deux hypothèses. Ces résultats nous ont donc conduit à effectuer une caractérisation écologique à une échelle spatiale fine des populations MIB et TOU et à étudier leur évolution phénotypique sur une dizaine de générations afin de déterminer la part relative de cette variation intra-population due à des processus sélectifs. Ces études portant sur les populations MIB et TOU feront l'objet du chapitre 3 de cette thèse.

La détection de relations significatives entre variation phénotypique et variation écologique nous a permis à la fois d'identifier (i) des traits phénologiques sous sélection, et (ii) les agents sélectifs potentiels agissant sur ces traits. Notre étude a démontré l'importance

de prendre en compte les processus non-sélectifs dans les études de relation phénotype-écologie et devraient donc nous encourager à la prudence dans l'attribution du caractère adaptatif à des traits phénotypiques en absence de contrôle pour des processus non-sélectifs. Le fait que nous ayons trouvé des relations entre facteurs écologiques et traits phénologiques après contrôle de l'effet de processus non-sélectifs renforce l'idée que ces traits sont sous sélection chez *A. thaliana*. Pour autant, le patron de sélection que nous observons semble très complexe, avec (i) des relations phénologie - écologie plus fortes à l'intérieur des régions qu'à l'échelle de la France, suggérant l'importance des agents sélectifs locaux dans le pattern de variation phénologique adaptative, et (ii) une diversité des agents sélectifs dépendante du trait, de la cohorte et de la région géographique considéré.

L'étendue de la variation phénologique observée à une échelle locale et l'identification de relations phénotype – écologie plus fortes au sein des régions qu'à une échelle géographique plus large, nous amène à nous poser la question suivante: pourquoi utiliser un échantillonnage mondial d'accessions pour identifier les bases génétiques associées à la variation naturelle adaptative d'un trait? En d'autres termes, plus modérés, l'utilisation d'un échantillonnage régional, voire local, d'accessions ne serait-il pas complémentaire à l'utilisation d'un échantillonnage mondial afin d'identifier les bases génétiques associées à la variation phénotypique adaptative ? En me basant sur le fait que nous n'ayons pas trouvé les mêmes bases génétiques à différentes échelles géographiques, je réponds par l'affirmative.

Par ailleurs, le fait que nous ayons identifié des agents sélectifs différents entre nos régions géographiques m'amène à me demander si les bases génétiques associées à la variation phénotypique adaptative sont les mêmes entre ces régions. Dans notre étude, seules les accessions bourguignonnes génotypées pour 214 kSNPs nous ont permis de réaliser des analyses de GWA mapping entre différentes populations naturelles (populations MIB et

TOU). Les résultats obtenus avec ce jeu de quelques accessions sont donc difficilement extrapolables aux autres régions françaises et m'empêchent de répondre à cette question.

Les résultats de cette étude concernant (i) l'échelle géographique de la variation phénotypique adaptative, et (ii) l'effet de l'échelle géographique sur l'identification des agents sélectifs et des bases génétiques, semble me confirmer la pertinence de l'utilisation d'un échantillonnage hiérarchique de populations afin d'identifier les bases génétiques de l'adaptation.

Cependant, les résultats trouvés dans cette étude se basent sur des données phénotypiques acquises en conditions contrôlées de serre et de ce fait, il est possible qu'ils ne reflètent pas une réalité biologique. En effet, dans son habitat naturel, les signaux perçus par une plante d'*A. thaliana* peuvent être bien plus nombreux et complexes que dans des conditions contrôlées de serre. Il m'a donc semblé important de replacer dans un contexte écologiquement réaliste l'étude précédente. Par ailleurs, notre étude en serre se concentrait sur la variation existant pour les traits phénologiques. Dans la nature, la sélection pourrait opérer de manière simultanée sur une multitude de traits physiologiques, morphologiques, phénologiques ou bien encore architecturaux. Il m'a donc semblé pertinent de travailler sur les stratégies phénotypiques observées dans les populations naturelles en analysant de manière conjointe plusieurs traits et leurs relations afin d'identifier de potentielles contraintes évolutives dues à des corrélations génétiques.

Notre étude indique que (i) l'échelle géographique de la variation naturelle adaptative, (ii) les agents sélectifs agissant sur les traits phénologiques et (iii) les bases génétiques associées à la variation naturelle sont fortement dépendants de la cohorte de germination considérée. Il m'est donc apparu pertinent de continuer à tenir compte de cet effet temporel, et

ceci à différentes échelles de temps (cohorte *vs*. année), afin d'étudier le rôle que joue la plasticité phénotypique dans l'adaptation chez *A. thaliana*.

Toutes ces perspectives font l'objet du chapitre 2 de cette thèse, où j'ai voulu étudier l'adaptation chez *A. thaliana* en me basant sur l'analyse conjointe d'un grand nombre de traits et de leurs plasticités mesurés dans un contexte écologique réaliste.

CHAPITRE 2 - REALISME ECOLOGIQUE, CONTEXTE MULTI-TRAITS ET INTEGRATION DE LA PLASTICITE PHENOTYPIQUE DANS L'ETUDE DE L'ADAPTATION: UNE TRANSITION VERS LA GENOMIQUE ECOLOGIQUE EVOLUTIVE

III. CHAPITRE 2 - REALISME ECOLOGIQUE, CONTEXTE MULTI-TRAITS ET INTEGRATION DE LA PLASTICITE PHENOTYPIQUE DANS L'ETUDE DE L'ADAPTATION: UNE TRANSITION VERS LA GENOMIQUE ECOLOGIQUE EVOLUTIVE :

1. Introduction

Dans ce chapitre, j'ai voulu (i) caractériser l'échelle spatiale à laquelle varient différents traits phénotypiques (morphologie, phénologie, architecture, acquisition de ressources, réallocation de ressources, survie…) et leurs plasticités dans des conditions écologiquement plus réalistes que celles contrôlées en serre, et (ii) identifier les patrons de sélection agissant sur les traits mesurés et leurs plasticités, en utilisant diverses approches (comparaison $F_{ST} - Q_{ST}$, relations phénotype – écologie et gradient de sélection génotypique).

Dans ce but, j'ai caractérisé la variation naturelle d'une vingtaine de traits phénotypiques existant au sein du jeu de 800 de familles utilisées dans le chapitre précédent. Les graines de ces 800 familles ont été semées à trois dates différentes sur un même terrain expérimental situé à l'Université de Lille 1. La date de germination pouvant grandement influencer la sélection opérant sur les traits post-germination au sein d'une même cohorte (Donohue *et al.* 2005 ; Wilczek *et al.* 2009), les graines ont été semées à deux dates différentes la même année (fin Août 2010 et fin Septembre 2010). Les dates de germination choisies encadrent la fenêtre de germination observée dans les populations naturelles d'*A. thaliana* situées dans le nord de la France. La sélection pouvant également grandement varier entre années (Siepielski *et al.* 2009), les graines des 800 familles ont aussi été semées à la même date sur deux années consécutives (fin Septembre 2010 et fin Septembre 2011). La

survie et la production totale de graines ont également été relevées pour chaque date de semis afin d'estimer la valeur sélective de chaque famille. La plasticité phénotypique exprimée par chaque famille a aussi été estimée pour chacun des traits entre cohortes (Août 2010 et Septembre 2010) et entre années (Septembre 2010 et Septembre 2011).

Cette étude a fait l'objet des deux manuscrits ci-joints :

R. VILLOUTREIX, C. GLORIEUX, B. BRACHI, J. BERGELSON, J. CUGUEN and F. ROUX. Evidence of selection acting on reproductive strategies in *Arabidopsis thaliana* at different spatial scales: when temporal heterogeneity in the field matters! *Soumis à Molecular Ecology.*

R. VILLOUTREIX, C. GLORIEUX, J. CUGUEN and F ROUX. Adaptive value and costs of phenotypic plasticity across germination cohorts and years in four regional sets of natural *Arabidopsis thaliana* families. *A soumettre une fois le précédent manuscrit accepté.*

Dans le premier manuscrit, je me suis attaché (i) à décrire la variation naturelle observée pour les stratégies reproductives et leurs plasticités et (ii) à étudier les patrons de sélection qui leur sont associés en utilisant deux approches complémentaires : comparaison $F_{ST} - Q_{ST}$ et relations phénotype – écologie. En effet, la quantité de totale de graines produites est souvent utilisée comme un estimateur de la valeur sélective chez *A. thaliana*. Pourtant, il a été montré que différentes lignées génétiques d'*A. thaliana* pouvaient produire un nombre équivalent de graines selon des patrons différents d'allocation de ressources (Roux *et al.* 2004). Il a par ailleurs été montré que ces patrons d'allocation de ressources et leurs plasticités pouvaient avoir des implications évolutives importantes dans la tolérance aux herbivores (Juenger & Bergelson 2000), dans la tolérance aux changements environnementaux (Shemesh & Novoplansky 2013) et dans la dispersion des graines (Wender *et al.* 2005). Le patron de

réallocation de ressources pouvant être autant la cible de la sélection naturelle que la quantité totale de graines produites, la description de ces patrons de réallocation et de leur plasticité ainsi que l'identification des régimes de sélection les affectant semblent donc être une piste à ne pas négliger dans l'optique d'identifier les bases génétiques de l'adaptation.

Alors que le premier manuscrit était focalisé uniquement sur les traits reproducteurs et la survie, je me suis attaché dans le deuxième manuscrit (i) à décrire la variation naturelle de 9 traits morphologiques, phénologiques et architecturaux et de leurs plasticités et (ii) à étudier les patrons de sélection qui leur sont associés par une approche de gradient de sélection génotypique. Dans ce manuscrit, je me suis également attaché à déterminer (i) si les traits sous sélection étaient identiques entre les trois dates de semis et entre différentes régions géographiques, (ii) si la plasticité phénotypique étaient sous sélection à différents pas de temps (cohorte *vs.* année) et dans différentes régions géographiques, et (iii) l'existence d'un potentiel coût de la plasticité phénotypique (en effet, de nombreux modèles théoriques prédisent une contre sélection de la plasticité phénotypique si celle-ci est coûteuse).

MANUSCRIT: EVIDENCE OF SELECTION ACTING ON REPRODUCTIVE STRATEGIES IN *ARABIDOPSIS THALIANA* AT DIFFERENT SPATIAL SCALES: WHEN TEMPORAL HETEROGENEITY IN THE FIELD MATTERS!

Evidence of selection acting on reproductive strategies in *Arabidopsis thaliana* at different spatial scales: when temporal heterogeneity in the field matters!

Romain Villoutreix[1], Cédric Glorieux[1], Benjamin Brachi[2], Joy Bergelson[2], Joël Cuguen[1], Fabrice Roux[1,3,4]*

[1]*Laboratoire Génétique et Evolution des Populations Végétales, UMR CNRS 8198, Université des Sciences et Technologies de Lille – Lille 1, F-59655 Villeneuve d'Ascq cedex France*

[2] *Department of Ecology and Evolution, University of Chicago, Chicago, Illinois 60637, USA.*

[3] *INRA, Laboratoire des Interactions Plantes-Microorganismes (LIPM), UMR441, F-31326 Castanet-Tolosan, France.*

[4] *CNRS, Laboratoire des Interactions Plantes-Microorganismes (LIPM), UMR2594, F-31326 Castanet-Tolosan, France.*

***Corresponding author:** Fabrice Roux

CNRS, Laboratoire des Interactions Plantes-Microorganismes (LIPM), UMR2594, F-31326 Castanet-Tolosan, France. E-mail: fabrice.roux@toulouse.inra.fr

Phone number: +33 (0)3 20 43 40 33 Fax number: +33 (0)3 20 43 69 79

Keywords: fitness-related traits, survival, ecology, adaptive differentiation, Q_{ST}-F_{ST} comparison, quantitative genetics

Running title: Detection of natural selection acting on reproductive strategies.

Summary

In so far as the pattern of resource allocation among reproductive traits may be as important as the total amount of resources allocated to reproduction, reproductive strategies offer a novel opportunity to study the genetics of local adaptation. Temporal and spatial variation in the strength and nature of selection in natural populations also calls for the incorporation of reaction norms in ecological genomic studies. Using a hierarchical sample of 49 stands located in four regions in France, measured under field conditions across two germination cohorts and two years, we aimed to (i) characterize natural genetic variation of individual reproductive traits, partitioning among reproductive traits, and their reaction norms, in the model species *Arabidopsis thaliana*, and (ii) test whether those traits are under natural selection. We considered two complementary approaches to measure selection, *i.e.* $Q_{ST} - F_{ST}$ comparisons and phenotype – ecology relationships. Extensive natural genetic variation was found for all traits, but it was dependent upon the spatial scale. Across germination cohorts, we observed little genetic variation in the reaction norms for total seed production, but extensive genetic variation in the reaction norms for resource allocation. The opposite pattern was observed across years, suggesting that plasticity in the trade-offs between resource acquisition and allocation depends on the time scale considered. The geographic region and environment also impact which traits were found to be under selection. Taken together, these results suggest that reproductive strategies in space and time should be considered when studying the genetics of local adaptation in *A. thaliana*. Selection acting on reaction norms further indicates that plasticity may have evolved independently of its expression in either environment, stressing the need to integrate genotype-by-environment interactions in the study of adaptive potential in *A. thaliana*.

Introduction

A central concept in life-history theory is resource allocation (Stearns 1992). Limitations in the resources available to an organism imply tradeoffs among traits, suggesting that the optimal allocation pattern differs with variation in selection pressures (reviewed in Weiner 2004). In plants, individuals may respond to their biotic or abiotic environment by modulating allocation to vegetative and reproductive traits, or allocation among reproductive traits (*i.e.* partitioning among different types of branches, between numbers of fruits and seeds per fruit, between seed quantity and quality). Such reproductive strategies are genotype-specific (Weiner 2004) and have evolutionary implications in tolerance to herbivory (Juenger & Bergelson 2000, Weinig *et al.* 2003), in environmental risks and production timing managements (Shemesh *et al.* 2012) and in seed dispersal (Wender *et al.* 2005). How resources are allocated among reproductive traits may therefore be as important as the amount of resources allocated to them.

A tremendous interest in finely mapping the genetic basis of local adaptation in *A. thaliana* has recently grown (Fournier-Level *et al.* 2011; Hancock *et al.* 2011; Gaut 2012), no doubt prompted by the powerful genomic resources publically available (Bergelson & Roux 2010). Building upon these initial studies, we propose two extensions. First, because genetic lines may produce the same number of seeds (an appropriate fitness proxy for a highly selfing species like *A. thaliana*) but with contrasted patterns of resource allocation (Roux *et al.* 2004; Roux & Reboud 2005; Paris *et al.* 2008), one should consider individual reproductive traits and seed partitioning among reproductive traits (hereafter named seed production ratios) when mapping the genetic bases of local adaptation. Second, fitness may vary from year to year in ecologically relevant conditions (Frenkel *et al.* 2008; Agren & Schemske 2012), suggesting that the reaction norms of reproductive traits may be as important as their expression in a

single environment. Identifying individual reproductive traits, seed production ratios and their reaction norms under natural selection may therefore help to map the genetic basis of reproductive strategies underlying a hierarchically structured phenotypic trait like total seed production (Bergelson & Roux 2010).

One approach to identify traits under selection is to perform multivariate genotypic selection gradient analysis to distinguish direct from indirect selection on individual phenotypic traits (Lande & Arnold 1983; Mitchell-Olds & Bergelson 1990; Rausher 1992). This method relies on an estimate of a fitness proxy and, therefore, cannot be applied to traits involved in fitness. This issue may be circumvented with two indirect approaches, a comparison of $Q_{ST} - F_{ST}$ and identification of phenotype – ecology relationships (Whitlock 2008, Hübner *et al.* 2012). In the $Q_{ST} - F_{ST}$ approach, the pattern of phenotypic and genetic differentiation is compared: if phenotypic differentiation among populations (Q_{ST}) is less than genetic differentiation (F_{ST}) then selection is considered uniform, whereas heterogeneous selection is suggested when $Q_{ST} > F_{ST}$. Equivalent values of Q_{ST} and F_{ST} suggest that trait divergence is the result of non-selective processes (Leinonen *et al.* 2013). These conclusions typically require independent validation because deviation from basic assumptions (such as dominance and epistasis; reviewed in Porcher *et al.* 2004) may lead to biased conclusions (Latta 1998; Le Corre & Kremer 2003). In the phenotype – ecology approach, correlations between plant phenotypes and environmental indices are sought. While simply correlational, this approach has the potential to identify both the traits under selection and the putative selective agents acting on them (Conner 2010). Furthermore, the effect of non-selective processes may be controlled, thereby avoiding spurious relationships between ecological factors and phenotypic variation (Brachi *et al.* 2013b).

In this study, we first aimed to describe natural variation in reproductive strategies by measuring survival, several reproductive traits, seed production ratios, and their reaction norms in *A. thaliana* across space and time in ecologically relevant conditions. Second, we aimed to identify reproductive traits, seed production ratios and their reaction norms under selection at different spatial scales by using both $Q_{ST} - F_{ST}$ comparisons and comparisons of phenotype – ecology relationships. Our approach involved a hierarchical sample of 800 families collected from 49 natural stands (defined as single patches of plants growing in relatively homogeneous ecological conditions) located in four climatically contrasted regions in France. In a previous study based on this set of 800 families, extensive natural variation for six phenological traits was observed across regions, stands of plants and families when plants were grown under greenhouse conditions (Brachi *et al.* 2013b). In this study, seeds were sown on three dates in the same common garden. Because the timing of germination is known to affect the expression of post-germination life-history traits in *A. thaliana* (Donohue 2002; Donohue *et al.* 2005; Wilczek *et al.* 2009), two germination cohorts were simulated within one year. In addition, natural accessions were grown in each of two consecutive years for the same germination cohort in order to explore fitness differences across years (Frenkel *et al.* 2008; Agren & Schemske 2012).

Materials and Methods

Plant material

The plant material used in this study has been described previously (Brachi *et al.* 2013b). Briefly, 800 individuals were collected from 49 natural stands located in four climatically contrasted regions of France (Brittany: oceanic, 11 stands; Burgundy: continental,

11 stands; Languedoc: Mediterranean, 16 stands and North: semi-oceanic, 11 stands). Seeds from individual plants constitute a "family". On average, 16.3 (SE: ±5.6) families were collected per stand, for a total of 200 families per region. Maternal effects were reduced by growing all families for one generation under controlled greenhouse conditions (16-h photoperiod, 20°C) at the University of Lille 1. Seven hundred and fifty-eight out of the 800 families were characterized genetically for 135 SNPs (Brachi *et al.* 2013b). The 49 stands were characterized for 25 biologically meaningful climatic variables, 14 edaphic factors and three intra- and inter-specific competition indices (Brachi *et al.* 2013b).

A set of 188 worldwide accessions (WA) and a set of 210 French accessions (FA) corresponding to the French RegMap (Horton *et al.* 2012) were also included in the experiment (Brachi *et al.* 2013b). These 398 accessions were measured for the purpose of another study and are therefore not considered here.

Experimental design

Seeds were sown at the University of Lille (Nord, France) on three dates (30th of August 2010, the 27th of September 2010 and the 26th of September 2011) to mimic *(i)* two distinct germination cohorts within the same year (late August 2010 and late September 2010) and *(ii)* the same germination cohort between two consecutive years (late September in 2010 and 2011). The two germination cohorts span the range of germination times in natural populations of *A. thaliana* in the north of France (personal observation, Fabrice Roux). The entire experiment of 7,416 plants was organized according to a split-plot design arranged as a randomized complete block design (RCBD) with three sowing dates (hereafter named 'August 2010', 'September 2010' and 'September 2011') nested within two experimental

blocks. Each 'block * sowing date' combination was represented by 19 arrays of 66 individual bottom-pierced wells (Ø4 cm, vol. ~38 cm3) (TEKU, JP 3050/66) filled with damp standard culture soil (007B Florafleur, France). Each 'block*sowing date' contained 1,236 plants with one replicate per family (n = 800), one replicate per natural accessions (n= 398) and the accession Bg-2, from Seattle (Washington, USA), placed in the same two positions within each array to control for micro-environmental variation within blocks (n = 38 = 19 trays * 2 replicates). In each 'block*sowing date' combination, the remaining 18 wells were left empty. At least five seeds were sown in each well. Seeds were stratified four days at 4°C in a cold chamber to promote germination. After the stratification treatment, arrays were treated to prevent dark-winged fungus gnats (Vectobac, 8mL per liter) and placed in a frost-free greenhouse without additional light or heating. To reduce micro-environmental variation, arrays were rotated daily in the cold chamber and in the greenhouse. For each of the three sowing dates, germination date was monitored daily for 17 days in the greenhouse. Germination rates were 95.7%, 96.4% and 96.3% for the 'August 2010', 'September 2010' and 'September 2011' sowing dates, respectively.

On day 18, wells with no germinated seeds received two seedlings of the same family but from a different well, and other wells were thinned to two seedlings. Flats were then transported outside to a common garden located at the University of Lille 1 (France). For each 'block*sowing date' combination, the 19 arrays were organized according to a grid of four columns and five lines (except the fourth column which had four lines). To facilitate root development, soil was tilled to allow arrays to be slightly buried. Seedlings were thinned to one seedling per well one week after placing the arrays in the common garden. Protection from herbivory by vertebrates, slugs and the aphid, *Myzus persicae,* was described in Brachi *et al.* (2010).

Plants were monitored every two or three days until all fruits were mature. The aboveground portion of each plant was harvested and stored at room temperature until phenotyping of reproductive traits and seed production ratios (see next subsection).

To characterize the climatic conditions experienced by plants, monthly data of precipitations and mean temperature were obtained from the Lille-Lesquin weather station located 6 km from the common garden experiment (https://espaceprofessionnels.meteofrance.com/espaceprofessionnels/accueil).

Survival, reproductive traits and seed production ratios

A total of 14 traits were measured. All plants that germinated but did not survive were counted as dead (survival = 0), whereas plants that were harvested were counted as alive (survival = 1).

We scored ten reproductive traits (Fig. 1). Total plant fitness was approximated by total fruit length (FITTOT) because the number of seeds in a fruit is highly correlated with fruit length (Roux *et al.* 2004). Seed production is a good proxy for fitness in a highly selfing annual species like *A. thaliana* (Platt *et al.* 2010). FITTOT was obtained by adding the fruit length produced on the primary shoot (FITSTEM), the primary branches on the primary shoot (FITPB) and the basal branches (FITBB). These estimates of fruit length were obtained by counting the number of fruits produced on each type of branches (FRUITSTEM, FRUITBB, FRUITPB) and multiplying these counts by an estimate of their corresponding fruit (or silique) length (SILSTEM, SILBB and SILPB), estimated as the average of three representative fruits.

We calculated three ratios corresponding to the percentage of seeds produced by one branch type as a function of the total amount of seed produced: RSTEM = FITSTEM / FITTOT, RPB = FITPB / FITTOT and RBB = FITBB / FITTOT.

Data Analysis

Natural variation and norms of reaction

We explored natural genetic variation of reproductive traits, seed production ratios and their reactions norms among our 800 families as a function of germination cohort within the same year (*i.e.* cohort effect) or year within the same germination cohort (*i.e.* year effect) using the following statistical model (Brachi *et al.* 2013):

$$Y_{ijklmc} = \mu_{trait} + block_i + sowing\ date_j + block_i*sowing\ date_j + region_k + stand_l\ (region_k) + family_m\ (stand_l\ (region_k)) + sowing\ date_j*region_k + sowing\ date_j*stand_j\ (region_k) + sowing\ date_j*\ family_m\ (stand_l\ (region_k)) + control_c\ (sowing\ date_j) + \varepsilon_{ijklmc} \quad (1)$$

In these models, '*Y*' is one of the quantitative traits, 'μ' is the overall phenotypic mean across environments and families; 'block' accounts for differences in micro-environment among the two experimental blocks; 'sowing date' corresponds to either the 'cohort effect' or the 'year effect'; 'region', 'stand' and 'family' measure the effect of three spatial scales in France, interaction terms involving 'sowing date' account for genetic variation in reaction norms among regions, among stands within regions and among families within stands; 'control' is a covariate accounting for array effects within blocks within each sowing date (the phenotypic

mean of two control replicates of Bg-2 per array); and 'ε' is the residual term. All factors were treated as fixed effects, except 'family', which was treated as a random effect. For fixed effects, terms were tested over their appropriate denominators for calculating *F*-values. Significance of the random effects was determined by likelihood ratio tests of model with and without these effects. For all traits except 'survival', model fitting was conducted using the PROC MIXED procedure in SAS 9.3 (SAS Institute Inc., Cary, North Carolina, USA). Because many families did not produce basal branches, we were not able to run this statistical model on SILBB. When necessary, raw data were either log transformed or Box-Cox transformed to satisfy the normality and equal variance assumptions of linear regression. For the binary trait 'survival', model fitting was conducted using the PROC GLIMMIX procedure in SAS 9.3 (SAS Institute Inc., Cary, North Carolina, USA). Convergence of the complete statistical model (1) for survival data was only obtained by excluding the terms 'family_m', 'family_m (stand_l (region_k))' and 'control_c (sowing date_j)'.

Within each sowing date (*i.e.* 'August 2010', 'September 2010' and 'September 2011'), the model described above was run excluding the terms involving 'sowing date_j'. For each sowing date, least-square means (LSmeans) for each trait (except survival) were obtained for each family using the following model in the *R* environment (Fox and Hong 2009).

$$Y_{imc} = \mu_{\text{trait}} + \text{block}_i + \text{family}_m + \text{control}_c + \varepsilon_{imc} \qquad (2)$$

Because *A. thaliana* is a highly selfing species, LSmeans correspond to genotypic values of families. For each trait, reaction norms were calculated as $\text{LSmean}_{\text{August2010}}$ - $\text{LSmean}_{\text{September2010}}$ (between cohorts) and $\text{LSmean}_{\text{September 2010}}$ - $\text{LSmean}_{\text{September2011}}$ (between years).

Heritability

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

$$Yij = \mu_{trait} + block_i + family_j + \varepsilon_{ij} \quad (3)$$

$$\text{as } H^2_{trait} = V_F / (V_F + V_R/n) \quad (4)$$

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

Following Zhang *et al.* (2012), broad-sense heritabilities for plasticity across germination cohorts or consecutive years ($H^2_{plasticity}$) were estimated using the following model based on variance components estimated by REML:

$$Y_{ijm} = \mu_{trait} + block_i + sowing\ date_j + block_i*sowing\ date_j + family_m + sowing\ date_j* family_m + \varepsilon_{ijm} \quad (5)$$

$$\text{as } H^2_{plasticity} = V_{F*E} / (V_{F*E} + V_F + V_E + V_R/n) \quad (6)$$

where V_{F*E} is the estimated variance component for the 'family-by-sowing-date' interaction effect, V_F is the estimated between family variance component, V_E is the estimated between sowing dates variance component, V_R is the residual variance and 'n' is the mean number of replicates per family.

For each trait, H^2_{trait} and $H^2_{plasticity}$ were estimated in five geographical units, *i.e* France (n = 800 families) and within each region (n = 200 families). Because the term '$control_c$' is a continuous variable, this term was not included in the variance component models (3) and (5), leading to potential downward-biased estimates of H^2_{trait} and $H^2_{plasticity}$. Significance of H^2_{trait}

and $H^2_{plasticity}$ estimates were assessed by testing the significance of the terms 'family$_m$' and 'sowing date$_j$* family$_m$' by fitting models (3) and (5) using the PROC MIXED procedure in SAS 9.3 (REML method), respectively.

Identification of traits under selection — F_{ST} - Q_{ST} comparison

Q_{ST} within each sowing date ($Q_{STtrait}$) and Q_{ST} for plasticity across germination cohorts or years ($Q_{STplasticity}$) were estimated for each trait (except survival) in five geographical units, *i.e.* in France, and within each region. Because the type of selection may depend on the spatial scale considered, e.g., no phenotypic differentiation among regions but high phenotypic differentiation among the 49 stands, Q_{ST} was estimated both among the regions and among stands in France. Because the type of selection may differ among regions, Q_{ST} among stands was also estimated within each region.

Based on variance components estimated by REML (PROC VARCOMP procedure in SAS 9.3, SAS Institute Inc.), $Q_{STtrait}$ was estimated using the following model:

$$Yij = \mu_{trait} + block_i + geographical\ scale_g + family_j\ (geographical\ scale_g) + \varepsilon_{ij} \quad (7)$$

$$\text{as } Q_{STtrait} = V_{GS} / (V_F + V_{GS}) \quad (8)$$

where V_{GS} is the estimated variance component for the 'geographical scale' effect (*i.e.* among regions or among stands in France, and among stands within each region) and V_F is the estimated variance component for the 'family' effect.

Following Banta *et al.* (2007), $Q_{STplasticity}$ was estimated using the following model based on variance components estimated by REML:

$$Y_{ijm} = \mu_{trait} + block_i + sowing\ date_j + block_i*sowing\ date_j + geographical\ scale_g + family_m + sowing\ date_j* geographical\ scale_g + sowing\ date_j* family_m\ (geographical\ scale_g) + \varepsilon_{ijm} \quad (9)$$

$$\text{as } Q_{STplasticity} = V_{GS*SD} / (V_{F*SD} + V_{GS*SD}) \quad (10)$$

where V_{GS*SD} is the estimated variance component for the 'geographical-scale-by-sowing-date' interaction effect and V_{F*SD} is the estimated variance component for the 'family-by-sowing-date' interaction effect.

$Q_{STtrait}$ and $Q_{STplasticity}$ were estimated only for traits with a significant estimate of H^2_{trait} and $H^2_{plasticity}$ at the geographical unit considered, respectively. Significance of $Q_{STtrait}$ and $Q_{STplasticity}$ estimates were assessed by testing the significance of the terms 'geographical $scale_g$' and 'sowing $date_j$ * geographical $scale_g$' by fitting models (7) and (9) using the PROC MIXED (REML method), respectively.

To test whether a trait is under uniform or heterogeneous selection, we compared Q_{ST} estimates to a genome-wide distribution of F_{ST} based on SNP data (Whitlock 2008; Edelaar *et al.* 2011; Lee & Mitchell-Olds 2013; Leinonen *et al.* 2013). Using the *R* package 'adegenet' (Jombart & Ahmed 2011), we first computed the distribution of F_{ST} in five geographical units (*i.e.* in France and within each region) based on the 758 families successfully genotyped for 135 SNPs (Brachi *et al.* 2013b). In order to make $Q_{ST} - F_{ST}$ comparison at the regional and stand scales in France, the distribution of F_{ST} was computed among regions and among the 49 stands in France.

Comparing Q_{ST} and F_{ST} estimates usually requires an estimate of the statistical error around Q_{ST} (Leinonen *et al.* 2013). Because (i) statistical errors around Q_{ST} are usually relatively large (Whitlock 2008) especially for low heritability quantitative traits like

reproductive traits, and (ii) high estimates of genetic differentiation in selfing species increases the likelihood of overlapping between Q_{ST} confidence intervals and genome-wide distribution of F_{ST}, we did not calculate statistical errors around Q_{ST} in this study. Instead, following Lee & Mitchell-Olds (2013), an empirical *P*-value was calculated based on the proportions of SNPs with F_{ST} higher than Q_{ST} when a Q_{ST} estimate was higher than the corresponding mean F_{ST} estimate. When a Q_{ST} estimate was lower than the corresponding mean F_{ST} estimate, an empirical *P*-value was calculated based on the proportions of SNPs with F_{ST} lower than Q_{ST}. We must therefore acknowledge that our results based on point estimates of Q_{ST} should be taken with extreme cautious and complemented by an alternative method, *i.e.* phenotype – ecology relationships.

Identification of traits under selection — phenotype - ecology relationships

The identification of the putative selective agents acting on phenotypic variation was performed using partial least square regressions (PLSR; Geladi & Kowalski 1986; Carrascal *et al.* 2009; Wagmann *et al.* 2010; Brachi *et al.* 2013b). We first ran PLSR between standardized ecological factors and standardized phenotypic medians per stand calculated from LSmeans obtained with equation (2). Because the demographic history of populations may generate spurious relationships between ecological factors and phenotypic variation (Stenoien *et al.* 2002, Roff & Mousseau 2005; Raabova *et al.* 2007; Mullen *et al.* 2009), we calculated LSmeans residuals from a model estimating the upper limit for the amount of phenotypic variation that could be explained by neutral processes. As previously described (Brachi *et al.* 2013b), we performed a principal component analysis (PCA) and included principal components (PCs) with eigenvalues ≥ 1 in the following multiple regression model:

$$Y_i = \mu_{trait} + PC1_i + \ldots + PCN_i + \varepsilon_i \quad (11)$$

where Y_i is a vector of trait LSmeans (or plasticity of traits); μ_{trait} is the mean of traits (or the mean of trait plasticity); $PC1_i + \ldots + PCN_i$ are the genetic principal components included in the model; and ε is the vector of phenotypic residuals of the model. We performed this procedure in five geographical units, *i.e.* France (n = 49 stands) and within each region. For a given combination of 'trait * geographical unit', we approximated the upper limit of the amount of phenotypic variation explained by non-selective processes using the r-square of the model (11).

In both analyses (*i.e.* with and without controlling non-selective processes), the optimal number of components included in the model was determined by leave-one-out cross-validation. Approximate t-tests based on jackknife variance estimates were used to test significance of regression coefficients. PLSR were run using the *R* package 'pls' (Mevik & Wehrens 2007) for each trait and reaction norm with a significant estimate of H^2_{trait} and $H^2_{plasticity}$ at the geographical unit considered.

Results

Climatic variation

Climate varied substantially across the two consecutive years (Fig. 2). In comparison to historical monthly averages from 1981-2009, the two years of our study were unusual. In 2010, the Autumn (August – November/December) was unusually cold and wet while Winter and Spring (December/January – June) were unusually warm and dry. In 2011, the Autumn (September – November) was unusually warm and dry while the Spring (in particular April)

was unusually wet and cold. Plants endured vernalizing temperatures (*i.e.* < 5°C in *A. thaliana*; Wilczek *et al.* 2009) in December 2010 and in February 2012.

Natural variation and norms of reaction at different spatial scales

Genetic variation in the reaction norms of survival across cohorts was only detected at the regional scale, however genetic variation in the reaction norms of survival across years was observed at both the regional and stand scales (Table 1). In particular, plants from Languedoc had one of the highest survival rates in September 2010 but the lowest average survival rate in September 2011 (Fig. 3).

To test whether genetic variation in reproductive traits, seed production ratios and their reaction norms, would depend on the time scale considered, we analyzed data across cohorts within the same year separately from data across years within the same cohort. First, in the 'cohort' treatment, a highly significant 'cohort' effect was detected for FITTOT, as well as most individual reproductive traits (with the exception of fruit length related traits, SILSTEM and SILPB) and seed production ratios (Table 1A, Fig. 3 and Fig. S1, Supporting information). Highly significant natural variation was observed at the three spatial scales tested, in particular at the stand scale (Table 1A, Fig. 3 and Fig. S1, Supporting information). Highly significant genetic variation in the reactions norms across cohorts was detected for most individual reproductive traits and seed production ratios (with the exception RSTEM) but not for FITTOT (Table 1A, Fig.3, Table S1 and Fig. S1, Supporting Information). This latter result suggests that genetic variation in the reaction norms of reproductive strategies did not produce genetic variation in the reaction norms for total seed production.

Second, in the 'year' treatment, a highly significant 'year' effect was detected for most individual reproductive traits and seed production ratios, but not for FITTOT (Table 1B, Fig.3, Table S1 and Fig. S1, Supporting Information). As previously observed in the cohort treatment, highly significant natural variation was detected at the three spatial scales tested, in particular at the stand scale (Table 1B, Fig. 3 and Fig. S1, Supporting information). Highly significant genetic variation in the reaction norms was observed at the regional and stand scales for FITTOT and most individual reproductive traits (Table 1B, Fig. 3 and Fig. S1, Supporting information) and to a lesser extent for seed production ratios (Table 1). Altogether, these results suggest genetic variation in the reaction norms across years for total resource allocation to reproduction (*i.e.* FITTOT) but not for resource allocation among the three types of branches.

Heritabilities

Of the 195 estimates of H^2_{trait} (3 sowing dates * 5 geographical units * 13 traits), 52.3% were significant (Table S2) with a mean estimate equal to 0.36 (SE: ± 0.13). While only one estimate of H^2_{trait} was significant for FITTOT (French scale in September 2010), 8.4 significant estimates of H^2_{trait} were on average detected for each individual reproductive trait and seed production ratio (Fig. 4 and Table S2, Supporting information). This result suggests that contrasting patterns of resource allocation among families can lead to the same total seed production.

Of the 130 estimates of $H^2_{plasticity}$ (2 treatments * 5 geographical units * 13 traits), 37.7% were significant (Table S2) with a mean estimate equal to 0.19 (SE: ± 0.07). Similarly to H^2_{trait}, only one estimate of $H^2_{plasticity}$ was found significant for FITTOT (French scale across

successive years) while four significant estimates of $H^2_{plasticity}$ were detected on average for each individual reproductive trait and seed production ratio (Table S2, Fig. 4). Estimates of $H^2_{plasticity}$ appeared lower than H^2_{trait} for all traits except FITTOT, FITPB and FRUITPB (Fig. 4).

Identification of traits under selection — F_{ST} - Q_{ST} comparison

Significance of Q_{ST} was estimated by comparing Q_{ST} point estimates to either side of genome-wide distributions of F_{ST} (Fig. S2, Supporting information). Of the 102 estimates of $Q_{STtrait}$ calculated for traits with a significant H^2_{trait} estimate, 47.1% were significantly different from the genome-wide distribution of F_{ST} (Table 2A). $Q_{STtrait}$ was significantly correlated with the *P*-values for trait divergence within each sowing date ($Q_{STtrait}$ – *P*-values for phenotypic divergence (equation 7): Pearson's correlation coefficient = 0.589, $P = 7.3 \times 10^{-11}$; Fig. S3, Supporting information). Of the 48 significant $Q_{STtrait}$ estimates, two thirds were detected either at the regional scale or at the stand scale in France (Table 2A). Heterogeneous selection was detected in 22.9% of traits, suggesting a prevalence of uniform selection acting on traits within each sowing date. In some cases, selection appeared to act differently across cohorts or years. For example, FRUITBB showed heterogeneous selection in August 2010 but uniform selection in September 2010 at the stand scale in France, and FRUITSTEM showed heterogeneous selection in September 2011 but uniform selection in September 2010 at the regional scale. In addition, closely related traits appear to undergo different types of selection; for example FRUITPB and SILPB showed heterogeneous and uniform selection, respectively, at the regional scale in September 2011 (Table 2A).

Of the 49 estimates of $Q_{STplasticity}$ calculated upon finding a significant estimate of $H^2_{plasticity}$, 59.2% were significantly different from the genome-wide distribution of F_{ST} (Table 2B). $Q_{STplasticity}$ was significantly correlated with the *P*-values for divergence in trait plasticity ($Q_{STplasticity}$ – *P*-values for divergence of phenotypic plasticity (equation 9): correlation coefficient of Pearson = 0.659, $P = 2.5 \times 10^{-7}$; Fig. S3, Supporting information). 63% of plasticity traits across years (mainly FITTOT and individual reproductive traits measured on the main stem and the primary branches) were found to be under heterogeneous selection, whereas most trait plasticities across cohorts within the same year were found to be under uniform selection (Table 2B). All plasticities of seed production ratios across years were found to be under uniform selection.

Interestingly, evidence of uniform or heterogeneous selection on trait plasticity did not necessarily translate into significant differences between Q_{ST} and F_{ST} for the traits measured within each sowing date. For example, this is the case for SILSTEM across cohorts in North and SILPB across years at the stand scale in France (Table 2).

Identification of traits under selection — phenotype - ecology relationships

The mean amount of phenotypic variation explained by genetic principal components was 26% (SE: ±12.3%) for reproductive traits and seed production ratios and 17% (SE: ±8.5%) for plasticity traits (Table S3, Supporting information). To test for traits under selection, we considered only phenotype-ecology relationships that were significant after controlling for non-selective processes. Although this approach may exclude true phenotype – ecology relationships that overlap with neutral genetic diversity, *i.e.* false negatives (Brachi et al. 2013b), it is conservative.

Models aimed at identifying PLSR components to explain phenotypic variation converged for 40.2% and 30.6% of all traits with significant H^2_{trait} (n = 102) and $H^2_{plasticity}$ (n = 49) estimates, respectively (Fig. 4 and 5). The percentage of phenotypic variation explained by ecological variation was higher within-regions than at the scale of France for both phenotypic traits (France: mean = 19.8% ± 9.4%; within-region: mean = 72.8% ± 30.6%; Fig. 4) and plasticity (France: mean = 11.9% (only one estimate); within-region: mean = 78.14% ± 24.5%; Fig. 5).

Five features characterize the relationships between phenotypic and ecological variation. First, significant phenotype – ecology relationships were specific to a given region. Second, even though the three types of ecological variables (climate, soil and competition) were associated with phenotypic traits (Fig. 4), only climate and edaphic variables were associated with phenotypic plasticity (Fig. 5). Third, some phenotypic traits were simultaneously associated with all three types of ecological variables, suggesting that a complex combination of selective agents may act on reproductive traits and seed production ratios (Fig. 4). Fourth, significant phenotype – ecology relationships depended on the sowing date considered. For example, the number of fruits on the main stem (FRUISTEM) was significantly associated with climatic and edaphic variables at the scale of France but only in September 2010 and September 2011 (Fig. 4). Fifth, as previously observed with $Q_{ST} - F_{ST}$ comparisons (see above), trait plasticity can be significantly associated with ecological variation even when there is no effect of that ecological variation within each sowing date (such as SILSTEM across cohorts in North and SILPB across years at the stand scale in France; Fig. 4 and 5).

The results of phenotype – ecology relationships that do not control for non-selective processes are provided in Fig. S4 and S5 (Supporting information).

Discussion

We described natural variation of reproductive strategies by measuring total seed production, several individual reproductive traits, seed production ratios and their associated reactions norms, in *A. thaliana* across space and time in ecologically relevant conditions. As previously observed for French stands of *A. thaliana* scored for phenological traits in greenhouse conditions (Le Corre 2005; Brachi *et al.* 2013b), extensive natural genetic variation was observed across spatial scales. The presence of within-stand genetic variation also suggests that natural sites of *A. thaliana* tend not to be represented by single accessions.

Contrasted reproductive strategies across germination cohorts and years

Under our ecologically relevant, experimental conditions, genetic variation in resource allocation within a sowing date can lead to the same total seed production among families in natural stands of *A. thaliana* (Table S1, Supporting Information). This is consistent with patterns observed contrasting Col-0 with a mutant in this background that is impaired in auxin response; these lines produce indistinguishable numbers of seeds, but with contrasting patterns of resource allocation (Roux & Reboud 2005; Paris *et al.* 2008).

We found that families maintained similar rankings across cohorts within the same year when indexed by total seed production, but not when indexed by individual reproductive traits or by seed production ratios. Since total seed production should correlate with resources available for reproduction, this indicates a lack of genetic variation in reaction norms across cohorts for resource acquisition but the presence of genetic variation in these reaction norms for resource allocation. There are two possible explanations for this pattern. First, the sensing of environmental cues differing between the two germination cohorts may have directly

triggered a genotypic-dependent allocation of resources. Alternatively, genetic variation in the reaction norms for resource allocation may be an indirect effect of genetic variation in the reaction norms for flowering time. In support of this second hypothesis, note that most plants flowered after winter when sown in late September 2010 while many plants started to flower before winter (*i.e.* in November) when sown in late August 2010 (our unpublished data). The early flowering observed in the early-autumn cohort was associated with apical meristem damage due to frost during winter, potentially leading to reallocation of resources from the main stem to primary and basal branches (Fig. 3). A similar plastic increase in basal branch number was observed in response to apical meristem damage inflicted by rabbits in a field experiment with recombinant inbred lines of *A. thaliana* (Weinig *et al.* 2003).

Plasticities across years revealed a different pattern. Here, we found changes in the ranking of families when indexed for total seed production, individual reproductive traits, and to a much lesser extent seed production ratios. This suggests extensive variation in the reaction norm of reproductive outputs but little genetic variation in the reaction norms for resources allocation among reproductive traits. Because edaphic conditions were similar across years, the genotypic-dependent differential acquisition of resources (translated into seed production) across years may have been triggered by (i) a change of a climate cue, (ii) a new interaction between climate and soil, and/or (iii) a biotic factor resulting from climate change between the two years (such as different pathogens attacking in September 2010 and September 2011).

Still, further field experiments are clearly needed to establish a potential link between the plasticity of resource acquisition/ allocation trade-off and the time scale considered (*i.e.* across cohorts, seasons, years…).

Identifying reproductive strategies under selection

Our results suggest that selection acts on all reproductive traits, seed production ratios and their reaction norms (except RSTEM) in at least one 'sowing date * geographical unit' combination, although the form of the test varies. This observation affirmed the need to not only focus on total seed production when studying the genetics of local adaptation in *A. thaliana*, but also to consider individual reproductive traits and seed production ratios across space and time. Furthermore, even when consistent results were obtained for the two indirect methods that we used, we must acknowledge that (i) the identification of phenotypic traits under selection is specific to our ecological conditions and (ii) other sources of evidence are needed to confirm that natural selection is acting on the traits identified in this study (Banta *et al.* 2007).

The main limitations of $Q_{ST} - F_{ST}$ and phenotype – ecology approaches have been fully described elsewhere (Le Corre & Kremer 2003; Porcher *et al.* 2004; Leinonen *et al.* 2013; Brachi *et al.* 2013b). Some important limitations include effects of dominance and epistasis that may lead to a deviation of Q_{ST} from the neutral expectation (Porcher *et al.* 2004) or uncharacterized ecological factors that could interfere with the ability to find correlates between ecology and plant phenotypes (Brachi *et al.* 2013b). Other explanations may nonetheless be advanced to explain why we failed to detect selection acting on reproductive traits and seed production ratios in a given 'sowing date * geographical unit' combination: (i) natural selection may mainly act on seed quality (not measured in this study) rather than on seed quantity, (ii) an apparent neutral trait may be integrative of two traits, each being under contrasted types of selection (*i.e.* uniform *vs.* heterogeneous), (iii) the geographical scale considered in our analyses may combine smaller geographic regions, each with contrasting types of selection.

Patterns of selection acting on reproductive strategies

While Q_{ST} estimates were found to generally exceed their corresponding F_{ST} in a meta-analysis (de Kort *et al.* 2012), we found a prevalence of uniform selection in this study (69.2% of cases where Q_{ST} was significantly different from F_{ST}). In *A. thaliana*, uniform selection has already been detected for hypocotyl elongation in Norwegian natural populations (Stenoien *et al.* 2002), seed dormancy in natural populations of Norway and Central Asia (Kronholm *et al.* 2012) and the mean number of fruits per plant in Spanish natural populations (Méndez-Vigo *et al.* 2013). This discrepancy with the meta-analysis study likely originates from the predominantly selfing breeding system of *A. thaliana* (Platt *et al.* 2010), leading to higher estimates of genetic differentiation than in outcrossing species (Le Corre & Kremer 2003) and thus potentially facilitating the identification of traits under uniform selection.

As previously observed for hypocotyl elongation in Norwegian natural populations of *A. thaliana* (Stenoien *et al.* 2002), many phenotypic traits found under uniform selection in the Q_{ST} - F_{ST} comparison were also associated with ecological variation. This suggests that reduced phenotypic differentiation relative to that expected by non-selective processes is not incompatible with local adaptation. For example, in August 2010, fruit length on the primary branches (SILPB) was under uniform selection in the North but was nonetheless associated with annual mean temperature, total nitrogen and competition with grasses. We suggest that the combination of Q_{ST} – F_{ST} and phenotype – ecology approaches may represent an opportunity to identify both phenotypic traits having close local optima in a given geographical region and the putative selective agents driven those local optima.

Geographical scale and environment-dependence of reproductive strategies under selection

In agreement with the adaptive divergence on quantitative traits in wild barley (Volis *et al.* 2005), selection acting on some traits was detected at the regional scale but not at the scale of stands within regions. In addition, traits under selection clearly depend on the geographic region considered, *i.e.* the genetic pool studied. This latter result is similar to what have been observed in *A. thaliana* for seed dormancy (Kronholm *et al.* 2012) and phenological traits scored on this same panel of 800 families under greenhouse conditions (Brachi *et al.* 2013b).

Few studies have explored whether our ability to detect selection acting on quantitative traits depends on the environmental conditions (Cano *et al.* 2004; Lind *et al.* 2010; Scheepens *et al.* 2010; Shama *et al.* 2011; Brachi *et al.* 2012; Rogell *et al.* 2012; Fournier-Level *et al.* 2013), especially over time (Siepielski *et al.* 2009). Our study indicates that the environmental conditions used for phenotyping quantitative traits influences whether selection is detected. Furthermore, in some cases multiple environments were required to detect selection since it was manifest in terms of plasticity but not within single assays. This result is in agreement with a Genome-Wide Association mapping study showing that the genetic basis of flowering time plasticity across two successive years was almost completely independent of the genetic basis of flowering time scored within each year (Brachi *et al.* 2013a). As previously argued in other species (Schlichting 1965), plasticity is a phenotypic trait by itself that may evolve independently of its expression in either environment. It remains to be determined whether the adaptive phenotypic plasticity that we detected can be related to fine environmental grain at both spatial and temporal scales, as for the common frog *Rana temporaria* in Sweden (Lind *et al.* 2010).

In conclusion, in the era of ecological genomics aimed at identifying the genetic basis of adaptive phenotypic natural variation in an ecological context (Bergelson & Roux 2010), it appears crucial to both consider a set of populations matching the geographical scale of adaptive divergence of reproductive strategies and to integrate genotype-by-environment interactions in the study of adaptive potential in *A. thaliana*.

Acknowledgments

Special thanks are given to Nathalie Faure, Stella Huynh and Benoit Bucher for their assistance during the field experiment. This study was supported by a PhD fellowship from the University of Lille 1 and a mobility grant from the Collège Doctoral Lille Nord de France to R.V.

References

Agren J, Schemske DW (2012) Reciprocal transplants demonstrate strong adaptive differentiation of the model organism Arabidopsis thaliana in its native range. *New Phytologist*, 194, 1112-1122.

Banta JA, Dole J, Cruzan MB, Pigliucci M (2007) Evidence of local adaptation to coarse-grained environmental variation in *Arabidopsis thaliana*. *Evolution*, 61, 2419-2432.

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

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

Brachi B, Faure N, Bergelson J, Cuguen J, Roux F (2013a) Genome-wide association mapping of flowering time in *Arabidopsis thaliana* in nature: genetics for underlying components and reaction norms across two successive years. *Acta Gallica Botanica: Botany Letters* (DOI: 10.1080/12538078.2013.807302).

Brachi B, Faure N, Horton M *et al.* (2010) Linkage and association mapping of *Arabidopsis thaliana* flowering time in nature. *PLoS genetics*, 6, e1000940.

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

Cano JM, Laurila A, Palo J, Merilä J (2004) Population differentiation in G matrix structure due to natural selection in *Rana temporaria*. *Evolution*, 58, 2013-2020.

Carrascal LM, Galvan I, Gordo O (2009) Partial least squares regressions as an alternative to current regression methods used in ecology. *Oikos*, 118, 618–690.

Conner JK (2010) Natural Selection in Plants 151 Years after The Origin: Introduction. *International Journal of Plant Sciences*, 171, 927–929.

Donohue K (2002) Germination timing influences natural selection on life-history characters in *Arabidopsis thaliana*. *Ecology*, 83, 1006–1016.

Donohue K, Dorn LA, Griffith C *et al.* (2005) The evolutionary ecology of seed germination of *Arabidopsis thaliana*: variable natural selection on germination timing. *Evolution*, 59, 758–770.

Edelaar P, Burraco P, Gomez-Mestre I (2011) Comparisons between Q_{ST} and F_{ST} — how wrong have we been? *Molecular Ecology*, 20, 4830-4839.

Fournier-Level A, Korte A, Cooper MD *et al.* (2011) A Map of Local Adaptation in *Arabidopsis thaliana*. *Science*, 334, 86–89.

Fournier-Level A, Wilczek AM, Cooper MD *et al.* (2013) Paths to selection on life history loci in different natural environments across the native range of *Arabidopsis thaliana*. *Molecular Ecology*, 22, 3552-3566.

Frenkel M, Jänkänpää, Moen J, Jansson S (2008) An illustrated gardener's guide to transgenic *Arabidopsis* field experiments. *New Phytologist*, 180, 545-555.

Gaut B (2012) *Arabidopsis thaliana* as a model for the genetics of local adaptation. *Nature Genetics*, 44, 115–121.

Geladi P, Kowalski BR (1986) Partial least-squares regression: a tutorial. *Analytica Chimica Acta*, 185.

Hancock AM, Brachi B, Faure N *et al.* (2011) Adaptation to Climate Across the *Arabidopsis thaliana* Genome. *Science*, 334, 83–86.

Horton MW, Hancock AM, Huang YS *et al.* (2012) Genomw-wide patterns of genetic variation in worldwide *Arabidopsis thaliana* accessions from the RegMap panel. *Nature*

Genetics, 44, 212-216.

Hübner S, Bdolach E, Eing-Gedy S *et al.* (2012) Phenotypic landscapes: phenological patterns in wild and cultivated barley. *Journal of Evolutionary Biology*, 26, 163-174.

de Kort H, Vandepitte K, Honnay O (2013) A meta-analysis of the effects of plant traits and geographical scale on the magnitude of adaptive differentiation as measured by the difference between Q_{ST} and F_{ST}. *Evolutionary ecology* (DOI 10.1007/s10682-012-9624-9)

Jombart, T, Ahmed, I (2011) *adegenet 1.3-1*: new tools of genome-wide SNP data. *Bioinformatics*, 27, 3070-3071.

Juenger, T, Bergelson J (2000) The evolution of compensation to herbivory in Scarlet gilia, *Ipomopsis aggregate*: herbivore-imposed natural selection and the quantitative genetics of tolerance. *Evolution*, 54, 764-777.

Kronholm I, Pico FX, Alonso-Blanco C, Goudet J, de Meaux J (2012) Genetic basis of adaptation in *Arabidopsis thaliana* local adaptation at the seed dormancy QTL *DOG1*. *Evolution*, 66, 2287-2302.

Lande R, Arnold SJ (1983) The measurement of selection on correlated characters. *Evolution*, 37, 1210–1226.

Latta RG (1998) Differentiation of allelic frequencies at quantitative trait loci affecting locally adaptive traits. *The American Naturalist*, 151, 283–292.

Le Corre V (2005) Variation at two flowering time genes within and among populations of *Arabidopsis thaliana*: comparison with markers and traits. *Molecular Ecology*, 14, 4181–4192.

Le Corre V, Kremer A (2003) Genetic variability at neutral markers, quantitative trait loci and trait in a subdivided population under selection. *Genetics*, 164, 1205-1219.

Lee CR, Mitchell-Olds T (2013) Complex trait divergence contributes to environmental niche differentiation in ecological speciation of *Boechera stricta*. *Molecular Ecology*, 22, 2204-2217.

Leinonen T, McCairns RJS, O'Hara RN, Merilä J (2013) Q_{ST}-F_{ST} comparisons: evolutionary and ecological insights from genomic heterogeneity. *Nature Reviews Genetics*, 14, 179-190.

Lind MI, Ingvarsson PK, Johansson H., Hall D, Johansson F (2010) Gene flow and selection on phenotypic plasticity in an island system of *Rana temporaria*. *Evolution*, 65, 684-697.

Méndez-Vigo B, Gomaa NH, Alonso-Blanco C, Pico FX (2013) Among- and within-population variation in flowering time of Iberian Arabidopsis thaliana estimated in field and glasshouse conditions. *New Phytologist*, 197, 1332-1343.

Mevik BH, Wehrens R (2007) The pls package: Principal component and partial least squares regression in R. *Journal of Statistical Software*, 18, 1–25.

Mitchell-Olds T, Bergelson J (1990) Statistical genetics in an annual plant, *Impatiens capensis*. II. Natural selection. *Genetics*, 124, 417-421.

Mullen LM, Vignieri SN, Gore JA, Hoekstra HE (2009) Adaptive basis of geographic variation: genetic, phenotypic and environmental differences among beach mouse populations. *Proceedings of the Royal Society of London, Series B: Biological Sciences*, 276, 3809-3818.

Paris M, Roux F, Bérard A, Reboud X (2008) The effects of the genetic background on herbicide resistance fitness cost and its associated dominance in *Arabidopsis thaliana*. *Heredity*, 101, 499-506.

Platt A, Horton M, Huang YS *et al.* (2010) The scale of population structure in *Arabidopsis thaliana*. *PLoS genetics*, 6, e10000843.

Porcher E, Giraud T, Goldringer I, Lavigne C (2004) Experimental demonstration of a causal relationship between heterogeneity of selection and genetic differentiation in quantitative traits. *Evolution*, 58, 1434-1445.

R Development Core Team (2010). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/

Raabova J, Münzbergova Z, Fischer M (2007) Ecological rather than geographic or genetic distance affects local adaptation of the rare perennial herb, *Aster amellus*. *Biological Conservation*, 139, 348-357

Rausher MD (1992) The measurement of selection on quantitative traits: biases due to environmental covariances between traits and fitness. *Evolution*, 46, 616-626.

Roff DA, Mousseau T (2005) The evolution of the phenotypic covariance matrix: evidence for selection and drift in *Melanoplus*. *Journal of Evolutionary Biology*, 18, 1104-1114.

Rogell B, Dannewitz J, Palm S *et al.* (2012) Strong divergence in trait means but not in plasticity across hatchery and wild populations of sea-run brown trout *Salmo trutta*. *Molecular Ecology*, 21, 2963-2976.

Roux F, Gasquez J, Reboud X (2004) The dominance of the herbicide resistance cost in several *Arabidopsis thaliana* mutant lines. *Genetics*, 166, 449-460.

Roux F, Reboud X (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, 101-110.

Scheepens JF, Frei ES, Stöcklin J (2010) Genotypic and environmental variation in specific leaf area in a widespread Alpine plant after transplantation to different altitudes. *Oecologia*, 164, 141-150.

Schlichting CD (1965) The evolution of phenotypic plasticity in plants. *Annual Review of Ecology and Systematics*, 17, 667-693.

Shama LNS, Campero-Paz M, Wegner KM, de Block M, Stoks R (2011) Latitudinal and voltinism compensation shape thermal reaction norms for growth rate. *Molecular Ecology*, 20, 2929-2941.

Shemesh H, Zaitchik B, Acuña T, Novoplansky A (2012) Architectural plasticity in a Mediterranean winter annual. *Plant Signaling & Behavior,* 7, 492-501.

Siepielski AM, DiBattista JD, Carlson SM (2009) It's about time: the temporal dynamics of phenotypic selection in the wild. *Ecology Letters*, 12, 1261–1337.

Stearns SC (1992) *The evolution of life histories*. Oxford University Press, Oxford.

Stenoien HK, Fenster CB, Tonteri A, Savolainen O (2005) Genetic variability in natural populations of *Arabidopsis thaliana* in northern Europe. *Molecular Ecology*, 14, 137-148.

Volis S, Yakubov B, Shulgina I, Ward D, Mendlinger S (2005) Distinguishing adaptive from nonadaptive genetic differentiation: comparison of Q_{ST} and F_{ST} at two spatial scales. *Heredity*, 95, 466-475.

Wagmann K, Hautekèete NC, Piquot Y *et al.* (2012) Seed dormancy distribution: explanatory ecological factors. *Annals of Botany*, 110, 1205-1219.

Weiner J (2004) Allocation, plasticity and allometry in plants. *Perspectives in Plant Ecology, Evolution and Systematics*, 6, 207-215.

Weinig C, Stinchcombe JR, Schmitt J (2003) Evolutionary genetics of resistance and tolerance to natural herbivory in *Arabidopsis thaliana*. *Evolution*, 57, 1270-1280.

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

Whitlock MC (2008) Evolutionary inference from Q_{ST}. *Molecular Ecology*, 17, 1885-1896.

Wilczek AM, Roe JL, Knapp MC *et al.* (2009) Effects of genetic perturbation on seasonal life history plasticity. *Science*, 323, 930-934.

Zhang YY, Fischer M, Colot V, Bossdorf O (2012) Epigenetic variation creates potential for evolution of plant phenotypic plasticity. *New Phytologist*, 197, 314-322.

Data Accessibility

Phenotypic data are available in the Dryad database: XXX

Figure Legends

Fig. 1 Reproductive traits measured in thus study. SILSTEM: mean silique length on the main stem, SILPB: mean silique length on the primary branches, SILBB: mean silique length on the basal branches, FRUITSTEM: number of fruits on the main stem, FRUITPB: number of fruits on the primary branches, FRUITBB: fruit number on the basal branches, FITSTEM: total silique length on the main stem, FITPB: total silique length on the primary branches, FITBB: total silique length on the basal branches, FITTOT: total silique length.

Fig. 2 Monthly precipitations and mean temperature during the growing period in the field experiment. A. Monthly precipitations (in mm). B. Monthly mean temperature (°C). White bars: 'August 2010' and 'September 2010' sowing dates, grey bars: 'September 2011' sowing date. Dotted lines stand for monthly precipitations and mean temperature values, based on averages calculated for the 1981-2009 period at the weather local station Lille-Lesquin located 6km from the common garden at the University of Lille 1.

Fig. 3 Natural variation of six phenotypic traits and their reaction norms at the scale of France and within each region. Open diamond: Brittany, open triangle: Burgundy, open square: Languedoc, open circle: North. At the scale of France, each value corresponds to the rate of survival within a region and the median of the LSmeans of 200 families per region for SURVIVAL and the five other traits, respectively. Within each region, each value corresponds to the rate of survival within stand and the median of the LS means of families per stand for SURVIVAL and the five other traits, respectively.

Fig. 4 Box-and-whisker plots of broad-sense heritabilities H^2 for 12 reproductive traits and seed production ratios. White box-and-whisker plots represent broad-sense heritabilities calculated for each 'sowing date * geographical unit' combination (n = 15 = 3 sowing dates * 5 geographical units (France + four regions)). Grey box-and-whisker plots represent broad-sense heritabilities of plasticity calculated for each 'treatment * geographical unit' combination (n = 10 = 2 treatments (cohort and year effects) * 5 geographical units (France + four regions)). Values above box-and-whisker plots indicate the number of significant broad-sense heritabilities (see Table S2).

Fig. 5 Identification of the putative selective agents acting on reproductive traits and seed production ratios in *A. thaliana* in five geographical units, after control for neutral genetic variation. Each line corresponds to the results of Partial Least Square Regression (PLSR) for a specific 'geographical unit (France and four regions) * sowing date ('August 2010', 'September 2010' and 'September 2011') * trait' combination. The 42 ecological variables characterizing the 49 stands (climate, soil and competition (Comp.)) correspond to the first 42 columns. For each column and line, the colored squares indicate significant regression coefficients (p-value < 0.05) and the colors represent the strength of the regression coefficient estimated between trait variation and ecological variation. The number of PLSR components retained after cross-validation ('axis'), the percentage of ecological variation explained by PLSR components ('Var X') and the percentage of trait variation explained by PLSR components ('Var Y') corresponds to the 43rd, 44th and 45th columns, respectively. PLSR were performed on standardized data, allowing comparison of regression coefficients among geographical units, sowing dates and traits. PLSR were run only for traits with significant broad-sense heritability. Traits with non-significant broad-sense heritability are marked by a

cross in the 'axis', 'Var X' and 'Var Y' columns. Empty cells in the 'axis', 'Var X' and 'Var Y' columns correspond to traits with significant broad-sense heritability but for which no convergence of PLSR models was reached.

Climatic variables: **Alt**, Altitude (m); **RH Wint**, **RH Spring**, **RH Summ** and **RH Fall** stand for relative humidity (%) in Winter, Spring, Summer and Fall, respectively; **Aridity**, Aridity; **Bio 1**, annual mean temperature (°C x 10); **Bio 2**, mean diurnal range (mean of monthly temperature range (max temp - min temp)); **Bio 3**, isothermality (Bio 2/Bio 7) (* 100); **Bio 4**, temperature seasonality (standard deviation *100); **Bio 5**, max temperature of warmest month (°C x 10); **Bio 6**, min temperature of coldest month (°C x 10); **Bio 7**, temperature annual range (Bio 5 - Bio 6); **Bio 8**, mean temperature of wettest quarter (°C x 10); **Bio 9**, mean temperature of driest quarter (°C x 10); **Bio 10**, mean temperature of warmest quarter (°C x 10); **Bio 11**, mean temperature of coldest quarter (°C x 10); **Bio 12**, annual precipitation (mm); **Bio 13**, precipitation of wettest month (mm); **Bio 14**, precipitation of driest month (mm); **Bio 15**, precipitation seasonality (coefficient of variation); **Bio 16**, precipitation of wettest quarter (mm); **Bio 17**, precipitation of driest quarter (mm); **Bio 18**, precipitation of warmest quarter (mm); **Bio 19**, precipitation of coldest quarter (mm).

Edaphic variables: OC, organic carbon ($g.kg^{-1}$); **N**, total nitrogen ($g.kg^{-1}$); **C/N**, carbon/nitrogen ratio; **SOM**, soil organic matter ($g.kg^{-1}$); **P2O5**, phosphorus (P_2O_5) ($g.kg^{-1}$); **Ca**, exchangeable calcium ($cmol+.kg^{-1}$); **Mg**, exchangeable magnesium ($cmol+.kg^{-1}$); **Na**, exchangeable sodium ($cmol+.kg^{-1}$); **K**, exchangeable potassium ($cmol+.kg^{-1}$); **Fe**, exchangeable iron ($cmol+.kg^{-1}$); **Mn**, exchangeable manganese ($cmol+.kg^{-1}$); **Al**, exchangeable aluminium ($cmol+.kg^{-1}$); **WHC**, soil water holding capacity ($ml.g^{-1}$); **pH**.

Competition variables: Herb, interspecific competition with herbs which are not grasses; **Grass**, interspecific competition with grasses; **Thal**, intraspecific competition.

Fig. 6 Identification of the putative selective agents acting on plasticity of reproductive traits and seed production ratios in *A. thaliana* in five geographical units, after control for neutral genetic variation. Each line corresponds to the results of Partial Least Square Regression (PLSR) for a specific 'geographical unit (France and four regions) * treatment ('cohort treatment' and 'year treatment') * trait' combination. The 42 ecological variables characterizing the 49 stands (climate, soil and competition (Comp.)) correspond to the first 42 columns. For each column and line, the colored squares indicate significant regression coefficients (p-value < 0.05) and the colors represent the strength of the regression coefficient estimated between variation of trait plasticity and ecological variation. The number of PLSR components retained after cross-validation ('axis'), the percentage of ecological variation explained by PLSR components ('Var X') and the percentage of variation of trait plasticity explained by PLSR components ('Var Y') corresponds to the 43rd, 44th and 45th columns, respectively. PLSR were performed on standardized data, allowing comparison of regression coefficients among geographical units, sowing date treatments and trait plasticities. PLSR were run only for plasticity of traits with significant broad-sense heritability. Plasticity of traits with non-significant broad-sense heritability are marked by a cross in the 'axis', 'Var X' and 'Var Y' columns. Empty cells in the 'axis', 'Var X' and 'Var Y' columns correspond to plasticity of traits with significant broad-sense heritability but for which no convergence of PLSR models was reached. See legend of Fig. 5 for a description of the 42 ecological variables.

Fig. 1

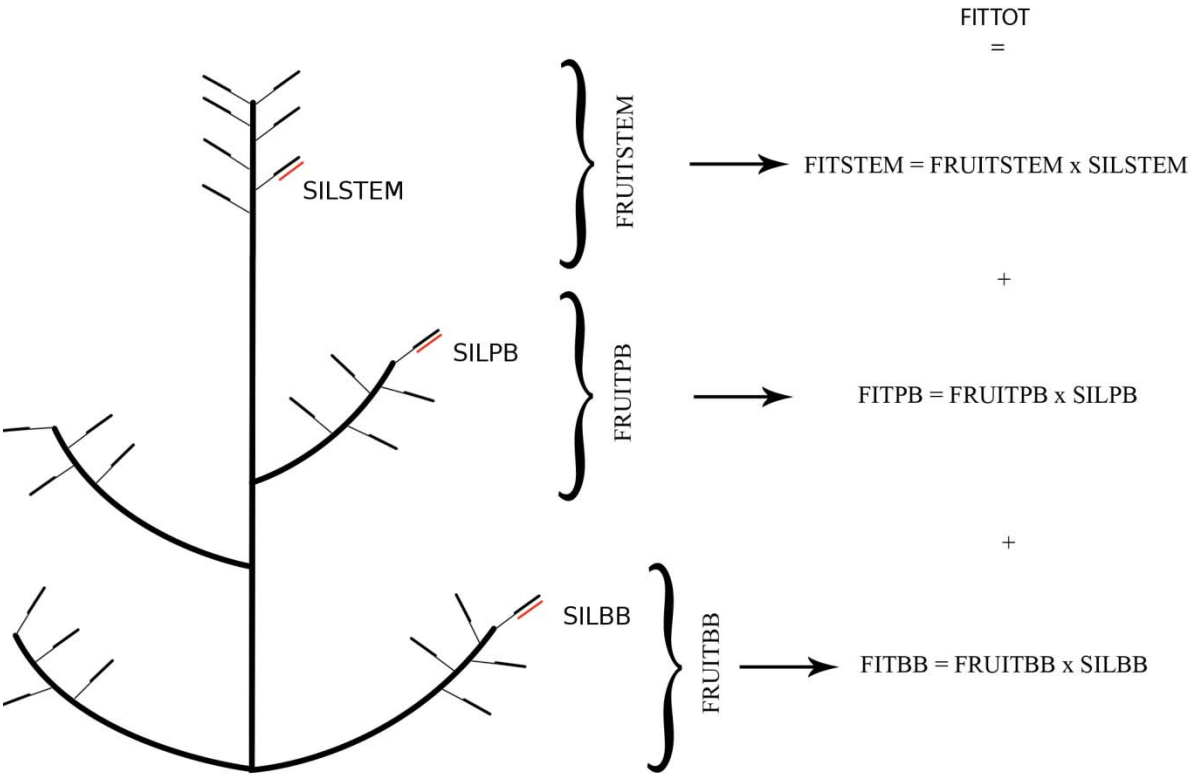


Fig. 2

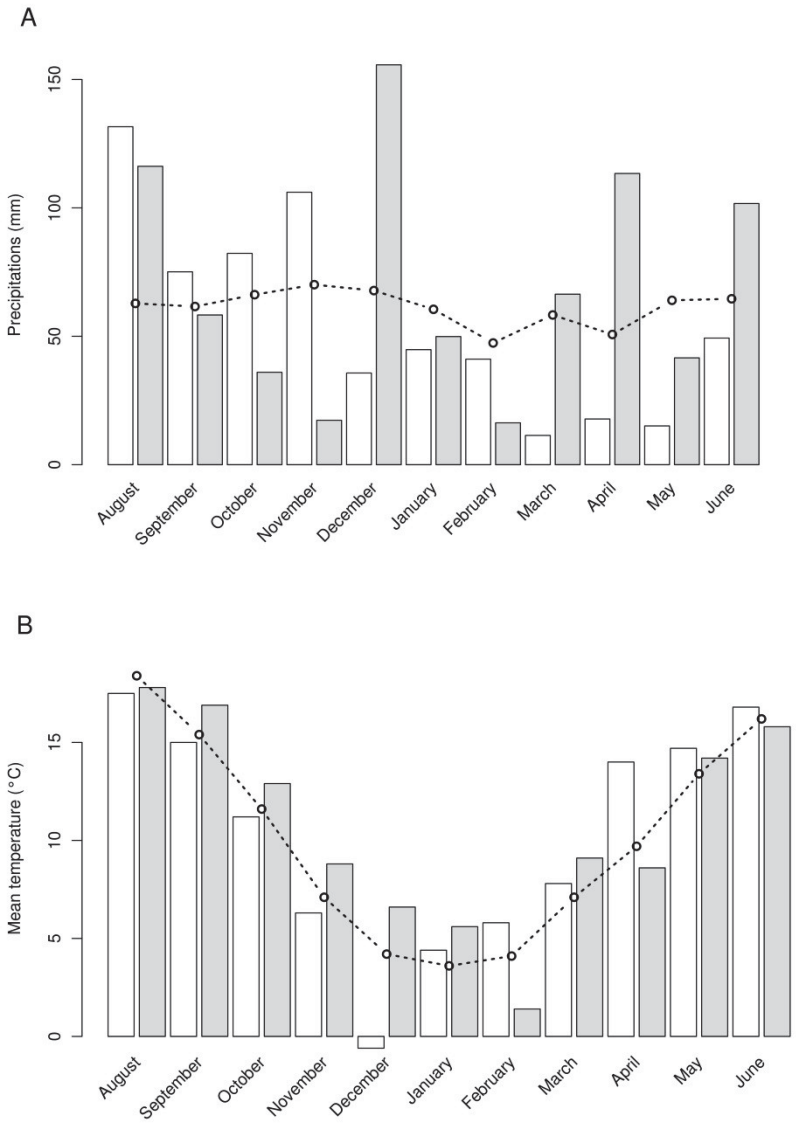


Fig. 3

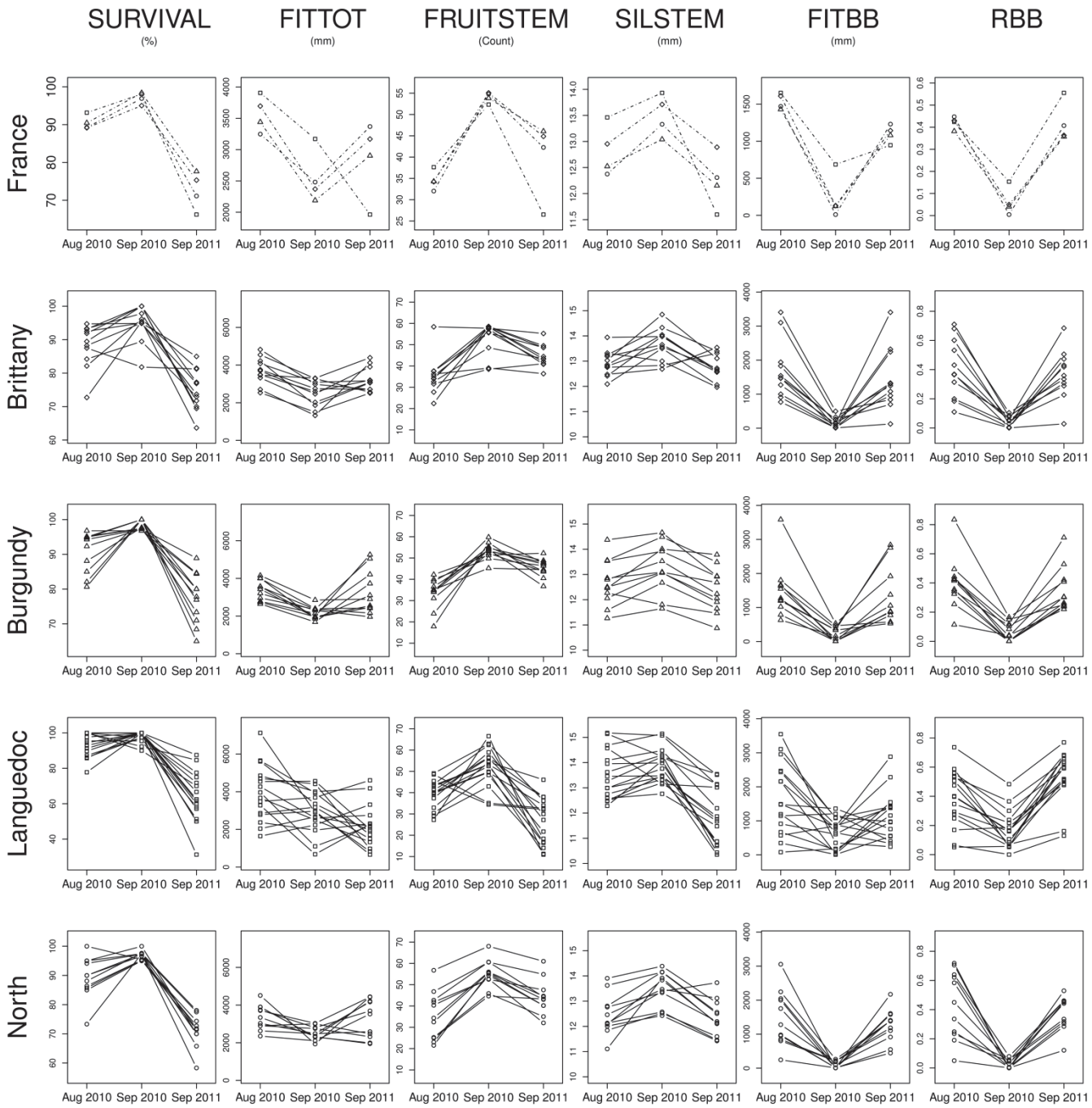


Fig. 4

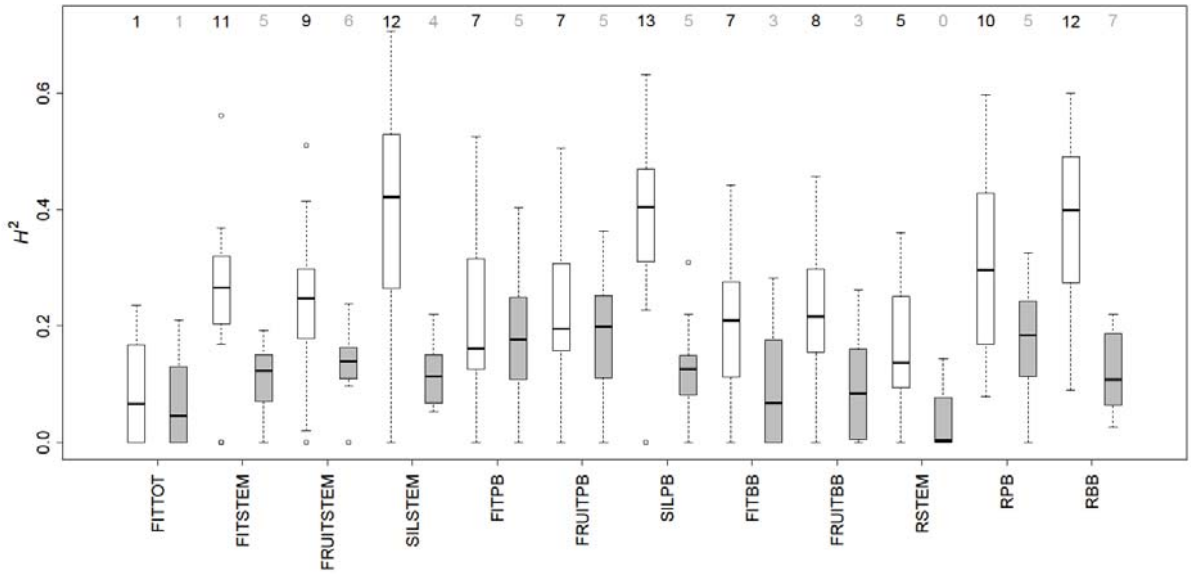


Fig. 5

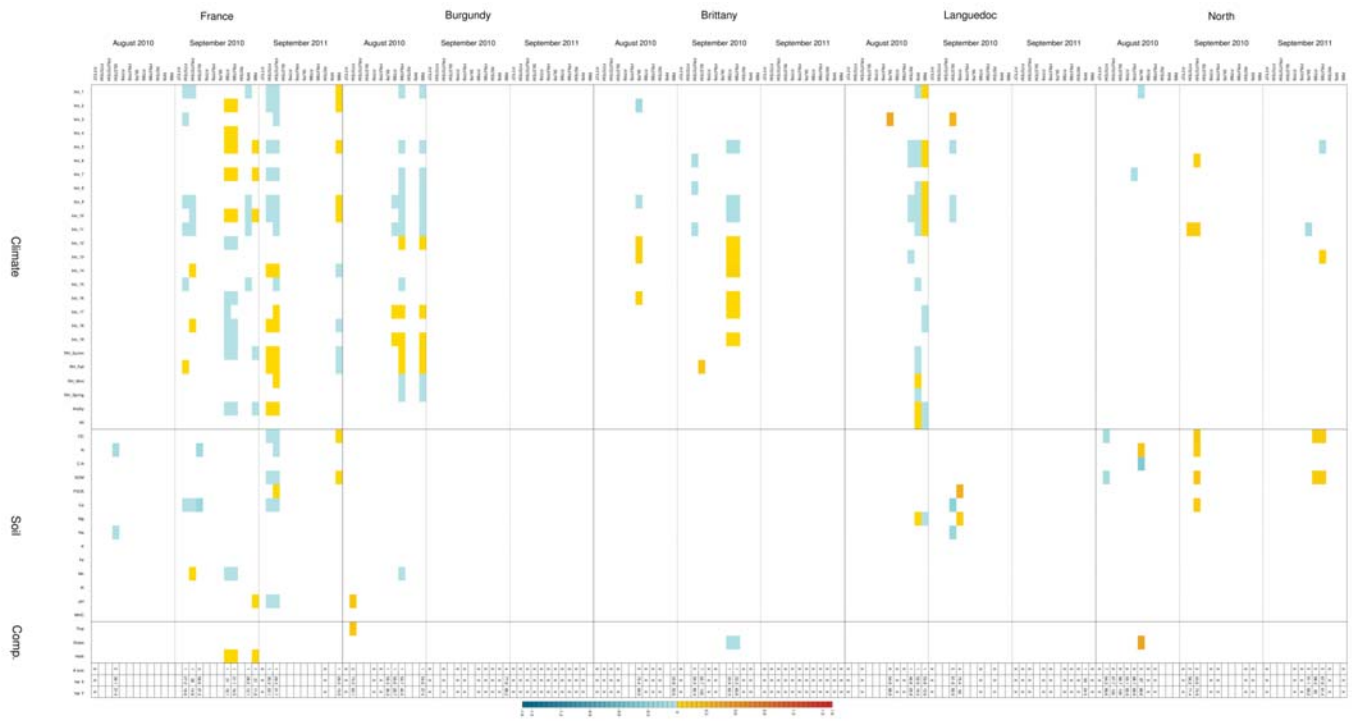


Fig. 6

Table Legends

Table 1 Natural variation of survival, reproductive traits and seed production ratios at three spatial scales, *i.e.* region, stand and family. A. Germination cohort effect within the same year ('August 2010' *vs.* 'September 2010'). B. Year effect within the same germination cohort ('September 2010' *vs.* 'September 2011'). Model random terms were tested with likelihood ratio tests (LRT) of models with and without these effects. Random effects are in italic. Because the variance associated with the model terms 'block*cohort' or 'block*year' are the correct error terms for testing the 'block' and 'cohort' or 'block' and 'year' effects, *F* value and *P*-value were not reported for the 'block*cohort' and 'block*year' terms in Tables 1A and 1B, respectively. *$0.05 > P > 0.01$, **$0.01 > P > 0.001$, *** $P < 0.001$. Bold stars indicate significant effect after Bonferroni correction. NE: not estimated.

Table 2 Trait divergence among regions and stands estimated by Q_{ST}. Q_{ST} values are only displayed for traits with significant broad-sense heritabilities (see Table S2). Boldface denotes significant Q_{ST}-F_{ST} differentiation at the 5% significance level. 'U' and 'H' stand for uniform and heterogeneous selection, respectively.

Table 1

A. Cohort treatment

Traits	survival		FITTOT		FITSTEM		FRUITSTEM		SILSTEM		FITPB		FRUITPB		SILPB		FITBB		FRUITBB		RSTEM		RPB		RBB	
Model terms	*F* or *LRT*	*P*	*F* or *LRT*	*P*	*F* or *LRT*	*P*	*F* or *LRT*	*P*	*F* or *LRT*	*P*	*F* or *LRT*	*P*	*F* or *LRT*	*P*	*F* or *LRT*	*P*	*F* or *LRT*	*P*	*F* or *LRT*	*P*	*F* or *LRT*	*P*	*F* or *LRT*	*P*	*F* or *LRT*	*P*
Block	0.03	ns	2.51	ns	6.42	*	2.74	ns	**20.68**	***	3.01	ns	3.69	ns	**10.33**	**	0.93	ns	1.16	ns	0.69	ns	0.48	n	0.01	ns
Cohort	0.01	ns	**11.01**	***	**71.55**	***	**50.99**	***	2.76	ns	**19.43**	***	**17.89**	***	0.59	ns	**198.76**	***	**42.15**	***	**62.44**	***	**32.59**	**	**295.06**	***
Region	**19.77**	***	**5.97**	***	2.84	*	0.54	ns	**19.83**	***	2.57	ns	1.68	ns	**12.13**	***	**5.03**	**	**9.32**	***	**5.09**	**	2.62	*	1.74	ns
Stand(Region)	**112.69**	***	**2.68**	***	**4.68**	***	**4.49**	***	**7.77**	***	**3.66**	***	**3.89**	***	**7.70**	***	**5.45**	***	**7.13**	***	**3.01**	***	**6.99**	***	**11.10**	***
Family(Stand(Region))	NE	NE	**10.10**	**	6.10	*	2.60	ns	**43.80**	***	**10.20**	**	**8.70**	**	**22.90**	***	**8.70**	**	**13.20**	***	5.70	*	4.20	*	5.20	*
Cohort*Region	**81.52**	***	0.36	ns	0.81	ns	1.07	ns	4.24	**	1.17	ns	1.22	ns	3.74	*	2.29	ns	**12.31**	***	2.16	ns	**6.36**	***	**9.98**	***
Cohort*Stand(Region)	0.43	ns	1.21	ns	**2.87**	***	**3.00**	***	**2.11**	***	**2.45**	***	**2.52**	***	**1.91**	***	**3.04**	***	**2.40**	***	1.14	ns	**4.30**	***	**6.59**	***
*Cohort*Family(Stand(Region))*	NE	NE	0.00	ns	4.20	*	4.20	*	6.60	*	1.50	ns	1.70	ns	7.30	**	0.00	ns	0.00	ns	0.20	ns	2.20	ns	**10.70**	**
Control Bg-2	NE	NE	0.58	ns	0.40	ns	0.46	ns	0.28	ns	3.89	*	4.24	*	3.07	*	0.79	ns	0.77	ns	1.53	ns	1.21	ns	0.99	ns

B. Year treatment

Traits	survival		FITTOT		FITSTEM		FRUITSTEM		SILSTEM		FITPB		FRUITPB		SILPB		FITBB		FRUITBB		RSTEM		RPB		RBB	
Model terms	*F* or *LRT*	*P*	*F* or *LRT*	*P*	*F* or *LRT*	*P*	*F* or *LRT*	*P*	*F* or *LRT*	*P*	*F* or *LRT*	*P*	*F* or *LRT*	*P*	*F* or *LRT*	*P*	*F* or *LRT*	*P*	*F* or *LRT*	*P*	*F* or *LRT*	*P*	*F* or *LRT*	*P*	*F* or *LRT*	*P*
Block	**Inf**	***	4.05	*	1.20	ns	0.51	ns	**12.25**	***	2.47	ns	2.47	ns	14.10	ns	1.01	ns	0.85	ns	0.30	ns	6.50	*	0.61	ns
Year	0.01	ns	0.08	ns	**36.22**	***	**21.10**	***	**14.01**	***	2.61	ns	2.61	ns	4.51	*	**186.61**	***	**202.43**	***	**38.08**	**	**42.44**	***	**255.69**	***
Region	**55.05**	***	3.39	*	**9.61**	***	**14.99**	***	**7.27**	***	**12.50**	***	**12.50**	***	2.78	*	1.15	ns	0.72	ns	3.54	*	**8.02**	***	**12.47**	***
Stand(Region)	**77.14**	***	**2.34**	***	**3.04**	***	**3.34**	***	**7.29**	***	**3.30**	***	**3.30**	***	**6.81**	***	**3.57**	***	**3.54**	***	**2.02**	***	**5.40**	***	**4.27**	***
Family(Stand(Region))	NE	NE	2.80	ns	4.60	*	3.00	ns	**25.80**	***	2.50	ns	2.50	ns	**25.60**	***	2.20	ns	3.40	ns	4.10	*	1.40	ns	2.10	ns
Year*Region	**92.58**	***	**17.06**	***	**18.10**	***	**16.35**	***	**19.21**	***	**11.38**	***	**11.38**	***	**16.80**	***	**6.62**	***	**7.01**	***	1.51	ns	**4.83**	**	3.38	*
Year*Stand(Region)	**7.8**	***	**2.67**	***	1.46	*	1.59	**	**2.13**	***	**1.83**	***	**1.83**	***	**2.18**	***	**2.74**	***	**2.64**	***	1.59	**	1.33	ns	1.64	**
*Year*Family(Stand(Region))*	NE	NE	0.00	ns	0.80	ns	1.30	ns	3.80	ns	1.20	ns	1.20	ns	2.30	ns	2.70	ns	4.50	*	0.00	ns	2.90	ns	6.50	*
Control Bg-2	NE	NE	0.94	ns	0.61	ns	0.42	ns	3.81	*	0.20	ns	0.20	ns	0.10	ns	0.99	ns	0.87	ns	0.07	ns	0.17	ns	0.85	ns

Table 2

A. Within each field experiment

		Traits											
		FITTOT	FITSTEM	FRUITSTEM	SILSTEM	FITPB	FRUITPB	SILPB	FITBB	FRUITBB	RSTEM	RPB	RBB
Field experiment	Geographical scale												
August 2010	Among regions		0.02	**0.00 U**	0.15	0.02	**0.00 U**	0.09	0.07	0.06	**0.00 U**	**0.00 U**	**0.00 U**
	49 stands		**0.49 U**	**0.49 U**	0.55	**0.44 U**	**0.45 U**	**0.49 U**	**1.00 H**	**1.00 H**	**0.25 U**	0.56	0.62
	Brittany					0.46	0.53	0.23				0.57	0.51
	Burgundy		0.39	0.33	0.55			0.61	0.44	0.49	0.21	0.66	0.52
	Languedoc		0.56		0.67	0.37	0.32	0.46			0.24	0.51	**0.87 H**
	North		0.50	0.52	0.36	0.38	0.38	**0.32 U**			0.39	0.56	0.80
September 2010	Among regions	0.29	0.02	**0.00 U**	0.13	0.10	0.06	0.26	0.30	0.28	**0.38 H**	0.07	0.29
	49 stands	0.57	**0.39 U**	**0.42 U**	**0.49 U**	0.59	0.61	0.61	**0.49 U**	**0.49 U**	**0.87 H**	0.64	0.65
	Brittany		0.31	0.43	0.44			0.55	**0.03 U**	**0.03 U**			**0.02 U**
	Burgundy		0.29		0.61			**0.88 H**					0.69
	Languedoc		0.39	0.53	0.32	0.56	0.59	0.32		0.80		0.59	0.78
	North		0.66	0.63	0.39			0.38	**0.06 U**	**0.08 U**			**0.17 U**
September 2011	Among regions		**0.53 H**	**0.52 H**	0.23	**0.40 H**	**0.40 H**	**0.00 U**	**0.00 U**	**0.00 U**		0.18	0.18
	49 stands		0.61	0.61	**0.78 H**	0.69	0.61	0.67	0.58	**0.49 U**		**0.50 U**	**0.42 U**
	Brittany												
	Burgundy												**0.02 U**
	Languedoc				0.80			0.65				0.82	
	North		**0.12 U**	**0.23 U**	**0.22 U**			**0.22 U**	**0.08 U**	**0.06 U**		**0.11 U**	

B. Plasticity

		Traits											
		FITTOT	FITSTEM	FRUITSTEM	SILSTEM	FITPB	FRUITPB	SILPB	FITBB	FRUITBB	RSTEM	RPB	RBB
Treatment	Geographical scale												
Cohort	Among regions		**0.00 U**	**0.00 U**	0.13	0.02	0.02	0.08				0.04	0.11
	49 stands		**0.42 U**	**0.43 U**	**0.39 U**	0.50	0.51	**0.46 U**				0.63	0.63
	Brittany		0.48	0.50		0.43	0.45	0.28					0.38
	Burgundy		0.35	0.23					0.34	0.39		**0.78 H**	0.46
	Languedoc					0.36	0.31	0.24				0.32	
	North		**0.32 U**	**0.29 U**	**0.15 U**								0.77
Year	Among regions	**0.58 H**	**0.51 H**	**0.49 H**	**0.42 H**	**0.46 H**	**0.43 H**	**0.51 H**	0.17	0.16		0.23	0.05
	49 stands	**1.00 H**	0.59	0.52	0.72	**0.85 H**	**0.75 H**	**0.94 H**	0.68	0.61		**0.32 U**	**0.27 U**
	Brittany												
	Burgundy												**0.04 U**
	Languedoc			0.73		**0.85 H**	0.76	0.52				0.44	
	North				**0.04 U**				**0.09 U**	**0.07 U**			**0.06 U**

SUPPORTING INFORMATION: EVIDENCE OF SELECTION ACTING ON REPRODUCTIVE STRATEGIES IN *ARABIDOPSIS THALIANA* AT DIFFERENT SPATIAL SCALES: WHEN TEMPORAL HETEROGENEITY IN THE FIELD MATTERS!

Fig. S1. Natural variation of seven fitness components and their reaction norms at the scale of France and within each region. Open diamond: Brittany, open triangle: Burgundy, open square: Languedoc, open circle: North. At the scale of France, each value corresponds to the median of the LSmeans of 200 families per region for seven fitness components. Within each region, each value corresponds to the median of the LS means of families per stand for the seven fitness components.

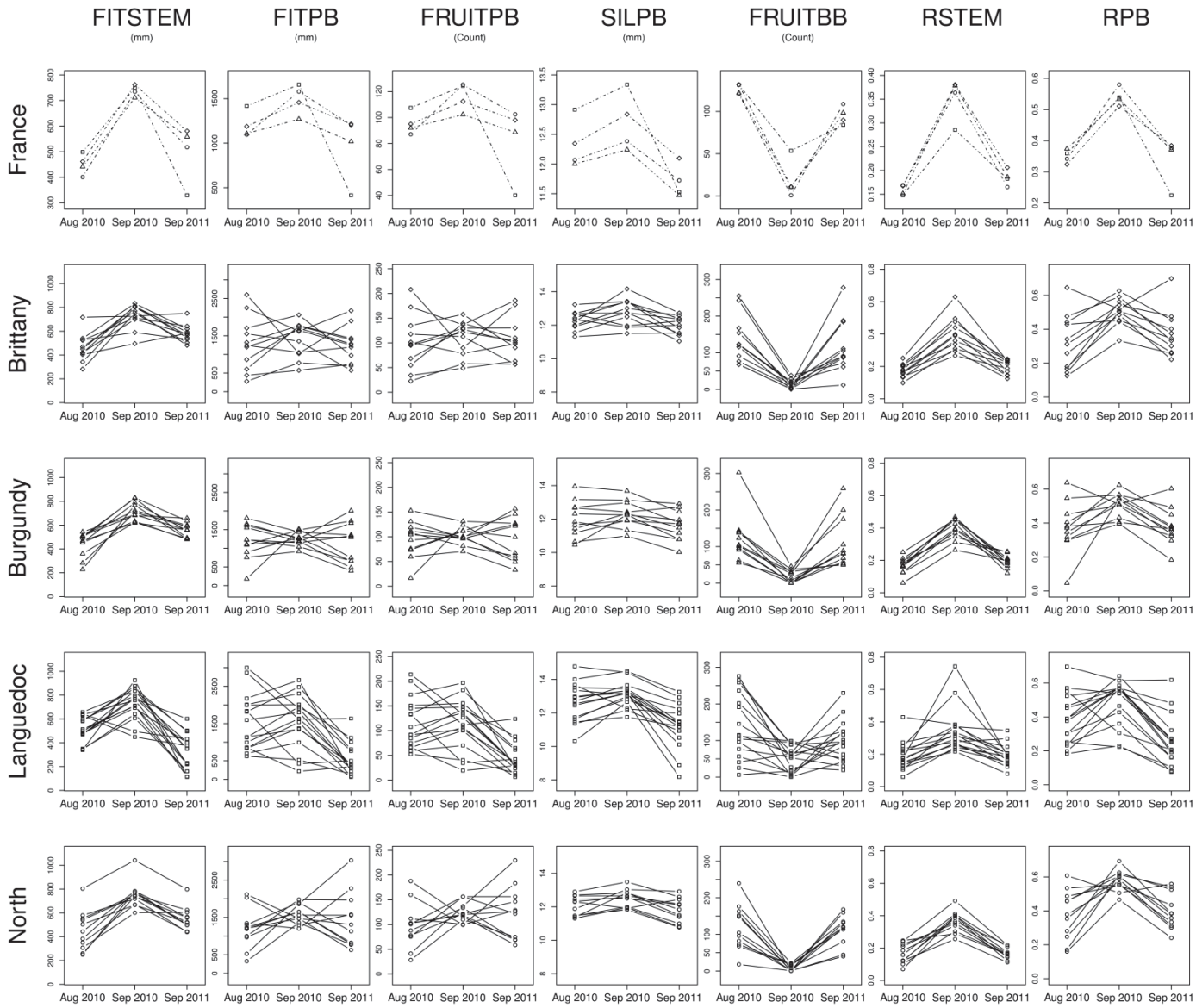


Fig. S2 Empirical SNP F_{ST} distributions among regions, among the 49 stands and among stands within each region. Solid and dashed lines indicate the upper and lower 5 % and 1% cut-offs, respectively. 'n' stands for the number of polymorphic loci.

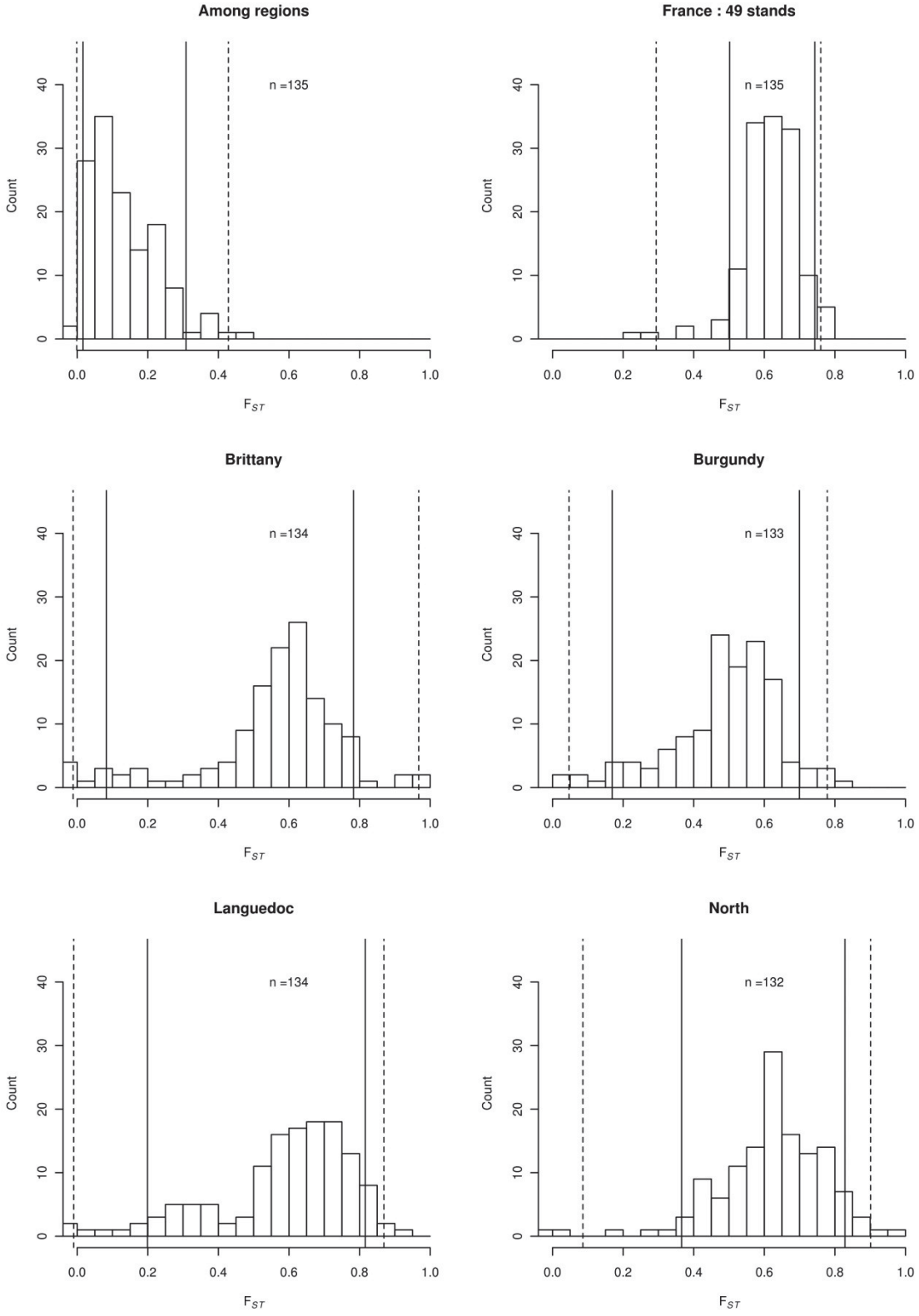


Fig. S3 Relationship (A) between trait Q_{ST} and negative $\log^{10}$ *P*-value for trait divergence among the geographical scales considered (either region or stands) within each sowing date (*i.e.* 'August 2010', 'September 2010' and 'September 2011') and (B) between Q_{ST} of trait plasticity and negative $\log^{10}$ *P*-value for divergence in trait plasticity among the geographical scales considered (either region or stands) within each treatment (*i.e.* cohort and year treatments). Dashed lines show the maximum number of digits after decimal dot available in in SAS 9.3 (SAS Institute Inc).

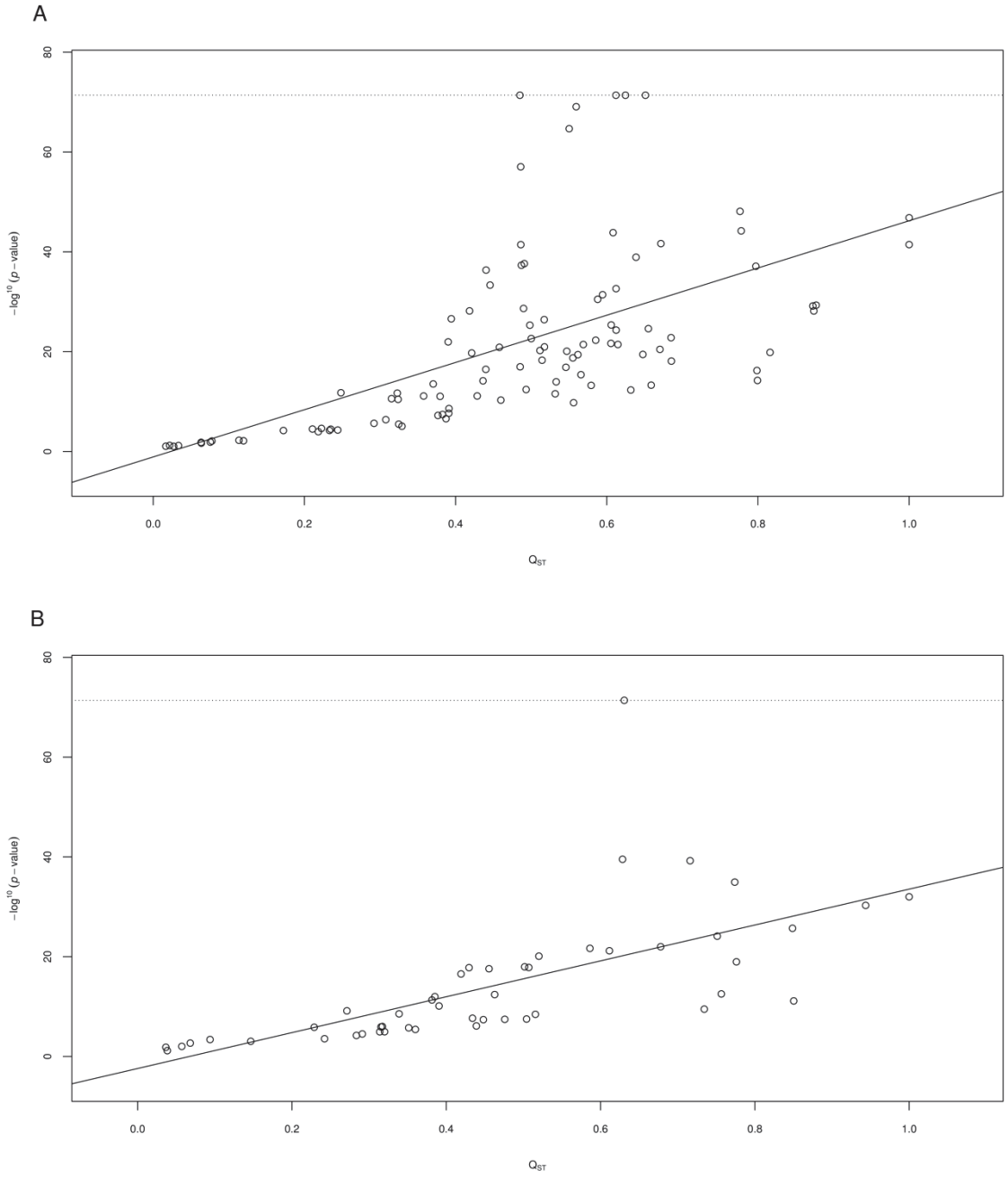
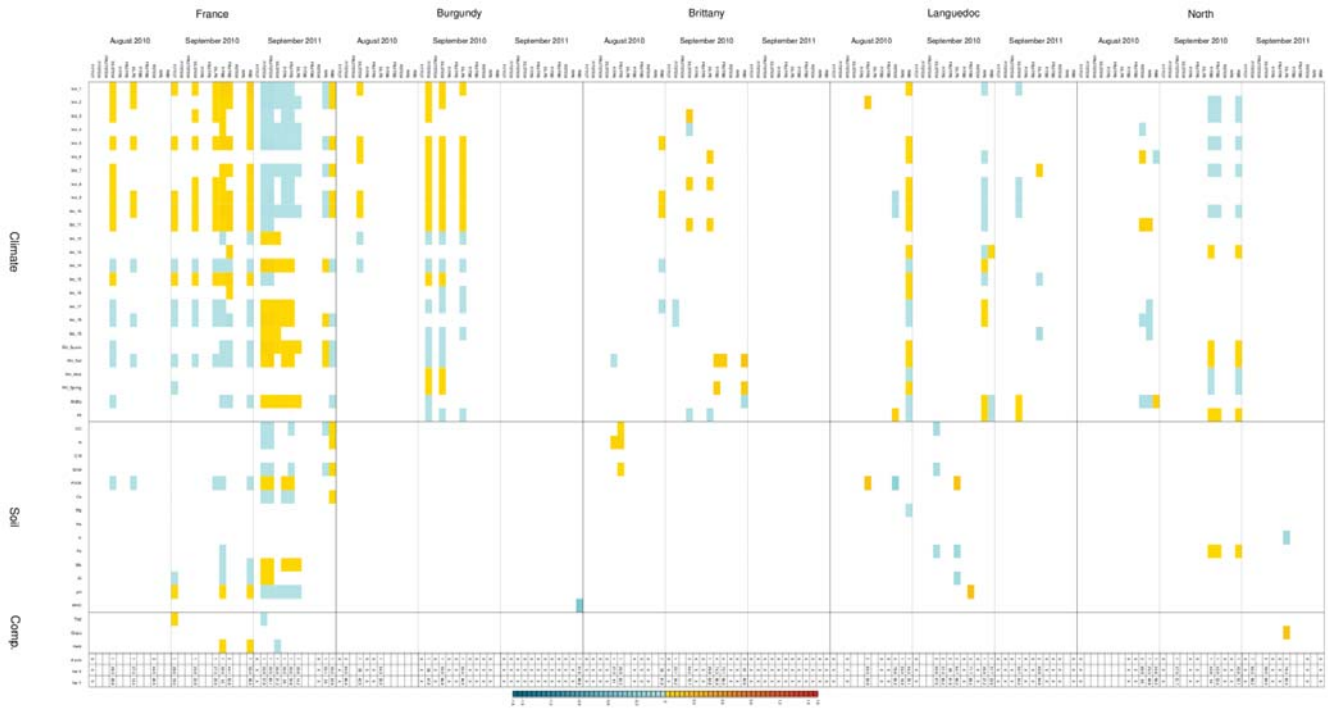


Fig. S4 Identification of the putative selective agents acting on fitness components in *A. thaliana* in five geographical units France, without control for neutral genetic variation. Each line corresponds to the results of Partial Least Square Regression (PLSR) for a specific 'geographical unit (France and four regions) * sowing date ('August 2010', 'September 2010' and 'September 2011') * fitness component' combination. The 42 ecological variables characterizing the 49 stands (climate, soil and competition (Comp.)) correspond to the first 42 columns. For each column and line, the colored squares indicate significant regression coefficients (p-value < 0.05) and the colors represent the strength of the regression coefficient estimated between fitness component variation and ecological variation. The last three columns correspond to the number of PLSR components retained after cross-validation ('axis'), the percentage of ecological variation explained by PLSR components ('Var X') and the percentage of fitness component variation explained by PLSR components ('Var Y') corresponds to the 43rd, 44th and 45th columns, respectively. PLSR were performed on standardized data, allowing comparison of regression coefficients among geographical units, sowing dates and fitness components. PLSR were run only for fitness components with significant broad-sense heritability. Fitness components with non-significant broad-sense heritability are marked by a cross in the 'axis', 'Var X' and 'Var Y' columns. Empty cells in the 'axis', 'Var X' and 'Var Y' columns correspond to fitness components with significant broad-sense heritability but for which no convergence of PLSR models was reached.

Fig. S4



Climatic variables: **Alt**, Altitude (m); **RH Wint**, **RH Spring**, **RH Summ** and **RH Fall** stand for relative humidity (%) in Winter, Spring, Summer and Fall, respectively; **Aridity**, Aridity; **Bio 1**, annual mean temperature (°C x 10); **Bio 2**, mean diurnal range (mean of monthly temperature range (max temp - min temp)); **Bio 3**, isothermality (Bio 2/Bio 7) (* 100); **Bio 4**, temperature seasonality (standard deviation *100); **Bio 5**, max temperature of warmest month (°C x 10); **Bio 6**, min temperature of coldest month (°C x 10); **Bio 7**, temperature annual range (Bio 5 - Bio 6); **Bio 8**, mean temperature of wettest quarter (°C x 10); **Bio 9**, mean temperature of driest quarter (°C x 10); **Bio 10**, mean temperature of warmest quarter (°C x 10); **Bio 11**, mean temperature of coldest quarter (°C x 10); **Bio 12**, annual precipitation (mm); **Bio 13**, precipitation of wettest month (mm); **Bio 14**, precipitation of driest month (mm); **Bio 15**, precipitation seasonality (coefficient of variation); **Bio 16**, precipitation of wettest quarter (mm); **Bio 17**, precipitation of driest quarter (mm); **Bio 18**, precipitation of warmest quarter (mm); **Bio 19**, precipitation of coldest quarter (mm).

Edaphic variables: OC, organic carbon ($g.kg^{-1}$); **N**, total nitrogen ($g.kg^{-1}$); **C/N**, carbon/nitrogen ratio; **SOM**, soil organic matter ($g.kg^{-1}$); **P2O5**, phosphorus (P_2O_5) ($g.kg^{-1}$); **Ca**, exchangeable calcium ($cmol+.kg^{-1}$); **Mg**, exchangeable magnesium ($cmol+.kg^{-1}$); **Na**, exchangeable sodium ($cmol+.kg^{-1}$); **K**, exchangeable potassium ($cmol+.kg^{-1}$); **Fe**, exchangeable iron ($cmol+.kg^{-1}$); **Mn**, exchangeable manganese ($cmol+.kg^{-1}$); **Al**, exchangeable aluminium ($cmol+.kg^{-1}$); **WHC**, soil water holding capacity ($ml.g^{-1}$); **pH**.

Competition variables: Herb, interspecific competition with herbs which are not grasses; **Grass**, interspecific competition with grasses; **Thal**, intraspecific competition.

Fig. S5 Identification of the putative selective agents acting on plasticity of fitness components in *A. thaliana* in five geographical units, without control for neutral genetic variation. Each line corresponds to the results of Partial Least Square Regression (PLSR) for a specific 'geographical unit (France and four regions) * treatment ('cohort treatment' and 'year treatment') * fitness component' combination. The 42 ecological variables characterizing the 49 stands (climate, soil and competition (Comp.)) correspond to the first 42 columns. For each column and line, the colored squares indicate significant regression coefficients (p-value < 0.05) and the colors represent the strength of the regression coefficient estimated between variation of fitness component plasticity and ecological variation. The last three columns correspond to the number of PLSR components retained after cross-validation ('axis'), the percentage of ecological variation explained by PLSR components ('Var X') and the percentage of fitness component variation explained by PLSR components ('Var Y') corresponds to the 43rd, 44th and 45th columns, respectively. PLSR were performed on standardized data, allowing comparison of regression coefficients among geographical units, sowing date treatments and fitness components. PLSR were run only for plasticity of fitness components with significant broad-sense heritability. Plasticity of fitness components with non-significant broad-sense heritability are marked by a cross in the 'axis', 'Var X' and 'Var Y' columns. Empty cells in the 'axis', 'Var X' and 'Var Y' columns correspond to plasticity of fitness components with significant broad-sense heritability but for which no convergence of PLSR models was reached. See legend of Fig. S4 for a description of the 42 ecological variables.

Fig. S5

Table S1 Natural variation of fitness components at three spatial scales (*i.e.* region, stand and family) within each sowing date. (A) 'August 2010' sowing date, (B) 'September 2010' sowing date and (C) 'September 2011' sowing date.

A. August 2010

Traits	survival	FITTOT	FITSTEM	FRUITSTEM	SILSTEM	FITPB	FRUITPB	SILPB	FITBB	FRUITBB	RSTEM	RPB	RBB
Model terms	*F* or LRT *P*	*F* or LRT *P*	*F* or LRT *P*	*F* or LRT *P*	*F* or LRT *P*	*F* or LRT *P*	*F* or LRT *P*	*F* or LRT *P*	*F* or LRT *P*	*F* or LRT *P*	*F* or LRT *P*	*F* or LRT *P*	*F* or LRT *P*
Block	0.03 ns	0.73 ns	2.00 ns	0.88 ns	**8.79** **	0.06 ns	0.10 ns	2.29 ns	0.53 ns	0.66 ns	0.06 ns	0.15 ns	0.05 ns
Region	**44.25** ***	1.79 ns	2.33 ns	0.97 ns	**11.30** ***	0.97 ns	0.41 ns	1.45 ns	1.27 ns	1.07 ns	2.32 ns	2.09 ns	1.39 ns
Stands(Region)	0.99 ns	**1.87** ***	**4.22** ***	**3.79** ***	**4.48** ***	**3.74** ***	**3.69** ***	**4.45** ***	**4.33** ***	**4.77** ***	**1.86** ***	**6.53** ***	**8.27** ***
Family(Stand(Region))	0.00 ns	0.00 ns	6.10 *	3.20 ns	**13.40** ***	**18.60** ***	**15.30** ***	**12.70** ***	0.00 ns	0.00 ns	**9.20** **	**9.10** **	**10.60** **
Control Bg-2	NE NE	0.64 ns	0.76 ns	0.76 ns	0.00 ns	5.17 *	6.19 *	3.72 ns	0.04 ns	0.06 ns	2.71 ns	1.94 ns	0.54 ns

B. September 2010

Traits	survival	FITTOT	FITSTEM	FRUITSTEM	SILSTEM	FITPB	FRUITPB	SILPB	FITBB	FRUITBB	RSTEM	RPB	RBB
Model terms	*F* or LRT *P*	*F* or LRT *P*	*F* or LRT *P*	*F* or LRT *P*	*F* or LRT *P*	*F* or LRT *P*	*F* or LRT *P*	*F* or LRT *P*	*F* or LRT *P*	*F* or LRT *P*	*F* or LRT *P*	*F* or LRT *P*	*F* or LRT *P*
Block	1.34 ns	3.56 ns	4.74 *	2.04 ns	**14.08** ***	3.69 ns	3.85 ns	**8.34** **	0.55 ns	0.52 ns	3.02 ns	6.43 *	0.07 ns
Region	0.00 ns	**5.57** ***	1.29 ns	0.48 ns	**15.38** ***	**4.88** **	**4.04** **	**21.21** ***	**15.17** ***	**15.06** ***	**5.26** **	**7.51** ***	**27.76** ***
Stands(Region)	0.08 ns	**2.35** ***	**3.37** ***	**3.63** ***	**7.57** ***	**3.37** ***	**3.69** ***	**6.01** ***	**2.61** ***	**2.89** ***	**2.64** ***	**4.23** ***	**7.23** ***
Family(Stand(Region))	**34.87** ***	1.60 ns	**12.00** ***	**11.40** ***	**67.40** ***	3.20 ns	3.00 ns	**28.90** ***	**9.50** **	**11.60** ***	0.50 ns	2.40 ns	**28.60** ***
Control Bg-2	NE NE	0.83 ns	0.09 ns	0.12 ns	0.99 ns	0.56 ns	0.45 ns	0.21 ns	3.24 ns	3.36 ns	0.00 ns	0.49 ns	3.52 ns

C. September 2011

Traits	survival	FITTOT	FITSTEM	FRUITSTEM	SILSTEM	FITPB	FRUITPB	SILPB	FITBB	FRUITBB	RSTEM	RPB	RBB
Model terms	*F* or LRT *P*	*F* or LRT *P*	*F* or LRT *P*	*F* or LRT *P*	*F* or LRT *P*	*F* or LRT *P*	*F* or LRT *P*	*F* or LRT *P*	*F* or LRT *P*	*F* or LRT *P*	*F* or LRT *P*	*F* or LRT *P*	*F* or LRT *P*
Block	1.39 ns	0.77 ns	0.02 ns	0.10 ns	2.18 ns	0.20 ns	0.11 ns	5.48 *	0.47 ns	0.37 ns	0.50 ns	1.00 ns	1.76 ns
Region	**4.69** **	**10.46** ***	**15.46** ***	**16.71** ***	**9.23** ***	**10.53** ***	**11.93** ***	2.72 *	0.82 ns	0.86 ns	0.60 ns	**5.18** **	**3.92** **
Stands(Region)	1.10 ns	**2.25** ***	1.56 *	**1.70** **	**3.49** ***	**1.89** ***	**1.82** **	**3.66** ***	**2.55** ***	**2.43** ***	1.15 ns	**2.74** ***	**1.77** **
Family(Stand(Region))	0.00 ns	0.00 ns	0.10 ns	0.10 ns	0.30 ns	0.80 ns	1.80 ns	3.70 ns	1.40 ns	2.90 ns	0.40 ns	6.40 *	2.30 ns
Control Bg-2	NE NE	0.97 ns	0.72 ns	0.44 ns	3.69 ns	0.01 ns	0.01 ns	0.00 ns	0.06 ns	0.01 ns	0.23 ns	0.00 ns	0.44 ns

Model random terms were tested with likelihood ratio tests (LRT) of models with and without these effects. Random effects are in italic. *$0.05 > P > 0.01$, **$0.01 > P > 0.001$, *** $P < 0.001$. Bold stars indicate significant effect after Bonferroni correction. NE: not estimated.

Table S2 Broad-sense heritability (H^2) estimates in five geographical units (A) for fitness components within each sowing date and (B) for plasticity of fitness components in the cohort and year treatments. *0.05 > P > 0.01, **0.01 > P > 0.001, *** P < 0.001.

A. Within field experiment

		Traits																					
		FITTOT																RSTEM		RPB		RBB	
			FITSTEM					FITPB				FITBB											
Field experiment	Geographical scale				FRUITSTEM		SILSTEM			FRUITPB		SILPB				FRUITBB							
August 2010	France	0.03	0.33	***	0.28	***	0.42	***	0.35	***	0.33	***	0.48	***	0.15	*	0.18	**	0.30	***	0.44	***	0.49	***
	Brittany	0.00	0.21		0.18		0.21		0.35	**	0.33	**	0.36	*	0.09		0.13		0.18		0.37	**	0.45	***
	Burgundy	0.00	0.27	*	0.25	*	0.45	***	0.15		0.16		0.45	***	0.44	***	0.46	***	0.34	**	0.44	***	0.60	***
	Languedoc	0.00	0.26	*	0.13		0.40	***	0.53	***	0.51	***	0.63	***	0.00		0.03		0.36	**	0.50	***	0.40	***
	North	0.09	0.56	***	0.51	***	0.45	***	0.29	*	0.29	*	0.28	*	0.21		0.23		0.31	*	0.42	***	0.49	***
September 2010	France	0.18	** 0.32	***	0.31	***	0.64	***	0.23	***	0.23	***	0.49	***	0.32	***	0.35	***	0.14	*	0.25	***	0.56	***
	Brittany	0.24	0.32	**	0.37	**	0.52	***	0.22		0.23		0.46	**	0.30	**	0.33	**	0.07		0.09		0.47	***
	Burgundy	0.07	0.29	*	0.20		0.71	***	0.14		0.16		0.34	**	0.19		0.19		0.05		0.17		0.33	**
	Languedoc	0.15	0.37	**	0.41	***	0.54	***	0.45	***	0.47	***	0.53	***	0.21		0.26	*	0.14		0.60	***	0.60	***
	North	0.03	0.26	*	0.26	*	0.66	***	0.00		0.00		0.40	***	0.37	**	0.37	***	0.13		0.08		0.39	***
September 2011	France	0.09	0.17	*	0.19	**	0.23	***	0.16	*	0.19	**	0.26	***	0.13	*	0.15	*	0.10		0.25	***	0.24	***
	Brittany	0.00	0.00		0.02		0.00		0.12		0.16		0.00		0.00		0.00		0.09		0.11		0.09	
	Burgundy	0.00	0.00		0.00		0.19		0.09		0.12		0.23		0.06		0.16		0.21		0.17		0.31	*
	Languedoc	0.19	0.20		0.17		0.31	*	0.10		0.11		0.45	**	0.23		0.22		0.00		0.30	*	0.24	
	North	0.20	0.27	*	0.28	*	0.30	*	0.16		0.18		0.34	**	0.25	*	0.25	*	0.12		0.33	*	0.22	

B. Plasticity

		Traits																					
		FITTOT																RSTEM		RPB		RBB	
			FITSTEM					FITPB				FITBB											
Treatment	Geographical scale				FRUITSTEM		SILSTEM			FRUITPB		SILPB				FRUITBB							
cohort	France	0.00	0.12	***	0.13	***	0.13	***	0.25	***	0.25	***	0.15	**	0.00		0.01		0.07		0.24	***	0.19	***
	Brittany	0.00	0.15	*	0.15	*	0.11		0.40	**	0.36	*	0.31	*	0.00		0.00		0.00		0.21		0.18	***
	Burgundy	0.03	0.16	*	0.15	*	0.10		0.18		0.20		0.00		0.23	***	0.23	***	0.08		0.33	**	0.22	***
	Languedoc	0.00	0.07		0.10		0.06		0.25	**	0.28	**	0.22	**	0.00		0.00		0.14		0.24	***	0.03	
	North	0.07	0.11	*	0.11	*	0.22	***	0.11		0.11		0.08		0.10		0.11		0.11		0.16		0.19	***
year	France	0.19	* 0.15	**	0.17	**	0.16	***	0.21	*	0.22	**	0.12	**	0.14	**	0.14	**	0.00		0.11	*	0.10	***
	Brittany	0.13	0.19		0.24		0.12		0.18		0.18		0.09		0.03		0.01		0.00		0.00		0.06	
	Burgundy	0.00	0.00		0.00		0.05		0.00		0.00		0.00		0.00		0.06		0.00		0.11		0.12	*
	Languedoc	0.21	0.06		0.11	*	0.07		0.16	*	0.20	*	0.13	*	0.28		0.26		0.00		0.12	*	0.06	
	North	0.13	0.13		0.16		0.15	*	0.05		0.09		0.14		0.18	**	0.16	**	0.01		0.21		0.10	*

Table S3 Percentage of phenotypic variation explained by neutral genetic diversity in five geographical units (A) for fitness components within each sowing date and (B) for plasticity of fitness components in the cohort and year treatments. Results are only shown for fitness components and plasticities of fitness components with a significant broad-sense heritability estimate (see Table S2). *0.05 > P > 0.01, **0.01 > P > 0.001, *** P < 0.001.

A. Within each field experiment

		Traits																					
		FITTOT		FITSTEM		FRUITSTEM		SILSTEM		FITPB		FRUITPB		SILPB		FITBB		FRUITBB		RSTEM		RPB		RBB	
Field experiment	Geographical scale																								
August 2010	France			0.15	***	0.15	***	0.23	***	0.14	***	0.12	***	0.19	***	0.15	***	0.17	***	0.09	***	0.20	***	0.25	***
	Brittany									0.21	***	0.22	***	0.20	**							0.23	***	0.28	***
	Burgundy			0.24	**	0.20	*	0.41	***					0.39	***	0.29	***	0.31	***	0.15		0.40	***	0.43	***
	Languedoc			0.28	***			0.46	***	0.30	***	0.27	***	0.39	***					0.20	**	0.35	***	0.45	***
	North			0.33	***	0.35	***	0.24	***	0.18	*	0.17	*	0.20	**					0.21	***	0.33	***	0.45	***
September 2010	France	0.12	***	0.14	***	0.13	***	0.39	***	0.14	***	0.13	***	0.32	***	0.15	***	0.15	***	0.12	***	0.14	***	0.30	***
	Brittany			0.18	**	0.24	***	0.45	***					0.31	***	0.16	**	0.17	**					0.15	*
	Burgundy			0.22	**			0.55	***					0.50	***									0.37	***
	Languedoc			0.26	***	0.28	***	0.45	***	0.33	***	0.34	***	0.36	***			0.40	***			0.42	***	0.64	***
	North			0.28	***	0.26	***	0.59	***					0.40	***	0.25	***	0.24	***					0.33	***
September 2011	France			0.16	***	0.18	***	0.17	***	0.11	***	0.12	***	0.19	***	0.09	**	0.09	**			0.14	***	0.11	***
	Brittany																								
	Burgundy																							0.19	
	Languedoc							0.37	***					0.46	***							0.42	***		
	North			0.12		0.16		0.16						0.26	**	0.13		0.13				0.12			

B. Plasticity

		Traits																					
		FITTOT		FITSTEM		FRUITSTEM		SILSTEM		FITPB		FRUITPB		SILPB		FITBB		FRUITBB		RSTEM		RPB		RBB	
Treatment	Geographical scale																								
cohort_effect	France			0.08	**	0.08	**	0.11	***	0.08	**	0.08	**	0.08	*							0.16	***	0.24	***
	Brittany			0.16	*	0.16	*			0.14		0.13		0.15										0.22	***
	Burgundy			0.27	***	0.24	**									0.25	***	0.26	***			0.35	***	0.38	***
	Languedoc									0.17	*	0.17	*	0.22	**							0.17	*		
	North			0.20	**	0.19	**	0.17	*															0.44	***
year_effect	France	0.14	***	0.13	***	0.12	***	0.18	***	0.11	***	0.11	***	0.15	***	0.11	***	0.11	***			0.06		0.07	
	Brittany																								
	Burgundy																							0.15	
	Languedoc					0.18				0.26	***	0.28	***	0.34	***							0.20	*		
	North							0.15								0.11		0.11						0.04	

MANUSCRIT: ADAPTIVE VALUE AND COSTS OF PHENOTYPIC PLASTICITY ACROSS GERMINATION COHORTS AND YEARS IN FOUR REGIONAL SETS OF NATURAL *ARABIDOPSIS THALIANA* FAMILIES

Adaptive value and costs of phenotypic plasticity across germination cohorts and years in four regional sets of natural *Arabidopsis thaliana* families

Romain Villoutreix[1], Cédric Glorieux[1], Joël Cuguen[1], Fabrice Roux[1,*]

[1]*Laboratoire Génétique et Evolution des Populations Végétales, UMR CNRS 8198, Université des Sciences et Technologies de Lille – Lille 1, F-59655 Villeneuve d'Ascq cedex France*

[3] *INRA, Laboratoire des Interactions Plantes-Microorganismes (LIPM), UMR441, F-31326 Castanet-Tolosan, France.*

[4] *CNRS, Laboratoire des Interactions Plantes-Microorganismes (LIPM), UMR2594, F-31326 Castanet-Tolosan, France.*

Corresponding author: Fabrice Roux

CNRS, Laboratoire des Interactions Plantes-Microorganismes (LIPM), UMR2594, F-31326 Castanet-Tolosan, France. E-mail: fabrice.roux@toulouse.inra.fr

Phone number: +33 (0)5 61 28 54 59 Fax number: +33 (0)5 61 28 50 61

Keywords: phenotypic plasticity, adaptation, cost, gradient of selection, temporal variation, reaction norms

Running title: Reaction norms across time

Summary

Phenotypic plasticity is a key factor in the process of adaptation in highly heterogeneous environments. Although numerous empirical studies were designed to study the adaptive value and cost of phenotypic plasticity in plants, the cost/benefit trade-off of plasticity in temporally heterogeneous natural environments has been rarely investigated. In this study, we aimed to study the adaptive value and cost of phenotypic plasticity across two time scales in *Arabidopsis thaliana*. Eight hundreds families collected from 49 stands located in four climatically contrasted regions in France were measured for nine traits related to resource acquisition, phenology and shoot architecture under field conditions across three sowing dates simulating two treatments: two germination cohorts within the same year and two years within the same germination cohort. In contrast to previous studies, we found a prevalence of signed values of plasticity across time for all traits (with the exception of flowering time), with families having opposite response across sowing dates. By performing genotypic selection gradients analyses based on total seed production, contrasted patterns of selection were observed between traits and their plasticities. While the nature and strength of selection acting on the nine traits were similar among the sowing dates and the four French regions, plasticities revealed a different and complex selective pattern which was dependent on interactions among treatment, geographical region and frequency of the focal sowing date. The fraction of significant cost was twice higher for the cohort than the year treatment. Only local costs were however detected, suggesting that evolution of phenotypic plasticity should not be constrained in our set of 800 families. Altogether, those results stress the need to integrate genotype-by-environment interactions measured in ecologically realistic conditions when performing ecological genomics.

Introduction

Phenotypic plasticity is the ability of a genotype to express different phenotypes depending on the biotic or abiotic environmental characteristics it encounters (Sultan 2000; Agrawal 2001). First considered as developmental noise or developmental instability, phenotypic plasticity is now widely accepted as a major component in evolutionary processes (Bradshaw 1965; Schlichting 1986; Scheiner 1993; Schlichting and Smith 2002; Pigliucci 2005), in particular in sessile organisms like plants (Sultan 1987; Sultan 2000; Alpert and Simms 2002; van Kleunen and Fischer 2005; Valladares *et al.* 2007). Because phenotypic plasticity may evolve over a short time scale (Reboud and Bell 1997; Kassen 2002, Stomp *et al.* 2008) and its high variability in many natural populations indicates potential for selection response (Schmitt *et al.* 1992; Dorn *et al.* 2000; Donohue *et al.* 2001; Agrawal *et al.* 2002; Donohue 2002; van Kleunen and Fischer 2001), understanding phenotypic plasticity may help to manage plant invasiveness (Richards *et al.* 2006; Matesanz *et al.* 2012) and predict species resilience in plant communities as well as crop response in the context of global change (Nicotra *et al.* 2010).

Phenotypic plasticity is predicted to be favored in highly heterogeneous environment if (i) selective pressures vary at the same scale than the response unit (Moran 1992; Sultan and Spencer 2002; Alpert and Simms 2002), (ii) the response of the organism is faster than the change in the environment (DeWitt *et al.* 1998; Alpert and Simms 2002), (iii) environmental variation is highly predictable (DeWitt *et al.*; Alpert and Simms 2002; Schlichting and Smith 2002; Tonsor *et al.* 2013), and (iv) the most plastic genotype has a higher relative fitness averaged across environments than the less plastic genotypes (van Kleunen and Fischer 2005). Despite its theoretical benefits, adaptive phenotypic plasticity is not as frequent as expected in nature. The evolution of adaptive phenotypic plasticity could be

impeded by diverse costs and limits (initially reviewed in DeWitt *et al.* 1998). Frequently considered costs are (i) that of maintaining the sensory and regulatory machinery required for plastic responses (DeWitt *et al.* 1998), and (ii) a correlation between plasticity and developmental instability, recently demonstrated in the model plant species *Arabidopsis thaliana* by growing a set of recombinant inbred lines (RILs) at four different nitrogen supply levels (Tonsor *et al.* 2013). One of the main limits concerns the unreliability of environmental cues, leading to a mal-adaptive plastic response (van Kleunen and Fischer 2005).

Although numerous empirical studies were designed to study the adaptive value and cost of phenotypic plasticity in plants, four points may still deserve further investigations. First, because plants are exposed to a great range of complex environmental cues in nature, performing experiments on phenotypic plasticity in ecologically relevant conditions would be complementary to controlled conditions where discrete levels of a given abiotic or biotic factor are tested (Baythavong and Stanton 2010; Bergelson and Roux 2010; Sultan 2010; Valladares *et al.* 2007). Second, theory predicts that phenotypic plasticity will be more favored by temporal environmental heterogeneity than by spatial environmental heterogeneity (Moran 1992). To our knowledge, the cost/benefit trade-off of plasticity in temporally heterogeneous natural environments was estimated in only few studies (Donohue *et al.* 2005). Third, because the strength and nature of selection acting on a specific trait may vary among natural populations (Kronholm *et al.* 2012; Brachi *et al.* 2013b), it would be worth testing whether a similar pattern of variation among different genetic pools can be observed for phenotypic plasticity (Villoutreix *et al.* submitted data). Fourth, while a very small cost may theoretically counteract the selection of phenotypic plasticity in presence of an optimal environmental heterogeneity (Moran 1992; Sultan and Spencer 2002), empirical evidence for a significant cost is still scarce. However, van Kleunen and Fischer (2005) argued that the

detection of cost of phenotypic plasticity may be improved by considering (i) signed values of plasticity rather than absolute ones and (ii) the covariance structure among traits, their plasticity and fitness using genotypic selection gradient analysis (Lande and Arnold 1983; Baythavong and Stanton 2010).

Here, we aimed to address these four issues by studying the adaptive value and cost of phenotypic plasticity across two time scales in *A. thaliana* using four regional sets of natural families. In a field experiment, nine traits related to resource acquisition, phenology and shoot architecture were measured for 800 families collected from 49 stands located in four climatically contrasted regions in France. Seeds of these 800 families were sown on three dates in the same common garden, simulating two germination cohorts within the same year and two climatically contrasted years within the same germination cohort (Villoutreix *et al.* submitted data). In this study, we first described natural genetic variation of the nine traits, their reaction norms (in particular the frequency of signed values of plasticity) and their interrelationships across cohorts and years. By performing genotypic selection gradients analyses based on total silique length as a proxy for fitness (previously described in Villoutreix *et al.* submitted data), we then tested whether (i) the nature and strength of selection acting on a specific trait depend on the sowing date and geographical region, and (ii) such a dependence for the selection pattern is also observed for plasticity. Finally, we investigated whether the relative frequency of cost of phenotypic plasticity depends on the time scale considered.

Materials and Methods

Plant material and experimental design

Plant material, growing conditions and experimental design of the field experiment have been thoroughly described in a previous study (Villoutreix *et al.* submitted data). Briefly, we used a set of 800 genotypic families sampled from 49 stands distributed across four French climatically contrasted regions (*i.e.* Brittany: oceanic, 11 stands; Burgundy: continental, 11 stands; Languedoc, Mediterranean, 16 stands and North: semi-oceanic, 11 stands). We phenotyped on average 16.3 (SE: ±5.6) families per stand, for a total of 200 families per region. A set of 188 worldwide accessions (WA) and a set of 210 French accessions (FA) corresponding to the French RegMap (Horton *et al.* 2012) were also included in the experiment (Brachi *et al.* 2013b). These 398 accessions have been measured for the purpose of another study. They were therefore discarded from further analyses in this study.

Seeds were sown at the University of Lille (Nord, France) on three dates (30th of August 2010, the 27th of September 2010 and the 26th of September 2011) to mimic (i) two distinct germination cohorts within the same year (late August 2010 and late September 2010) and (ii) the same germination cohort between two consecutive years (late September 2010 and late September 2011). The entire experiment of 7,416 plants followed a split-plot design arranged as a randomized complete block design (RCBD) with three sowing dates (hereafter named ‘August 2010’, ‘September 2010’ and ‘September 2011’) nested within two experimental blocks. Each ‘block*sowing date’ combination was represented by 19 arrays of 66 individual wells, for a total of 1,236 plants with one replicate per family (n = 800), one replicate per natural accessions (n= 398) and a control worldwide accession Bg-2 placed in the same two positions within each array to control for micro-environmental variation within

blocks (n = 38 = 19 trays * 2 replicates). In each 'block*sowing date' combination, the remaining 18 wells were left empty. The 'August 2010', 'September 2010', 'September 2011' sowing dates stopped on the 27th of May 2011, the 27th of May 2011 and the 10th of June 2012, respectively.

Phenotypic characterization

During the growing period, plants were monitored (i) daily for germination date (opening of both cotyledons) from the sowing date to 13 days after sowing and (ii) every two or three days for bolting date (inflorescence distinguishable from the leaves at a size < 5 mm), flowering date (appearance of the first open flower) and date of maturation of the last fruit. Based on these four phenological dates, four phenological traits were calculated (Brachi *et al.* 2012). Germination time (GERM) was measured as the number of days between the end of the stratification treatment (*i.e.* four days after the sowing date; see Villoutreix *et al.* submitted data) and germination date. Bolting time (BT), flowering interval (INT) and the reproductive period (RP) were scored as the time interval between germination date and bolting date, between bolting date and flowering date and between flowering date and date of maturation of the last fruit, respectively. Because climatic conditions fluctuated across the two years of this study, BT, INT and RP were scaled in photothermal units (PTU) using a phenological model integrating both photoperiod length and temperature as described in Brachi *et al.* (2010). At bolting, the maximum diameter of the rosette measured at the nearest millimeter was used as a proxy for plant size (DIAM; Weinig *et al.* 2006). After maturation of the last fruit, the aboveground portion was harvested and stored at room temperature until phenotyping of the following traits described as non-collinear and related to adaptation in *A. thaliana* (Reboud *et al.* 2004, Weinig *et al.* 2005): the number of basal branches (RAMBB), the number of

primary branches with siliques on the main stem (RAMPB_S), the height from soil to the first silique on the main stem (H1S) and maximum height of the plant (HMAX). Plant fitness was approximated by total fruit length (FITTOT) which has been shown to be a good proxy of lifetime fitness for a selfing annual like *A. thaliana* (Roux *et al.* 2004, Roux *et al.* 2005). FITTOT has been thoroughly analyzed in a previous study (Villoutreix *et al.* submitted data) and will be used here for genotypic selection gradients (see below).

Data Analysis

Natural variation and norms of reaction

Using the 800 families, we studied natural variation of nine traits (GERM, BT, INT, RP, DIAM, RAMBB, RAMPB_S, H1S and HMAX) and their reaction norms using the following hierarchical linear mixed model (HLMM) as previously described in Villoutreix *et al.* (submitted data):

$$Y_{ijklmn} = \mu_{trait} + block_i + sowing\ date_j + block_i*sowing\ date_j + region_k + stand_l\ (region_k) + family_m\ (stand_l\ (region_k)) + sowing\ date_j*region_k + sowing\ date_j*stand_j\ (region_k) + sowing\ date_j*\ family_m\ (stand_l\ (region_k)) + control_n\ (sowing\ date_j) + \varepsilon_{ijklmc} \quad (1)$$

In these models, '*Y*' is one of the nine phenotypic traits, 'μ' is the overall mean; 'block' accounts for differences in micro-environment among the two experimental blocks; 'sowing date' corresponds to either the 'cohort' (*i.e.* two cohorts within the same year: 'August 2010' *vs* 'September 2010' sowing dates) or 'year' effect (*i.e.* two years within the same germination cohort: 'September 2010' *vs.* 'September 2011' sowing dates); 'region',

'stand' and 'family' measure the effect of three spatial scales in France, *i.e.* regions, stands within regions and families within stands; interaction terms involving the 'sowing date' term account for genetic variation in reaction norms among regions, among stands within regions and among families within stands; 'control' is a covariate accounting for array effects within blocks within each sowing date (phenotypic mean of two control replicates Bg-2 per array); and 'ε' is the residual term. All factors were treated as fixed effects, except the terms involving 'family' which were treated as random effects. For fixed effects, terms were tested over their appropriate denominators for calculating *F*-values. Significance of the random effects was determined by likelihood ratio tests of model with and without these effects. Model fitting was conducted using the PROC MIXED procedure in SAS 9.3 (SAS Institute Inc., Cary, North Carolina, USA). Raw data were either log transformed or Box-Cox transformed to satisfy the normality and equal variance assumptions of linear regression.

Natural variation within each sowing date was studied by running model (1) excluding the terms involving 'sowing date$_j$'.

Estimation of family genotypic means

Because *A. thaliana* is a highly selfing species, the genotypic means (*i.e.* family means) were estimated for each trait within each sowing date by extracting the LSmeans of the following model (lm function; R 2.12.1 environment) and treating all factors as fixed effects:

$$Y_{ijc} = \mu_{trait} + block_i + family_j + control_c + \varepsilon_{ijc} \qquad (2)$$

Genotypic means for FITTOT were previously estimated in Villoutreix *et al.* (submitted data).

Estimation of phenotypic plasticity

Plasticity of a given family was calculated as the difference of the family genotypic mean between two sowing dates. For each of the nine traits, plasticities across germination cohorts were calculated as $LSmean_{August2010}$ - $LSmean_{September2010}$. Similarly, plasticities across years were calculated as $LSmean_{September\ 2010}$ - $LSmean_{September2011}$. Because families did not respond in the same direction across two sowing dates (see Results section), we used signed values of phenotypic plasticity instead of absolute values of phenotypic plasticity (as advised by van Kleunen and Ficher, 2005).

Relationships among traits

To test if relationships among traits were conserved between pairs of regions within each sowing date and across sowing dates in a given treatment (*i.e.* cohort and year effect), we first estimated a Spearman correlation matrix based on family genotypic means for each 'geographical unit (n = 5 = France and four regions) * sowing date (n = 3)' combination. We then performed non-parametric Mantel tests of the Spearman correlation between two matrices. The significance of the Mantel statistics was evaluated by 10,000 permutations (vegan 1.17-10 package in R 2.12.1 environment). Because a significant correlation between two correlation matrices may still be different from one (*i.e.* relationships among traits are not completely conserved), we then performed non-parametric Mantel tests of the correlation between two dissimilarity matrices. Dissimilarity matrices were obtained by calculating the dissimilarity coefficients between each pair of traits. Dissimilarity coefficients are signed coefficients and were calculated as follow:

For positive rho: dissimilarity coefficient = 1 – rho (3)

For negative rho: dissimilarity coefficient = -1 – rho (4)

The significance of the Mantel statistics was evaluated by 10,000 permutations (vegan 1.17-10 package in R 2.12.1 environment).

The same procedure was used to test whether relationships among plasticities was conserved between pairs of regions in a given treatment and between the two treatments in a given region.

Genotypic gradients of selection within each sowing date

To test for linear and non-linear selection acting on the nine traits in each 'geographical unit* sowing date' combination, genotypic selection analyses based on family genotypic means calculated within each sowing date were performed with the following polynomial regression (lm function; R 2.12.1 environment):

$$\text{Relative fitness}_i = \mu_{\text{relative fitness}} + \text{trait1}_i + \text{trait1}_i^2 + \ldots + \text{trait9}_i + \text{trait9}_i^2 + \varepsilon_i \quad (5)$$

Where 'Relative fitness' is the relative fitness within each 'geographical unit * sowing date' combination calculated as the fitness estimate divided by the mean fitness estimate within that combination, 'μ' is the constant, 'trait1' to 'trait9' correspond to the nine phenotypic traits standardized within each 'geographical unit * sowing date' combination, and ε is the residual term. Linear selection gradients were estimated by linear partial regression coefficients whereas quadratic regression coefficients were doubled to correctly estimate non-linear selection gradients (Stinchcombe *et al.* 2008).

In order to discriminate between curvilinear and stabilizing/disruptive selection acting on a specific trait, the point of null derivative was calculated individually for each trait as follow:

$$Point\ of\ null\ derivative\ = \frac{-a}{2b} \quad (6)$$

where 'a' and 'b' correspond to the linear and quadratic regression coefficients of the trait in equation (5), respectively. If (i) the point of null derivative was contained within the 95% interval genotypic distribution of the trait of interest and (ii) 'b' was significantly positive and negative, selection acting on a trait was considered disruptive and stabilizing, respectively. If the point of null derivative was not contained within the 95% interval genotypic distribution of the trait of interest, selection was considered curvilinear.

For each geographical unit, differences in linear and non-linear selection gradients across cohorts and years were tested using analyses of covariance (ANCOVA; R 2.12.1 environment). Significant 'trait * sowing date' and 'trait² * sowing date' interactions indicate varying directional and non-linear selection between sowing dates, respectively. Similarly, for each sowing date, differences in linear and quadratic selection gradients among the four regions were tested using analyses of covariance (ANCOVA; R 2.12.1 environment). Significant 'trait * region' and 'trait² * region' interactions indicate varying directional and non-linear selection among regions, respectively.

Adaptive value of plasticity

To test for linear and non-linear selection acting on the nine trait plasticities in each 'geographical unit (n = 5) * treatment (n = 2 = cohort and year)' combination, genotypic selection analyses based on family genotypic means of plasticity were performed with the following polynomial regression (lm function; R 2.12.1 environment):

$$\text{Relative mean fitness}_i = \mu_{\text{relative mean fitness}} + {}_{\text{pl}}\text{trait1}_i + {}_{\text{pl}}\text{trait1}_i^2 + \ldots + {}_{\text{pl}}\text{trait9}_i + {}_{\text{pl}}\text{trait9}_i^2 + \varepsilon_i \qquad (7)$$

where 'Relative mean fitness' is the relative fitness within each 'geographical unit * treatment' combination calculated as the fitness mean over sowing dates divided by the mean of the fitness mean over sowing dates within that combination, 'μ' is the constant, '${}_{\text{pl}}$trait1' to '${}_{\text{pl}}$trait9' correspond to the nine phenotypic plasticities standardized within each 'geographical unit * treatment' combination, and ε is the residual term. The method described in the 'Genotypic gradients of selection within each sowing date' subsection was used to discriminate between curvilinear and stabilizing/disruptive selection acting on a specific plasticity.

Because the adaptive value of phenotypic plasticity may depend on the frequencies with which a plant may experience one of the two environments (Gomulkiewicz and Kirkpatrick 1992, van Kleunen and Fischer 2005, Donohue 2005), the mean fitness over sowing dates in a given treatment was weighted for each family by the frequency to encounter one of the two environments (ranging from 1% to 99% with an increment of 1% in August

2010 in the cohort treatment and in September 2010 in the year treatment). For each simulated frequency, equation (7) was then run to estimate linear and quadratic selection gradients.

Cost of plasticity

To test for a cost of plasticity in each focal sowing date, we applied for each 'geographical unit * sowing date' combination the same formula described in DeWitt *et al.* (1998) in a given treatment, but expressed in a multiple trait context (lm function; R 2.12.1 environment):

$$\text{Relative focal fitness}_i = \mu_{\text{relative focal fitness}} + \text{trait1}_i + \text{trait1}_i^2 + {}_{\text{pl}}\text{trait1}_i + {}_{\text{pl}}\text{trait1}_i^2 + \ldots + {}_{\text{pl}}\text{trait9}_i + \text{trait9}_i^2 + \text{trait9}_i + {}_{\text{pl}}\text{trait9}_i^2 + \varepsilon_i \quad (8)$$

Where 'Relative focal fitness' is the relative fitness in the focal sowing date within each 'geographical unit * sowing date' combination calculated as the fitness estimate in the focal sowing date divided by the mean fitness estimate in the focal environment within that combination, 'μ' is the constant, 'trait1' to 'trait9' correspond to the nine phenotypic traits in the focal sowing date standardized within each 'geographical unit * sowing date' combination, '$_{\text{pl}}$trait1' to '$_{\text{pl}}$trait9' correspond to the nine phenotypic plasticities across sowing dates in the focal treatment standardized within each 'geographical unit * treatment' combination, and ε is the residual term. The method described in the 'Genotypic gradients of selection within each sowing date' subsection was used to discriminate between curvilinear and stabilizing/disruptive cost of plasticity.

Results

Natural variation and norms of reaction at different spatial scales

Extensive natural variation for the nine traits was detected within each sowing date at the three spatial scales, in particular at the stand scale with all traits being on average significantly different among stands within a region (Fig. 1, Table 1, Table S1, Supporting information). In a given treatment (*i.e.* cohort and year effect), the significance of phenotypic differences among regions and among families within stands was dependent on the sowing date considered, such as INT at the regional scale and RAMPB_S at the family scale in the cohort treatment or BT at the regional scale and DIAM at the family scale in the year treatment (Table S1, Supporting information).

Highly significant genetic variation was detected for plasticity across cohorts and years at different spatial scales (Fig. 1, Table 1). Strong genotype-by-environment interactions were often coupled with a non-significant cohort or year effect, indicating that families had opposite response across sowing dates in a given treatment (Fig. 1, Table 1). More precisely, values corresponding to the 95% confidence intervals of phenotypic plasticity distribution were of opposite signs for all traits in each 'geographical unit (n = 5) * treatment (cohort and year)' combination, with the exception of germination time in Burgundy in the year treatment and bolting time in all combinations.

Effect of region and sowing date on relationships among traits and among plasticities

For each pair of regions, trait correlation matrices and plasticity correlation matrices were significantly similar within each sowing date and within each treatment, respectively (*i.e.* Spearman correlation between two matrices > 0; Table S2, Supporting information). At

the same time, ~55% of the correlations between two trait correlation matrices and ~58% of the correlations between two plasticity correlation matrices were significantly different from 1 (Table S2, Supporting information). The latter result suggests that relationships among traits and among plasticities were not completely conserved among regions.

For each geographical unit (*i.e.* France and within each region), trait correlation matrices were similar across cohorts and years. At the same time, 80% of the correlations between two trait correlation matrices were significantly different from 1, suggesting that relationships among traits were not completely conserved across cohorts and years (Table S2, Supporting Information). Plasticity correlation matrices were similar across treatments for each geographical unit (with the exception of the North region), but only one correlation between two plasticity correlation matrices was not significantly different from 1. This later result suggests that relationships among plasticities were not completely conserved between the two treatments (Table S2, Supporting Information).

Genotypic gradients of selection within each sowing date

For each 'geographical unit * sowing date' combination, a large portion of relative fitness variation can be explained by the nine phenotypic traits measured in this study, as indicated by high R^2 values (mean = 0.76, SE: ± 0.06; Fig. 2).

For the branching traits RAMBB and RAMPB_S, significant strong directional selection for a higher number of basal branches and primary branches was detected in each 'geographical unit * sowing date' combination (Fig. 2). Significant directional selection was frequently detected for BT, DIAM and HMAX. Earlier bolting, bigger rosette diameter at bolting and higher maximal plant height were favored in nine, eleven and thirteen of the

fifteen ‘geographical unit * sowing date’ combinations, respectively (Fig. 2). Rosette diameter at bolting was frequently detected under disruptive selection, suggesting the existence of two phenotypic optima for this trait. Selection for earlier germination, shorter interval between bolting and flowering, longer and intermediate reproductive period and intermediate height from soil to the first silique on the main stem was specific to few ‘geographical * sowing date’ combinations (Fig. 2).

Significant differences in selection intensity across cohorts were detected at the Bonferonni threshold for (i) RAMBB, with selection for a higher number of basal branches being stronger in September 2010 than in August 2010 in France and Brittany, and (ii) BT, with selection for earlier flowering being stronger in August 2010 than in September 2010 in North (Fig. 3A, Table S3A, Supporting information). Variation in selection intensity across years was only detected for HMAX, with selection for taller plants being stronger in September 2011 than in September 2010 in France, Burgundy and Languedoc (Fig. 3B, Table S3B, Supporting information). Altogether, these results suggest a general pattern of selection which is similar across the three sowing dates.

At the Bonferonni threshold, no significant variation in linear and quadratic selection gradients was detected among regions (Table S4, Supporting information), suggesting similar intensities of selection acting on the nine traits across France.

Adaptive value of phenotypic plasticity

Across the 45 ‘geographical unit (n = 5) * plasticity trait (n = 9)’ combinations, the fraction of significant linear and/or non-linear selection detected in at least one simulated frequency of environment was 71.2% and 57.8% in the cohort and year treatments,

respectively (Fig. 4). No clear pattern of selection acting on phenotypic plasticity can be described, with the intensity and nature of selection acting on phenotypic plasticity being dependent on a complex interaction among the identity of the plastic trait, geographical unit, frequency to encounter one of the two environments and treatment.

The sign of selection gradients was rarely kept constant across the entire range of simulated environmental frequencies, concerning only six (two) and three (one) linear (quadratic) selection gradients in the cohort and year treatments, respectively. In Languedoc, families having more primary branches and a shorter height from soil to the first silique on the main stem in August 2010 than in September 2010 were favored in the cohort treatment, whatever the frequency to encounter the August 2010 environment (Fig. 5A). Similarly, families having a bigger rosette diameter at bolting in Burgundy and accumulating much more photothermal units for bolting in North in September 2010 than in September 2011 were favored in the year treatment, whatever the frequency to encounter the September 2010 environment (Fig. 5B).

In most cases, the sign of directional and quadratic selection gradients varied across the entire range of simulated environmental frequencies from significant to non-significant (Fig. 4). We also observed extreme cases of directional selection having opposite signs between extremes simulated frequencies, highlighting the importance to consider the frequency in which the families occurs in each environment. In the cohort treatment in Burgundy, directional selection favored families having smaller and bigger rosette diameter in August 2010 than in September 2010 when the frequency to encounter the August 2010 environment was 1% and 99%, respectively (Fig. 6A). In the year treatment, directional selection favored families from Brittany being shorter and taller and families from North having less and more basal branches in September 2010 than in September 2010 when the

frequency to encounter the September 2010 environment was 1% and 99%, respectively (Fig. 6B).

Cost of plasticity

Across all the 90 'geographic unit (n = 5) * focal sowing date (n = 2) * plasticity trait (n = 9)' combinations, the fraction of significant linear and/or non-linear cost was 24.4% in the cohort treatment and 12.2% in the year treatment (Fig. 7). A significant stabilizing cost of plasticity was detected for bolting time in Brittany in September 2010 and for height from soil to the first silique in Languedoc in September 2011; whereas a significant disruptive cost of plasticity was detected for the number of basal branches in France in September 2010 and for the maximum height in Languedoc in September 2011. The intensity and shape of cost associated with plasticity was dependent on a complex interaction among plastic trait, geographical unit, treatment and sowing date within a treatment. In most cases, cost of plasticity for a given trait was only detected in only one of the two sowing dates in a given treatment. For example, in the cohort treatment, cost of plasticity for bolting time in Burgundy was only significant in August 2010, with families that accumulated less photothermal units for bolting in August 2010 than in September 2010 experiencing a cost in comparison with families that accumulated much more photothermal units for bolting (Fig. 8A). Similarly, in the year treatment, cost of plasticity for plant maximum height in Brittany was only significant in September 2010, with families that were shorter in September 2010 than in September 2011 experiencing a cost in comparison with families that were taller (Fig. 8B).

In few cases, the relationship between plasticity and fitness within a focal sowing date was of opposite sign between two sowing dates in a given treatment. For example, plasticity

for the number of primary branches on the main stem across cohorts in Brittany was significantly negatively and positively correlated with fitness in August 2010 and September 2010, respectively (Fig. 8A), with families that had less primary branches on the main stem in August 2010 than in September 2010 experienced an advantage in August 2010 but a cost in September 2010. Altogether, these results indicate that the cost experienced by a family was dependent on both the sign and intensity of plasticity it expressed and the focal environment.

Discussion

By replacing the phenotyping of 800 *A. thaliana* families collected from four French regions in ecologically realistic conditions, we studied the adaptive value and cost of phenotypic plasticity across two time scales. Because the timing of germination is known to affect the expression of post-germination life-history traits in *A. thaliana* (Donohue 2002; Donohue *et al.* 2005; Wilzeck *et al.* 2009; Huang *et al.* 2010), two germination cohorts were simulated within one year. Because yearly variation for flowering time was observed for worldwide accessions of *A. thaliana* in ecologically relevant conditions (Brachi *et al.* 2010; Brachi *et al.* 2013a), plants were grown for two consecutive years within the same germination cohort.

Prevalence of signed values of plasticity across time

As previously observed for the same set of 800 *A. thaliana* families scored for seed production related traits in the same field experiment (Villoutreix *et al.* submitted data), extensive natural genetic variation was observed across spatial scales for the nine traits related

to resource acquisition, phenology and shoot architecture. In addition, reaction norms strongly differed among stands, suggesting that the plastic response of the nine traits to either germination cohorts or years can evolve. In *A. thaliana*, natural genetic variation for temporal plasticity has already been detected for germination related traits scored on a RIL family in two dispersal seasons in a common garden located in Rhode Island (Donohue *et al.* 2005), flowering time scored on 473 worldwide accessions in two seasonal plantings (*i.e.* Spring and Summer) in each of two simulated climates (Spain and Sweden; Li *et al.* 2010), and for flowering time and the height of the main stem (but not for the number of branches) scored in an outbred population in two simulated germination cohorts (Spring and Fall; Scarcelli *et al.* 2007).

Relationships among traits and among plasticities were highly conserved among the four geographical regions, though they were often different from one. This latter result is consistent with a previous study in *A. thaliana* demonstrating that the pattern of co-variation among three phenological and vegetative growth related traits varies among geographical region at the continental scale (Debieu *et al.* 2013). Interestingly, relationships among traits appeared less conserved between sowing dates in a given treatment. This weak conservation is even more obvious for relationships among plasticities between the cohort and year treatments, which strengthens the idea that plasticity is a phenotypic trait by itself that may evolve not only independently of its expression in a given environment (Schlichting 1986; Schmitt et al. 1992; van Kleunen and Fischer 2005), but also independently from other plastic traits.

While in most studies all genotypes respond to an environmental change in the same direction (reviewed in van Kleunen and Fischer 2005), we found a prevalence of signed values of plasticity across time for all traits (with the exception of flowering time), with

families having opposite response across sowing dates within a given treatment. This discrepancy with previous studies may originate from (i) the high number of families collected from 49 contrasted ecological natural stands (Brachi *et al.* 2013b), (ii) the diversity of traits measured in our study, and (iii) complex abiotic and biotic cues sensed by plants under our ecologically relevant, experimental conditions. Three non-exclusive hypotheses may be advanced to explain the presence of opposite response among families. First, variation in magnitude of plastic response without inversion of family ranks may be associated with an equality of trait mean but varying genetic variance between two environments (*i.e.* scale effect; Lacaze *et al.* 2009). Because inversion in the family ranking across sowing dates (*i.e.* cross-over effect; Lacaze *et al.* 2009) was often observed in our study (Fig. 1), alternative hypotheses should be considered. Second, environmental cues sensed by families increasing trait value from one environment to a second environment may be different from environmental cues sensed by families decreasing trait value. If plastic response is adaptive, it would suggest that opposite plastic responses evolved in response to different selective agents. Third, some families may have not been exposed in their natural habitats to environmental cues sensed in this study. In turn, this reliability limit may have led to an inappropriate plastic response (DeWitt *et al.* 1998; Tufto 2000). In either of the latter two cases, we suggest that opposite plastic responses would be associated to different genetic bases.

Contrasted patterns of selection between traits and their plasticities

In agreement with previous studies performed in ecologically realistic conditions or simulating climate in growth chambers (Weinig *et al.* 2003; Weinig *et al.* 2006; Korves *et al.* 2007; Scarcelli *et al.* 2007; Fournier-Level *et al.* 2013), the main traits under selection across

the three sowing dates were related to resource acquisition (rosette diameter at bolting) and shoot architecture. Although the four French geographical regions are climatically contrasted (Brachi *et al.* 2013b), the nature and strength of selection acting on the nine traits were similar among the four French regions. This result suggests that the selective pattern may result from fine-scale variation of non-climate factors, like edaphic conditions and plant-plant interactions whose variance was mostly partitioned among stands within regions (Brachi *et al.* 2013b).

Based on an interconnected mapping population of 117 Recombinant Inbred Lines (RILs), the influence of bolting time on branch number was found to differ among four field sites across the native European range of *A. thaliana* (Fournier-Level *et al.* 2013). Variation in selection intensity on some traits also differed across cohorts and years in our study. Altogether, it appears that both spatial and temporal environmental variation may contribute to maintenance of natural genetic variation in *A. thaliana*, even at a small spatial scale.

Plasticities across cohorts and years revealed a different and complex selective pattern. The nature and strength of selection were dependent on an interaction among the identity of the plastic trait, geographical region and treatment. This result suggests that the traits involved in the response to an environmental cue may depend on the temporal environmental grain of this cue. Given the prevalence of signed values of plasticity detected in this study, significant associations between plasticity and mean fitness over sowing dates indicates that plasticity may be either adaptive or maladaptive. At the same time, the adaptive value of plasticity was strongly dependent on the frequency with which a plant may experience one of two sowing dates. Although the adaptive value of phenotypic plasticity is theoretically very dependent on the relative frequency of selective environments (Gomulkiewicz and Kirkpatrick 1992), this environmental frequency dependence has rarely been empirically tested (Donohue *et al.* 2005). For plastic traits having opposite sign in response to an environment, the extreme case

of directional selection that reverses between extreme simulated frequencies should promote the persistence of natural variation for opposite plasticities in *A. thaliana* in geographical regions where the frequency to experience one of the two sowing dates varies among stands.

Does cost of plasticity depend on the time scale?

Theoretical work indicates that the evolution of plasticity should be countered in presence of global costs (*i.e.* costs expressed in multiple environments) but not in presence of local costs (*i.e.* cost expressed in only one environment; *e.g.* Weining *et al.* 2006; Sultan and Spencer 2002). In our study, we failed to detect global cost of plasticity. In contrast, local cost of plasticity was sometimes associated with different signs of plasticity between the two focal environments (Fig. 8), suggesting that evolution of phenotypic plasticity should not be constrained in our set of 800 families. We must however acknowledge that (i) the identification of cost of phenotypic plasticity is specific to our ecological conditions and (ii) identified costs may result from the effect of unmeasured correlated traits on fitness.

While the relative frequency of cost of phenotypic plasticity was close to the one estimated in review of 207 analyses (van Kleunen and Fischer 2005), it was twice as much as important in the cohort treatment, suggesting that cost of plasticity may depend on the time scale. A possible explanation for this pattern may stem from genes associated with different intrinsic genetic costs. For example, when *A. thaliana* seeds germinate in late August, natural flowering time variation is potentially associated with the gene *FRIGIDA* (*FRI*; Wilczeck *et al.* 2009; our unpublished results) which has putative pleiotropic effects on leaf and shoot architecture related traits as well as on flowering/reproductive periods ratio (Scarcelli *et al.* 2007; Fournier-Level *et al.* 2013, Brachi *et al.* 2013b). When *A. thaliana* seed germinate in

late September, *TWIN SISTER OF FT* (*TSF*) but not *FRI* did associate significantly with flowering time variation in a Genome-Wide Association mapping study (Brachi *et al.* 2010; Brachi *et al.* 2013a). So far, no negative pleitropic effects have been identified for *TSF* (Roux *et al.* 2006). Based on a pooled population sequencing (Pool-Seq) approach, the on-going genomic characterization of the 49 stands will soon reveal whether the genomic regions associated with costly plastic traits overlap with the genomic regions associated with pleiotropic effects in a single sowing date.

Acknowledgments

Special thanks are given to Nathalie Faure, Stella Huynh, Benoit Bucher and Angélique Bourceaux for their assistance during the field experiments. This study was supported by a PhD fellowship from the University of Lille 1 to R.V.

References

Agrawal, A., 2001. Phenotypic plasticity in the interactions and evolution of species. Science 294:321-326.

Agrawal, A. A., J. K. Conner, M. T. J. Johnson, and R. Wallsgrove. 2002. Ecological genetics of an induced plant defense against herbivores: additive genetic variance and costs of phenotypic plasticity. Evolution 56:2206-2213.

Alpert, P., and E. L. Simms. 2002. The relative advantages of plasticity and fixity in different environments: when is it goof for a plant to adjust? Evol. Ecol. 16:285-297.

Baythavong, B. S., and M. L. Stanton. 2010. Characterizing selection on phenotypic plasticity in response to natural environmental heterogeneity. Evolution 64:2904-2920.

Bergelson, J., and F. Roux. 2010. Towards identifying genes underlying ecologically relevant traits in *Arabidopsis thaliana*. Nat. Rev. Genet. 11:867-879.

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:e32069.

Brachi, B., N. Faure, J. Bergelson, J. Cuguen, and F. Roux. 2013a. Genome-wide association mapping of flowering time in *Arabidopsis thaliana* in nature: genetics for underlying components and reaction norms across two successive years. Acta Bot. Gallica (DOI:10.1080/12538078.2013.807302)

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 Genet. 6:e1000940.

Brachi, B., R. Villoutreix, N. Faure, N. Hautekèete, Y. Piquot, M. Pauwels, D. Roby, J. Cuguen, J. Bergelson, and F. Roux. 2013b. Investigation of the geographical scale of adaptive phenological variation and its underlying genetics in *Arabidopsis thaliana*. Mol. Ecol. 22:4222-4240.

Bradshaw, A. D. 1965. Evolutionary significance of phenotypic plasticity in plants. Adv. Genet. 13:115-155.

Debieu, M., C. Tang, B. Stich, T. Sokosek, S. Effgen, E. Josephs, J. Schmitt, M. Nordborg, M. Koornneef, and J. de Meaux. 2013. Co-variation between seed dormancyn, growth rate and flowering time changes with latitude in *Arabidopsis thaliana*. PLoS One 8:e61075.

DeWitt, T. J., A. Sih, and D. S. Wilson. 1998. Costs and limits of phenotypic plasticity. Trends Ecol. Evol. 13:77-81.

Donohue, K. 2002. Germination timing influences natural selection on life-history characters in *Arabidopsis thaliana*. Ecology 83:1006-1016.

Donohue, K., L. Dorn, C; Griffith, E.-S. Kim, A. Aguilera, C. R. Polisetty, and J. Schmitt. 2005. The evolutionary ecology of seed germination of *Arabidopsis thaliana*: variable natural selection on germination timing. Evolution 59:758-770.

Donohue, K., E. H. Pyle, D. Messiqua, M. S. Heschel, and J. Schmitt. 2001. Adaptive divergence in plasticity in natural populations of *Impatiens capensis* and its consequences for performance in novel habitats. Evolution 55:692-702.

Dorn, L. A., E. H. Pyle, and J. Schmitt. 2000. Plasticity to light cues and resources in *Arabidopsis thaliana*: testing for adaptive value and costs. Evolution 54:1982-1994.

Fournier-Level, A., A. M. Wilczek, M. D. Cooper, J. L. Roe, J. Anderson, D. Eaton, B. T. Moyers, R. H. Petipas, R. N. Schaeffer, B. Pieper, M. Reymond, M. Koornneef, S. M.

Welch, D. L. Remington, and J. Schmitt. 2013. Paths to selection on life history loci in different natural environments across the native range of *Arabidopsis thaliana*. Mol. Ecol. 22:3552-3566.

Gomulkiewicz, R., and M. Kirkpatrick. 1992. Quantitative genetics and the evolution of reaction norms. Evolution 46:390-411.

Huang, X., J. Schmitt, L. Dorn, C. Griffith, S. Effgen, S. Takao, M. Koornneef, and K. Donohue. The earliest stages of adaptation in an experimental plant population: strong selection on QTLs for seed dormancy. Mol. Ecol. 19:1335-1351.

Kassen, R. 2002. The experimental evolution of specialists, generalists, and the maintenance of diversity. J. Evolution. Biol. 15:173-190.

Korves, T. M., K. J. Schmid, A. L. Caicedo, C. Mays, J. R. Stinchcombe, M. D. Purugganan, and J. Schmitt. 1997. Fitness effects associated with the major flowering time gene *FRIGIDA* in *Arabidopsis thaliana* in the field. Am. Nat. 169:E141-E157.

Kronholm, I., F. X. Pico, C. Alonso-Blanco, J. Goudet, and J. de Meaux. 2012. Genetic basis of adaptation in *Arabidopsis thaliana*: local adaptation at the seed dormancy QTL *DOG1*. Evolution 66: 2287-2302.

Lacaze, X., P. M. Hayes, and A. Korol. 2009. Genetics of phenotypic plasticity: QTL analysis in barley, *Hordeum vulgare*. Heredity 102:163-173.

Lande, R., and S. J. Arnold. 1983. The measurement of selection on correlated characters. Evolution 37:1210-1226.

Li, Y., Y. Huang, J. Bergelson, M. Nordborg, and J. O. Borevitz. 2010. Association mapping of local climate-sensitive quantitative trait loci in *Arabidopsis thaliana*. P. Natl. Acad. Sci. USA 107:21199-21204.

Matesanz, S., T. Horgan-Kobelski, and S. E. Sultan. 2012. Phenotypic plasticity and population differentiation in an ongoing species invasion. PLoS One 7:e44955.

Moran, N. 1992. The evolutionary maintenance of alternative phenotypes. Am. Nat. 139:971-989.

Nicotra, A. B., O. K. Atkin, S. P. Bonser, A. M. Davidson, E. J. Finnegan, U. Mathesius, P. Poot, M. D. Purugganan, C. L. Richards, F. Valladares, and M. van Kleunen. 2010. Plant phenotypic plasticity in a changing climate. Trends Plant Sci. 15:684-692.

Pigliucci, M. 2005. Evolution of phenotypic plasticity: where are we going now? Trends Ecol. Evol. 20: 481-486.

Reboud, X., and G. Bell. 1997. Experimental evolution in *Chlamydomonas*. III. Evolution of specialist and generalist types in environments that vary in space and time. Heredity 78:507-514.

Richards, C. L., O. Bossdorf, N. Z. Muth, J. Gurevitch, and M. Pigliucci. 2006. Jack of all trades, master of some? On the role of phenotypic plasticity in plant invasions. Ecol. Lett 9:981-993.

Roux, F., C. Camilleri, A. Bérard, and X. Reboud. 2005. Multigenerational versus single generation studies to estimate herbicide resistance fitness cost in *Arabidopsis thaliana*. Evolution 59:2264-2269.

Roux, F., J. Gasquez, and X. Reboud. 2004. The dominance of the herbicide resistance cost in several *Arabidopsis thaliana* mutant lines. Genetics 166:449-460.

Roux, F., P. Touzet, J. Cuguen, and V. Le Corre. 2006. How to be early flowering: an evolutionary perspective ? Trends Plant Sci. 11:375-381.

Scarcelli, N., J M. Cheverud, B. A. Schaal, and P. X. Kover. 2007. Antagonistic pleiotropic effects reduce the potential adaptive value of the *FRIGIDA* locus. P. Natl. Acad. Sci. USA 104:16986-16991.

Scheiner, S. M. 1993. Genetics and evolution of phenotypic plasticity. Annu. Rev. Ecol. Syst 24:35-68.

Schlichting, C. D. 1986 The evolution of phenotypic plasticity in plants. Annu. Rev. Ecol. Syst. 17:667-693.

Schlichting, C. D., and H. Smith. 2002. Phenotypic plasticity: linking molecular mechanisms with evolutionary outcomes. Evol. Ecol. 16:189-211.

Schmitt, J., J. Niles, and R. D. Wulff. 1992. Norms of reaction of seed traits to maternal environments in *Plantago lanceolata*. Am. Nat. 139:451-466.

Stinchcombe, J. R., A. F. Agrawal, P. A. Hohenlohe, S. J. Arnold, and M. W. Blows. 2008. Estimating nonlinear selection gradients using quadratic regression coefficients: double or nothing? Evolution 62: 2435-2440.

Stomp, M., M. A. van Dijk, H. M. J. van Overzee, M. T. Wortel, C. A. M. Sigon, M. Egas, H. Hoogveld, H. J. Gons, and J. Huisman. 2008. The time scale of phenotypic plasticity and its impact on competition in fluctuating environments. Am. Nat. 172:169-185.

Sultan, S. E. 1987. Evolutionary implications of phenotypic plasticity in plants. Evol. Biol. 21:127-178.

Sultan, S. E. 2000. Phenotypic plasticity for plant development, function and life history. Trends Plant Sci. 5:537-542.

Sultan, S. E. 2010. Plant developmental response to the environment: eco-devo insights. Curr. Opin. Plant Biol./ 13:96-101.

Sultan, S. E., and H. G. Spencer. 2002. Metapopulation structure favors plasticity over local adaptation. Am. Nat. 160:271-283.

Tonsor, S. J., T. W. Elnaccash, and S. M. Scheiner. 2013. Developmental instability is genetically correlated with phenotypic plasticity, constraining heritability, and fitness. Evolution 67:2923-2935.

Tufto, J. 2000. The evolution of plasticity and nonplastic spatial and temporal adaptations in the presence of imperfect environmental cues. Am. Nat. 156:121-130.

Valladares, F., E. Gianoli, and J. M. Gomez. 2007. Ecological limits to plant phenotypic plasticity. New Phytol. 176:749-763.

van Kleunen, M., and M. Fischer. 2001. Adaptive evolution of plastic foraging responses in a clonal plant. Ecology 82:3309-3319.

van Kleunen, M., and M. Fischer. 2005. Constraints on the evolution of adaptive phenotypic plasticity in plants. New Phytol. 166:49-60.

Villoutreix, R., C. Glorieux, B. Brachi, J. Bergelson, J. Cuguen, and F. Roux. Evidence of selection acting on reproductive strategies in *Arabidopsis thaliana* at different spatial scales: when temporal heterogeneity in the field matters! (submitted data).

Weinig, C., J. Johnston, M. German, and L. M. Demink. 2006. Local and global costs of adaptive plasticity to density in *Arabidopsis thaliana*. Am. Nat. 167:826-836.

Weinig, C., J. R. Stinchcombe, and J. Schmitt. 2003. Evolutionary genetics of resistance and tolerance to natural herbivory in *Arabidopsis thaliana*. Evolution 57:1270-1280.

Wilczek, A. M., J. L. Roe, M. C. Knapp, M. D. Cooper, C. Lopez-Gallego, L. J. Martin, C. D. Muir, S. Sim, A. Walker, J. Anderson, J.F. Egan, B. T. Moyers, R. Petipas, A. Giakountis, E. Charbit, G. Coupland, S. M. Welch, and J. Schmitt. 2009. Effects of genetic perturbation on seasonal life history plasticity. Science 323:930-934.

Figure Legends

Fig. 1 Natural variation of nine phenotypic traits and their reaction norms at three spatial scales (*i.e.* region, stand and family). Open diamond: Brittany, open triangle: Burgundy, open square: Languedoc, open circle: North. At the scale of France, each value corresponds to the median of the LSmeans of 200 families per region for nine phenotypic traits. Within each region, each value corresponds to the LS means of families per stand for the nine phenotypic traits.

Fig. 2 Genotypic selection analysis with linear (trait) and non-linear (trait²) selection gradients for nine phenotypic traits in each 'geographical unit * sowing date' combination. The colored squares for traits indicate significant regression coefficients (p-value < 0.05). Red and blue colors represent the strength of the positive and negative regression coefficients, respectively. 'D' and 'S' stand for disruptive and stabilizing selection, respectively. R^2: portion of relative fitness explained by the nine traits.

Fig. 3 Differences in linear and non-linear genotypic selection gradients across sowing dates. (A) 'August 2010' *vs.* 'September 2010' sowing dates. (B) 'September 2010' *vs.* 'September 2011' sowing dates. Black squares, blue circles and red diamonds correspond to the families for 'August 2010', 'September 2010' and 'September 2011' sowing dates, respectively. For illustration purposes, a polynomial regression including both linear and quadratic terms was first run including all traits but the trait of interest. Then, a second polynomial regression including the linear and quadratic terms associated with the trait of interest was run on the residual fitness of the first polynomial regression. The fitted lines were drawn using the

parameters from this second polynomial regression. Traits are expressed in standardized values.

Fig. 4 Linear and non-linear genotypic selection gradients for plasticity of nine phenotypic traits for each 'geographical unit * treatment' combination across a range of simulated environmental frequencies. Black squares, circles, peak-up triangles and peak-down triangles stand for plasticity of germination, bolting, bolting to flowering interval and reproductive period, respectively. Blue squares stand for plasticity of rosette diameter at bolting. Red squares, circles, peak-up triangles and peak-down triangles stand for plasticity of number of basal branches, number of primary branches on the main stem, height from soil to the first silique on the main stem and maximum height, respectively. Beta (linear) and gamma (non-linear) selection coefficients are displayed in distinct graphics.

Fig. 5 No change in linear selection gradients across the entire range of simulated environmental frequencies. (A) Cohort treatment. (B) Year treatment. Dotted lines display the 95 % interval of plasticity values. For illustration purposes, a polynomial regression including both linear and quadratic terms was first run including all traits but the trait of interest. Then, a second polynomial regression including the linear and quadratic terms associated with the trait of interest was run on the residual fitness of the first polynomial regression. The black lines were drawn using the parameters from this second polynomial regression. Traits are expressed in standardized values. Red line displays null raw plasticity.

Fig. 6 Dependence on simulated environmental frequencies for genotypic selection gradients of plasticity at 1 % and 99% simulated environmental frequencies. (A) Cohort treatment with 1% and 99% simulated frequencies to encounter 'August 2010'. (B) Year treatment with 1% and 99% simulated frequencies to encounter 'September 2010'. Dotted lines display the 95 % interval of plasticity values. For illustration purposes, a polynomial regression including both linear and quadratic terms was first run including all traits but the trait of interest. Then, a second polynomial regression including the linear and quadratic terms associated with the trait of interest was run on the residual fitness of the first polynomial regression. The black lines were drawn using the parameters from this second polynomial regression. Traits are expressed in standardized values. Red line displays null raw plasticity.

Fig. 7 Linear and non-linear genotypic cost gradients of plasticity of nine phenotypic traits for each 'geographical unit * sowing date' combinations. The colored squares for traits indicate significant regression coefficients (p-value < 0.05). Red and blue colors represent the strength of the positive and negative regression coefficients, respectively. 'D' and 'S' stand for disruptive and stabilizing selection, respectively.

Fig. 8 Variation of genotypic cost gradients of plasticity across sowing dates within the (A) cohort and (B) year treatments. Dotted lines display the 95 % interval of plasticity values. For illustration purposes, a polynomial regression including both linear and quadratic terms was first run including all traits but the trait of interest. Then, a second polynomial regression including the linear and quadratic terms associated with the trait of interest was run on the residual fitness of the first polynomial regression. The black lines were drawn using the

parameters from this second polynomial regression. Traits are expressed in standardized values. Red line displays null raw plasticity.

Fig. 1

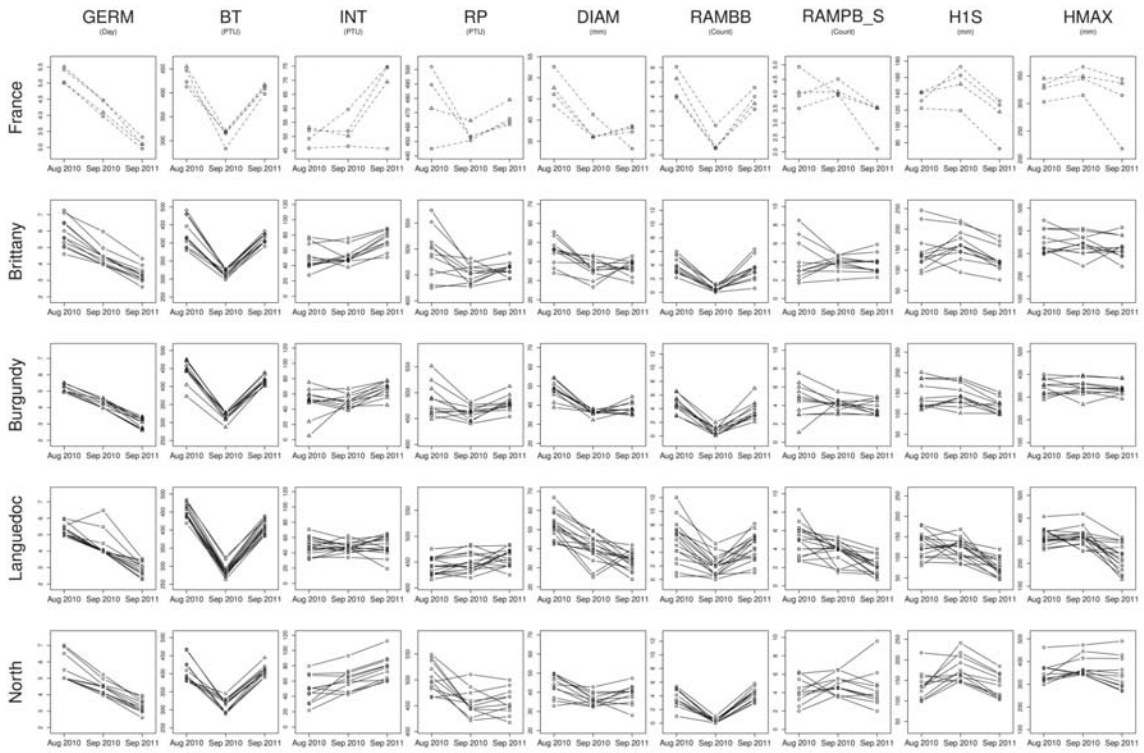


Fig. 2

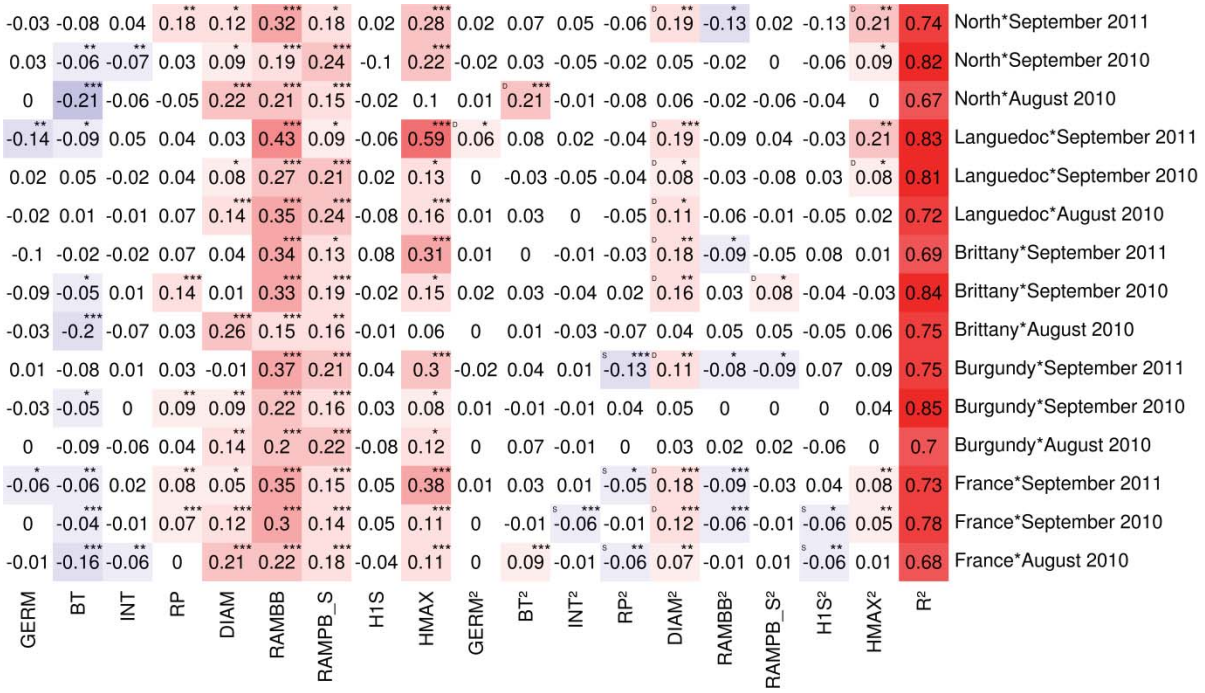


Fig. 3

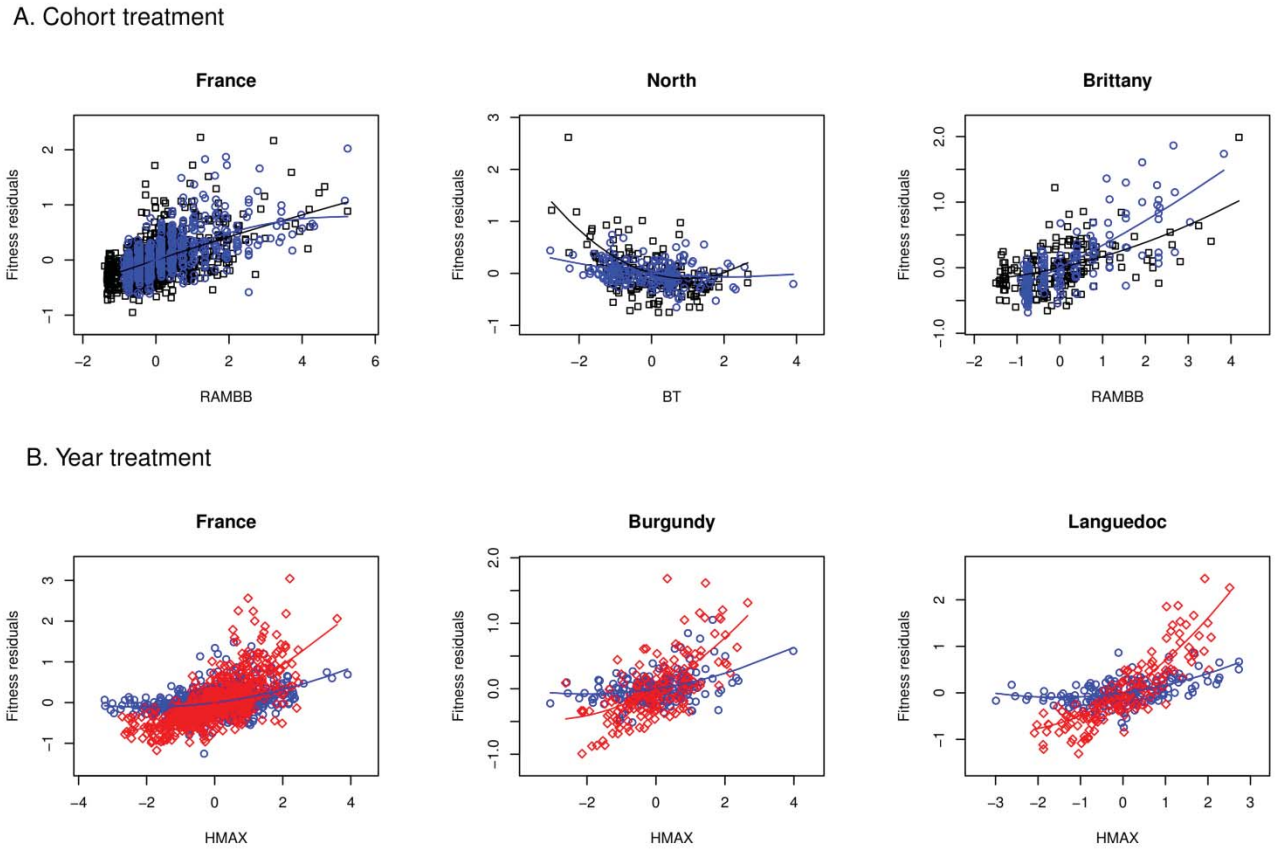


Fig. 4

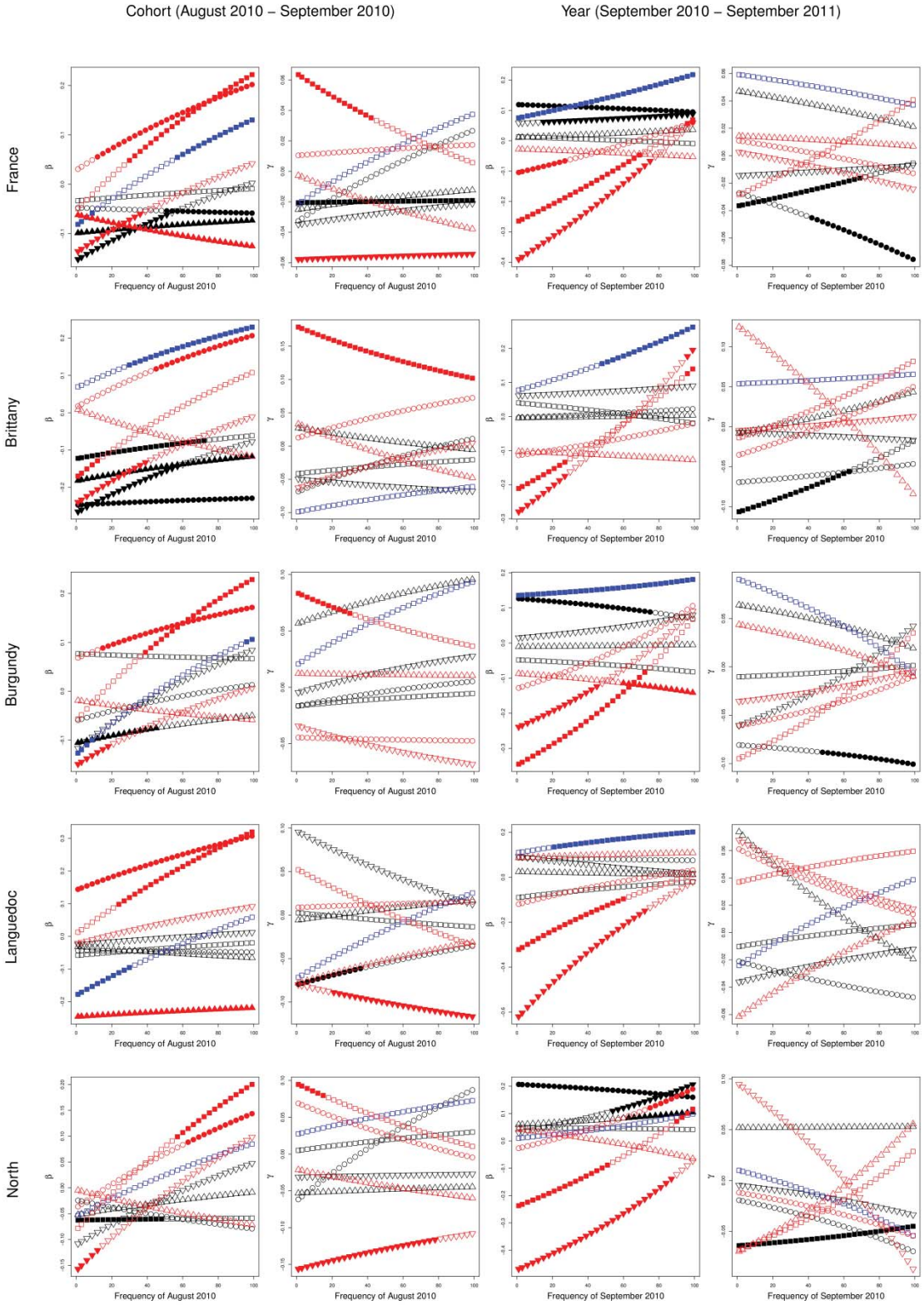


Fig. 5

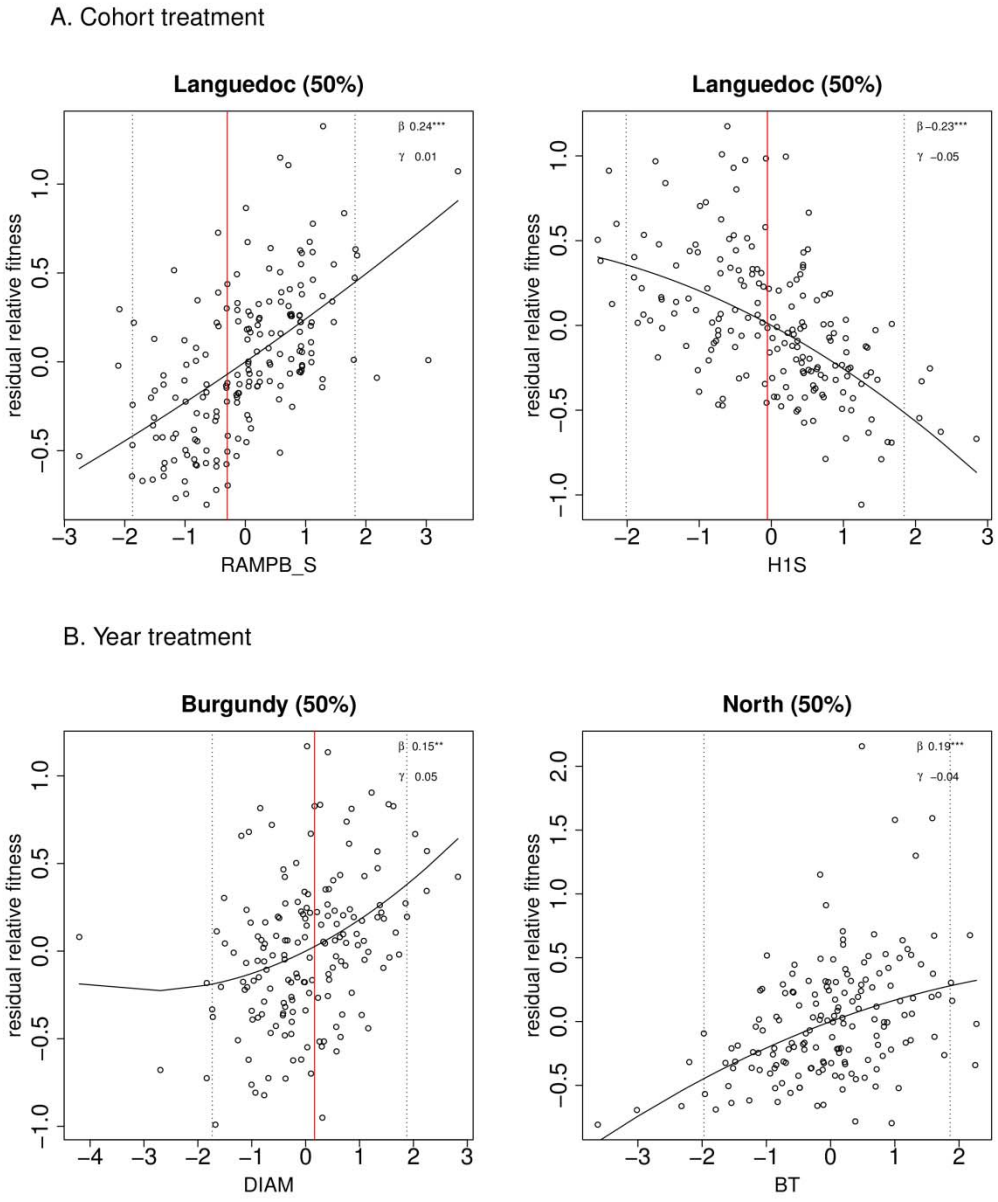


Fig. 6

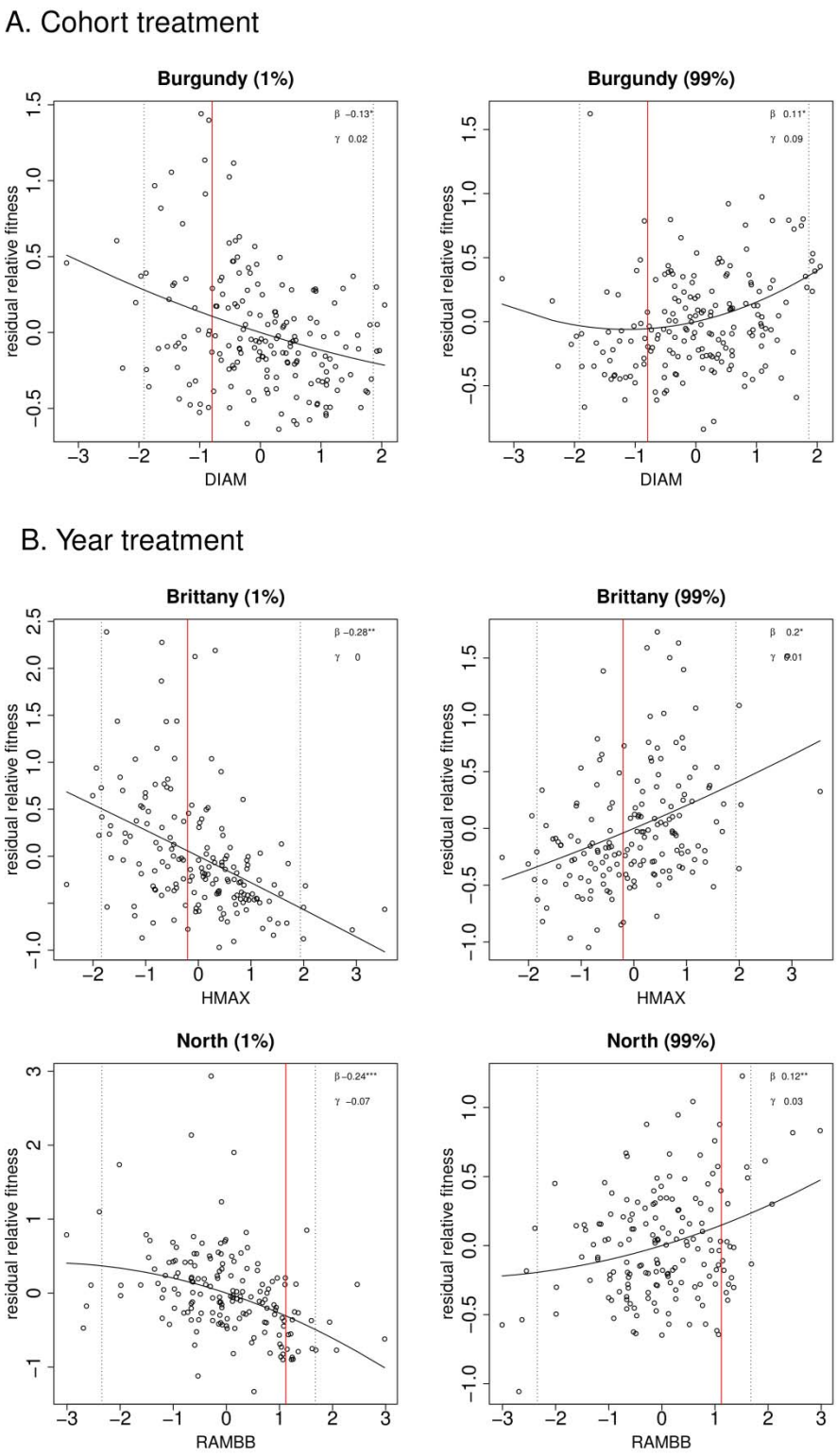


Fig. 7

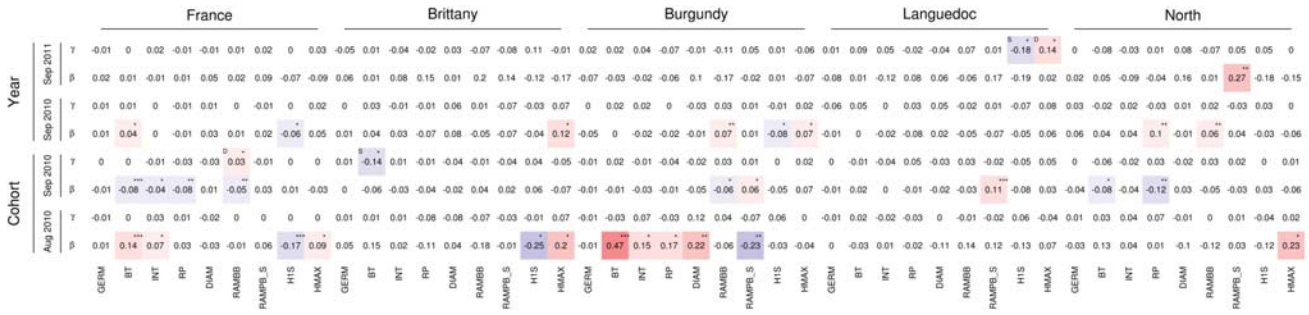


Fig. 8

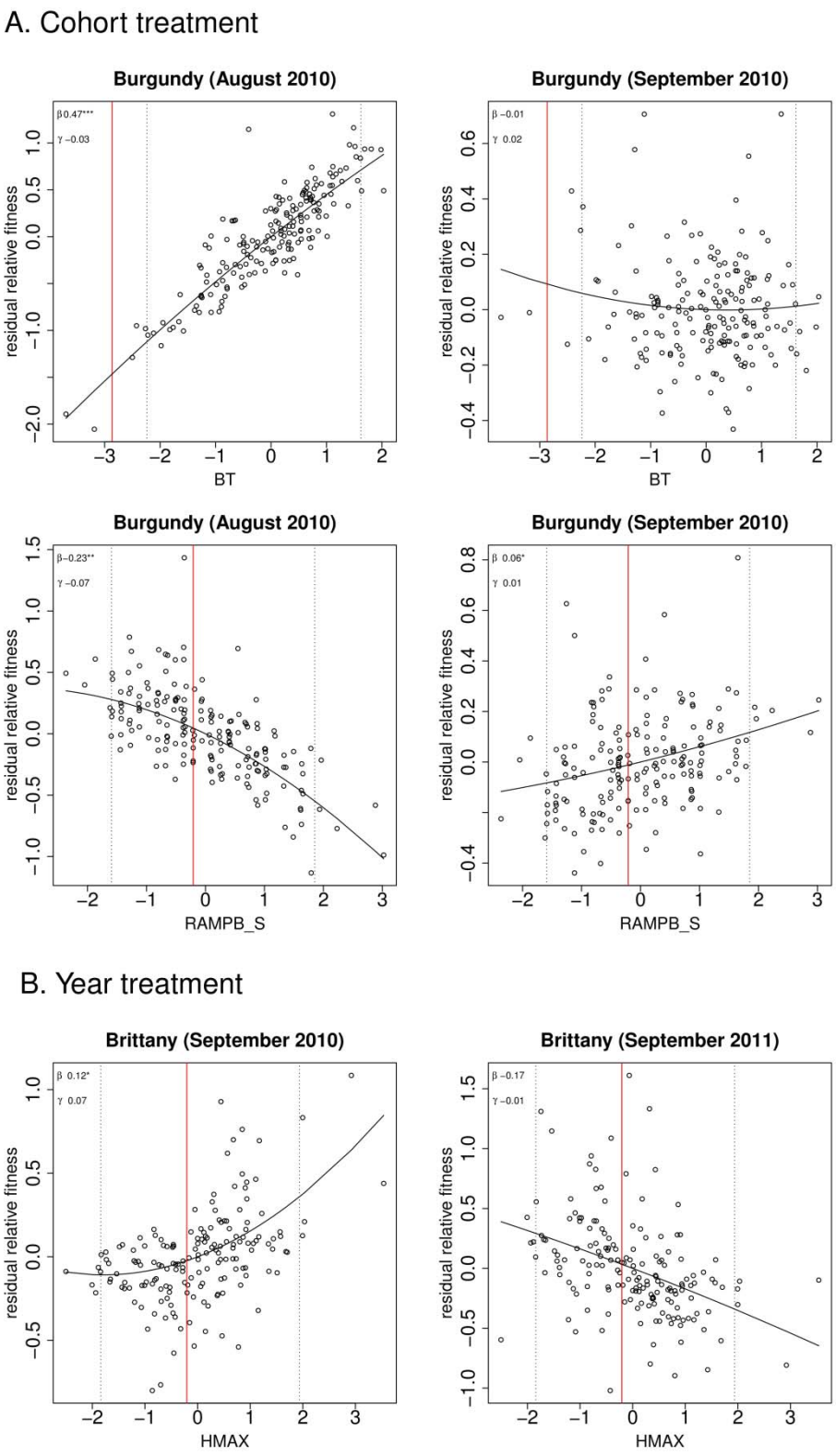


Table 1 Natural variation of nine phenotypic traits at three spatial scales, *i.e.* region, stand and family. A. Germination cohort effect within the same year ('August 2010' *vs.* 'September 2010'). B. Year effect within the same germination cohort ('September 2010' *vs.* 'September 2011').

A. Cohort treatment

	Traits																	
	GERM		BT		INT		RP		DIAM		RAMBB		RAMPB_S		H1S		HMAX	
Model terms	*F* or *LRT*	*P*	*F* or *LRT*	*P*	*F* or *LRT*	*P*	*F* or *LRT*	*P*	*F* or *LRT*	*P*	*F* or *LRT*	*P*	*F* or *LRT*	*P*	*F* or *LRT*	*P*	*F* or *LRT*	*P*
Block	1.57	ns	**31.83**	***	0.62	ns	0.19	ns	0.01	ns	1.27	ns	5.73	*	5.95	*	0.17	ns
Cohort	5.56	*	**20.47**	***	0.02	ns	0.14	ns	**16.52**	***	**229**	***	1.96	ns	4.78	*	8.1	**
Region	**5.08**	**	**17.35**	***	2.3	ns	**25.01**	***	**16.77**	***	**24.92**	***	**5.44**	***	**30.76**	***	**17.62**	***
Stand(Region)	**5.87**	***	**13.61**	***	**10.74**	***	**8**	***	**3.5**	***	**8.69**	***	**5.61**	***	**11.73**	***	**8.46**	***
Family(Stand(Region))	**169.2**	***	**34**	***	1.9	ns	**12**	***	5	*	**8**	**	**11**	***	**16.9**	***	**19.8**	***
Cohort*Region	**6.56**	***	**109.8**	***	**9.29**	***	**26.61**	***	2.21	ns	1.95	ns	**8.09**	***	**27.92**	***	3.53	*
Cohort*Stand(Region)	**1.69**	**	**11.63**	***	**7.44**	***	**4.93**	***	1.62	**	**3.6**	***	**3.91**	***	**4.36**	***	1.59	**
*Cohort*Family(Stand(Region))*	2.1	ns	**65.2**	***	**35.1**	***	**21.5**	***	0	ns	0	ns	5.4	*	4.1	*	0	ns
Control Bg-2	**6.68**	**	**6.38**	**	0.09	ns	2.29	ns	0.31	ns	1.58	ns	0.96	ns	1.17	ns	**5.76**	**

B. Year treatment

	Traits																	
	GERM		BT		INT		RP		DIAM		RAMBB		RAMPB_S		H1S		HMAX	
Model terms	*F* or *LRT*	*P*	*F* or *LRT*	*P*	*F* or *LRT*	*P*	*F* or *LRT*	*P*	*F* or *LRT*	*P*	*F* or *LRT*	*P*	*F* or *LRT*	*P*	*F* or *LRT*	*P*	*F* or *LRT*	*P*
Block	1.07	ns	0.95	ns	0.55	ns	5.12	ns	1.39	ns	0.21	ns	2.32	ns	8.77		9.07	ns
Year	**32.16**	***	2.92	ns	7.24	**	3.1	ns	0.13	ns	**227.2**	***	1.2	ns	**13.9**	***	5.41	*
Region	**4.41**	**	**44.23**	***	**90.62**	***	**8.04**	***	0.07	ns	**17.59**	***	**23.34**	***	**84.28**	***	**47.75**	***
Stand(Region)	**5.21**	***	**11.43**	***	**12.89**	***	**5.16**	***	**2.75**	***	**6.06**	***	**5.53**	***	**9.23**	***	**6.29**	***
Family(Stand(Region))	**145.1**	***	**33.6**	***	**9.1**	**	3.6	ns	3.3	ns	1.4	ns	**9.9**	**	**27**	***	**15**	***
Year*Region	2.7	*	**30.02**	***	**16.96**	***	1.65	ns	**16.54**	***	1.59	ns	**7.82**	***	2.81	*	**15.2**	***
Year*Stand(Region)	**2.09**	***	**2.53**	***	**2.12**	***	**2.13**	***	**2.36**	***	**2.65**	***	1.26	ns	1.63	**	**2.23**	***
*Year*Family(Stand(Region))*	0.4	ns	**10.6**	**	1.3	ns	0	ns	0.9	ns	**12.3**	***	2.1	ns	0	ns	0	ns
Control Bg-2	**33.8**	***	**12.54**	***	1.17	ns	1.39	ns	0.74	ns	3.36	*	1.75	ns	0.04	ns	0.8	ns

Model random terms were tested with likelihood ratio tests (LRT) of models with and without these effects. Random effects are in italic. *$0.05 > P > 0.01$, **$0.01 > P > 0.001$, *** $P < 0.001$. Bold stars indicate significant effect after Bonferroni correction. NE: not estimated.

SUPPORTING INFORMATION: ADAPTIVE VALUE AND COSTS OF PHENOTYPIC PLASTICITY ACROSS GERMINATION COHORTS AND YEARS IN FOUR REGIONAL SETS OF NATURAL *ARABIDOPSIS THALIANA* FAMILIES

Table S1 Natural variation of nine phenotypic traits at three spatial scales (*i.e.* region, stand and family) within each sowing date.

A. August 2010

	Traits															
	GERM		BT		INT		RP		DIAM		RAMBB		RAMPB_S		H1S		HMAX	
Model terms	*F* or LRT	*P*	*F* or LRT	*P*	*F* or LRT	*P*	*F* or LRT	*P*	*F* or LRT	*P*	*F* or LRT	*P*	*F* or LRT	*P*	*F* or LRT	*P*	*F* or LRT	*P*
Block	2.67	ns	**11.1**	***	3.88	*	**14.88**	***	0.95	ns	0.97	ns	2.63	ns	1.06	ns	2.31	ns
Region	**16.8**	***	**39.9**	***	1.21	ns	**34.57**	***	**10.2**	***	**10.2**	***	2.83	*	**4.1**	**	**8.15**	***
Stands(Region)	**7.57**	***	**12.3**	***	**6.89**	***	**6.85**	***	**2.64**	***	**5.81**	***	**4.52**	***	**6.51**	***	**5.98**	***
Family(Stand(Region))	**107**	***	**99.2**	***	**25**	***	**58.9**	***	0.1	ns	0	ns	**18.6**	***	14.1	***	3.1	ns
Control Bg-2 (Block)	**9.87**	**	4.53	*	0.06	ns	2.74	ns	0.48	ns	0	ns	1.16	ns	3.21	ns	**11.4**	***

B. September 2010

	Traits															
	GERM		BT		INT		RP		DIAM		RAMBB		RAMPB_S		H1S		HMAX	
Model terms	*F* or LRT	*P*	*F* or LRT	*P*	*F* or LRT	*P*	*F* or LRT	*P*	*F* or LRT	*P*	*F* or LRT	*P*	*F* or LRT	*P*	*F* or LRT	*P*	*F* or LRT	*P*
Block	3.41	ns	**38.9**	***	0.19	ns	4.19	*	1.43	ns	0.24	ns	3.47		6.65	*	**15.8**	***
Region	2.15	ns	**125**	***	**45.4**	***	2.27	ns	**9.27**	***	**33.1**	***	**16.4**	***	**76.6**	***	**16.1**	***
Stands(Region)	**4.91**	***	**14.3**	***	**11**	***	**5.58**	***	**2.64**	***	**7.12**	***	**4.77**	***	**10.4**	***	**5.35**	***
Family(Stand(Region))	**129**	***	**76.1**	***	**21**	***	4.2	*	1	ns	6.6	*	3.6	ns	**10**	**	**6.8**	**
Control Bg-2 (Block)	3.87	*	**8.26**	**	1.61	ns	0.97	ns	0.21	ns	**8.14**	**	0.48	ns	0.01	ns	0.08	ns

C. September 2011

	Traits															
	GERM		BT		INT		RP		DIAM		RAMBB		RAMPB_S		H1S		HMAX	
Model terms	*F* or LRT	*P*	*F* or LRT	*P*	*F* or LRT	*P*	*F* or LRT	*P*	*F* or LRT	*P*	*F* or LRT	*P*	*F* or LRT	*P*	*F* or LRT	*P*	*F* or LRT	*P*
Block	**50**	***	0	ns	3.15	ns	**9.44**	**	0.17	ns	0.82	ns	0.06	ns	**9.65**	**	**20.6**	***
Region	**10.6**	***	1.91	ns	**53.1**	***	**5.42**	**	**6.42**	***	3.58	*	**13.6**	***	**32.4**	***	**34.5**	***
Stands(Region)	**4.74**	***	**4.82**	***	**5.95**	***	**2.53**	***	**2.31**	***	**2.95**	***	**2.66**	***	**3.41**	***	**3.72**	***
Family(Stand(Region))	**72.9**	***	**12.5**	***	0.9	ns	0	ns	**7.1**	**	**9.7**	**	**11.6**	***	0.6	ns	0.2	ns
Control Bg-2 (Block)	**121**	***	**12**	***	0.99	ns	1.86	ns	1.23	ns	1.5	ns	2.32	ns	0.03	ns	0.99	ns

Model random terms were tested with likelihood ratio tests (LRT) of models with and without these effects. Random effects are in italic. *$0.05 > P > 0.01$, **$0.01 > P > 0.001$, *** $P < 0.001$. Bold stars indicate significant effect after Bonferroni correction. NE: not estimated.

Table S2 Pairwise comparison of correlations matrices of nine phenotypic traits and nine phenotypic plasticities (A) between regions within each sowing date or treatment and (B) between sowing dates or treatments date within each geographical unit. Non-significant correlation between two dissimilarity matrices indicates that two trait correlation matrices or two plasticity correlation matrices are significantly different from one.

A. Pairwise comparison between regions

			Sowing dates or treatments																	
			August 2010				September 2010				September 2011				Cohort treatment				Year treatment			
			similarity		dissimilarity		similarity		dissimilarity		similarity		dissimilarity		similarity		dissimilarity		similarity		dissimilarity	
Pairwise regions			rho	*P*	rho	*P*	rho	*P*	rho	*P*	rho	*P*	rho	*P*	rho	*P*	rho	*P*	rho	*P*	rho	*P*
Brittany	-	Burgundy	0.89	***	0.67	***	0.92	***	0.46	*	0.87	***	0.28	*	0.89	***	0.65	***	0.95	***	0.58	***
Brittany	-	Languedoc	0.76	***	-0.01		0.84	***	0.47	**	0.74	***	0.11		0.79	***	0.25		0.79	***	-0.10	
Brittany	-	North	0.93	***	0.42	**	0.90	***	0.40	*	0.85	***	0.12		0.92	***	0.54	**	0.91	***	0.16	
Burgundy	-	Languedoc	0.79	***	-0.11		0.78	***	0.26		0.73	***	0.29	*	0.79	***	0.38	*	0.75	***	0.19	
Burgundy	-	North	0.88	***	0.03		0.90	***	0.44	**	0.89	***	0.10		0.90	***	0.71		0.92	***	0.25	
Languedoc	-	North	0.78	***	0.28		0.84	***	0.09		0.88	***	0.65		0.73	***	0.37	*	0.77	***	0.31	

B. Pairwise comparison between sowing dates or treatments

	Pairwise correlations											
	Aug 2010 - Sep 2010				Sep 2010 - Sep 2011				Cohort - Year			
	similarity		dissimilarity		similarity		dissimilarity		similarity		dissimilarity	
Geographical unit	*rho*	*P*	*rho*	*P*	*rho*	*P*	*rho*	*P*	*rho*	*P*	*rho*	*P*
France	0.79	***	0.48	**	0.75	***	-0.26		0.50	*	0.19	
Brittany	0.59	*	0.23		0.64	**	-0.17		0.49	*	0.24	
Burgundy	0.60	**	0.17		0.72	***	-0.06		0.54	**	0.25	
Languedoc	0.77	***	-0.05		0.65	**	-0.16		0.88	***	0.36	*
North	0.77	***	0.41	**	0.72	***	0.09		0.38		0.26	

*$0.05 > P > 0.01$, **$0.01 > P > 0.001$, *** $P < 0.001$.

Table S3 Variation in linear and non-linear genotypic selection gradients across sowing dates within each geographical unit in the (A) cohort and (B) year treatments.

A. Cohort treatment	Geographical unit									
	France		Brittany		Burgundy		Languedoc		North	
Model terms	F	*P*	F	*P*	F	*P*	F	*P*	F	*P*
Cohort	3.93	*	0.07		0.12		0.14		2.52	
GERM	0.17		1.96		0.12		0.05		0.11	
BT	**18.05**	***	0.18		1.10		1.03		**11.45**	***
INT	4.16	*	0.41		1.35		0.51		4.57	*
RP	6.32	*	8.29	**	4.21	*	3.63		0.01	
DIAM	**108.98**	***	**15.18**	***	**12.05**	***	9.47	**	**16.59**	***
RAMBB	**236.47**	***	**62.89**	***	**49.07**	***	**82.36**	***	**36.10**	***
RAMPB_S	**105.81**	***	**27.09**	***	**33.47**	***	**34.95**	***	**35.78**	***
H1S	0.02		0.13		0.34		0.68		1.14	
HMAX	**30.89**	***	6.02	*	8.66	**	**13.35**	***	10.31	**
GERM²	0.01		0.17		0.13		0.53		0.05	
BT²	0.50		0.78		0.03		0.25		5.97	*
INT²	7.04	**	0.85		0.07		1.27		0.92	
RP²	3.39		0.07		0.61		3.33		1.35	
DIAM²	**38.12**	***	7.14	**	1.49		6.47	*	1.57	
RAMBB²	**10.98**	***	0.67		0.02		1.19		0.86	
RAMPB_S²	0.02		3.55		0.03		1.59		0.48	
H1S²	9.59	**	1.11		0.25		0.06		1.16	
HMAX²	3.20		0.20		0.34		2.81		1.03	
Cohort * GERM	0.31		0.31		0.08		0.45		0.13	
Cohort * BT	2.90		1.18		0.04		0.09		**11.23**	***
Cohort * INT	3.03		1.70		1.18		0.14		0.05	
Cohort * RP	3.63		2.86		1.15		0.37		0.15	
Cohort * DIAM	0.41		2.68		0.11		0.38		1.21	
Cohort * RAMBB	**23.55**	***	**29.64**	***	7.97	**	0.50		0.98	
Cohort * RAMPB_S	0.10		1.56		0.02		0.04		2.46	
Cohort * H1S	6.48	*	0.00		2.03		1.93		0.10	
Cohort * HMAX	1.12		0.55		1.24		0.98		0.37	
Cohort * GERM²	0.12		0.21		0.01		0.40		0.36	
Cohort * BT²	3.71		0.61		0.26		1.51		0.95	
Cohort * INT²	5.35	*	0.23		0.01		1.23		0.69	
Cohort * RP²	0.59		0.83		0.71		0.11		0.21	
Cohort * DIAM²	10.25	**	3.65		0.54		0.05		0.00	
Cohort * RAMBB²	9.93	**	0.17		0.00		0.23		0.42	
Cohort * RAMPB_S²	0.39		1.42		0.01		1.36		0.63	
Cohort * H1S²	0.03		0.00		0.39		1.03		0.09	
Cohort * HMAX²	1.70		1.62		0.32		0.82		1.00	

*$0.05 > P > 0.01$, **$0.01 > P > 0.001$, *** $P < 0.001$. Bold stars indicate significant effect after Bonferroni correction.

B. Year treatment	Geographical unit									
	France		Brittany		Burgundy		Languedoc		North	
Model terms	F	*P*	F	*P*	F	*P*	F	*P*	F	*P*
Year	2.22		0.19		5.00	*	0.10		0.95	
GERM	3.89	*	6.95	**	0.29		1.37		0.00	
BT	8.79	**	0.00		2.50		3.83		2.95	
INT	0.36		0.09		0.00		0.04		0.56	
RP	**21.97**	***	10.33	**	5.05	*	2.02		8.40	**
DIAM	**23.74**	***	0.35		1.18		4.84	*	6.65	*
RAMBB	**390.74**	***	**105.18**	***	**98.75**	***	**140.69**	***	**55.76**	***
RAMPB_S	**68.41**	***	**18.67**	***	**28.83**	***	**29.56**	***	**22.80**	***
H1S	7.47	**	0.92		1.41		0.44		0.63	
HMAX	**88.19**	***	**20.15**	***	**25.74**	***	**36.74**	***	**15.79**	***
GERM²	0.12		0.20		0.69		5.01	*	0.18	
BT²	0.19		0.38		0.24		0.00		2.50	
INT²	3.85	*	0.47		0.01		1.89		0.01	
RP²	2.98		0.08		1.46		2.77		0.70	
DIAM²	**87.50**	***	**15.69**	***	6.89	**	**16.86**	***	5.63	*
RAMBB²	**21.22**	***	0.02		0.21		3.08		2.23	
RAMPB_S²	0.91		0.49		0.87		0.90		0.07	
H1S²	0.29		0.18		1.05		0.02		2.52	
HMAX²	**10.34**	**	0.07		3.92	*	**11.82**	***	5.91	*
Year * GERM	4.75	*	0.71		0.04		3.07		0.54	
Year * BT	0.75		0.11		0.60		1.06		2.78	
Year * INT	3.50		0.00		0.07		1.99		2.72	
Year * RP	0.01		0.45		0.35		0.14		4.77	*
Year * DIAM	2.29		0.02		3.03		0.11		0.08	
Year * RAMBB	0.01		4.74	*	0.06		0.40		1.90	
Year * RAMPB_S	0.18		0.70		0.46		8.49	**	0.57	
Year * H1S	0.01		1.27		0.30		0.74		0.06	
Year * HMAX	**26.44**	***	2.86		**10.73**	**	**15.49**	***	0.38	
Year * GERM²	0.23		0.01		1.27		4.67	*	0.53	
Year * BT²	2.04		0.36		0.58		3.74		0.32	
Year * INT²	6.39	*	0.19		0.15		2.79		1.45	
Year * RP²	0.52		0.61		6.96	**	0.84		0.08	
Year * DIAM²	2.21		0.29		1.01		0.93		1.79	
Year * RAMBB²	1.87		0.86		0.27		0.03		0.06	
Year * RAMPB_S²	0.07		2.72		1.01		3.06		0.03	
Year * H1S²	5.01	*	1.43		0.84		0.63		0.26	
Year * HMAX²	0.04		0.20		0.58		0.05		0.20	

*0.05 > *P* > 0.01, **0.01 > *P* > 0.001, *** *P* < 0.001. Bold stars indicate significant effect after Bonferroni correction.

Table S4 Variation in linear and non-linear genotypic selection gradients among regions within each sowing date.

Interactions with regions						
	Treatment					
	August 2010		September 2010		September 2011	
Model terms	F	*P*	F	*P*	F	*P*
Region	2.84	*	**9.24**	***	1.04	
GERM	0.17		0.30		4.48	*
BT	**11.20**	***	**10.66**	**	7.37	**
INT	4.67	*	2.96		0.68	
RP	0.51		**18.52**	***	**12.23**	***
DIAM	**64.75**	***	**11.40**	***	3.03	
RAMBB	**105.48**	***	**218.59**	***	**232.90**	***
RAMPB_S	**84.82**	***	**114.73**	***	**25.65**	***
H1S	4.54	*	0.01		0.20	
HMAX	**21.80**	***	**31.48**	***	**92.06**	***
GERM²	0.40		0.13		1.28	
BT²	6.06	*	0.03		2.26	
INT²	0.19		6.14	*	0.39	
RP²	4.76	*	0.09		7.94	**
DIAM²	6.97	**	**16.69**	***	**36.06**	***
RAMBB²	0.03		0.06		**16.94**	***
RAMPB_S²	0.00		0.02		0.27	
H1S²	3.83		0.17		0.00	
HMAX²	0.75		5.84	*	**10.72**	**
Region * GERM	0.06		1.73		0.93	
Region * BT	1.58		1.80		0.42	
Region * INT	0.20		1.47		0.34	
Region * RP	0.22		2.49		1.81	
Region * DIAM	3.05	*	0.62		1.12	
Region * RAMBB	1.67		4.48	**	0.91	
Region * RAMPB_S	2.37		1.63		0.44	
Region * H1S	1.03		0.77		0.53	
Region * HMAX	1.58		1.43		3.09	*
Region * GERM²	0.13		0.61		0.88	
Region * BT²	2.68	*	1.77		0.26	
Region * INT²	0.08		0.89		0.23	
Region * RP²	0.83		1.59		1.38	
Region * DIAM²	0.85		1.22		0.36	
Region * RAMBB²	1.60		0.86		1.07	
Region * RAMPB_S²	1.04		3.39	*	1.28	
Region * H1S²	0.14		0.66		2.03	
Region * HMAX²	0.34		2.17		1.16	

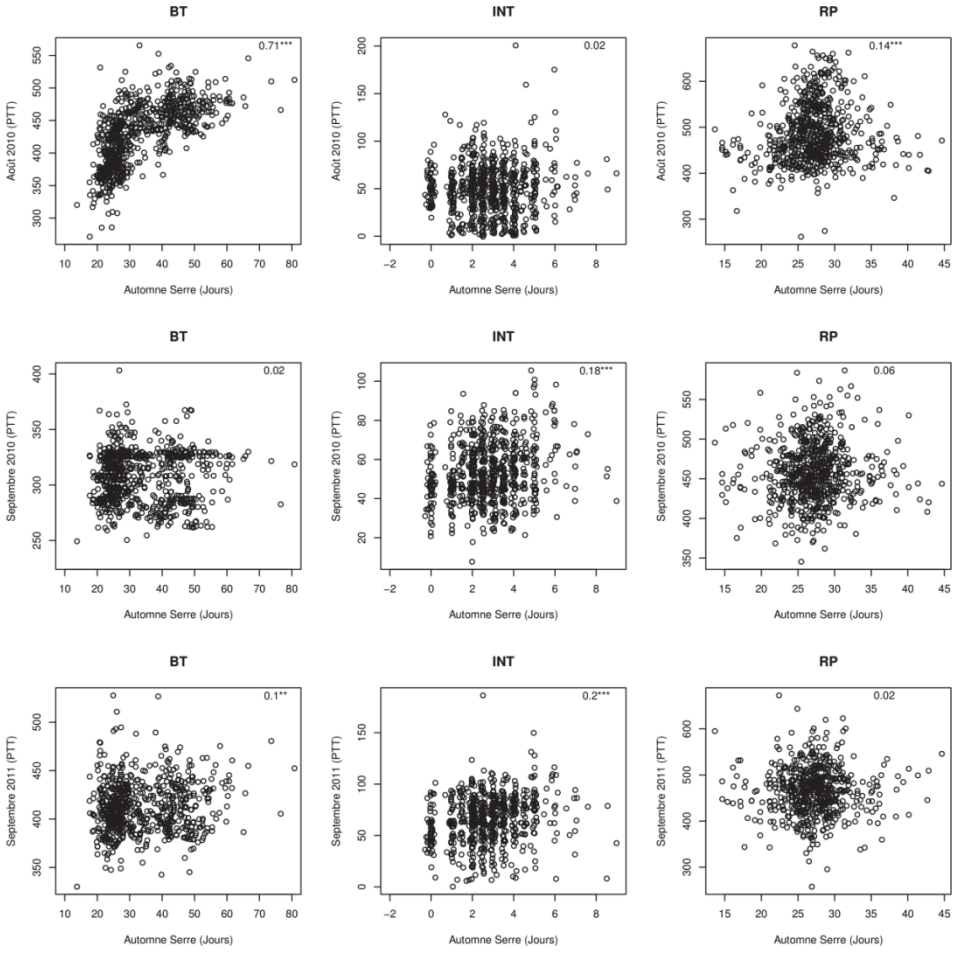
*$0.05 > P > 0.01$, **$0.01 > P > 0.001$, *** $P < 0.001$. Bold stars indicate significant effect after Bonferroni correction.

CHAPITRE 2 – CONCLUSIONS ET PERSPECTIVES

2. Conclusions et perspectives

Comme précédemment observé pour les 6 traits phénologiques mesurés en conditions contrôlées de serre, une forte variation génétique naturelle a été observée à différentes échelles spatiales (entre région, entre populations à l'intérieur des régions et entre familles à l'intérieur des populations) pour les 20 traits et leurs plasticités mesurés sur le terrain expérimental de l'Université de Lille. Les signaux perçus par une plante d'*A. thaliana* étant certainement plus nombreux et complexes dans un milieu naturel que dans des conditions contrôlées de serre, j'avais pensé que la relation phénotypique entre les données obtenues en serre et les données obtenues sur un terrain expérimental devait être relativement faible voire nulle. Afin de vérifier cette hypothèse, j'ai calculé les coefficients de corrélation génétique entre les données obtenues dans le traitement Automne simulé en serre (chapitre1) et les trois traitements du terrain expérimental (Août 2010, Septembre 2010 et Septembre 2011) pour trois traits phénologiques communs à tous ces environnements. Les corrélations génétiques inter-environnements ne dépassent pas 0.2 et sont souvent proches de 0, à l'exception de la date de montaison mesurée entre les conditions de serre (Automne) et le traitement Août 2010 sur le terrain expérimental (Figure 6). Ce résultat est en accord avec (seulement) deux études précédentes effectuées chez *A. thaliana* reportant de faibles coefficients de corrélation génétique pour la date de floraison entre des conditions contrôlées de serre et des conditions écologiquement réalistes de terrain expérimental (utilisation de RILs : Weinig *et al.* 2002 ; utilisation d'accessions naturelles : Bergelson & Roux 2010). L'étude de la variation phénotypique naturelle dans des conditions écologiquement réalistes apparaît donc comme un complément indispensable aux études menées en conditions contrôlées de serre. Il est certain que pouvoir contrôler des facteurs environnementaux dans une serre permet (i) de tester l'effet d'un facteur environnemental donné sur la réponse phénotypique des plantes (ceci est d'autant

plus important si le facteur environnemental testé correspond à un agent sélectif potentiel identifié par des études de relations phénotype – écologie), et (ii) de répéter les expériences plusieurs fois. Cependant, dans la nature, les conditions biotiques et abiotiques peuvent être très changeantes dans l'espace mais aussi dans le temps. Dans le cas de l'identification des bases génétiques associées à la variation naturelle adaptative, ne vaudrait-il donc pas mieux se focaliser sur des phénotypes exprimés dans des environnements naturels et mesurés plusieurs années de suite?



<u>Figure 6 :</u> Corrélations génétiques entre les données phénologiques obtenues pour les 800 familles françaises en serre et sur un terrain expérimental de l'Université de Lille. La valeur du *rho* de Spearman ainsi que sa significativité sont indiquées pour chaque cas étudié. * :$P < 0.05$; ** : $P < 0.01$; *** : $P < 0.05$. BT : bolting time, INT : interval between bolting time and flowering time, RP : length of the reproductive period.

Après avoir compté plus de 1 567 000 siliques sur l'ensemble des plantes de cette expérimentation réalisée sur un terrain expérimental, pour une longueur cumulée totale de siliques supérieure à 19,835 km, j'ai pu :

- mettre en évidence qu'une même production de graines entre deux familles pouvait être obtenue à partir d'une variation naturelle des traits reproducteurs sous-jacents. Ce résultat laisse penser qu'il est nécessaire de décomposer des traits très intégrateurs tels que la production de graines, afin de pouvoir identifier des processus sélectifs agissant de manière différente sur les traits sous-jacents.

- identifier de nombreux traits phénotypiques potentiellement sous sélection, notamment les traits liés à l'allocation des ressources entre traits reproducteurs. Ce dernier résultat suggère que les stratégies de réallocation des ressources peuvent être aussi importantes que la quantité totale de graines produites dans l'étude de l'adaptation chez *A. thaliana*.

- constater que les patrons de sélection pouvaient être fortement dépendants de la date de semis (aussi bien entre cohortes qu'entre années) et de la région géographique considérée. Ce constat s'applique principalement pour la plasticité de tous les traits mesurés dans cette étude.

- suggérer que la plasticité phénotypique puisse être un trait à part entière pouvant évoluer indépendamment de son expression dans un environnement donné, mais également indépendamment d'autres traits plastiques. En effet, les relations entre les plasticités des différents traits semblent peu conservées entre traitements, suggérant que peu de contraintes génétiques existent au sein des 800 familles utilisées dans cette thèse.

En conclusion, bien que l'acquisition de données phénotypiques apparaisse de plus en plus comme le facteur limitant principal en génomique écologique, elle s'avère une étape indispensable dans l'étude de l'adaptation d'un organisme. Cette étape doit comprendre l'acquisition de données pour de nombreux traits phénotypiques (y compris la valeur sélective en la décomposant en différent traits), et ceci dans plusieurs environnements écologiquement réalistes à partir d'un échantillonnage hiérarchique de populations naturelles. Une fois les traits adaptatifs et leurs échelles de variation naturelle correspondantes identifiés, des projets visant à identifier les régions génomiques associées à cette variation naturelle peuvent être ensuite envisagés.

Une des perspectives de ce travail sera de transiter vers une étude de génomique écologique évolutive, en identifiant les bases génétiques de l'adaptation chez *A. thaliana* en profitant notamment de la caractérisation phénotypique et écologique réalisée au cours de cette thèse (Figure 7). Ce travail a déjà été initié lors de ma seconde année de thèse avec le séquençage génomique des 49 populations, en bulkant les ADN par population selon une approche Pool-Seq (pooled population sequencing). L'analyse bio-informatique des données de séquençage par Sophie Gallina (responsable de la plate-forme de bio-informatique du laboratoire GEPV) a permis d'estimer récemment les fréquences alléliques le long du génome de tous les polymorphismes identifiés entre les 49 populations.

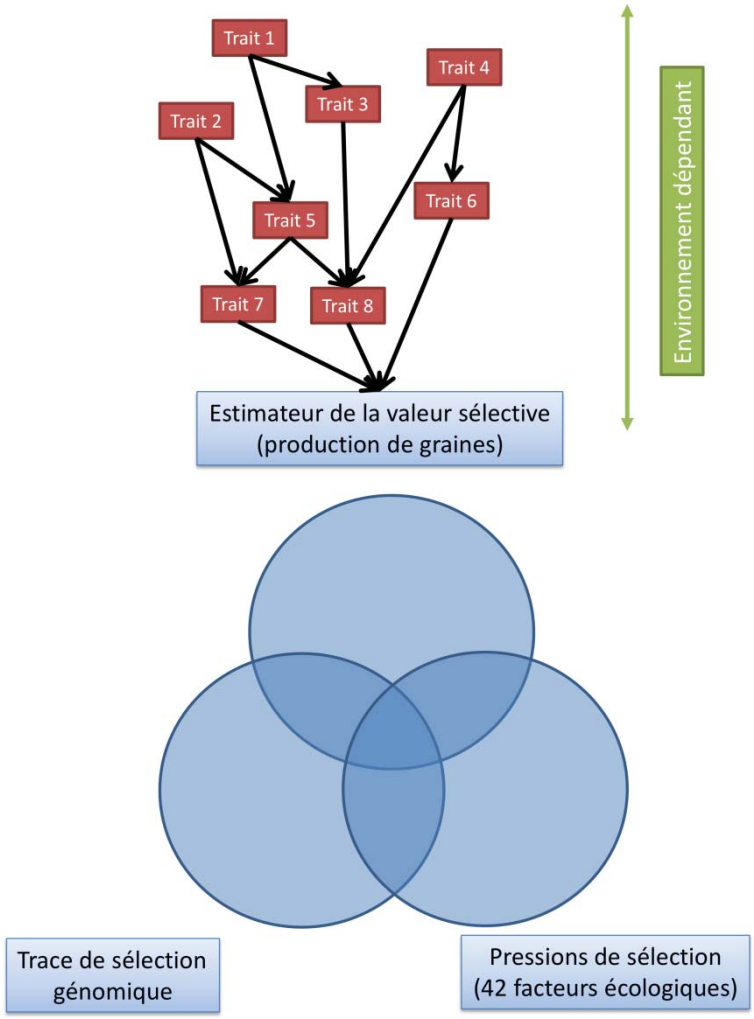


Figure 7 : Perspectives en génomique écologique évolutive. Nous chercherons les régions génomiques présentant des traces de sélection, associées à la variation d'un trait (ou d'une plasticité phénotypique) reliée à la fitness et à des variables écologiques.

A partir des données génomiques, écologiques et phénotypiques obtenues pour les 49 populations françaises, nous essaierons de répondre aux questions suivantes :

- Est-il possible d'identifier des régions génomiques qui sont à la fois associées à une variation naturelle phénotypique, à des agents sélectifs potentiels et présentant des traces de sélection ?

- Si oui, les régions génomiques identifiées sont-elles identiques à différentes échelles spatiales (France *vs.* régions) ? et entre les 4 régions géographiques?
- Si les régions génomiques identifiées ne sont pas identiques entre les régions géographiques, peut-on établir un lien entre les fonctions biologiques associées à ces régions génomiques et l'identité des agents sélectifs potentiels ?
- Les régions génomiques associées à la plasticité pour un trait donné sont-elles différentes des régions génomiques associées à ce même trait mesuré dans les environnements focaux ?
- Si oui, l'intensité et les traces de sélection associées aux régions génomiques identifiées sont-elles différentes entre les traits et leurs plasticités?

CHAPITRE 3 - DYNAMIQUE EVOLUTIVE DE POPULATIONS NATURELLES *D'ARABIDOPSIS THALIANA* SUR UNE COURTE ECHELLE DE TEMPS

IV. CHAPITRE 3 - DYNAMIQUE EVOLUTIVE DE POPULATIONS NATURELLES D'*ARABIDOPSIS THALIANA* SUR UNE COURTE ECHELLE DE TEMPS :

1. Introduction

Les deux chapitres précédents se sont focalisés sur la variation phénotypique présente sur un ensemble de 49 populations françaises récoltées au printemps 2009. Les résultats trouvés dans ces chapitres ne reflètent peut-être qu'une image instantanée des différents processus sélectifs et non-sélectifs agissant sur *A. thaliana*. Aurions-nous trouvé les mêmes résultats si les populations avaient été récoltées une année différente ?

Par ailleurs, étant donné la très forte diversité phénotypique et génétique observée dans de nombreuses populations, on peut se demander si cette diversité se maintient au cours du temps : Si oui, comment l'expliquer ? Si non, quelle est sa dynamique temporelle ? L'absence de données empiriques sur la dynamique temporelle au sein des populations naturelles d'*A. thaliana* nous empêche pour l'instant de répondre à ces questions. Il m'a donc paru important de caractériser cette dynamique temporelle, en utilisant les deux populations naturelles TOU et MIB décrites comme très polymorphes au niveau phénotypique et génétique dans le chapitre 1.

Ainsi, deux nouveaux échantillonnages ont été réalisés pour la population TOU (en 2007 et 2010) initialement récolté en 2002 et un nouvel échantillonnage a été réalisé pour la population MIB (en 2011) initialement récoltée en 2002. Une caractérisation phénotypique de ces différentes générations a été réalisée sur le terrain expérimental de Lille1 pour de nombreux traits phénotypiques (morphologie, phénologie, architecture, acquisition de

ressources, réallocation de ressources, survie…) afin de vérifier si une évolution phénotypique d'origine génétique s'était produite dans ces populations naturelles. Des échantillons de sol ont également été prélevés dans ces deux populations afin de vérifier si la diversité phénotypique observée dans ces deux populations pouvait être due à une adaptation micro-locale au sol.

Cette étude a fait l'objet :

- d'un manuscrit sur la population TOU :

R. VILLOUTREIX, V. LE CORRE, E. BARON, L. AMSELLEM and F. ROUX. Rapid phenotypic evolution and fine-grained spatial variation in a natural population of *Arabidopsis thaliana*: a resurrection study. *En préparation.*

Dans ce manuscrit, je me suis attaché à (i) caractériser la variation phénotypique existant pour de nombreux traits phénotypiques (ii) et à déterminer si une évolution phénotypique avait eu lieu en moins de 8 générations. La caractérisation du sol de cette population que j'ai réalisée avec l'aide d'Etienne Baron (doctorant de l'équipe « Changements globaux : de la génomique écologique à l'écologie des communautés ») m'a permis d'identifier l'existence d'une forte variabilité des conditions édaphiques dans cette population à une échelle très fine (de l'ordre de 20 mètres). Il m'est donc apparu nécessaire de déterminer si une part de la variation phénotypique observée, et de son évolution, était due à des phénomènes d'adaptation micro-locale à différentes conditions édaphiques.

- d'un rapport sur la population MIB :

Dans ce rapport, je me suis attaché à (i) caractériser la variation phénotypique existant pour de nombreux traits phénotypiques (ii) et à déterminer si une évolution phénotypique

avait eu lieu en moins de 9 générations. Là encore, j'ai effectué une caractérisation du sol, ce qui m'a permis de mettre en évidence une homogénéité des conditions édaphiques dans cette population. 52 accessions de la population MIB ayant été génotypées pour 214 kSNPs, j'ai entrepris d'identifier les bases génétiques associées à la variation phénotypique par une approche de GWA mapping. Plusieurs développements méthodologiques étant encore nécessaires afin de comprendre la dynamique temporelle observée au sein de cette population, ce rapport est une ébauche réalisée afin de présenter les résultats préliminaires en notre possession.

MANUSCRIT: RAPID PHENOTYPIC EVOLUTION AND FINE-GRAINED SPATIAL HETEROGENEITY IN A NATURAL POPULATION OF *ARABIDOPSIS THALIANA*: A RESURRECTION STUDY.

Rapid phenotypic evolution and fine-grained spatial heterogeneity in a natural population of *Arabidopsis thaliana*: a resurrection study.

Romain Villoutreix[1], Valérie Le Corre[2], Etienne Baron[1], Laurent Amsellem[1], Fabrice Roux[1,3,4,*]

[1]*Laboratoire Génétique et Evolution des Populations Végétales, UMR CNRS 8198, Université des Sciences et Technologies de Lille – Lille 1, F-59655 Villeneuve d'Ascq cedex, France*

[2] *INRA, UMR 1347 Agroécologie, F-21000 Dijon, France*

[3] *INRA, Laboratoire des Interactions Plantes-Microorganismes (LIPM), UMR441, F-31326 Castanet-Tolosan, France*

[4] *CNRS, Laboratoire des Interactions Plantes-Microorganismes (LIPM), UMR2594, F-31326 Castanet-Tolosan, France.*

Corresponding author: Fabrice Roux

CNRS, Laboratoire des Interactions Plantes-Microorganismes (LIPM), UMR2594, F-31326 Castanet-Tolosan, France.

E-mail: fabrice.roux@toulouse.inra.fr

Phone number: +33 (0)5 61 28 54 59 Fax number: +33 (0) 5 61 28 50 61

Keywords: adaptive traits, edaphic conditions, common garden, phenotypic evolution, *haldanes*, fitness

Running title: Phenotypic evolution across a fine-scale edaphic variation.

Summary

Because both rapid phenotypic evolution and spatial variation in relative fitness have been documented in natural plant populations, one may wonder whether the rate of contemporary microevolution depends on a fine-scale spatial ecological heterogeneity. Based on 313 seed families collected over eight generations along a 350m transect in a local French population of the annual selfing species *Arabidopsis thaliana*, we set up a resurrection study by growing all those families under common conditions in a field experiment. Population growth as well as rapid phenotypic evolution for survival and a suite of seven traits related to phenology, resource acquisition, architecture and seed dispersion were observed over the first five generations, followed by a quasi-stasis over the last three generations. Rates of evolutionary change expressed in *haldanes* are consistent with rates reported in previous studies on plants. Soil characterization along the transect indicated large edaphic variation at a spatial scale less than 20 meters, that often exceeds the range of edaphic variation found among natural populations at a small geographical scale. This larger spatial scale of edaphic variation with respect to the passive mode and short distance of seed dispersal may promote fine-scale selective heterogeneity in this population. Finally, phenotypic evolution was found to depend on sodium concentration for survival and organic carbon / total nitrogen ratio for the length of the reproductive period, suggesting that the rate of contemporary microevolution can depend on a fine-grained spatial ecological heterogeneity.

Introduction

Rapid phenotypic evolution in natural populations has been documented in many plant and animal species and can occur on a timescale of decades or even years (Hendry and Kinnison 1999, Kinnison and Hendry 2001, Bone and Farres 2001, Reznick and Ghalambor 2001). Response to anthropogenically induced environmental change like heavy-metal tolerance and herbicide resistance are classic examples of contemporary microevolution in plants (Holt and Le Baron 1990, Jasienuk *et al.* 1996, Bone and Farres 2001). More recently, an increased number of observations of rapid phenotypic evolution have been reported for invasive species (Bossdorf *et al.* 2005, Sax *et al.* 2007, Buswell *et al.* 2011, Sultan *et al.* 2012) or in response to global climate change (Fitter and Fitter 2002, Franks *et al.* 2007, Miller-Rushing and Inouye 2009, Anderson *et al.*2012, Novy *et al.* 2012). Because phenotypic evolution monitored *in-situ* may result from phenotypic plasticity and/or adaptive evolution, a promising and powerful approach called the "resurrection paradigm" recently emerged as a mean to disentangle the relative roles of plasticity and adaptive evolution in response to an environmental change (Franks *et al.* 2008). By growing under common environmental conditions genotypes collected in many populations over several generations, this approach can (i) quantify evolutionary change over a short time scale and (ii) test whether the rate of contemporary microevolution vary among populations (Franks *et al.* 2007, Sultan *et al.* 2012).

In addition to phenotypic evolution, spatial variation in relative fitness has been highly documented in natural plants populations (Primack and Kang 1989, Kingslover *et al.* 2001). In few studies, genotype-by-environment interactions for fitness have even been demonstrated to occur at a very small spatial scale, *i.e.* at the scale of few meters (or even few centimeters) within a plant population (Kalisz 1986, Stratton 1994, Stratton 1995, Stratton and Bennington 1998). Local adaption results from an interplay between environment grain, dispersal and the

intensity of reversal of relative fitness of genotypes among selective habitats (Slatkin 1973, Nagylaki 1975, Hedrick 1986). In the case of two selective habitats each with a different phenotypic optimum, local adaptation will be favored when (i) the environmental grain with respect to dispersal leads to a reduction of effective gene flow between the two types of habitats, and (ii) the rank of relative fitness of genotypes adapted to either of the two habitats is reversed between the two habitats. Local adaptation at a very small spatial scale within a natural population may therefore take place if (i) the dispersal distance is less than the environmental selective spatial grain, and (ii) the reversal of relative fitness among selective habitats is high (Roux *et al.* 2008). Small scale variation of soil nutrients, competitive effects among individual neighbors and soil biota have been suggested as putative ecological causes of fine-grained spatial selective heterogeneity (Turkington And Harper 1979, Stratton 1994, Sherrard and Maherali 2012). Because plant populations may face both a moving phenotypic optimum over time and ecological variation at a very small spatial scale, one may wonder whether the pace of phenotypic evolution within a population depend on a fine-scale spatial ecological heterogeneity.

The mouse-ear cress annual selfing species, *Arabidopsis thaliana*, encounters a wide diversity of ecological conditions across its worldwide distribution (Mitchell-Olds and Schmitt 2006, Montesinos *et al.* 2009, Brachi *et al.* 2013b) and recently emerged as a model plant in evolutionary ecological genomics (Bergelson and Roux 2010, Hancock *et al.* 2011, Fournier-Level *et al.* 2012, Gaut 2012, Lowry 2012). Natural populations of *A. thaliana* have been long considered to be not much genetically and phenotypically variable (likely due its selfing rate close to 98%). However, recent studies challenged this view. Many natural populations have been described to be highly variable at the genetic and phenotypic levels (Le Corre 2005, Platt *et al.* 2010, Bomblies *et al.* 2010, Montesinos *et al.* 2009, Lundemo *et al.*

2009, Brachi *et al.* 2013b, Huard-Chauveau *et al.* 2013, Mendez-Vigo *et al.* 2013). Despite this genetic and phenotypic variation available at a small geographical scale, little is known about the pace of adaptive dynamics in natural populations of *A. thaliana* (Picó 2012). Because edaphic variation at the scale of few meters has been described within some French populations (Brachi *et al.* 2013b), such populations represent therefore a unique opportunity to test whether phenotypic evolution depends on a small spatial scale of ecological variation.

In this study, we set up a resurrection study by first collecting genotypes over eight generations along a transect of 350m in a single population located in France (Burgundy). All genotypes were then grown under common conditions in a field experiment. Based on genotypes collected the first generation, this local population was found to be highly diverse at both the genetic level (based on the genotyping of 149 SNPs; Platt *et al.* 2010) and the phenotypic level for six phenological traits scored in greenhouse conditions (Brachi *et al.* 2013b) and quantitative resistance to the bacterial vascular pathogen *Xanthomonas campestris* pv. *campestris* (Huard-Chauveau *et al.* 2013). Here, we address the following questions:

1. Do we observe rapid phenotypic evolution over eight generations?
2. What is the level of edaphic variation observed within the TOU-A population in comparison to the level of edaphic variation observed among natural populations from the same geographical region?
3. Does the rate of contemporary microevolution within a population depend on a fine-scale selective heterogeneity?

Materials and Methods

Plant material

In this study, we focused on the population TOU-A located under a 350m electric fence separating two permanent meadows (Fig. 1A) in the village of Toulon sur Arroux (Burgundy, France). Seeds from individual plants were collected in 2002 (TOU-A-2002), 2007 (TOU-A-2007) and 2010 (TOU-A-2010) according to a sampling scheme allowing us to take into account the density of *A. thaliana* plants along the transect: (1) from the starting point of the transect (Fig. 1A), walk along the transect until a plant is found and collect seeds from this plant, (2) if this plant is at the beginning of a patch, then collect seeds from plants located every 50cm along this patch, (3) else, walk along the transect until a new plant is found and collect seeds from this plant. According to this sampling scheme, seeds of 83, 115 and 115 individual plants were collected in 2002, 2007 and 2010, respectively (Fig. 1B). Seeds collected from those 313 individual plants constitute seed families.

Seeds from the 83 families collected in 2002 have been sent in 2003 to Joy Bergelson (University of Chicago, USA) for a project on the scale of population structure in *A. thaliana* (Platt *et al.* 2010). For this project, seed production for each TOU-A-2002 family was performed by growing one plant of each family under controlled greenhouse conditions.

Maternal effects of the 313 families were reduced by growing one plant of each family for one generation under controlled greenhouse conditions (16-h photoperiod, 20°C) in early 2011 at the University of Lille 1. For this purpose, we planted seeds produced at the University of Chicago for the families from the TOU-A-A2 population and seeds collected in the field for the families from the TOU-A-2007 and TOU-A-2010 populations.

Experimental design

Seeds were sown at the University of Lille (Nord, France) on the 22th September 2011 to mimic the main natural germination cohort observed in natural populations of *A. thaliana* in the North of France. The field experiment was organized in two blocks. Each block was represented by 15 arrays of 66 individual wells (Ø4 cm, vol. ~38 cm3) (TEKU, JP 3050/66) filled with damp standard culture soil (Huminsubstrat N3, Neuhaus). Each block corresponded to an independent randomization of 969 plants with three replicates per family (n = 313) and a control worldwide accession Bg-2 placed in the same two positions within each array (n = 30 = 15 arrays * 2 replicates). In each block, the remaining 21 wells were left empty. A minimum of five seeds were sown in each well. Seeds were stratified four days at 4°C in a cold chamber to promote germination. After the stratification treatment, arrays were preventively treated against dark-winged fungus gnats (Vectobac, 8mL per liter) and placed in a frost-free greenhouse for 17 days to protect seeds from rainfall but otherwise mimic outdoor conditions (no additional light or heating). To reduce micro-environmental variation, arrays were rotated daily in the cold chamber and in the greenhouse. Germination date was monitored daily for 17 days in the greenhouse (see below). On day 18, wells with no germinated seeds received extra-seedlings of the same family but from a different well, and other wells were thinned to two seedlings, such that two seedlings were present in each well when flats were transported outside to a common garden located at the University of Lille 1 (France). For each block, the 15 arrays were organized according to a grid of three columns and five lines. To facilitate root development, soil was tilled to allow arrays to be slightly buried. Because the bottom of the wells was pierced, roots were able to reach the soil easily. Seedlings were thinned to one seedling per well one week after placing the arrays in the

common garden. Plants were protected from herbivory by vertebrates, slugs and *Myzus persicae* as described in Brachi *et al.* (2010).

Phenotypic characterization

Thirteen phenotypic traits related to phenology (n = 4), resource acquisition (n = 2), architecture and seed dispersion (n = 5), survival (n=1) and total seed production (n = 1) were scored in this study.

During the growing period, plants were monitored daily for germination date (opening of both cotyledons) from when seeds were sown to 17 days after the stratification treatment and every 2 or 3 days for bolting date (inflorescence distinguishable from the leaves at a size < 5 mm), flowering date (appearance of the first open flower) and date of maturation of the last fruit. Based on these four phenological dates, four phenological traits were calculated (Brachi *et al.* 2012). Germination time (GERM) was measured as the number of days between the end of the stratification treatment and germination date. Bolting time (BT), flowering interval (INT) and the reproductive period (RP) were scored as the time interval between germination date and bolting date, between bolting date and flowering date and between flowering date and date of maturation of the last fruit, respectively. BT, INT and RP were scaled in photothermal units (PTU) using a phenological model integrating both photoperiod length and temperature as described in Brachi *et al.* (2010).

Two proxies for resources acquisition were measured on each plant. Rosette surface area (AREA) was measured using a nondestructive approach, by imaging every tray 35 days after transferring the trays from the greenhouse to the common garden, and using a Canon digital camera (model EOS 500D). Each image of an array was submitted to a four-step

treatment (Fig. S1, Supporting Information) allowing to get 1938 individual images ((313 families * 6 replicates) + (30 arrays * 2 control plants)). AREA were then determined with an image processing technique using the ImageJ software (version 4.01; Universal Imaging Corporation, West Chester, PA, http://rsbweb.nih.gov/ij/index.html). At bolting, the maximum diameter of the rosette measured at the nearest millimeter was used as a proxy for plant size (DIAM; Weinig *et al.* 2006).

After maturation of the last fruit, the aboveground portion was harvested and stored at room temperature until phenotyping of the following architectural and seed-dispersal related traits described as non-collinear and related to adaptation in *A. thaliana* (Reboud *et al.* 2004): the number of basal branches (RAMBB), the number of primary branches with siliques on the main stem (RAMPB_S), the height from soil to the first silique on the main stem (H1S) and height of the main stem (HSTEM), the average length between two fruits on the main stem (INTERNOD = (HSTEM – H1S)/(number of fruits on the main stem -1)).

All plants that germinated but did not survive were counted as dead (SURVIVAL = 0). Harvested plants with a germination date were counted as alive (SURVIVAL = 1). Total seed production was approximated by total fruit length (FITTOT) which has been shown to be a good proxy of lifetime fitness for a selfing annual like *A. thaliana* (Roux *et al.* 2004).

Edaphic characterization

A sample of the 5-cm upper soil layer was collected at 83 positions scattered along the transect in 2010 (Fig. 1B). These 83 samples correspond to the positions of the 83 families collected in 2002. Soil samples were transferred to the greenhouse and air-dried. Soil samples were then stored in the laboratory at room temperature. As described in Brachi *et al.* (2013b),

each soil sample was characterized for 14 edaphic factors: pH, maximal water holing capacity (WHC), content of total nitrogen (N) and organic carbon (C), C/N ratio, content of soil organic matter (SOM), concentrations of P_2O_5, K, Ca, Mg, Mn, Al, Na and Fe. Three samples were not further considered in this study due to outliers for concentration of either Na, Al or P_2O_5. Fe concentration was excluded from further analyses due to a lack of variation among the remaining 80 samples. The set of remaining 13 edaphic variables was pruned based on the pairwise Spearman correlations of the variables (Fig. S2, Supporting Information) so that no two variables had a Spearman *rho* greater than 0.8. In cases where variables were strongly correlated with one another, the variable with the most obvious link to the ecology *A. thaliana* was selected. The final set of 10 edaphic variables considered in this study was N, C/N ratio, pH, WHC, P_20_5, K, Mg, Mn, Na and Al.

Soil characteristics of families from years 2007 and 2010 were extrapolated from the 80 samples described above. For a given family, soil characteristics were defined as identical to the soil characteristics of the nearest position among the 80 samples with an upper limit of ten meters, allowing a successful extrapolation for 96.5% and 93.9% of the families collected in 2007 and 2010, respectively.

Because (i) edaphic conditions may have evolved between 2002 and 2010 and (ii) our extrapolation approach may have induced spurious soil characteristics for families collected in 2007 and 2010, significant phenotypic evolution in relation with edaphic conditions should be conservative.

Data analysis

Natural variation and phenotypic evolution

We explored natural variation of all phenotypic traits (with the exception of SURVIVAL) using the following statistical model (PROC MIXED procedure, REML method, SAS 9.3, SAS Institute Inc):

$$Y_{ijklmc} = \mu_{trait} + block_i + year_j + block_i*year_j + family_k\ (year_j)) + block_i*\ family_k\ (year_j)) + control_c + \varepsilon_{ijkc} \quad (1)$$

In these models, '*Y*' is one of the fitness components, 'μ' is the overall phenotypic mean; 'block' accounts for differences in micro-environment among the two experimental blocks; 'year' corresponds to effect of the three sampling years; 'family' measure the effect of families within year; interaction terms involving the 'block' term account for genetic variation in reaction norms of years and families among the two blocks; 'control' is a covariate accounting for array effects within blocks within each sowing date (the phenotypic mean of two control replicates per array was used as a covariate); and 'ε' is the residual term. Because natural variation in rosette surface area measured at a given date may indirectly result from natural variation in germination time, GERM was also added as a covariate when running model (1) for the phenotypic trait 'AREA'. All factors were treated as fixed effects, except 'family' which was treated as a random effect. For fixed effects, terms were tested over their appropriate denominators for calculating *F*-values. Significance of the random effects was determined by likelihood ratio tests of model with and without these effects. When necessary, raw data were either log transformed or Box-Cox transformed to satisfy the normality and

equal variance assumptions of linear regression. For the binary trait 'survival', model fitting was conducted using the PROC GLIMMIX procedure in SAS 9.3 (SAS Institute Inc., Cary, North Carolina, USA). Significant difference between two years was assessed after Tukey's multiple comparisons tests.

For each year, least-square means (LSmeans) for each phenotypic trait were obtained for each family using the following model in the *R* environment (R Development Core Team 2010).

$$Y_{imc} = \mu_{\text{trait}} + \text{block}_i + \text{family}_m + \text{control}_c + \varepsilon_{imc} \qquad (2)$$

Because *A. thaliana* is a highly selfing species, LSmeans correspond to genotypic values of families.

To test if relationships among traits were conserved among the three years, we first estimated a Spearman correlation matrix based on family genotypic means for each year. We then performed non-parametric Mantel tests of the Spearman correlation between two matrices. The significance of the Mantel statistics was evaluated by 10,000 permutations (vegan 1.17-10 package in R 2.12.1 environment).

For traits that showed a significant year effect, rates of evolutionary change in *haldanes* were calculated as described in Hendry And Kinnison (1999) based on genotypic values of families. Because only one generation was observed every year between 2002 and 2010, we consider eight generations in the calculation of *haldanes* values.

Phenotypic evolution in relationship with edaphic conditions

To test whether phenotypic evolution was dependent on edaphic conditions, an analysis of covariance (ANCOVA) was performed based on genotypic values of families for each trait using the following model (R 2.12.1 environment):

$$\text{trait}_{ij} = \mu_{\text{relative trait}} + \text{year}_j + \text{edaphic1}_i + \text{edaphic1}_i^2 + \ldots + \text{edaphic10}_i + \text{edaphic10}_i^2 + \text{year}_j * \text{edaphic1}_i + \text{year}_j * \text{edaphic1}_i^2 + \ldots + \text{year}_j * \text{edaphic10}_i + \text{year}_j * \text{edaphic10}_i^2 + \varepsilon_{ij} \quad (3)$$

where 'trait' correspond to the standardized values of one of the 13 traits, 'μ' is the constant, 'edaphic1_i' to 'edaphic10_i' correspond to the ten standardized edaphic variables, 'year_j' corresponds to effect of the three sampling years, and ε_{ij} is the residual term. Significant '$\text{year}_j * \text{edaphic}_i$' and '$\text{year}_j * \text{edaphic}_i^2$' interactions indicate varying linear and non-linear relationship between trait and edaphic variable across years, respectively.

Within each year, relationship between each phenotypic trait and the ten ecological variables was studied by running model (3) excluding the terms involving 'year_j':

$$\text{trait}_{ij} = \mu_{\text{relative trait}} + \text{edaphic1}_i + \text{edaphic1}_i^2 + \ldots + \text{edaphic10}_i + \text{edaphic10}_i^2 + \varepsilon_{ij} \quad (4)$$

In order to discriminate between curvilinear and parabolic relationship between a phenotypic trait and an ecological variable, the point of null derivative was calculated individually for each ecological variable as follow:

$$Point\ of\ null\ derivative\ = \frac{-a}{2b} \quad (5)$$

where ‘a’ and ‘b’ correspond to the linear and quadratic regression coefficients of the ecological variable in equation (4), respectively. If (i) the point of null derivative was contained within the 95% interval distribution of the ecological variable of interest and (ii) ‘b’ was significantly positive and negative, the relationship between a phenotypic trait and an ecological variable was considered as a positive and negative parabole, respectively. If the point of null derivative was not contained within the 95% interval distribution of the ecological variable, the relationship between a phenotypic trait and an ecological variable was considered curvilinear.

Results

Phenotypic evolution

According to our sampling scheme, the density of the population TOU-A increased between 2002 and 2007 (in particular in the first 50 meters of the transect), but remained stable between 2007 and 2010. The spatial distribution of the plants has mainly changed in the 170m – 350m section of the transect, with gaps partly but not strictly overlapping over the three years.

Highly significant genetic variation was on average observed within each year for the 13 traits measured in this study (Fig. 2, Table 1). Rapid phenotypic evolution (mainly between 2002 and 2007) was observed for survival and seven traits related to phenology, resource acquisition, architecture, seed dispersion (Table 1). Phenotypic evolution was observed

towards a higher survival rate, a later germination, a longer interval between bolting and flowering time, a shorter reproductive period, a smaller rosette surface area, an increase of the height from soil to the first silique on the main stem, a shorter average length between two fruits on the main stem and smaller number of basal branches (Fig. 2, Fig. 3).

Trait correlation matrices were significantly similar for each pair of years (2002-2007: Spearman *rho* = 0.73, $P < 0.001$; 2007-2010: Spearman *rho* = 0.86, $P < 0.001$; 2002-2010: Spearman *rho* = 0.71, $P < 0.001$).

For the eight traits that showed a significant year effect, rates of evolutionary change expressed in *haldanes* was consistently higher between 2002 and 2007 than between 2002 and 2010 (Table 2). For six of the eight traits, the sign of evolution rate shifted between 2002-2007 and 2007-2010 comparisons (Table 2). Absolute evolution rate between 2002 and 2007 ranged from 0.043 *haldanes* for the average length between two fruits on the main stem to 0.411 *haldanes* for germination time (mean = 0.138).

Edaphic variation along the transect

A significant variation was observed for each of the ten edaphic variables considered in this study, with a variation up to more than 20-fold for concentrations in magnesium and aluminum (Fig. 4). Edaphic variation was often observed at a very short spatial scale, like for total nitrogen, pH, water holding capacity and concentrations in P_20_5, K and Al in the first 20 meters of the transect.

Interestingly, we found that edaphic variation was sometimes even higher within the population TOU-A than across 11 natural populations of *A. thaliana* collected in the same geographical region than TOU-A (Brachi *et al.* 2013b). For example, the range of

concentrations in magnesium, potassium and aluminum was larger within the population TOU-A than across the 11 populations from Burgundy (Fig. 5). Variation in total nitrogen within the population TOU-A represented 60.7% of the variation in total nitrogen observed across the 11 populations (Fig. 5).

Phenotypic evolution in relationship with edaphic factors

Highly significant 'year * edaphic variable' interactions have been found between survival and concentration in sodium and between reproductive period and C/N ratio (Table 3). The linear relationship between survival and sodium shifted from negative to positive across years, with a phenotypic evolution for a higher survival rate in areas along the transect with the highest sodium concentrations (Fig. 6A). Interestingly, the relationship between the length of reproductive period and C/N ratio shifted across years from a significant positive linear relationship to a significant positive parabolic relationship (Fig. 6B).

Discussion

Rapid (?) phenotypic evolution

In agreement with the expectation that fast directional selection over a short time scale interspersed by periods of no apparent evolutionary change should be observed in most populations (Kinnison and Hendry 2001), our resurrection study indicates a significant phenotypic evolution in a local population of *A. thaliana* over less than five generations, followed by a quasi-stasis over three generations. This result suggests that the rates of evolutionary change could be underestimated, not to say non-significant, if the time points of

sample collecting are too dispersed in comparison to the time scale of environmental change (Hendry and Kinnison 1999). The observed decrease in the rate of evolutionary change may indicate either that the local adaptive peak of the population is approached or that rate of the environmental change varied over time (Bone and Farres 2001).

Eight traits involved in different biological processes were found to evolve over a short-time scale, which is consistent to empirical studies showing that adaptive response to an environmental change involves a suite of traits rather than a single trait (Ghalambor *et al.* 2007). Survival (but not fecundity) increased between 2002 and 2007 and remained stable between 2007 and 2010. Interestingly, population growth approximated by the density of plants along the transect was also found to follow the same temporal pattern, suggesting that an increase in survival rate contributed to population growth in the TOU-A population.

The rates of evolutionary changes found in the TOU-A population fall in the low end of the distribution of *haldanes* values reviewed in plants (Bone and Farres 2001). One exception with a higher rate of evolutionary change concerns germination timing. Because only one generation of growing in controlled greenhouse conditions has been performed to reduce maternal effects of the families collected in 2007 and 2010 (in contrast two generations for the families collected in 2002), we cannot rule out the possibility of grand-maternal effects that still persist for those families.

Whether the term ‘rapid’ should be used for the phenotypic evolution observed for the TOU-A population is on open question. Because the term ‘rapid’ has often be used to describe the rates of evolutionary change (Kinnison and Hendry 2001), Bone and Farrer (2001) suggested to use ‘rapid’ only for the exceptional evolutionary changes, *i.e.* rate of evolutionary change in the right tail of the distribution of *haldanes* values. We should therefore be cautious in using the term ‘rapid’ in our resurrection study. On the other hand,

most studies estimating the rate of evolutionary change in plants focused on populations that may have faced strong selective pressures, like a known (generally anthropogenically induced) environmental change (*e.g.* herbicides), colonization of a new environment (*e.g.* invasive species) or artificial selection. In this study, the choice of a temporal tracking of the TOU-A population was first based on the high genetic and phenotypic diversity present in 2002 rather than on an obvious environmental change. We might speculate that the heat wave in 2003 in France (https://espaceprofessionnels.meteofrance.com/espaceprofessionnels/accueil) led to a drastic directional selection on seed dormancy. Although less than 50 plants of the TOU-A population produced seeds in 2003, hundreds to thousands of plants however produced seeds in spring 2004 (Fabrice Roux, personal observation). The presence of this high number of plant producing seeds in 2004 (even in areas along the transect where no plant producing seeds was observed in spring 2003) likely resulted from the presence of a seed bank as previously observed in many other natural populations of *A. thaliana* (Lundemo *et al.* 2009, Picó 2012). It would be worth testing whether seed dormancy increased between 2002 and 2007. Another hypothetical environmental change would be related to an increased intensity in plant competition. In agreement with this hypothesis, we observed an evolution towards (i) an increase in stem elongation and a shorter reproductive period which both confer a fitness advantage among plants in crowded settings (Dudley and Schmitt 1996, Weinig *et al.* 2006, Brachi *et al.* 2012), and (ii) a longer time interval between bolting and flowering whose genes associated with natural variation in a set of worldwide accessions of *A. thaliana* are related to pathways involved in the shade avoidance syndrome (Brachi *et al.* 2013a). Monitoring the micro-evolutionary phenotypic dynamics across a range of intensities in interspecific competition (with species encountered by *A. thaliana* in the TOU-A community) in artificial

populations initiated by families collected in 2002 would allow testing whether an increased intensity in plant competition may have driven the observed phenotypic evolution in the TOU-A population.

Impact of fine-grained edaphic variation on phenotypic evolution

While edaphic variation was found to vary at a small geographical scale in *A. thaliana*, *i.e.* among natural populations in a French region (Brachi *et al.* 2013b), large edaphic variation was also found at the within-population scale at the scale of less than 20 meters. Given that, in *A. thaliana*, seed dispersal in mainly passive (Weinig *et al.* 2006) and mean dispersal distance of seeds is highly correlated to the height of the plants (Wender *et al.* 2005), *i.e.* ~32cm in our study (Fig. 2), dispersal distance of most seeds may not exceed the spatial scale of edaphic variation observed in this study. If families showed genotype-by-environment interactions for fitness across different types of soil along the transect, we may expect a larger edaphic grain with respect to seed dispersal to promote micro-local adaptation within the TOU-A population. Two observations may support this expectation. First, by sowing a bulk mixture of seeds of three highly phenotypically differentiated genotypes of *A. thaliana* in a field, Stratton and Bennington (1996) found that the final spatial distribution of those genotypes was consistent with natural selection acting at the scale of nearly 50cm. Second, phenotypic evolution was found to depend on micro-scale edaphic variation for survival and the length of reproductive period in our study. While selection for a higher survival rate in areas along the transect with the highest sodium concentrations seems biologically meaningful, it is hard to come up with a physiological explanation for a rate of evolutionary change that may depend on the organic carbon / total nitrogen ratio for the length of the reproductive period. There is also a possibility that the edaphic-dependence evolution of the

length of reproductive period results also from a correlated selection on another phenotypic trait that we did not measure. Nonetheless, the putative edaphic-dependence phenotypic evolution described in this study suggests that the rate of contemporary microevolution within a population may depend on a fine-scale selective heterogeneity.

We cannot rule out that other selective agents in the TOU-A population have a spatial scale of variation different from the spatial scale of variation for edaphic factors. For example, the spatial scale of functional species composition of plant communities may be different from the spatial scale of edaphic variation, with abundance and identity of neighboring species that may vary at a smaller scale than edaphic variation. Given the diversity of selective agents acting simultaneously on a suite of traits and its putative associated complexity of overlapping environmental grains, one may want to also consider the contribution of adaptive plasticity in phenotypic evolution in a spatially and temporally heterogeneous environment (Ghalambor *et al.* 2007). Recently, rapid adaptive evolution within populations of the invasive plant species *Polygonum cespitosum* was suggested to be facilitated by an evolutionary change in adaptive plasticity (Sultan *et al.* 2012). An *in-situ* 'reciprocal transplant – resurrection' study involving the growing of families collected in 2002 and 2010 on three soil types in combination with an interspecific competition treatment is currently underway to tease apart the relative contribution of putative adaptive phenotypic plasticity and fine-scale selective heterogeneity in phenotypic evolution observed in the TOU-A population.

Acknowledgments

Special thanks are given to Cédric Glorieux, Nathalie Faure, Stella Huynh, Cyprien Moriss and Angélique Bourceaux for their assistance during the field experiments. This study was supported by a PhD fellowship from the University of Lille 1 to R.V.

References

Anderson JT, Inouye DW, McKinney AM, Colauti RI, Mitchell-Olds T. 2012. Phenotypic plasticity and adaptive evolution contribute to advancing flowering phenology in response to climate change. *Proceedings of the Royal Society B* 279: 3483-3852.

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

Bomblies K, Yant L, Laitinen RA, Kim S-T, Hollister JD, Warthmann N, Fitz J, Weigel D. 2010. Local-scale patterns of genetic variability , outcrossing, and spatial structure in natural stands of *Arabidopsis thaliana*. *PLoS Genetics* 6: e1000890.

Bone E, Farres A. 2001. Trends and rates of microevolution in plants. *Genetica* 112-113: 165-182.

Bossdorf O, Auge H, Lafuma L, Rogers WA, Siemann E, Prati D. 2005. Phenotypic and genetic differentiation between native and introduced populations. *Oecologia* 144: 1-11.

Brachi B, Aimé C, Glorieux C, Cuguen J, Roux F. 2012. Adaptive value of phenological traits in stressful environments : predictions based on seed production and Laboratory Natural Selection. *PLoS One* 7: e32069.

Brachi B, Faure N, Bergelson J, Cuguen J, Roux F. 2013a. 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* (DOI:10.1080/12538078.2013.807302)

Brachi B, Faure N, Horton M, Flahauw E, Vazquez A, Nordborg M, Bergelson J, Cuguen J, Roux F. 2010. Linkage and association mapping of *Arabidopsis thaliana* flowering time in nature. *PLoS Genet*ics 6: e1000940.

Brachi B, Villoutreix R, Faure N, Hautekèete N, Piquot Y, Pauwels M, Roby D, Cuguen J, Bergelson J, Roux F. 2013b. Investigation of the geographical scale of adaptive phenological variation and its underlying genetics in *Arabidopsis thaliana*. *Molecular Ecology* 22: 4222-4240.

Buswell JM, Moles AT, Hartley S. Is rapid evolution common in introduced plant species. *Journal of Ecology* 99: 214-224.

Dudley SA, Schmitt J. 1996. Testing the adaptive plasticity hypothesis: density-dependent selection on manipulated stem length in *Impatiens capensis*. *The American Naturalist* 147: 445-465.

Fitter AH, Fitter RSR. 2002. Rapid changes in flowering time in British plants. *Science* 296: 1689-1691.

Fournier-Level A, Korte A, Cooper MD, Nordborg M, Schmitt J, Wilczek AM. 2011. A map of local adaptation in *Arabidopsis thaliana*. *Science* 334: 86–89.

Franks SJ, Sim S, Weis AE. 2007. Rapid evolution of flowering time by an annual plant in response to a climate fluctuation. *Proceedings of National Acadely of Sciences of the USA* 104: 1278-1282.

Franks SJ, Avise JC, Bradshaw WE, Conner JK, Etterson JR, Mazer SJ, Shaw RG, Weis AE. 2008. The resurrection initiative: storing ancestral genotypes to capture evolution in action. *BioScience* 58: 870-873.

Gaut B. 2012. *Arabidopsis thaliana* as a model for the genetics of local adaptation. *Nature Genetics* 44: 115–121.

Ghalambor CK, McKay JK, Carroll SP, Reznick DN. 2007. Adaptive versus non-adaptive phenotypic plasticity and the potential for contemporary adaptation in new environments. *Functional Ecology* 21: 394-407.

Hancock AM, Brachi B, Faure N, Horton MW, Jarymowycz LB, Sperone FG, Toomajian C, Roux F, Bergelson J. 2011. Adaptation to Climate Across the *Arabidopsis thaliana* Genome. *Science* 334: 83–86.

Hedrick PW. 1986. Genetic polymorphism in heterogeneous environments, ten years later? *Annual Review of Ecology and Systematics* 17: 535-566.

Hendry AP, Kinnison MT. 1999. The pace of modern life: measuring rates of contemporary microevolution. *Evolution* 53: 1637-1653.

Holt JS, Le Baron HM. 1990. Significance and distribution of herbicide resistance. *Weed Technology* 4: 141-149

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

Jasienuk M, Brûlé-Babel AL, Morrison IN. 1994. Inheritance of trifluralin resistance in green foxtail (Setaria viridis). *Weed Science* 42: 123-127.

Kalisz S. 1986. Variable selection on the timing of germination in *Collinsia verna* (Scrophulariaceae). *Evolution* 40: 479-491.

Kingsolver JG, Hoekstra HE, Hoekstra JM, Berrigan D, Vignieri SM, Hill CE, Hoang A, Gibert P, Beerli P. 2001. The strength of phenotypic selection in natural populations. *The American Naturalist* 157: 245-261.

Kinnison MT, Hendry AP. 2001. The pace of modern life II: from rates of contemporary microevolution to pattern and process. *Genetica* 112-113: 145-164.

Lê S, Josse J, Husson F. 2008. FactoMineR: An *R* Package for Multivariate Analysis. *Journal of Statistical Software*. 25: 1-18.

Le Corre V. 2005. Variation at two flowering time genes within and among populations of *Arabidopsis thaliana*: comparison with markers and traits. *Molecular Ecology* 14: 4181–4192.

Lowry DB. 2012. Local adaptation in The model plant. *New Phytologist* 194: 888-890.

Lundemo S, Falahati-Anbaran M, Stenoien HK. 2009. Seed banks cause elevated generation times and effective population sizes of *Arabidopsis thaliana* in northern Europe. *Molecular Ecology* 18: 2798-2811.

Méndez-Vigo B, Gomaa NH, Alonso-Blanco C, Picó FX. 2013. Among- and within-population variation in flowering time in Iberian *Arabidopsis thaliana* estimated in field and glasshouse conditions. *New Phytologist* 197: 1332-1343.

Miller-Rushing AJ, Inouye DW. 2009. Variation in the impact of climate change in flowering phenology and abundance: an examination of two pairs of closely related wildflowers species. *American Journal of Botany* 96: 1821-1829.

Mitchell-Olds T, Schmitt J. 2006. Genetic mechanisms and evolutionary significance of natural variation in *Arabidopsis*. *Nature* 441: 947-952.

Montesinos A, Tonsor SJ, Alonso-Blanco C, Picó FX. 2009. Demographic and genetic patterns of variation among populations of *Arabidopsis thaliana* from contrasting native environments. *PLoS One* 4: e7213.

Nagylaki T. 1975. Conditions for the existence of clines. *Genetics* 80: 595-615.

Novy A, Flory SL, Hartman JM. 2013. Evidence for rapid evolution of phenology in an invasive species. *Journal of Evolutionary Biology* 26: 443-450.

Picó FX. 2012. Demographic fate of *Arabidopsis thaliana* cohorts of autumn- and spring-germinated plants along an altitudinal gradient. *Journal of Ecology* 100: 1009-1018.

Platt A, Horton M, Huang YS, Li Y, Anastasio AE, Mulyati NW, Agren J, Bossdorf O, Byers D, Donohue K, Dunning M, Holub EB, Hudson A, Le Corre V, Loudet O, Roux F, Warthman N, Weigel D, Rivero L, Scholl R, Nordborg M, Bergelson J, Borevitz JO. 2010. The scale of population structure in *Arabidopsis thaliana*. *PLoS Genetics* 6: e10000843.

Primack RB, Kang H. 1989. Measuring fitness and natural selection in wild plant populations. *Annual Review of Ecology and Systematics* 20: 367-396.

Reboud X, Le Corre V, Scarcelli N, Roux F, David JL, Bataillon T, Camilleri C, Brunel D, McKhann H. 2004. Natural variation among accessions of *Arabidopsis thaliana*: Beyond the flowering date, what morphological traits are relevant to study adaptation?

pp. 135–142 in Plant Adaptation: Molecular Biology and Ecology, edited by QC Cronk, J Whitton and IEP Taylor. NRC Research Press, Ottawa, Canada.

Reznick DN, Ghalambor CK. 2001. The population ecology of contemporary adaptations: what empirical studies reveal about the conditions that promote adaptive evolution. *Genetica* 112-113: 183-198.

Roux F, Gasquez J, Reboud X. 2004. The dominance of the herbicide resistance cost in several *Arabidopsis thaliana* mutant lines. *Genetics* 166: 449-460.

Roux F, Paris M, Reboud X. 2008. Delaying weed adaptation to herbicide by environmental heterogeneity: a simulation approach. *Pest Management Science* 64: 16-29.

Sax DF, Stachowicz JJ, Brown JH, Bruno JF, Dawson MN, Gaines SD, Grosberg RK, Hastings A, Hold RD, Mayfield MM, O'Connor MI, Rice WR. 2007. Ecological and evolutionary insights from species invasion. *Trends in Ecology and Evolution* 22: 465-471.

Sherrard ME, Maherali H. 2012. Local adaptation across a fertility gradient is influenced by soil biota in the invasive grass, *Bromus inermis*. *Evolutionary Ecology* 26: 529-544.

Slatkin M. 1973. Gene flow and selection in a cline. *Genetics* 75: 733-756.

Stratton DA. 1994. Genotype-by-environment inetarctions for fitness of *Erigeron annuus* show fine-scale selective heterogeneity. *Evolution* 48: 1607-1618.

Stratton DA. 1995. Spatial scale of variation in fitness of *Erigeron annuus*. *The American Naturalist* 146: 608-624.

Stratton DA, Bennington CC. 1996. Measuring spatial variation in natural selection using randomly-sown seeds of *Arabidopsis thaliana*. *Journal of Evolutionary Biology* 9: 215-228.

Stratton DA, Bennington CC. 1998. Fine-grained spatial and temporal variation in selection does not maintain genetic variation in *Erigeron annuus*. *Evolution* 52: 678-691.

Sultan SE, Horgan-Kobelski T, Nichols LM, Riggs CE, Waples RK. 2012. A resurrection study reveals rapid adaptive evolution within populations of an invasive species. *Evolutionary applications* 6: 266-278.

Turkington R, Harper JL. 1979. The growth, distribution, and neighbour relationships of *Trifolium repens* in a permanent pasture. IV. Fine-scale biotic differentiation. *Journal of Ecology* 67: 245-254.

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

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

Figure legends

Fig. 1 TOU-A population. (A) Photograph showing the habitat type. The population is located under a 350m electric fence separating two permanent meadows. (B) Position of families for which seeds have been collected in 2002, 2007 and 2010. Position of soil samples collected in 2010.

Fig. 2 Violin plots (*i.e.* box-and-whisker plot overlaid with a kernel density plot) of phenotypic evolution of 13 traits between 2002 and 2010. Different letters indicate different significant values at $P = 0.05$ after Tukey's multiple comparisons tests. FITTOT: total silique length, GERM: germination time, BT: bolting time, INT: interval between bolting and flowering, RP: reproductive period, Area: rosette surface area, DIAM: rosette diameter at bolting, H1S: height from soil to the first silique on the main stem, HSTEM: height of the main stem, INTERNOD: length of inflorescence between two fruits on the main stem, RAMBB: number of basal branches, RAMPB_S: number of primary branches on the main stem.

Fig. 3 Year effect in a phenotypic space. (A) Factor loading plot resulting from principal components analysis procedure. Factor 1 and factor 2 explained 27.78% and 15.54% of total phenotypic variance. (B) Position of the 313 families in the 'Factor1 – Factor 2' phenotypic space. Red circles, blue circles and black diamonds correspond to the genotypic values of families collected in 2002, 2007 and 2010, respectively. The principal component analysis

(PCA) was run on the 13 standardized traits using the *R* package 'FactoMineR' (Lê *et al.* 2008).

Fig. 4 Edaphic variation across the transect. The dotted lines correspond to the positions along the transect every 10 meters.

Fig. 5 Box-and-whiskers plots of edaphic variation in 11 stands from Burgundy and 80 TOU-A samples.

Fig. 6 Relationship between traits and edaphic variables across years. A. Survival and concentration in Na. B. Length of reproductive period (RP) and C/N ratio. Red circles, blue circles and black diamonds correspond to the genotypic values of families collected in 2002, 2007 and 2010, respectively. For each year, a polynomial regression including both linear and quadratic terms was run using raw data of the phenotypic trait and the corresponding ecological variable. For illustration purposes, the lines were drawn using the parameters from this polynomial regression. Beta (linear) and gamma (non-linear) values correspond to partial linear and quadratic regression coefficients obtained from a polynomial regression of the standardized values of phenotypic trait against the standardized values of the 10 ecological variables.

Fig. 1

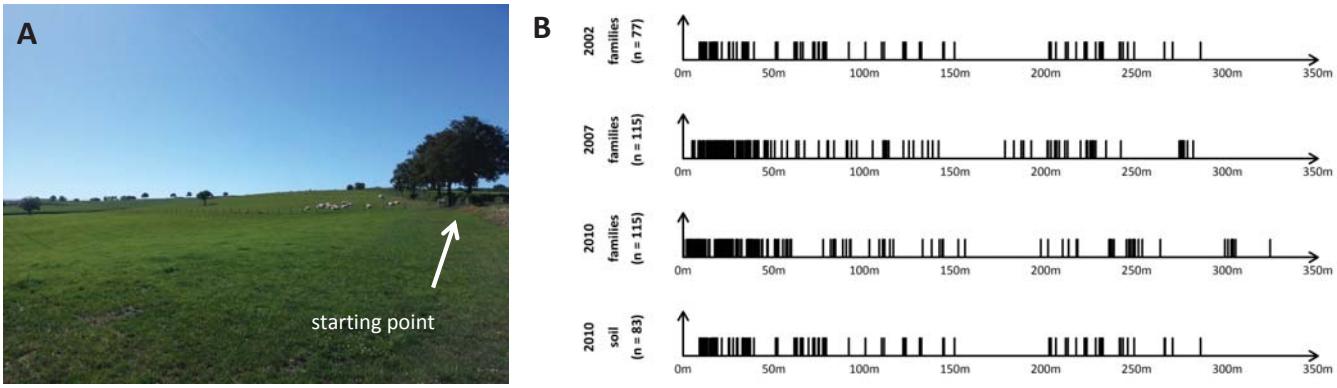


Fig. 2

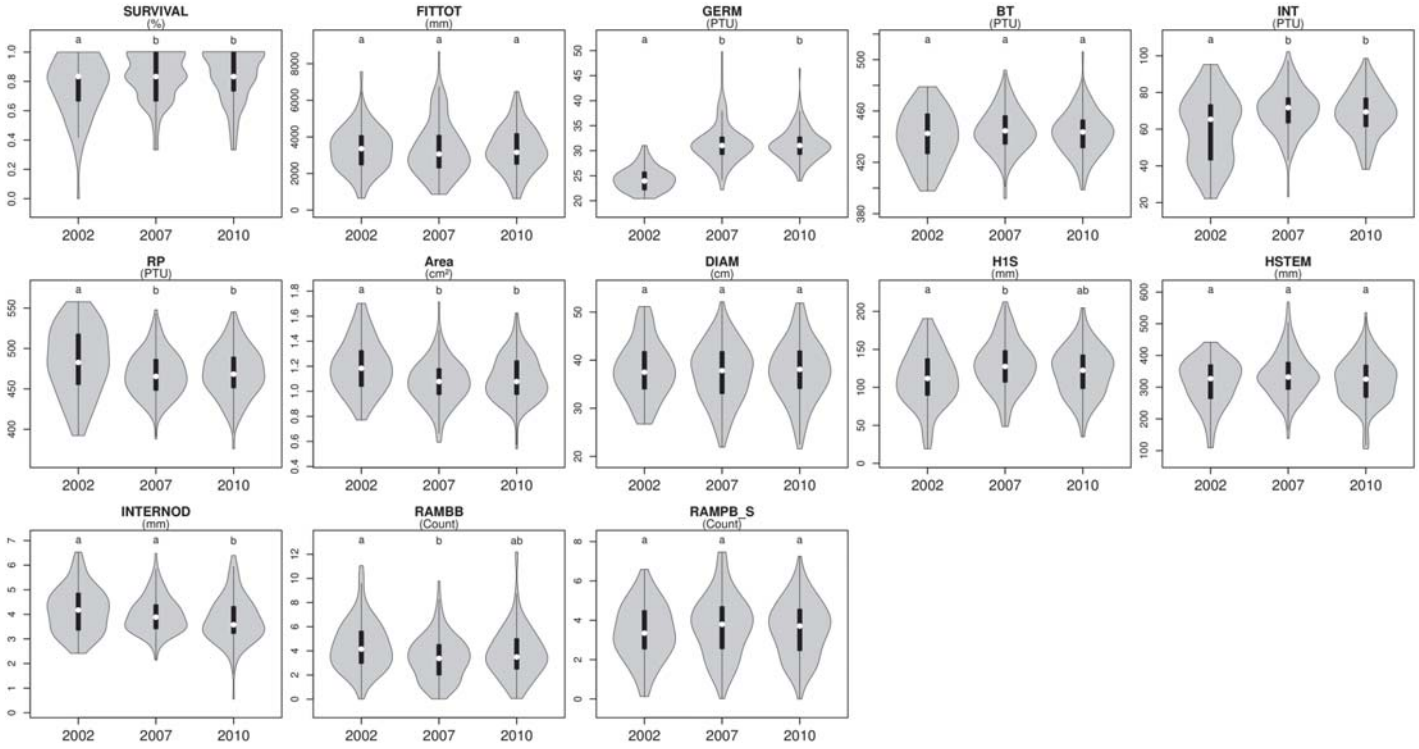


Fig. 3

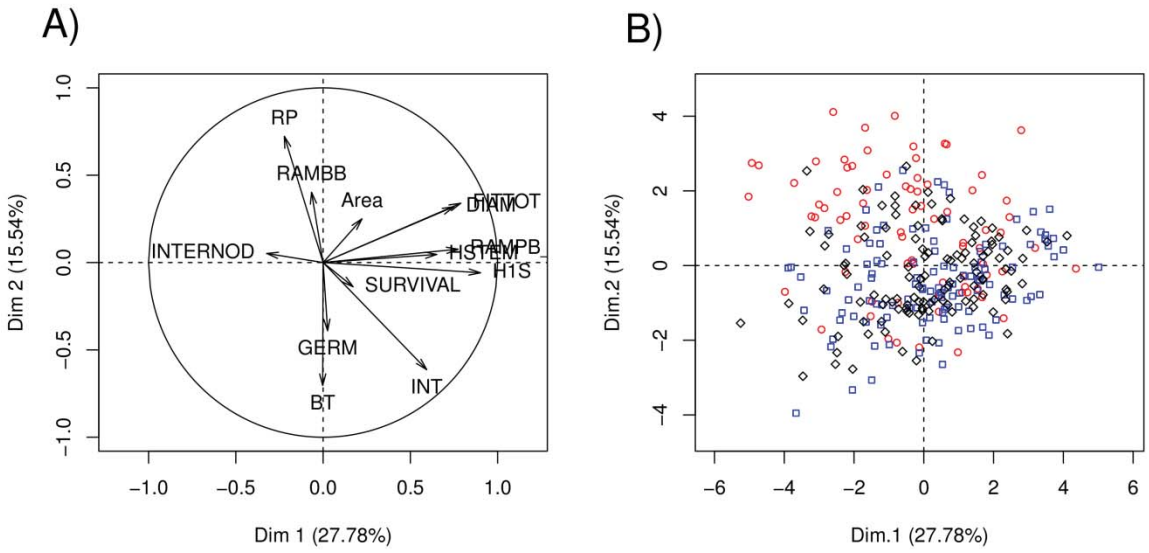


Fig. 4

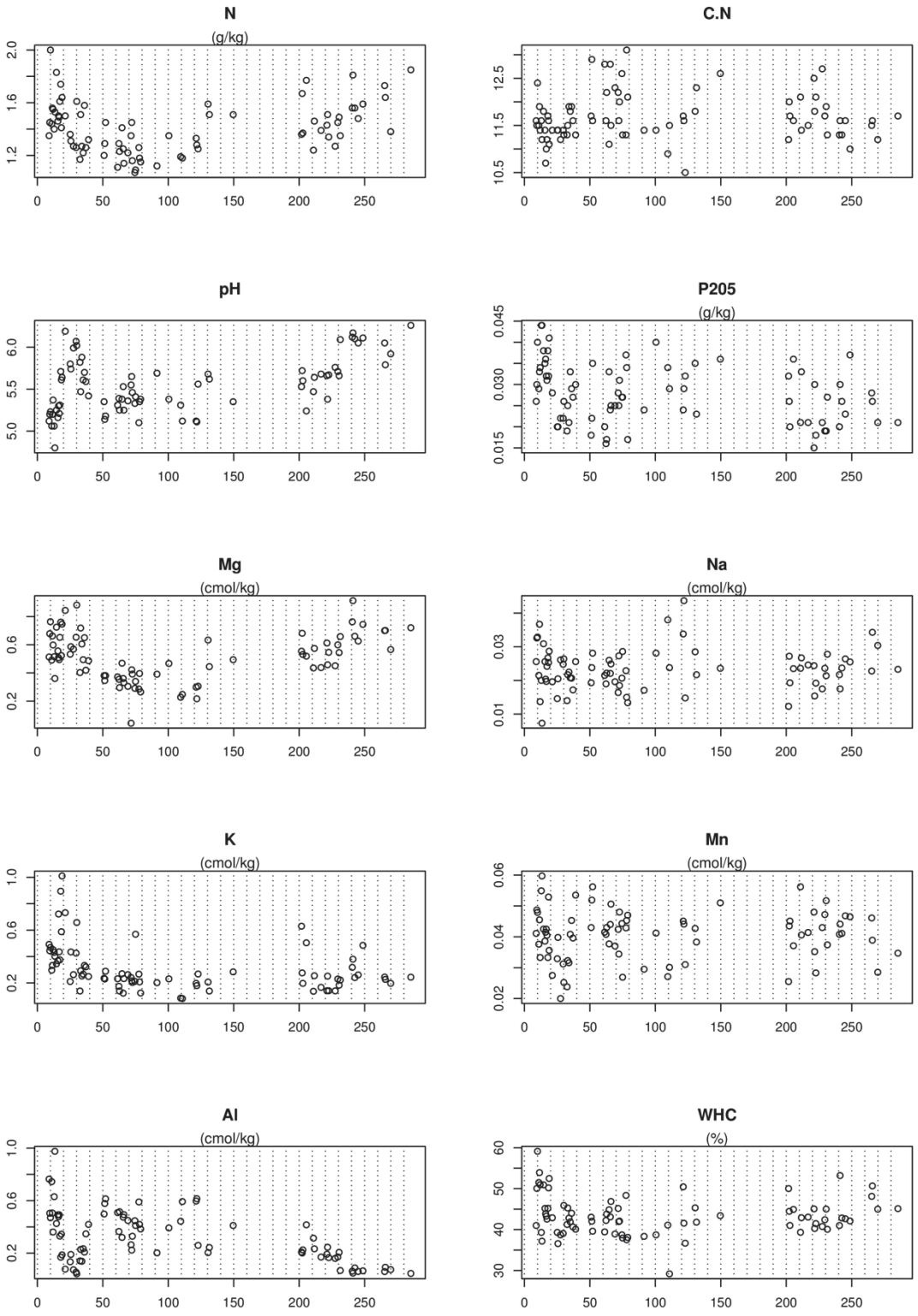


Fig. 5

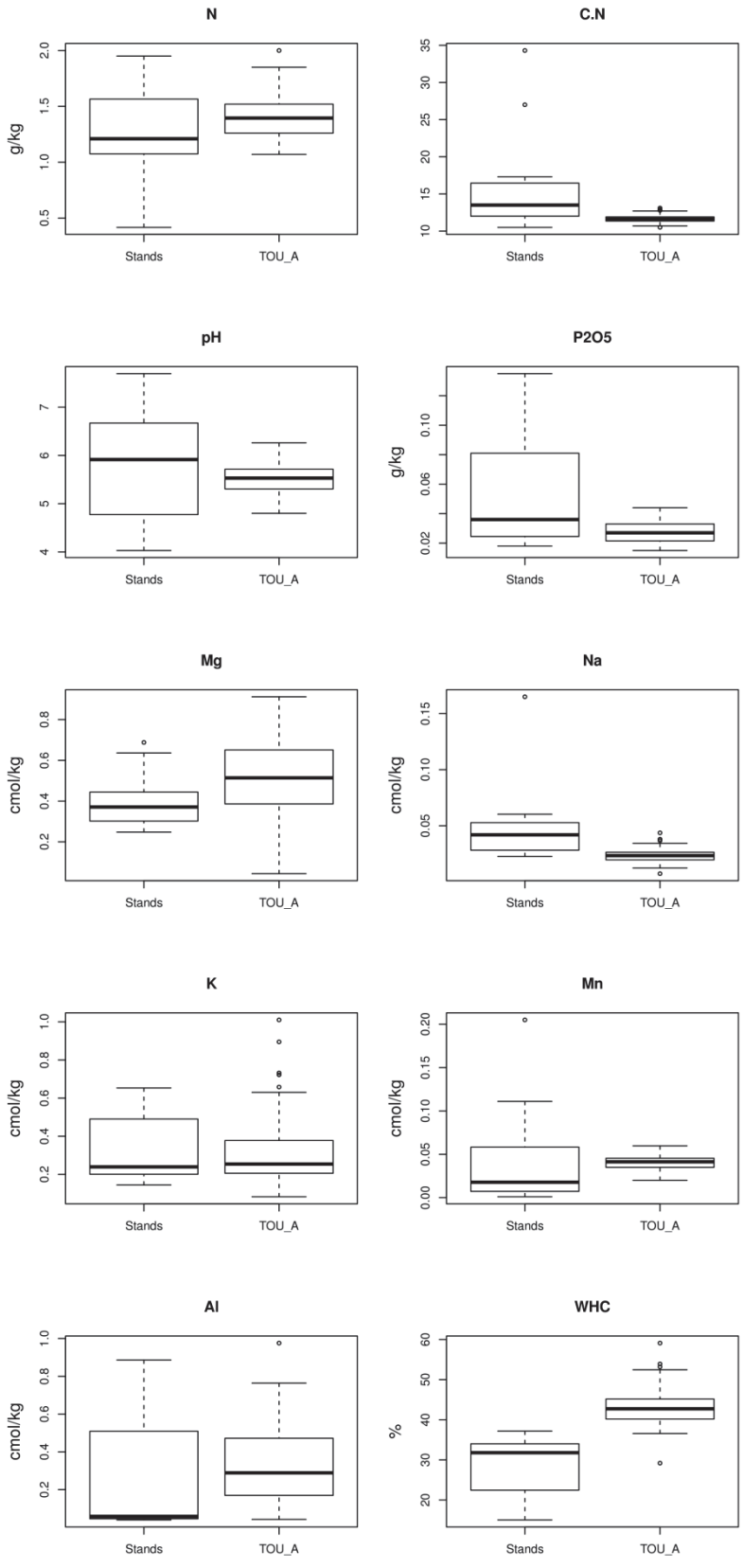


Fig. 6

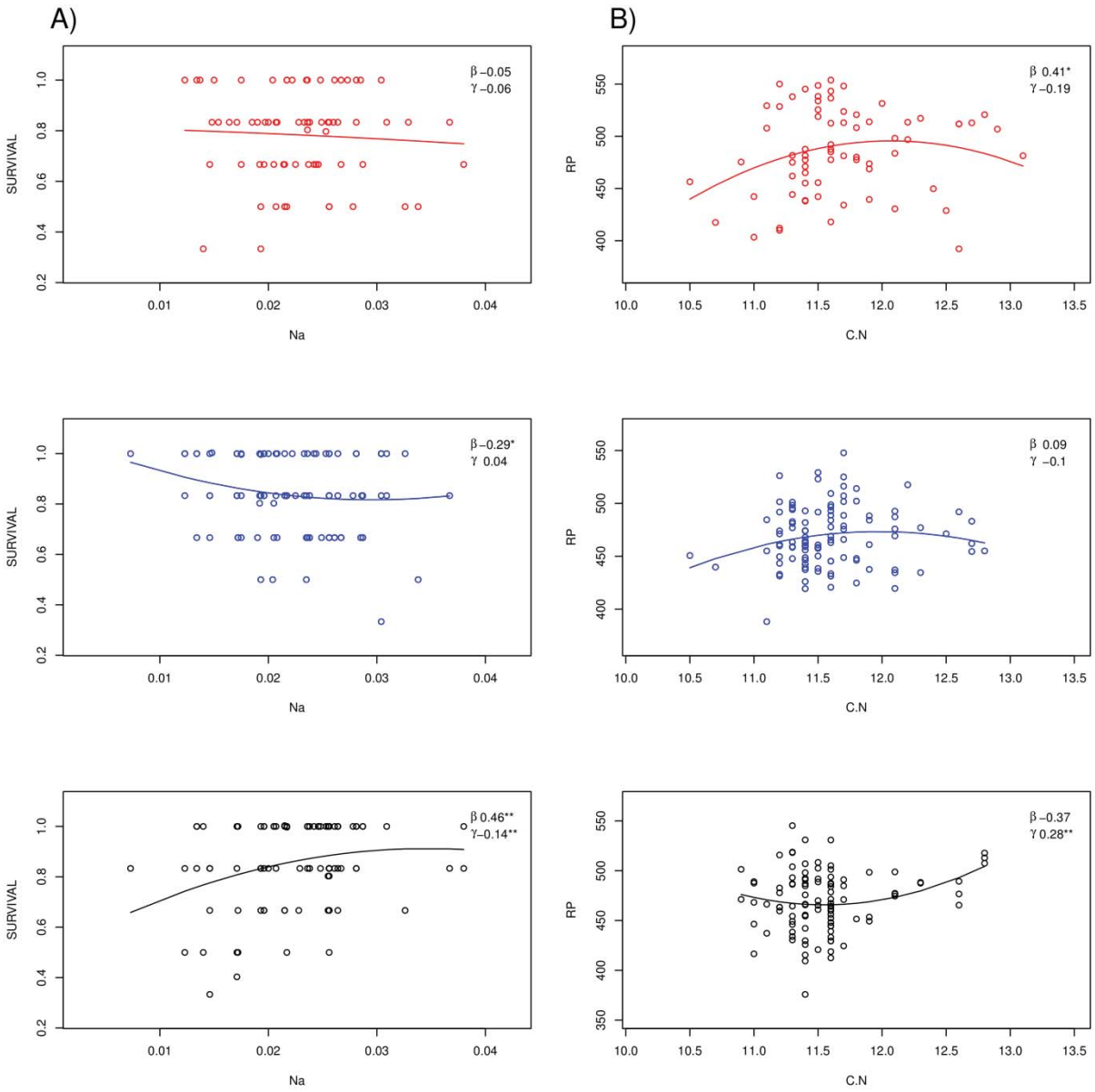


Table 1 Natural variation of 13 phenotypic traits. FITTOT: total silique length, GERM: germination time, BT: bolting time, INT: interval between bolting and flowering, RP: reproductive period, Area: rosette surface area, DIAM: rosette diameter at bolting, H1S: height from soil to the first silique on the main stem, HSTEM: height of the main stem, INTERNOD: length of inflorescence between two fruits on the main stem, RAMBB: number of basal branches, RAMPB_S: number of primary branches on the main stem.

Traits	survival		FITTOT		GERM		BT		INT		RP		AREA		DIAM		H1S		HSTEM		INTERNOD		RAMPB_S		RAMBB	
Model terms	*F* or *LRT*	*P*	*F* or *LRT*	*P*	*F* or *LRT*	*P*	*F* or *LRT*	*P*	*F* or *LRT*	*P*	*F* or *LRT*	*P*	*F* or *LRT*	*P*	*F* or *LRT*	*P*	*F* or *LRT*	*P*	*F* or *LRT*	*P*	*F* or *LRT*	*P*	*F* or *LRT*	*P*	*F* or *LRT*	*P*
Block	3.8	ns	2.1	ns	0.3	ns	**70.0**	***	6.3	*	0.6	ns	**41.1**	***	0.7	ns	0.4	ns	4.7	*	**8.6**	**	**11.7**	***	0.0	ns
Year	**4.9**	**	0.4	ns	**134.6**	***	2.9	ns	**14.0**	***	**6.6**	**	**6.6**	**	0.3	ns	**5.2**	**	2.2	ns	**6.8**	**	1.1	ns	**5.3**	**
Block*Year	2.2	ns	4.1	*	**16.1**	***	0.1	ns	0.2	ns	0.9	ns	0.0	ns	2.8	ns	**5.0**	**	4.0	*	3.3	*	3.9	*	0.5	ns
Family(Year)	**6.8**	**	**8.3**	**	**64.2**	***	**63.8**	***	**74.2**	***	**41.0**	***	**31.9**	***	**40.8**	***	**43.8**	***	**29.5**	***	**37.7**	***	**19.9**	***	**17.7**	***
*Block*Family(Year)*	0.0	ns	1.1	ns	0.0	ns	0.1	ns	0.0	ns	0.0	ns	0.0	ns	0.0	ns	0.0	ns	0.0	ns	0.0	ns	0.0	ns	0.0	ns
Control	NE	NE	0.3	ns	**38.3**	***	0.0	ns	1.2	ns	3.4	ns	**123.8**	***	0.9	ns	0.3	ns	1.6	ns	0.0	ns	0.6	ns	0.6	ns

Model random terms were tested with likelihood ratio tests (LRT) of models with and without these effects. Random effects are in italic. *0.05 > *P* > 0.01, **0.01 > *P* > 0.001, *** *P* < 0.001. Bold stars indicate significant effect after Bonferroni correction.

Table 2 Evolution rates expressed in *haldanes*.

Trait	2002-2007	2002-2010	2007-2010
SURVIVAL	0.080	0.056	0.024
GERM	0.411	0.295	-0.017
INT	0.139	0.073	-0.043
RP	-0.097	-0.051	0.023
AREA	-0.135	-0.062	0.062
H1S	0.097	0.032	-0.082
INTERNOD	-0.043	-0.051	-0.081
RAMBB	-0.103	-0.034	0.073

Table 3 Phenotypic evolution in relation with edaphic conditions. FITTOT: total silique length, GERM: germination time, BT: bolting time, INT: interval between bolting and flowering, RP: reproductive period, AREA: rosette surface area, DIAM: rosette diameter at bolting, H1S: height from soil to the first silique on the main stem, HSTEM: height of the main stem, INTERNOD: length of inflorescence between two fruits on the main stem, RAMBB: number of basal branches, RAMPB_S: number of primary branches on the main stem.

	Traits																					
	SURVIVAL		FITTOT		GERM		BT		INT		RP		DIAM		AREA		H1S		HSTEM		INTERNOD		RAMBB		RAMPB_S	
Model terms	*F*	*P*	*F*	*P*	*F*	*P*	*F*	*P*	*F*	*P*	*F*	*P*	*F*	*P*	*F*	*P*	*F*	*P*	*F*	*P*	*F*	*P*	*F*	*P*	*F*	*P*
Year	1.10		0.18		16.68	***	0.37		0.51		2.15		0.41		1.65		1.09		2.65		0.65		0.54		0.33	
N	0.95		1.71		0.45		2.75		0.10		0.06		1.68		0.15		0.54		1.12		0.48		0.66		0.41	
C.N	0.02		0.29		0.29		1.73		1.68		8.61	**	0.04		0.03		0.02		0.06		0.39		0.02		0.04	
pH	1.06		1.20		0.66		0.00		1.22		0.12		1.24		0.59		0.43		1.32		0.02		0.06		0.09	
P_2O_5	1.07		0.06		0.27		1.46		0.86		1.96		0.11		5.20	*	0.07		0.75		0.10		2.64		0.08	
Mg	0.70		1.12		1.56		1.26		2.07		0.33		2.52		3.69		1.80		0.01		1.88		0.01		1.06	
Na	0.05		0.15		0.04		1.28		0.03		0.16		0.42		1.78		0.05		0.68		0.01		0.15		0.11	
K	0.78		0.68		0.11		1.90		0.05		0.13		0.04		1.73		0.05		1.28		1.12		0.22		0.33	
Mn	0.08		0.40		0.55		3.98	*	0.14		2.79		0.85		0.70		0.03		0.20		0.19		0.55		0.23	
Al	1.91		0.27		0.22		0.44		1.00		0.02		0.10		1.63		0.08		1.92		0.00		0.04		0.37	
WHC	0.40		0.36		0.51		0.87		1.88		0.37		3.32		0.92		3.91	*	1.17		0.12		0.05		0.71	
N^2	0.00		1.07		0.18		2.61		4.62	*	1.58		0.93		2.86		0.53		0.00		0.81		0.68		0.99	
$C.N^2$	0.10		1.90		0.26		1.48		2.47		5.72	*	0.43		0.01		0.53		0.49		0.28		0.00		0.01	
pH^2	0.19		0.40		0.12		0.20		0.07		0.60		1.12		0.01		0.29		0.78		2.17		1.02		0.32	
$P_2O_5^2$	0.14		0.36		1.39		5.98	*	7.26	**	11.00	**	0.06		0.01		0.94		0.03		3.86		0.15		0.75	
Mg^2	0.33		0.04		0.01		1.34		0.52		0.05		0.57		0.75		1.52		4.76	*	0.13		3.16		0.95	
Na^2	0.34		0.00		0.06		0.88		1.63		2.27		3.02		0.93		0.70		0.63		0.26		0.09		0.11	
K^2	1.23		0.23		0.49		3.82		2.11		3.51		0.46		0.00		0.11		0.22		2.33		0.09		0.09	
Mn^2	2.44		0.65		0.04		1.90		1.33		4.30	*	0.88		0.25		0.44		3.63		2.99		0.00		0.04	
Al^2	0.00		0.00		0.65		0.00		2.55		0.34		0.02		0.77		0.01		0.57		4.78	*	0.60		0.88	
WHC^2	0.46		0.01		0.44		2.49		1.36		0.41		0.10		0.60		2.73		0.18		0.36		0.76		1.12	
Year:N	1.43		2.04		0.47		0.47		1.09		0.12		1.34		0.88		1.49		1.72		0.00		0.27		0.36	
Year:C.N	0.14		0.31		0.36		1.04		1.20		4.09	*	0.33		0.49		1.07		0.29		1.66		0.07		0.20	
Year:pH	0.76		0.23		1.04		0.02		0.72		0.48		1.01		0.05		0.06		0.58		0.12		0.07		0.62	
Year:P205	1.42		0.00		1.14		1.61		0.31		1.66		0.39		0.34		0.96		0.91		1.09		1.65		0.84	
Year:Mg	3.16	*	0.91		1.04		0.54		3.45	*	2.28		0.35		1.91		0.59		0.42		0.07		0.32		0.17	
Year:Na	5.77	**	0.06		1.68		1.18		3.42	*	0.05		0.20		3.39	*	0.66		0.58		0.18		0.86		0.41	
Year:K	1.48		0.56		0.22		1.52		1.74		0.47		0.48		1.90		0.07		0.34		1.19		0.40		0.72	
Year:Mn	0.12		0.30		0.79		0.62		0.02		0.74		2.14		0.04		0.39		0.04		2.30		0.02		0.05	
Year:Al	1.40		0.72		0.46		0.12		1.28		0.09		1.62		0.29		0.41		1.38		0.04		0.11		0.34	
Year:WHC	0.68		0.76		0.52		1.13		0.58		2.82		1.40		0.31		0.36		0.31		0.28		2.15		0.08	
$Year:N^2$	0.08		0.36		0.44		2.80		0.88		2.39		0.14		0.83		0.49		0.03		0.63		0.75		0.44	
$Year:C.N^2$	0.28		0.94		0.69		1.80		1.52		7.52	***	0.69		0.66		0.78		0.59		4.55	*	0.23		0.44	
$Year:pH^2$	0.32		1.37		0.09		0.02		0.14		0.15		0.09		2.29		0.19		0.09		2.82		0.34		0.32	
$Year:P205^2$	0.03		0.26		0.71		2.68		1.44		4.32	*	0.62		0.35		0.04		0.17		1.35		0.67		0.37	
$Year:Mg^2$	0.14		1.15		0.08		0.54		1.47		0.33		0.10		0.84		0.26		0.77		0.58		0.71		0.58	
$Year:Na^2$	0.39		0.04		0.51		0.44		1.44		1.24		0.35		0.75		0.27		0.28		0.42		0.38		0.10	
$Year:K^2$	0.38		1.84		0.09		2.24		1.99		1.08		1.86		1.29		0.57		0.20		2.56		0.22		0.04	
$Year:Mn^2$	0.54		0.97		1.80		0.25		0.33		1.58		0.24		1.05		0.16		1.20		3.78	*	0.04		1.03	
$Year:Al^2$	0.03		1.90		0.46		0.22		0.80		0.21		0.93		2.25		0.54		1.01		3.43	*	0.29		0.30	
$Year:WHC^2$	0.16		1.08		0.24		0.82		1.33		0.59		0.16		0.77		2.42		0.75		1.60		0.53		1.31	

*$0.05 > P > 0.01$, **$0.01 > P > 0.001$, *** $P < 0.001$.

SUPPORTING INFORMATION: RAPID PHENOTYPIC EVOLUTION AND FINE-GRAINED SPATIAL HETEROGENEITY IN A NATURAL POPULATION OF *ARABIDOPSIS THALIANA*: A RESURRECTION STUDY.

Fig. S1 A four-step treatment to obtain individual images. The second step was performed using the software Adobe Photoshop CS3 Extented (version 10.0) and the software ImageMagick (version 7:6.6.2.6-1ubuntu4.2). The software ImageJ was used (i) to manually select plants individually in the third step and (ii) to measure rosette surface area in the fourth step. Cutting an image of tray into 66 individual images in the third step was performed using a custom Perl script.

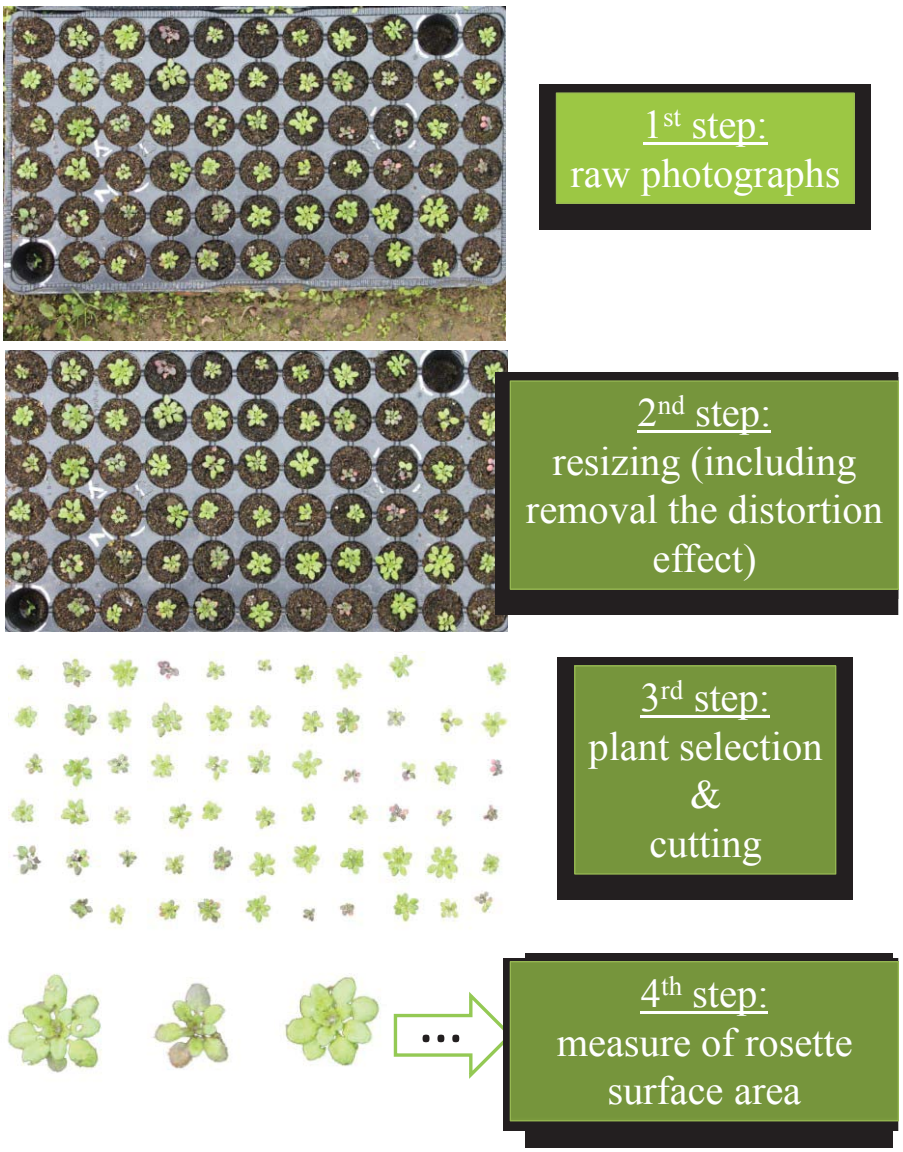


Fig. 2 Scatterplot matrix among 13 edaphic variables measured in the population TOU-A. Up and down values in each case above the diagonal correspond to Pearson and Spearman correlation coefficients, respectively. Values in brackets correspond to the 95% confidence intervals of the Pearson correlation coefficient. Red and blue values stand for positive and negative correlation coefficients, respectively. ns: non-significant, *0.05 > *P* > 0.01, **0.01 > *P* > 0.001, *** *P* < 0.001.

RAPPORT: UNE POPULATION NATURELLE D'*ARABIDOPSIS THALIANA* TRES DIVERSIFIEE PHENOTYPIQUEMENT ET GENETIQUEMENT SANS SIGNE APPARENT D'EVOLUTION PHENOTYPIQUE SUR UNE COURTE ECHELLE DE TEMPS: UNE ETUDE DE RESURRECTION

Une population naturelle d'*Arabidopsis thaliana* très diversifiée phénotypiquement et génétiquement sans signe apparent d'évolution phénotypique sur une courte échelle de temps: une étude de résurrection.

Introduction

Dans le manuscrit précédent, une évolution phénotypique rapide ainsi qu'une hétérogénéité édaphique importante sur une échelle spatiale parfois inférieure à 20 mètres ont été observées au sein de la population TOU-A. Cette population est située à l'interface de deux prairies permanentes et semble peu perturbée d'une année sur l'autre. Afin de tester si ces observations sont généralisables à d'autres populations naturelles d'*A. thaliana*, je me suis intéressé à une autre population située dans un milieu très différent de celui rencontré par la population TOU-A. En effet, la population MIB est située à l'interface d'une parcelle cultivée et d'une bande enherbée et subit des perturbations régulières, avec une rotation des cultures sur 3 ans et l'application de pesticides associée. La population MIB ayant été récoltée en 2002 et 2011, je me suis demandé si l'évolution phénotypique observée était cohérente avec les prédictions établies à partir des données acquises sur les plantes de 2002. Des familles génétiques échantillonnées en 2002 au sein de cette population MIB ayant aussi été caractérisées génétiquement pour 214k SNPs, je me suis demandé s'il était possible d'identifier les bases génétiques associées à la variation phénotypique par une approche de GWA mapping ; et si oui, quelles fonctions biologiques étaient associées à ces gènes.

Matériel et Méthodes

Matériel végétal

La population étudiée MIB est située sur la commune de Mirebeau-sur-Lèze (Côte-d'Or, France), à l'interface entre une parcelle cultivée et une bande enherbée (Figure 1a).

Figure 1 : La population MIB. a) Habitat de la population correspondant à l'interface entre une parcelle cultivée et une bande enherbée (photographie prise le 21 janvier 2011). b-g) Photographies de plantes d'*A. thaliana* situées le long de l'interface le 21 janvier 2011. h-m) Photographies de plantes d'*A. thaliana* lors de l'expérimentation réalisée sur un terrain expérimental de l'Université de Lille 1 (le 17 novembre 2011).

Les graines de 74 plantes ont été échantillonnées au printemps 2002 par Valérie Le Corre (Pôle GEAPSI, Equipe/thème Adaptation des adventices aux agroécosystèmes, INRA Dijon) tous les 1 à 2 mètres selon un transect de 100m située à l'interface parcelle cultivée – bande enherbée. En janvier 2011, 80 plantes ont été échantillonnées par Fabrice Roux (GEPV,

Université de Lille 1) suivant le même transect effectué par Valérie Le Corre en 2002. Les plantes échantillonnées en 2011 ont été ramenées en conditions contrôlées de serre (20°C, 16H de photopériode, Université de Lille 1) afin que ces plantes terminent leur cycle de vie. Les graines récoltées sur le terrain en 2002 et en serre en 2011 constituent des familles génétiques par plante échantillonnée. En 2002, la culture dans la parcelle correspondait à de la moutarde ; alors que du blé était cultivé sur cette parcelle en 2011.

Des graines des 74 familles échantillonnées en 2002 ont été envoyées en 2003 à Joy Bergelson (Department of Ecology & Evolution, University of Chicago) pour un projet sur l'échelle de la structure des populations chez *A. thaliana* (Platt *et al.* 2010). Pour ce projet, une production de graines a été effectuée pour chaque famille à partir d'une seule plante ayant poussé dans des conditions contrôlées de serre. Suite à ce projet, 52 des 74 familles ont été génotypées pour 214 kSNPs (Horton *et al.* 2012).

Pour notre étude, une production de graines dans des conditions contrôlées de serre (20°C, 16H de photopériode, Université de Lille 1) a été effectuée à la fin du printemps 2011 pour réduire les effets maternels des 154 familles. Une plante par famille issue des graines produites à l'Université de Chicago pour les familles échantillonnées en 2002 et une plante par famille issue des graines récoltées sur les plantes échantillonnées en 2011 a été mise en culture pour obtenir un lot de graines par famille génétique.

Design expérimental

Les graines de ces différentes familles ont été semées à l'Université de Lille 1 le 22 septembre 2011, ce qui correspond à la principale cohorte de germination observée dans les populations naturelles d'*A. thaliana* dans la région Nord de la France. L'expérience a été

organisée en deux blocs expérimentaux. Chaque bloc expérimental comprenait sept plaques de 66 alvéoles percées en leur fond (Ø4 cm, vol. ~38 cm3) (TEKU, JP 3050/66), remplies avec un substrat de culture standard (Huminsubstrat N3, Neuhaus). Chacun de deux blocs contenait un total de 462 plantes (154 familles * 3 réplicats) réparties de manière aléatoire. Un minimum de cinq graines ont été semées dans chaque alvéole. Afin de promouvoir leur germination, les graines semées ont été soumis à un traitement de stratification de 4 jours en chambre froide à une température de 4°C. A la fin du traitement de stratification, un traitement préventif contre les mouches du terreau (Vectobac, 8mL par Litre) a été effectué avant de placer les plaques alvéolées en serre champ pour une durée de 17 jours, ceci afin d'être au plus proche des conditions de photopériode et de température de l'extérieur tout en évitant la possibilité que la pluie ne mélange les graines des différentes accessions (aucun supplément en lumière ou température n'a été réalisé durant cette période). Une rotation journalière des plateaux a été réalisée dans la chambre froide et dans la serre champ afin de limiter la variation micro-environnementale. La date de germination des familles a été mesurée chaque jour durant les 17 jours où les plateaux ont été placés en serre champ. Le 18ème jour, un repiquage de plantules dans les alvéoles sans germination a été effectué à partir de plantes de la même famille ayant germé dans d'autres alvéoles. Ce même jour, (i) un démariage a été effectué afin de ne laisser que deux plantules de la même famille par alvéole et (ii) les plaques alvéolées ont été placées sur le terrain expérimental de Lille 1 (France) situé à une cinquante de mètres de la serre champ. Les sept plaques alvéolées de chaque bloc ont été disposées sur 2 colonnes et 4 lignes (seulement trois lignes pour la deuxième colonne). Afin de faciliter le développement racinaire des plantules, le sol a été préalablement biné à l'installation des plateaux. Les plantes ont également été protégées des herbivores vertébrés, des limaces et des pucerons (*Myzus persicae*) (description disponible dans Brachi *et al.* 2010).

Caractérisation phénotypique

Tout au long de l'expérience, les plantes ont été observées tous les jours en serre champ afin de relever la date de germination (ouverture des deux cotylédons) et tous les deux ou trois jours sur le terrain expérimental afin de relever la date de montaison (inflorescence distinguable entre les feuilles avec une hauteur supérieur à 5 mm), la date de floraison (apparition de la première fleur ouverte). Grace à ces trois date phénologiques, trois traits phénologiques ont été calculés (Brachi *et al.* 2012). Le temps de germination (GERM) a été mesuré comme le nombre de jours s'étant écoulé entre la fin du traitement de stratification et la date de germination. Le temps de montaison (BT) et l'intervalle entre montaison et floraison (INT) ont été mesurés comme le nombre de jours s'étant écoulé entre la date de germination et la date de montaison et entre la date de montaison et la date de floraison. GERM, BT et INT ont été transformés en unités photo-thermales (PTU) grâce à un modèle phénologique intégrant à la fois la photopériode et la température (voir Brachi *et al.* 2010 pour plus de détails).

Deux mesures estimant l'acquisition de ressources ont été réalisées sur chaque plante. La surface de la rosette (AREA) a été mesurée en utilisant une méthode non destructive, en photographiant les plateaux (35 jours après avoir transférer les plateaux de la serre froide sur le terrain expérimental) à l'aide d'un appareil photographique numérique de marque Canon (modèle EOS 500D). Le traitement de chaque photographie s'est effectué en quatre étapes (voir manuscrit sur la population TOU-A, Villoutreix *et al.* 2013 *en préparation*) permettant d'obtenir au final 924 images individuelles de plantes (154 familles * 6 réplicats). AREA a ensuite été mesurée informatiquement en utilisant le logiciel ImageJ (version 4.01; Universal Imaging Corporation, West Chester, PA, http://rsbweb.nih.gov/ij/). A montaison, le diamètre

maximum de la rosette a également été mesuré (DIAM) car ce trait s'avère être un bon estimateur de la taille des plantes d'*A. thaliana* (Weinig *et al.* 2006).

Après sénescence de toutes les plantes, la partie aérienne des plantes a été récoltée le 6 juin 2012 puis stockée à température ambiante en attendant que des traits architecturaux décrits comme importants pour la dispersion des graines et dans l'adaptation locale soient mesurés (Reboud *et al.* 2004, Wender *et al.* 2005) : le nombre de branches basales (RAMBB), le nombre de branches primaires sur la tige principale (RAMPB_S), la hauteur entre le sol et la première silique sur la tige principale (H1S) et la hauteur de la tige principale (HSTEM).

Toute plante ayant germé mais n'ayant pas survécu durant l'expérience a été comptée comme morte (SURVIVAL=0). Toute plante récolté ayant une date de germination a été comptée comme vivante (SURVIVAL =1). La production totale de graine a été estimée en mesurant la longueur de fruit totale (FITTOT) qui semble être un bon estimateur de la valeur sélective chez les plantes annuelles autogames tel qu'*A. thaliana* (Roux *et al.* 2004).

Caractérisation édaphique

Sept échantillons de sol ont été prélevés sur une profondeur de 5cm le long du transect. Ces échantillons de sol ont été transférés en serre afin d'être séchés à l'air libre, pour ensuite être stockés à température ambiante dans le laboratoire. Quatorze variables édaphiques ont été mesurés sur ces échantillons afin d'en déterminer les qualités agronomiques (voir Brachi *et al.* 2013) : pH, capacité maximale de rétention en eau (WHC), concentration en carbone (C) et azote total (N), rapport C/N, concentration en matière organique, et concentrations en P_2O_5, K, Ca, Mg, Mn, Al, Na and Fe. Peu de variation a été constatée entre nos sept échantillons pour l'ensemble des variables édaphiques, suggérant une homogénéité édaphique du milieu.

Analyse des données

Variation naturelle et évolution phénotypique

Afin d'explorer la variation phénotypique pour tous les traits phénotypiques mesurés dans cette étude (sauf la survie) et étudier leur potentielle évolution, nous avons utilisé la modèle statistique suivant (PROC MIXED procedure, REML method, SAS 9.3, SAS Institute Inc) :

$$Y_{ijklm} = \mu_{trait} + bloc_i + année_j + bloc_i*année_j + famille_k\ (année_j)) + bloc_i*\ famille_k\ (année_j)) + \varepsilon_{ijkc} \qquad (1)$$

Dans ces modèles, *'Y'* est un des 10 traits phénotypiques quantitatifs, 'μ' est la moyenne phénotypique globale ; le terme 'bloc' permet de tenir compte des différences micro-environnementales entre les deux blocs ; le terme 'année' teste les différences entre les deux années d'échantillonnage ; le terme 'famille' mesure l'effet des différentes familles génétiques au sein de chaque année d'échantillonnage ; les interactions impliquant le terme 'bloc' permettent de tenir compte de la variation génétique des normes de réaction des familles et des deux années entre les deux blocs ; le terme 'ε' correspond à l'erreur résiduelle du modèle. Une part de la variation phénotypique observée pour l'aire de la rosette pouvant résulter d'une variation de la date de germination, GERM a été rajouté comme co-variable dans le modèle (1) pour le trait AREA. Tous les termes du modèle ont été traités comme des effets fixes à l'exception des termes incluant l'effet 'famille' qui ont été traités comme aléatoires. Les effets fixes ont été testés en divisant le terme correspondant par le dénominateur approprié afin de calculer une *F*-value. La significativité des effets aléatoires a été testé pour chaque terme en comparant, grâce à un test de probabilité (likelihood test), le modèle complet avec un modèle ne contenant pas ce terme. Les données ont été log

transformées ou Boc-Cox transformées si besoin était, afin de satisfaire les hypothèses de normalité et d'égalité des variances associées aux régressions linéaires. La survie étant un trait binaire, le modèle (1) a été testé en utilisant la procédure PROC GLIMMIX dans SAS 9.3 (SAS Institute Inc., Cary, North Carolina, USA).

Au sein de chaque génération, les moyennes des moindres carrés (LSmeans) ont été calculées pour chaque famille en utilisant le modèle statistique suivant (R Development Core Team 2010):

$$Y_{im} = \mu_{\text{trait}} + \text{bloc}_i + \text{famille}_m + \varepsilon_{imc} \quad (2)$$

A. thaliana étant une espèce majoritairement autogame, les LSmeans calculées par cette méthode correspondent aux valeurs génotypiques des familles.

Prédictions de l'évolution phénotypique

Trois méthodes ont été utilisées pour prédire la micro-évolution phénotypique au sein de la population MIB à partir des données accumulées sur les plantes de la génération 2002 (Morrissey *et al.* 2010) :

1. Identité de Robertson-Price (Robertson-Price Identity) selon la formule suivante :

$$\Delta z = \sigma_a(z,w) \quad (3)$$

Où Δz est la différence de moyenne attendue du phénotype z entre deux générations successives, w est la fitness relative, et $\sigma_a(z,w)$ correspond à la covariance génétique additive entre z et w. Etant donné le régime de reproduction fortement autogame d'*A. thaliana*, la variance génétique est difficilement décomposable entre variance additive et variance épistatique dans notre étude. Nous avons donc remplacé le terme $\sigma_a(z,w)$

par le terme $\sigma_g(z,w)$ correspondant à la covariance génétique entre z et w. Dans cette étude, l'identité de Robertson-Price a été calculée à partir de valeurs génotypiques obtenues aussi bien pour les traits phénotypiques que le proxy de fitness FITTOT.

2. Univariate breeder's equation selon la formule suivante :

$$R = h^2 S \qquad (4)$$

Où R est la réponse à la sélection d'un trait phénotypique par génération, h^2 correspond à l'héritabilité au sens-strict, et S est le différentiel de sélection correspondant à la covariance phénotypique entre un trait et la fitness relative. Là encore, nous ne pouvons pas distinguer la variance additive de la variance épistatique, nous amenant à remplacer h^2 par H^2, l'héritabilité au sens large calculée comme suit dans notre étude (PROC VARCOMP procedure in SAS 9.3, SAS Institute Inc.) :

$$Yij = \mu_{trait} + \text{bloc}_i + \text{famille}_j + \varepsilon_{ij} \qquad (5)$$

$$\text{avec } H^2_{trait} = V_F / (V_F + V_R/n) \qquad (6)$$

où V_F est la composante de la variance estimée entre les familles, V_R est la variance résiduelle et 'n' est le nombre moyen de réplicats par famille.

3. Multivariate breeder's equation selon la formule suivante :

$$\Delta z = G\beta \qquad (7)$$

Où Δz est le vecteur de changement des valeurs moyennes de traits entre deux générations successives, β correspond au vecteur de gradients de sélection génotypique pour les traits, et G correspond à la matrice de variance-covariance

génétique. Dans cette étude, les estimations de β ont été réalisées à partir d'une régression multiple suivante en utilisant les moyennes génotypiques de chaque famille :

$$\text{Fitness relative}_i = \mu_{\text{fitness relative}} + \text{trait1}_i + \ldots + \text{trait9}_i + \varepsilon_i \quad (8)$$

Où les termes 'trait1' jusqu'à 'trait9' correspondent aux 9 traits non standardisés (GERM, BT, INT, AREA, DIAM, H1S, HSTEM, RAMBB et RAMPB_S).

Nous avons utilisé deux méthodes d'analyse univariée (Robertson-Price identity et univariate breeder's equation) qui sont relativement différentes dans leurs variables d'entrée. Alors que l'univariate breeder's equation implique la mesure de la relation phénotypique entre un trait et la fitness (*i.e.* sélection) et la relation entre un trait et les gènes (*i.e.* héritabilité), l'identité de Robertson-Price implique l'analyse conjointe de la variation de fitness avec la génétique du trait. Si la covariance phénotypique entre un trait et la fitness est de signe identique aux covariances génétiques et environnementales, les prédictions d'évolution phénotypique seront similaires entre l'identité de Robertson-Price et l'univariate breeder's equation (Morrissey *et al.* 2010). Par contre si la covariance génétique est de signe opposé à la covariance environnementale, l'évolution phénotypique prédite par l'univariate breeder's equations sera de signe opposé à l'évolution phénotypique prédite par l'identité de Robertson-Price. En d'autres termes, un trait héritable avec une covariance phénotypique avec la fitness peut ne pas évoluer si cette covariance phénotypique ne résulte que d'une covariance environnementale. L'identité de Robertson-Price apparaît donc comme un meilleur prédicteur de l'évolution phénotypique que l'univariate breeder's equation.

Cependant, les prédictions établies pour un trait donné peuvent être erronées si d'autres sources de covariance existent entre ce trait et la fitness, impliquant de considérer les prédictions d'évolution phénotypique d'un trait dans un contexte multi-traits en utilisant la multivariate breeder's equation.

Identification des bases génétiques sous-jacentes à la variation naturelle

A partir de 52 familles échantillonnées en 2002 et génotypées pour 214 kSNPs, une approche GWA mapping a été effectuée à partir du modèle mixte suivant (package EMMAX ; Kang *et al.* 2010) pour chaque SNP avec une fréquence de l'allèle mineur supérieure à 10%:

$$Y = \beta X + u + \varepsilon \qquad (9)$$

Où Y est le vecteur de phénotypes, X est le vecteur de génotypes, β est l'effet fixe du SNP testé, $u \sim N_n(0, \sigma_g^2 K)$ et $\varepsilon \sim N_n(0, \sigma_e^2 I)$ sont les effets aléatoires capturant la variance due à la structure des populations et à l'environnement, respectivement. La matrice d'apparentement génétique K tient compte de la relation génétique entre les 52 familles basée sur les 214k SNPs. La structure des populations n'ayant que très peu d'effets sur les associations phénotype – génotype sur les 6 traits phénologiques mesurés en serre sur ces mêmes familles (chapitre I, Brachi *et al.* 2013), l'analyse GWA mapping a aussi été effectuée sans tenir compte de la matrice d'apparentement K. Pour déterminer s'il y avait un excès de faibles valeurs de probabilité d'association phénotype - SNP en tenant compte du grand nombre de tests effectués, la distribution des valeurs de probabilité d'association observées à partir du modèle linéaire (9) a été comparée à une distribution nulle obtenue en permutant 100 fois les valeurs phénotypiques entre les individus et en exécutant le même modèle linéaire. Les

résultats ont été visualisés par Quantile-Quantile plots (QQ) et les intervalles de confiance ont été estimés à partir des permutations.

Résultats

Pour tous les traits phénotypiques mesurés dans cette étude (à l'exception de la survie), une importante variation génétique a été observée en moyenne à chaque génération. Aucune évolution phénotypique n'a été détectée à l'exception de la date de germination, avec des familles germant en moyenne plus tardivement en 2011 qu'en 2002 (Figure 2, Tableau 1).

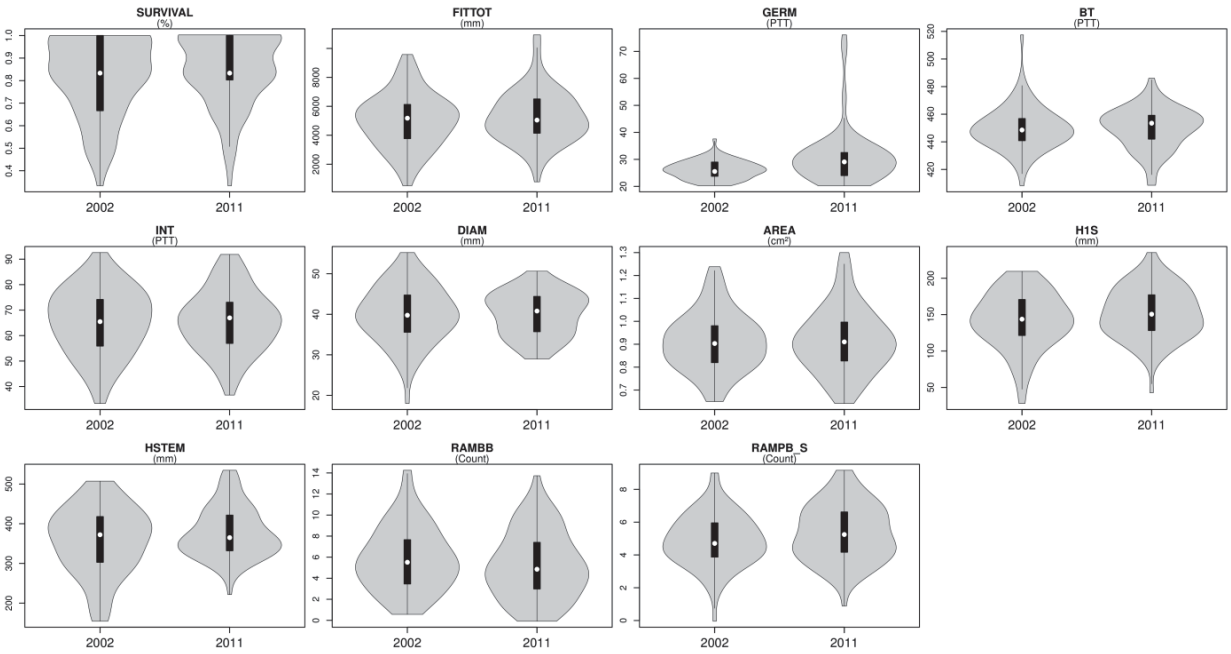


Figure 2: Evolution phénotypique observée entre les deux années d'échantillonnage pour la survie, la production de graines, 3 traits phénologiques (GERM, BT et INT), 2 traits d'accumulation de ressources (AREA et DIAM), et 4 traits architecturaux (H1S, HSTEM, RAMBB et RAMPB_S). Les violin-plots présentés sont basés sur les moyennes génotypiques calculées pour chaque famille au sein de chaque année d'échantillonnage.

Tableau 1: Etude de la variation naturelle pour la survie, la production de graines, les traits phénologiques, les traits d'accumulation de ressources et les traits architecturaux suivant les deux années d'échantillonnage. Les effets aléatoires sont en italique. *0.05 > *P* > 0.01, **0.01 > *P* > 0.001, *** *P* < 0.001. Les termes en gras indiquent les effets restant significatifs après une correction de Bonferroni. N.E. : Non estimé.

	Traits																					
	SURVIVAL		FITTOT		GERM		BT		INT		DIAM		AREA		H1S		HSTEM		RAMBB		RAMPB_S	
Effet	*F ou Chi²*	*P*	*F ou Chi²*	*P*	*F ou Chi²*	*P*	*F ou Chi²*	*P*	*F ou Chi²*	*P*	*F ou Chi²*	*P*	*F ou Chi²*	*P*	*F ou Chi²*	*P*	*F ou Chi²*	*P*	*F ou Chi²*	*P*	*F ou Chi²*	*P*
Bloc	1.74		0.07		**13.78**	***	0.27		0.04		5.5	*	**33.38**	***	0.03		0.04		4.16	*	1.12	
Année	1.18		0.37		**14.71**	***	0.06		4.14	*	0.01		0		2.39		1.89		1.42		2.61	
Bloc * Année	0		1.82		1.9		0.77		0.76		1.61		0.91		2.2		3.72		0		1.31	
Famille(Année)	2.43		**10.2**	**	**98.4**	***	**32.6**	***	**9.5**	**	**8.2**	**	**11**	***	**17.2**	***	**20.6**	***	**32.9**	***	**8**	**
*Bloc * Famille(Année)*	0		0		0.7		0.5		0		0		0.1		0		0		0		0	
GERM	N.E.	N.E.	N.E.	N.E.	N.E.	N.E.	N.E.	N.E.	N.E.	N.E.	N.E.	N.E.	**29.38**	***	N.E.	N.E.	N.E.	N.E.	N.E.	N.E.	N.E.	N.E.

Tableau 2 : Prédiction de l'évolution phénotypique à partir de deux méthodes d'analyse univariée et d'une méthode d'analyse multivariée. '*P*' correspond aux valeurs de probabilités associées à l'identité de Roberston-Price, à l'héritabilité pour l'univariate breeder's equation, et aux gradients de sélection génotypique pour la mutivariate breeder's equation. Les moyennes génotypiques attendues 3 et 9 générations après 2002 ont été estimées à partir de la multivariate breeder's equation. *0.05 > *P* > 0.01, **0.01 > *P* > 0.001, *** *P* < 0.001.

	UNIVARIEE							MULTIVARIEE			Moyenne genotypique			
	Robertson-Price Identity			Breeder's equation				Multivariate breeder's equation			Observé		Attendu	
Trait	*I*		*P*	*S*	*H²*	*P*	*R*	*β*	*P*	*R*	MIB - 2002	MIB - 2011	3 gen	9 gen
GERM	-0.005	PTU/gen	ns	-0.077	0.5	***	-0.039 PTU/gen	-0.012	ns	-0.061 PTU/gen	26.17	30.86	25.99	25.62
BT	-0.768	PTU/gen	ns	-2.219	0.72	***	-1.598 PTU/gen	0.002	ns	-0.761 PTU/gen	449.78	449.87	447.49	442.93
INT	1.112	PTU/gen	*	1.763	0.56	***	0.987 PTU/gen	-0.003	ns	1.237 PTU/gen	65.49	65.70	69.20	76.63
DIAM	1.996	mm/gen	***	4.583	0.55	***	2.521 mm/gen	0.022	***	2.088 mm/gen	40.03	40.07	46.30	58.83
AREA	0.015	cm²/gen	**	0.021	0.47	***	0.010 cm²/gen	0.006	ns	0.016 cm²/gen	0.91	0.92	0.96	1.06
RAMBB	0.107	branches/gen	ns	0.918	0.6	***	0.551 branches/gen	0.048	***	0.074 branches/gen	5.87	5.25	6.09	6.53
RAMPB_S	0.383	branches/gen	***	0.823	0.22	ns	0.181 branches/gen	0.076	**	0.406 branches/gen	4.90	5.33	6.11	8.55
H1S	8.967	mm/gen	***	16.555	0.48	***	7.946 mm/gen	0.000	ns	9.784 mm/gen	142.24	152.29	171.59	230.29
HSTEM	19.594	mm/gen	***	36.346	0.63	***	22.898 mm/gen	0.002	**	21.655 mm/gen	357.24	377.71	422.21	552.14

L'héritabilité pour la production totale de graines est de 0.48 pour l'année 2002 ($P < 0.001$). Les deux méthodes d'analyse univariée sont cohérentes quant au signe et à l'intensité de la réponse à la sélection, suggérant une covariance génétique et une covariance environnementale de même signe (Tableau 2). Bien que le signe de certaines covariances génétiques entre deux traits soit opposé au sens de la sélection directionnelle (Tableaux 2 et 3) comme pour le nombre de branches basales avec la hauteur de la tige principale (Figure 3), les réponses prédites à partir de la multivariate breeder's equation sont de signes équivalents et d'intensités similaires à ceux obtenus à partir des analyses univariées, suggérant que la covariance génétique entre un trait donné et la fitness est relativement indépendante de la sélection agissant sur un trait génétiquement corrélé mesuré dans cette étude.

	FITTOT	GERM	BT	INT	DIAM	Area	RAMBB	RAMPB_S	H1S	HSTEM
FITTOT	0.138	.	.	.	.	.	.	.	.	.
GERM	-0.053	11.173	.	.	.	.	.	.	.	.
BT	-0.768	-22.971	251.277	.	.	.	.	.	.	.
INT	1.112	-1.269	18.233	168.023	.	.	.	.	.	.
DIAM	1.996	1.294	-21.123	33.318	50.344	.	.	.	.	.
Area	0.015	-0.113	0.215	0.284	0.287	0.016	.	.	.	.
RAMBB	0.107	0.330	3.488	-8.475	-3.454	0.042	9.222	.	.	.
RAMPB_S	0.383	0.793	-9.567	7.323	7.868	0.038	-1.554	2.488	.	.
H1S	8.967	8.355	-171.229	210.659	177.150	0.743	-42.993	46.162	1487.337	.
HSTEM	19.594	5.761	-230.523	402.658	352.300	1.875	-101.567	84.499	2755.971	7042.337

Tableau 3 : Matrice de variance-covariance génétique entre les 10 traits quantitatifs.

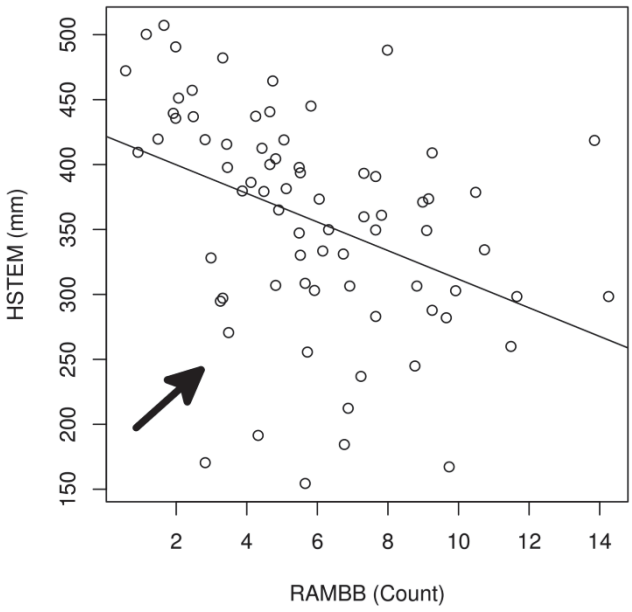


Figure 3 : Covariance génétique négative entre le nombre de branches basales et la hauteur de la tige principale. La flèche indique le sens de la direction directionnelle attendue.

Les prédictions établies à partir de la multivariate breeder's equation sur 9 générations (si nous supposons que chaque année correspond à une génération entre 2002 et 2011) contrastent fortement avec l'absence d'évolution phénotypique pour la majorité des traits (Tableau 2). Etant donné une rotation des cultures de type blé-orge-colza/moutarde classiquement observée dans la région où MIB est localisée (Valérie Le Corre, communication personnelle), nous pouvons supposer que les plantes d'*A. thaliana* sont mortes suite à des traitements herbicides anti-dicotylédones les années de culture de blé ou d'orge, entraînant une absence de variation génétique pour la fitness. Ainsi, nous aurions 3 générations (correspondant aux 3 années où du colza ou de la moutarde a été cultivé) où la sélection a pu opérer. Malgré ce nombre de générations plus petit, les moyennes phénotypiques observées en 2011 restent toujours éloignées, voire avec une évolution de signe opposé, aux prédictions établies sur 3 générations.

Aucune région génomique n'a été significativement associée à la variation phénotypique observée dans cette population en 2002 (voir Figure 4 pour illustration avec la date de montaison).

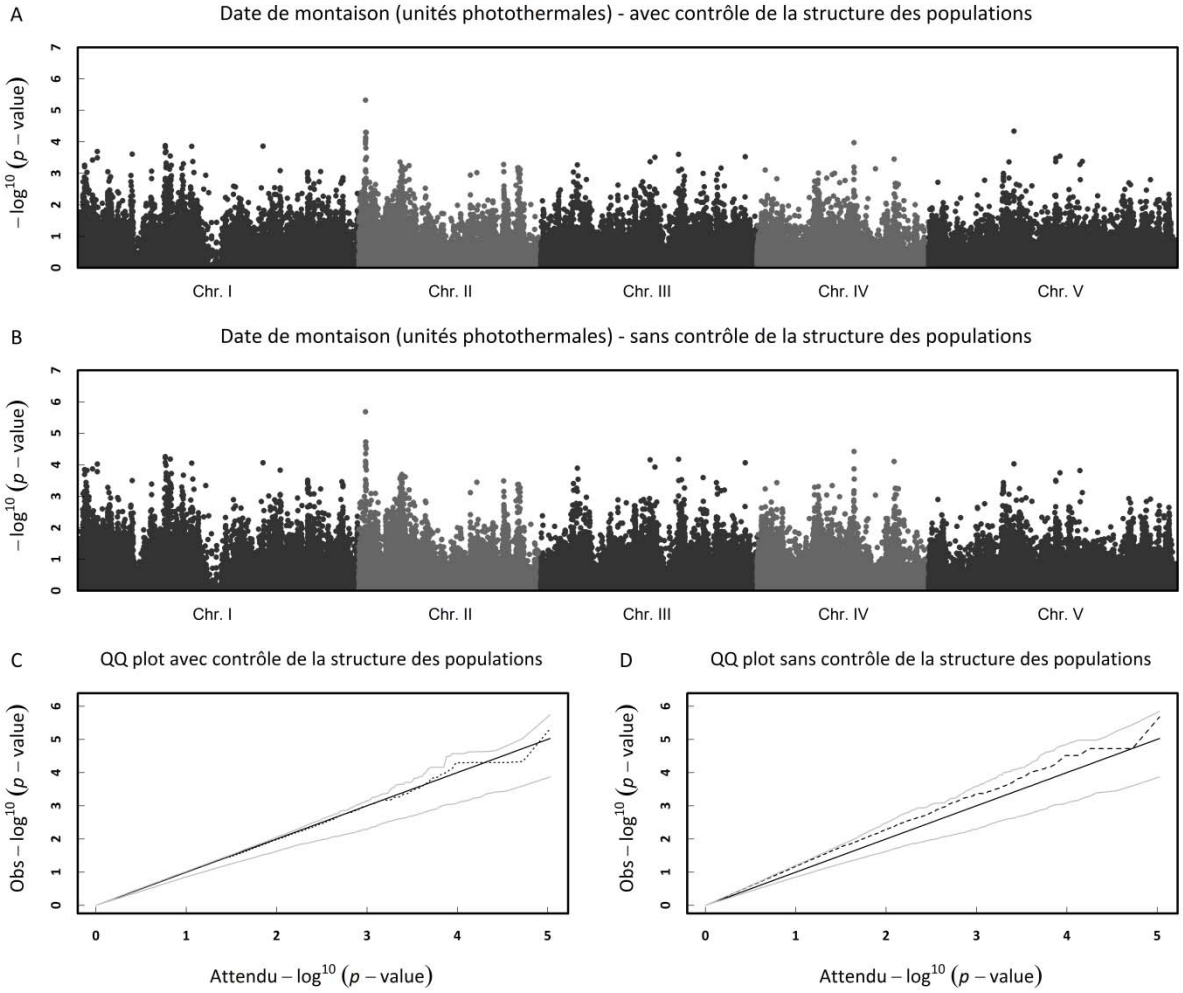


Figure 4 : Etude de GWA mapping au sein de la population MIB échantillonnée en 2002 pour la date de montaison. Manhattan plots représentant les associations entre la date de montaison et les SNPs ayant une fréquence relative de l'allèle mineur supérieure à 10% (n = 106 261 SNPs) en présence (A) ou absence (B) de contrôle pour la structure des populations (axe des abscisses : position physique des 106 261 SNPs le long des 5 chromosomes). Quantile-Quantile plots en présence (C) ou absence (D) de contrôle pour la structure des populations. La ligne pleine noire représente la distribution neutre uniforme des probabilités d'association SNP – phénotype. La ligne en tirets représente la distribution observée des probabilités d'association SNP – phénotype. Les lignes pleines grises représentent les intervalles de confiance obtenus après 100 permutations du phénotype entre les individus.

Discussion

Malgré une variation génétique significative d'un proxy de fitness mesuré dans cette étude et des covariances génétiques trait – fitness elles-aussi significatives, l'absence d'évolution phénotypique entre 2002 et 2011 au sein de la population MIB (à l'exception de la date de germination) contraste avec les prédictions de changement évolutif basées sur trois méthodes. Comment expliquer cette absence d'évolution phénotypique ?

1. Proxy de fitness

Nous avons utilisé la production totale de graines comme une approximation de la valeur sélective. Ce proxy de fitness peut paraître adéquat pour une espèce avec un taux d'autogamie proche 98% (Platt *et al.* 2010). Mais ce taux n'est qu'une moyenne basée sur l'analyse de plusieurs dizaines de populations naturelles. Le taux d'allogamie de la population MIB a été évalué à 8% (Platt *et al.* 2010 ; Valérie Le Corre, communication personnelle), suggérant que la fitness mâle devrait aussi être prise en compte dans notre étude. Par ailleurs, comme évoqué dans le manuscrit 1 du chapitre 2 'Réalisme écologique, contexte multi-traits et intégration de la plasticité dans l'étude de l'adaptation : une transition vers la génomique écologique évolutive' de cette thèse, les stratégies reproductives peuvent être aussi importantes que la quantité totale de graines produites par une plante. Il serait donc intéressant de tester si la variation phénotypique observée au sein de l'année 2002 est associée à différentes stratégies reproductives.

2. Contraintes génétiques et traits sous sélection non mesurés

Le signe et l'intensité de la réponse à la sélection prédits par la multivariate breeder's equation sont cohérents avec les prédictions obtenues avec deux méthodes d'analyse mutivariée, suggérant que les contraintes génétiques sont peu présentes dans notre jeu

de données. Pour autant, des contraintes génétiques avec des traits phénotypiques non mesurés peuvent exister et ainsi ralentir l'évolution phénotypique des traits mesurés dans cette étude. A l'inverse des expériences de sélection artificielle où le phénotype sous sélection est connu, la multitude de traits pouvant être sous sélection dans les populations naturelles rend difficilement envisageable de tous les identifier, les mesurer et les intégrer dans une analyse de prédiction d'évolution phénotypique.

3. Hétérogénéité temporelle du milieu et évolution de la plasticité

Les 3 méthodes d'analyse utilisées dans cette étude reposent toutes sur des hypothèses de base, à savoir une démographie constante de la population, des générations discrètes et des conditions environnementales constantes (Morrissey *et al.* 2013). La présence potentielle d'une banque de graines dans la population MIB entraînerait un chevauchement des générations, ralentissant ainsi l'intensité de la sélection. De même, la population MIB étant située dans un environnement anthropique régulièrement perturbé, l'optimum phénotypique pourrait être très variable d'une année sur l'autre, permettant ainsi de maintenir une diversité phénotypique et génétique. D'un autre côté, un milieu très hétérogène dans le temps devrait entraîner une sélection de génotypes plastiques (Moran 1992). Il serait donc intéressant de mesurer les populations MIB échantillonnées en 2002 et 2011 dans différents environnements rencontrés par cette population au cours de son évolution, afin de tester si la plasticité a été sélectionnée au cours du temps.

4. Environnement de mesure phénotypique

Nous avons montré dans le deuxième manuscrit du chapitre 2 'Réalisme écologique, contexte multi-traits et intégration de la plasticité dans l'étude de l'adaptation : une transition vers la génomique écologique évolutive' que l'intensité de la sélection pouvait varier entre deux cohortes au sein d'une même année ou bien entre deux

années consécutives au sein d'une même cohorte. Par ailleurs, une variabilité plus importante *in-situ* (Figure 1b à 1g) que sur un terrain expérimental de l'Université de Lille 1 (Figure 1h à 1m) a été observée pour la forme de la rosette et la forme des feuilles. Ainsi, les résultats obtenus dans cette étude ne sont peut-être spécifiques qu'à l'environnement du terrain expérimental de l'Université de Lille 1 et pourraient donc ne pas être révélateur des covariances génétiques observées *in-situ*. Si le terrain expérimental de l'Université de Lille 1 est trop éloigné écologiquement du milieu naturel de la population MIB, la variance génétique naturelle observée dans cette étude ne résulterait-elle pas de l'expression de mutations phénotypiquement neutres *in-situ* ? La détection des bases génétiques associées à la variation naturelle par une approche de GWA mapping étant d'autant plus aisée si cette variation naturelle est adaptative (Bergelson & Roux 2010), l'absence d'identification de régions génomiques associées à la variation naturelle de la population MIB mesurée sur notre terrain expérimental serait cohérente avec cette hypothèse. Nous ne pouvons pas cependant écarter un manque de puissance dans notre étude de GWA mapping due au nombre réduit de familles (n = 52) utilisées.

Une approche indirecte pour identifier les traits phénotypiques sous sélection serait d'identifier les gènes qui présentent des traces de sélection génomique. Pour cela, nous avons extrait l'ADN des 154 familles MIB, puis regrouper des ADN par année de récolte afin d'effectuer un séquençage génomique massif en bulk (approche Pool-Seq). Après un traitement bio-informatique des données de séquençage par Sophie Gallina (responsable de la plate-forme de bio-informatique au laboratoire GEPV), les fréquences alléliques le long du génome ont été estimées pour tous les polymorphismes détectés. Un scan génomique des traces de sélection au sein de chaque année (D de Tajima par exemple) et entre les années (F_{ST} temporel par exemple) est prévu prochainement.

Bibliographie

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

Brachi, B., C. Aimée, et al. (2012) "Adaptive Value of Phenological Traits in Stressful Environments: Predictions Based on Seed Production and Laboratory Natural Selection." PLoS ONE **7**: e32069.

Brachi, B., N. Faure, et al. (2010). "Linkage and association mapping of *Arabidopsis thaliana* Flowering time in nature." PLoS Genetics **6**: e1000940.

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

Kang, H. M., N. A. Zaitlen, et al. (2008) "Efficient controlof population structure in model organism association mapping. " Genetics **178**: 1709–1723.

Horton, M., A. Hancock, et al. (2012). "Genome-wide patterns of genetic variation in worldwide Arabidopsis thaliana accessions from the RegMap panel." Nature Genetics **44**: 212-216.

Moran, N. (1992). "The evolutionary maintenance of alternative phenotypes." American Naturalist **139**: 971-989.

Morrissey, M. B., L. E. B. Kruuk, et al. "The danger of applying the breeder's equation in observational studies of natural populations." Journal of Evolutionary Biology **23**: 2277-2288.

Reboud, X., V. Le Corre, et al. (2004) "Natural variation among accessions of Arabidopsis thaliana: beyond the flowering date, what morphological traits are relevant to study adaptation?" Plant Adaptation: Molecular Genetics and Ecology :135-142.

Platt, A., M. Horton, et al. (2010). "The scale of population structure in *Arabidopsis thaliana*." PLoS Genetics **6**: E1000843.

Roux, F., J. Gasquez, et al. (2004). "The dominance of the herbicide resistance cost in several Arabidopsis thaliana mutant lines." Genetics **166**: 449-460.

Siepielski, A. M., J. D. DiBattista, et al. (2009). "It's about time: the temporal dynamics of phenotypic selection in the wild." Ecology Letters **12**: 1261–1276.

Weinig, C., J. Johnston, et al. (2006). "Local and global costs of adaptive plasticity to density in Arabidopsis thaliana." American Naturalist **167**: 826-836.

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

CHAPITRE 3 – CONCLUSIONS ET PERSPECTIVES

2. Conclusions et perspectives

Des patrons contrastés d'évolution ont été observés entre les deux populations MIB et TOU. En effet, la population TOU, située dans un milieu spatialement hétérogène pour le sol et peu perturbé dans le temps, présente une 'rapide' évolution phénotypique pour de nombreux traits. En revanche, la population MIB, située dans un milieu spatialement homogène pour le sol et potentiellement très perturbée dans le temps, ne présente pas d'évolution phénotypique pour les traits que nous avons mesurés (à l'exception de la date de germination), et ceci contrairement à nos prédictions. Pour la population MIB, plusieurs hypothèses ont été formulées afin d'expliquer la différence entre évolution prédite et évolution observée : l'existence de contraintes génétiques, la présence d'une hétérogénéité temporelle très fine favorisant la sélection de génotypes plastiques, la possibilité que la production totale de graines ne soit pas un bon estimateur de la valeur sélective, ou encore la possibilité que le phénotype exprimé dans nos conditions environnementales ne reflète pas le phénotype exprimé au sein de la population.

Bien que ces résultats portent sur des populations très polymorphes pouvant ne pas refléter le fonctionnement de populations naturelles moins polymorphes, il apparait néanmoins qu'une évolution phénotypique rapide d'origine génétique puisse avoir lieu au sein d'une population naturelle d'*A. thaliana.* Les résultats obtenus avec les 49 populations françaises récoltées la même année dans les chapitres précédents pourraient donc être erronés si certaines populations récoltées n'étaient pas à l'équilibre. Afin de vérifier la pérennité des populations naturelles d'*A thaliana* sur une courte échelle de temps, une visite des 11 populations de Bourgogne utilisées dans cette thèse a été effectuée en novembre 2012, soit 3 années après leur récolte initiale. Sur ces onze populations, nous avons observé la présence de plantes dans seulement sept d'entre elles. Il serait donc intéressant de ré-échantillonner ces différentes populations en récoltant les individus germés, mais également en effectuant des

prélèvements de sol ce qui nous permettrait d'estimer la diversité génétique présente dans la banque de graines. Bien que la diversité génétique présente dans la banque de graines puisse ne pas refléter la diversité génétique présente dans la population de plantes germées, ce procédé d'échantillonnage nous permettrait de déterminer si ces populations peuvent être pérennes et si elles peuvent évoluer malgré l'absence constatée de plantes certaines années. Une caractérisation phénotypique des plantes germées nouvellement récoltées nous permettrait également de vérifier les résultats obtenus avec les plantes récoltées en 2009 et/ou d'observer une évolution temporelle de ces populations.

Les différences constatées pour la population MIB entre évolution prédite et évolution observée me laisse penser qu'il est nécessaire, afin d'étudier l'évolution phénotypique des populations naturelles, et plus particulièrement les processus d'adaptation locale, de replacer les génotypes dans leur milieu d'origine. J'ai initié, en collaboration avec Etienne Baron (doctorant au sein de l'équipe « Changements globaux : de la génomique écologique à l'écologie des communautés »), ce travail pour la population TOU. Une expérimentation *in situ* associant étude de résurrection (incluant toutes les familles échantillonnées en 2002 et 2010) et transplantation réciproque sur 3 types de sols au sein de la population TOU a été mise en place en septembre 2012 et s'est terminée en juillet 2013. Afin de déterminer si l'évolution phénotypique observée dans la population TOU dépend aussi d'une adaptation locale à une variation édaphique ou à un changement de communauté végétale associé à cette variation édaphique, les transplantations réciproques ont été effectuées avec et sans ajout d'un compétiteur. Le pâturin annuel, *Poa annua*, a été choisi comme compétiteur dans cette expérience car il semblait dominer fortement la communauté végétale de cette population en 2010. Un séquençage génomique de toutes les familles échantillonnées en 2002 a d'ores et déjà été effectué, et un séquençage génomique de toutes les familles échantillonnées en 2010 est prévu en 2014. Comme initialement proposé pour une étude de génomique écologique

évolutive basée sur les 49 populations françaises, nous espérons que cette caractérisation génomique associée à la caractérisation phénotypique et écologique de cette population nous permettra de répondre aux questions suivantes :

- Est-il possible d'identifier des régions génomiques qui sont à la fois associées à une variation naturelle phénotypique, à des agents sélectifs potentiels tels que le sol et présentant des traces de sélection?
- Si oui, les régions génomiques identifiées sont-elles identiques entre les 3 types de sols ? et entre les deux traitements de compétition ?
- Les régions génomiques associées à la plasticité exprimée entre les 3 types de sol ou les deux traitements de compétition pour un trait donné sont-elles différentes des régions génomiques associées à ce même trait mesuré dans un type de sol ou un traitement de compétition donné?
- Si oui, l'intensité et les traces de sélection associées aux régions génomiques identifiées sont-elles différentes entre les traits et leurs plasticités?

CONCLUSIONS GENERALES ET PERSPECTIVES

V. CONCLUSIONS GENERALES ET PERSPECTIVES :

En combinant différentes disciplines que sont l'écologie, l'évolution et la génétique quantitative, ma thèse a porté sur l'étude de l'échelle spatiale et temporelle de l'adaptation chez *A. thaliana*, avec un intérêt particulier sur la prise en compte de la plasticité dans cette étude. Une synthèse du matériel biologique utilisée, des expériences réalisées, des approches abordées et des perspectives envisagées est illustrée dans la figure 8.

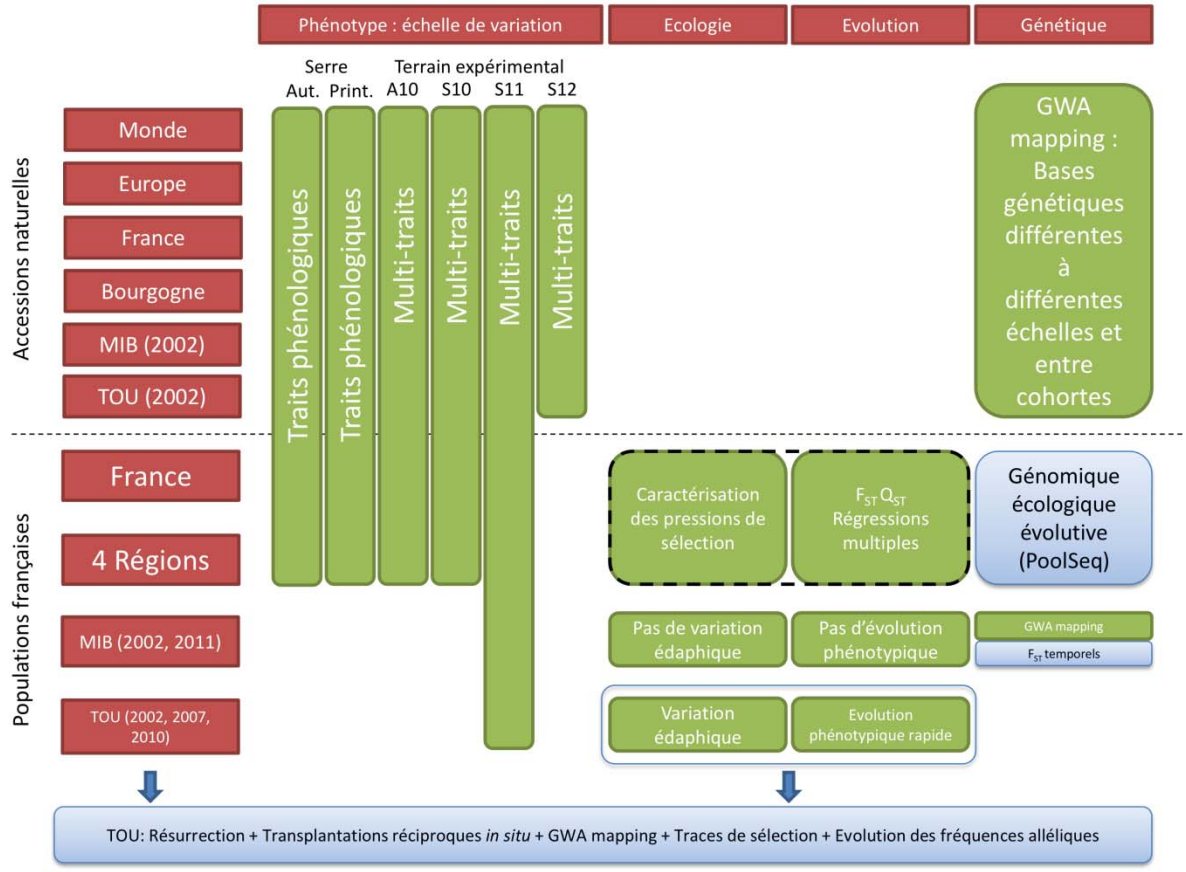


Figure 8 : Synthèse des différentes études menées durant cette thèse et principales perspectives. En rouge : matériel biologique utilisé et disciplines abordées pendant la thèse. En vert : expériences et analyses réalisées pendant cette thèse (A10 : Août 2010, S10 : Septembre 2010, S11 : Septembre 2011, S12 : Septembre 2012 non présenté dans ce mémoire de thèse). En bleu : perspectives.

Bien que les études menées dans cette thèse ne prennent pas en compte un certain nombre de facteurs biotiques importants (tels que les pathogènes, les herbivores ou encore la compétition interspécifique) aussi bien au niveau de la description de l'habitat des populations naturelles qu'au niveau des designs expérimentaux mis en œuvre pour l'étude de la plasticité, plusieurs conclusions peuvent tout de même être tirées de mes travaux. En effet, des expérimentations et des analyses menées durant cette thèse sur des jeux d'accessions ou de familles génétiques issus d'échantillonnages réalisés à différentes échelles géographiques (Figure 8), il ressort qu'une importante variabilité existe pour de nombreux traits phénotypiques à une échelle spatiale large comme le monde ou la France mais également à une échelle spatiale très fine notamment au sein de certaines populations. Bien qu'une part importante de cette variation puisse être attribuée à des processus non sélectifs, une part non négligeable de celle-ci serait due à des processus d'adaptation locale. Le patron d'adaptation locale révélé par différentes méthodes (comparaison Q_{ST} –F_{ST}, relation phénotype-écologie, gradients de sélection, étude de résurrection) est très complexe et semble être la résultante de pressions de sélection emboitées variant à différentes échelles, aussi bien au niveau de la France qu'au sein d'une population, et agissant sur différents traits ou plasticités. L'utilisation d'un jeu d'accessions ou de familles génétiques issus d'échantillonnages à différentes échelles, leur caractérisation phénotypique pour de nombreux traits et plasticités ainsi que leur caractérisation écologique pour de nombreux facteurs variant à différentes échelles semblent donc conjointement nécessaires à l'identification de phénomènes d'adaptation locale. Bien que ce constat puisse paraître une évidence, il est en opposition avec les études récentes réalisées afin d'identifier les bases génétiques de l'adaptation chez *A. thaliana*. En effet, ces études tendent à maximiser la diversité génétique présente dans l'échantillon analysé et n'utilisent qu'une accession par population (Todesco *et al.* 2010 ; Baxter *et al.* 2010 ; Hancock *et al.* 2011 ; Fournier-Level *et al.* 2011). Ces études négligent donc la variabilité des

processus sélectifs pouvant agir à différentes échelles géographiques et plus particulièrement la variabilité pouvant exister au sein des populations naturelles d'*A. thaliana*, ce qui pourrait freiner voir empêcher la détection des bases génétiques de l'adaptation.

L'environnement de mesure phénotypique influe grandement sur les conclusions obtenues dans les différentes analyses. Afin de proposer des scénarios évolutifs fiables il apparait donc nécessaire d'effectuer ces mesures dans des environnements réalistes, l'idéal étant certainement le milieu d'origine. Ce travail a été initié pour la population TOU afin de vérifier si une adaptation micro-locale au sol existe bien dans cette population. Dans le cas des 49 populations, un design plus rigoureux que celui utilisé dans cette thèse aurait été d'effectuer des transplantations réciproques dans chacune des 49 populations récoltées. On comprendra aisément que les moyens humains et matériels nécessaire à la mise en place d'une telle expérimentation ne sont pas à notre disposition. De plus, les populations locales n'excédant parfois pas plus d'un ou deux mètres carrés, la mise en place d'une expérience *in situ* dans ces populations parait impossible. A l'heure de la génomique, notre capacité à découvrir les bases génétique de l'adaptation pourrait donc être fortement limitée par notre capacité à effectuer les mesures phénotypiques dans les environnements adéquats. Pour autant, bien que potentiellement imparfaite, cette caractérisation phénotypique reste nécessaire dans l'étude de la génomique écologique évolutive. En effet, notre capacité à identifier les bases génétiques de l'adaptation dépendra fortement de notre capacité à faire le lien entre variation phénotypique, variation écologique et variation génétique afin de pouvoir distinguer les différenciations génétiques réellement dues à des phénomènes d'adaptation locale des différenciations génétiques dues à des processus non-sélectifs.

REFERENCES BIBLIOGRAPHIQUES

VI. REFERENCES BIBLIOGRAPHIQUES :

Abbott, R. J. and M. F. Gomes (1989). "Population genetic structure and outcrossing rate of Arabidopsis thaliana." Heredity **62**: 411-418.

Agrawal, A. (2001). "Phenotypic plasticity in the interactions and evolution of species." Science **294**: 321-326.

Aranzana, M. J., S. Kim, et al. (2070) " Genome-Wide Association Mapping in Arabidopsis Identifies Previously Known Flowering Time and Pathogen Resistance Genes ". PLoS Genetics **1**: e60

Atwell, S., Y. Huang, et al. (2010). "Genome-wide association study of 107 phenotypes in Arabidopsis thaliana inbred lines." Nature **465**: 627-631.

Baxter, I., J. N. Brazelton, et al. (2010)" A Coastal Cline in Sodium Accumulation in Arabidopsis thaliana Is Driven by Natural Variation of the Sodium Transporter AtHKT1;1" PLoS Genetics**6**: e1001193

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

Bomblies, K., L. Yant, et al. (2010) Local-scale patterns of genetic variability, outcrossing, and spatial structure in Natural stands of *Arabidopsis thaliana*. PLoS Genetics **6:** e1000890

Brachi, B., N. Faure, et al. (2010). "Linkage and association mapping of *Arabidopsis thaliana* Flowering time in nature." PLoS Genetics **6**: e1000940.

Bradshaw, A. D. (1965). "Evolutionary significance of phenotypic plasticity in plants." Advances in Genetics **13**: 115-165.

Chao, D., A. Silva (2012). " Genome-Wide Association Studies Identify Heavy Metal ATPase3 as the Primary Determinant of Natural Variation in Leaf Cadmium in Arabidopsis thaliana" Plos Genetics **8**: e1002923.

Cipollini, D., C. Purrington, et al. (2003). "Costs of induced responses in plants." Basic and Applied Ecology **4**: 77-89.

Cirulli, E. T. and D. G. Goldstein (2010). "Uncovering the roles of rare variants in common disease through whole-genome sequencing." Nature Genetics **11**: 416-425.

Cookson, W., L. Liang, et al. (2009). "Mapping complex disease traits with global gene expression." Nature Reviews Genetics **10**: 184-194.

Debat, V. and P. David (2001). "Mapping phenotypes: canalization, plasticity and developmental stability." Trends in Ecology & Evolution.**16**:555-561.

Dewitt, T.J., A. Sih, et al. (1998) "Cost and limits of phenotypic plasticity". Trends in Ecology and Evolution **13**: 77-81.

Donohue, K., L. Dorn, et al. (2005). "Niche construction through germination cueing: Life-history responses to timing of germination in *Arabidopsis thaliana*." Evolution **59**: 771-785.

Donohue, K., D. Messiqua, et al. (2000). "Evidence of adaptive divergence in plasticity: density- and site-dependent selection on shade-avoidance responses in Impatiens capensis." Evolution **54**: 1956-1968.

Feder, M. E. and T. Mitchell-Olds (2003). "Evolutionary and ecological functional genomics." Nature Reviews Genetics **4**: 649-655.

Fournier-Level, A., A. Korte, et al. (2011). "A map of local adaptation in Arabidopsis thaliana." Science **334**: 86-89.

Fox, J. and J. Hong (2009). "Effect Displays in R for Multinomial and Proportional-Odds Logit Models: Extensions to the effects Package." Journal of Statistical Software **32**: 1-25.

Frenkel, M., H. Johansson Jänkänpää, et al. (2008). "An illustrated gardener's guide to transgenic Arabidopsis field experiments." New Phytologist **180**: 545-555.

Gaut, B. (2012). "Arabidopsis thaliana as a model for the genetics of local adaptation." Nature Genetics **44**: 115–116.

Goddard, M. and B. Hayes (2009). "Mapping genes for complex traits in domestic animals and their use in breeding programmes." Nature Reviews. Genetics **10**: 381-391.

Haldane, J. B. (1948). "The theory of cline." Journal of Genetics **48**: 277-284.

Hancock, A., B. Brachi, et al. (2011). "Adaptation to climate across the Arabidopsis thaliana genome." Science **334**: 83-86.

Horton, M., A. Hancock, et al. (2012). "Genome-wide patterns of genetic variation in worldwide Arabidopsis thaliana accessions from the RegMap panel." Nature Genetics **44**: 212-216.

Huard-Chauveau, C., L. Perchepied et al. (2013) "An Atypical Kinase under Balancing Selection Confers Broad-Spectrum Disease Resistance in Arabidopsis" PLos Genetics **9**: e1003766.

Juenger, T. E. and J. Bergelson (2000). "Factors limiting rosette recruitment in scarlet gilia, Ipomopsis aggregata: seed and microsite limitation." Oecologia **123**: 358-363.

Kim, S., V. Plagnol, et al. (2007). "Recombination and linkage disequilibrium in Arabidopsis thaliana." Nature genetics **39**: 1151-1155.

Kimura, M. (1955). "Stochastic processes and distribution of gene under natural selection." Cold Spring Harbor Symposia on Quantitative Biology **20**: 33-53.

Lande, R. & Arnold, S. J. (2009). "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 : comparaison with markers and traits." Molecular Ecology **14**: 4181-4192.

Li, Y., Y. Huang et al. (2010) " Association mapping of local climate-sensitive quantitative trait loci in Arabidopsis thaliana" PNAS **107**: 21199-21204

Mitchell-Olds, T. and J. Schmitt (2006). "Genetic mechanisms and evolutionary significance of natural variation in Arabidopsis." Nature **441**: 947-952.

Moran, N. (1992). "The evolutionary maintenance of alternative phenotypes." American Naturalist **139**: 971-989.

Nemri, A., S. Atwell, et al. (2010)" Genome-wide survey of Arabidopsis natural variation in downy mildew resistance using combined association and linkage mapping". PNAS **107**: 10302-10307.

Nordborg M, J.O. Borevitz (2002) "The extent of linkage disequilibrium in Arabidopsis thaliana". Nature Genetics **30**: 190–193.

Nordborg, M. , T. T. Hu, et al. (2005) "The Pattern of Polymorphism in Arabidopsis thaliana" PLoS Biology **3**: e196.

Orsini, L., R. Andrew, et al. (2013). "Evolutionary Ecological Genomics." Molecular Ecology **22**: 527-531.

Pagny , G., P. S. Paulstephenraj, et al.(2012). " Family-based linkage and association mapping reveals novel genes affecting Plum pox virus infection in Arabidopsis thaliana". New Phytologist **196**: 873-886.

Platt, A., M. Horton, et al. (2010). "The scale of population structure in *Arabidopsis thaliana*." PLoS Genetics **6**: E1000843.

Poveda, K., I. Steffan-Dewenter, et al. (2003). "Effects of below- and above-groung herbivores on plant growth, flower visitation and seed set." Oecologia **135**: 601-605.

Rafalski, J. (2010). "Association genetics in crop improvement." Current opinion in plant biology **13**: 174-180.

Rathcke, B. and E. P. Lacey (1985). "Phenological patterns of terrestrial plants." Annual Review of Ecology, Evolution and Systematics **16**: 179-214.

Rosas, U., A., C.-Jaramillo, et al. (2013) "Integration of responses within and across Arabidopsis natural accessions uncovers loci controlling root systems architecture" PNAS **110**: 15133-15138.

Roux, F., J. Gasquez, et al. (2004). "The dominance of the herbicide resistance cost in several Arabidopsis thaliana mutant lines." Genetics **166**: 449-460.

Sardansa, J., F. Rodàb, et al. (2005). "Effects of water and a nutrient pulse supply on Rosmarinus officinalis growth, nutrient content and flowering in the field." Environmental and Experimental Botany **53**: 1–11.

Schlichting, C. D. (1986). "The evolution of phenotypic plasticity in plants." Annual Review of Ecology and Systematics **17**: 667-693.

Shemesh, H. and A. Novoplansky (2013). "Branching the risks: architectural plasticity and bet-hedging in Mediterranean annuals." Plant biology **15**: 1001-1012

Shindo, C., G. Bernasconi, et al. (2007). "Natural genetic variation in Arabidopsis: tools, traits and prospects for evolutionary ecology." Annals of Botany **99**: 1043-1054.

Siepielski, A. M., J. D. DiBattista, et al. (2009). "It's about time: the temporal dynamics of phenotypic selection in the wild." Ecology Letters **12**: 1261–1276.

Steinger, T., B. Roy, et al. (2003). "Evolution in stressful environments II: adaptive value and costs of plasticity in response to low light in Sinapis arvensis." Journal of Evolutionary Biology **16**: 313-323.

Stinchcombe, J. R., C. Weinig, et al. (2004). "A latitudinal cline in flowering tine in *Arabidopsis thaliana* modulated by the flowering time gene FRIGIDA." PNAS **101**: 4712-4717.

Sultan, S. and H. Spencer (2002). "Metapopulation structure favors plasticity over local adaptation." The American Naturalist **160**: 271-283

Sultan, S. E. (2000). "Phenotypic plasticity for plant development, function a life history." Trends in Plant Science **5**: 537-542.

Swarbreck, D., C. Wilks, et al. (2008). "The Arabidopsis Information Resource (TAIR): gene structure and function annotation." Nucleic acids Research **36**: 1009-1014.

Todesco, M., S. Balasubramanian, et al.(2010) "Natural allelic variation underlying a major fitness trade-off in Arabidopsis thaliana". Nature **465**: 632-636.

Tonsor, S., T. Elnaccash, et al. (2013). "Developmental instability is genetically correlated with phenotypic plasticity, constraining heritability and fitness." Evolution **67**: 2923-2935.

Ungerer, M. C., L. C. Johnson, et al. (2008). "Ecological genomics: understanding gene and genome function in the natural environment." Heredity **100**: 178-183.

van Kleunen, M. and M. Fischer (2001). "Adaptive evolution of plastic foraging responses in a clonal plant." Ecology **82**: 3309-3319.

van Kleunen, M. and M. Fischer (2005). "Constraints on the evolution of adaptive phenotypic plasticity in plants." The New Phytologist **166**: 49-60.

van Kleunen, M., M. Fischer, et al. (2000). "Costs of plasticity in foraging characteristics of the clonal plant Ranunculus reptans." Evolution **54**: 1947-1955.

Weigel, D. and M. Nordborg (2005). "Natural variation in Arabidopsis. How do we find the causal genes?" Plant physiology **135**: 567-568.

Weinig, C., M. C., Ungerer, et al. (2002). "Novel loci control variation in reproductive timing in Arabidopsis thaliana in natural environments." Genetics **162**: 1875-1884.

Weinig, C. and L. Delph (2001). "Phenotypic plasticity early in life constrains developmental responses later." Evolution **55**: 930-936.

Weinig, C., J. Johnston, et al. (2006). "Local and global costs of adaptive plasticity to density in Arabidopsis thaliana." American Naturalist **167**: 826-836.

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

Wilczek, A., J. Roe, et al. (2009). "Effects of genetic perturbation on seasonal life history plasticity." Science **323**: 930-934.

Zhao, K., M. J., Aranzana (2007) "An Arabidopsis Example of Association Mapping in Structured Samples" Plos Genetics **3**:1.

Echelle spatiale et temporelle de l'adaptation chez *Arabidopsis thaliana* : intégration de la plasticité phénotypique

Un défi majeur de la biologie moderne reste l'identification des bases génétiques sous-jacentes à la variation phénotypique naturelle observée à l'intérieur d'une espèce et plus particulièrement l'identification des bases génétiques de l'adaptation. Cette thèse fait partie intégrante d'un projet de génomique écologique évolutive chez *A. thaliana*. A partir d'un échantillonnage hiérarchique et/ou temporel de populations naturelles, je me suis attaché dans cette thèse à (i) faire une caractérisation écologique (climat, sol et compétition) faisant défaut chez cette espèce modèle en génomique. (ii) caractériser la variation phénotypique existant à différentes échelles spatiales, en mesurant de nombreux traits phénotypiques et leurs plasticités en conditions contrôlées de serre mais aussi sur un terrain expérimental de l'Université de Lille 1, (iii) identifier les traits sous sélection en utilisant trois approches : comparaison $F_{ST} - Q_{ST}$, relations phénotype – écologie et gradients de sélection génotypiques, (iv) identifier les agents sélectifs potentiellement responsables de ces variations phénotypiques adaptatives et (v) et identifier les bases génétiques associées à la variation naturelle par une approche de GWA mapping. Des expérimentations menées durant cette thèse, il ressort qu'une importante variabilité existe pour de nombreux traits phénotypiques à une échelle spatiale large comme le monde ou la France mais également à une échelle spatiale très fine, notamment au sein de certaines populations. Bien qu'une part importante de cette variation puisse être attribuée à des processus non sélectifs, une part non négligeable de celle-ci serait due à des processus d'adaptation locale. Le patron d'adaptation locale révélé par les différentes méthodes utilisées est très complexe et semble être la résultante de pressions de sélection emboitées variant à différentes échelles, aussi bien au niveau de la France qu'au sein d'une population, et agissant sur différents traits ou plasticités. En accord avec ce patron, les bases génétiques associées à la variation naturelle phénotypique semblent être très dépendantes de l'échelle géographique considérée.

Spatial and temporal scale of adaptation in *Arabidopsis thaliana*: Integrating phenotypic plasticity in the study of evolution

A major challenge in modern biology remains the identification of the genetic basis of natural phenotypic variation, especially the identification of the genetic basis of adaptation. This thesis is embedded in a project on ecological evolutionary genomics in *A. thaliana*. Based on a hierarchically spatial and/or temporal sampling of natural populations, I aimed at (i) making an ecological characterization (climate, soil and competition) lacking for this model plant in genomics, (ii) characterizing the extent of phenotypic variation existing at different spatial scales, by measuring many traits and their phenotypic plasticity in controlled greenhouse conditions and in an experimental field at the University of Lille 1, (iii) identifying the traits under selection using three approaches : $F_{ST} - Q_{ST}$ comparisons, phenotype-ecology relationships and genotypic gradients of selection, (iv) identifying the selective agents potentially responsible for adaptive phenotypic variation, and (v) identifying the genetic basis associated with natural variation by a GWA mapping approach. The experiments conducted during this thesis suggest that a significant amount of variation exists for many phenotypic traits at a large spatial scale (World or France), but also at a very fine scale (even within many populations). While a significant part of this variation may have been shaped by non-selective processes, the other part of this variation is suggested to have been shaped by local adaptation. The pattern of local adaptation revealed by the different methods is very complex and appears to be the result of nested selection pressures varying at different geographic scales, (France scale and within-population scale), and acting on different traits or plasticities. In agreement with this pattern, the genetic basis associated with phenotypic natural variation was shown to be highly dependent on the geographical scale considered.