



**HAL**  
open science

# A European ecological and evolutive history of black locust (*Robinia pseudoacacia* L.)

Xavier Bouteiller

► **To cite this version:**

Xavier Bouteiller. A European ecological and evolutive history of black locust (*Robinia pseudoacacia* L.). Life Sciences [q-bio]. Université de Bordeaux, 2018. English. NNT: . tel-02789657

**HAL Id: tel-02789657**

**<https://hal.inrae.fr/tel-02789657>**

Submitted on 5 Jun 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

# Une histoire écologique et évolutive du robinier faux-acacia (*Robinia pseudoacacia* L.) depuis son introduction en Europe

Xavier Bouteiller

► **To cite this version:**

Xavier Bouteiller. Une histoire écologique et évolutive du robinier faux-acacia (*Robinia pseudoacacia* L.) depuis son introduction en Europe. Biodiversité et Ecologie. Université de Bordeaux, 2018. Français. <NNT: 2018BORD0162>. <tel-01926405>

**HAL Id: tel-01926405**

**<https://tel.archives-ouvertes.fr/tel-01926405>**

Submitted on 19 Nov 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



# THÈSE

Présentée à

**L'UNIVERSITÉ de BORDEAUX**

ÉCOLE DOCTORALE SCIENCES ET ENVIRONNEMENTS

**Par Xavier Bouteiller**

Pour obtenir le grade de DOCTEUR

SPÉCIALITÉ :

Écologie évolutive, fonctionnelle et des communautés

**Une histoire écologique et évolutive du robinier faux-acacia  
(*Robinia pseudoacacia* L.) depuis son introduction en Europe.**

**A European ecological and evolutionary history of black locust  
(*Robinia pseudoacacia* L.)**

Soutenue le 28 septembre 2018

M. Thomas Guillemaud	DR, INRA Sophia Antipolis	Rapporteur
M. Éric Petit	DR, INRA Rennes	Rapporteur
M. Gérard Largier	Ing., CBN des Pyrénées et de Midi-Pyrénées	Examineur
M. Leopoldo Sanchez Rodriguez	DR, INRA Orléans	Examineur / Président du jury
Mme Annabel Porté	DR, INRA Bordeaux	Directrice de thèse
Mme Stéphanie Mariette	CR, INRA Bordeaux	Co-directrice de thèse

*Il est démontré, disait-il, que les choses ne peuvent être autrement : car tout étant fait pour une fin, tout est nécessairement pour la meilleure fin. Remarquez bien que les nez ont été faits pour porter des lunettes ; aussi avons-nous des lunettes. Les jambes sont visiblement instituées pour être chaussées, et nous avons des chausses. [...].Par conséquent, ceux qui ont avancé que tout est bien ont dit une sottise : il fallait dire que tout est au mieux.*

*Voltaire, Candide ou l'Optimisme, 1759*

# Remerciements

Lorsqu'il y a six ans, j'ai repris le chemin des bancs de l'université de biologie, en parallèle à mon activité d'enseignant, après un premier cursus de physicien, j'étais loin d'imaginer où cela aller me mener. Pendant deux ans, j'ai mené une double vie d'enseignant et d'étudiant. J'ai fini par me consacrer entièrement au Master de biologie, effectué à l'université de Rouen, avant de rejoindre Bordeaux pour me lancer dans cette thèse.

Toute cette expérience a été particulièrement intense et enrichissante. J'ai singulièrement apprécié la variété du travail effectué, depuis les missions d'échantillonnage à l'étranger, les 8000 km parcourus sur les routes américaines et les 3000 km en Europe du Nord, en passant par le travail d'expérimentation, avec les 5000 robiniers semés que nous avons suivi un par un et enfin jusqu'à l'analyse des données pour laquelle je me suis passionné et qui m'a ouvert un autre univers. Je souhaite remercier toutes les personnes qui m'ont accompagné et soutenu durant ces années.

Je remercie tout d'abord Annabel Porté et Stéphanie Mariette qui ont été des directrices de recherche idéales. Elles ont toujours su me conseiller pour m'orienter dans mes recherches et si ce travail a pu voir le jour c'est bien grâce à leurs enseignements. Les nombreuses discussions que nous avons eues tout au long de cette thèse ont largement contribué à alimenter ma réflexion. Merci encore de votre disponibilité et de la confiance que vous m'avez accordée.

Je remercie également tous les membres du jury qui me font l'honneur de relire ce travail, Gérard Largier, du CBN des Pyrénées et de Midi-Pyrénées, Leopoldo Sanchez Rodriguez, de l'INRA d'Orléans, Thomas Guillemaud, de l'INRA de Sophia Antipolis et Éric Petit de l'INRA de Rennes

Ce projet n'aurait pas pu se réaliser sans l'aide des nombreuses personnes qui ont échantillonné des populations de robinier à travers toute l'Europe, de l'Espagne à la Pologne, et les Etats-Unis. A ce titre je remercie particulièrement Marcela van Loo, de l'université de Vienne en Autriche, Gideon Bradburd et Paul Bloese, de l'université du Michigan aux États-Unis qui m'ont fourni une aide logistique indispensable. J'ai particulièrement apprécié votre accueil et les conversations que nous avons pu avoir sur nos pays respectifs. Un grand merci

également à Dennis Born, Nicolas Cheval, Marianne Correard, Olivier De Thier, Christos Galanos, Pablo Goikoetxea, Sébastien Irola, Marcin Klisz, Emilio Laguna, Ludivine Lassois, Roman Longaner, Antonin Martinik, Coralie Mengal, Arnaud Monty, Chrisotphe Orazio, Patrick Pastuszka, Joan Pino, Alexandra Quenu, Volker Schneck, Cindy Verdu, Tomasz Wojda, Rafael Zas.

Je tiens également à adresser mes remerciements à Rémy Petit qui m'accueilli au sein de l'unité de recherche dont il avait la charge ainsi que tous les membres des équipes Génétique et Ecologie des Populations et Écologie et Génomiques fonctionnelles, auxquelles j'ai appartenu, pour leur accueil chaleureux.

J'ai particulièrement apprécié l'ambiance du laboratoire oscillant entre moments conviviaux et discussions de fond. Merci à toutes les personnes avec lesquelles j'ai partagé des moments chaleureux. Merci à Hermine A., Frédéric B., Marta B.G., Gaëlle C., Sylvain D., Thibaut F., Santiago G.M., Anne-Isabelle et Laurent L., Marina de M.V., Alix S., Corinne V.

Je garde également en souvenir les conversations que j'ai eues avec Myriam H. sur les sentiers de Pierroton lors de nos sessions de running et ainsi que les séances de fractionné sur la fameuse grande ligne droite. Je crois que j'ai épuisé toutes mes excuses disponibles pour éviter les prochains marathons !

Merci également à tous mes collègues enseignants avec lesquels je suis intervenu : Marie-Lise B., Hugues B., Emmanuel C., Jean-Christophe D., Virgil F., Xavier N., Laure V. J'ai toujours souhaité m'investir dans l'enseignement et je suis heureux d'avoir pu intervenir dans vos équipes.

C'est l'occasion pour moi de remercier également tous les enseignants qui, sans qu'ils ne le sachent, ont eu une influence considérable sur mon parcours et que je cite par ordre chronologique d'apparition : Mme Leclerc, Alain Mariette, Sylvain Le Brozec et Pierre Margerie.

Je remercie également tous les autres doctorants qui m'ont accompagné durant ce travail : Juliette A., qui commence, Thomas C., le calcul de  $Q_{ST}$  par des modèles *Zero Inflated Negative Binomial* n'aura bientôt plus de secret pour nous, Alexandre C., merci pour la formation qui m'a permis d'atteindre les sommets, Tania F., Homero G., Myriam G., Adib O. et Charlie P.

Merci à tous les membres de la plateforme génome transcriptome de Bordeaux Émilie C., Adline D., Erwan G. pour leur aide dans le génotypage des individus. Merci aussi à Patrick L. pour son aide dans le travail de laboratoire.

Merci à Régis B. et Raphaël S. pour leur aide précieuse dans la configuration des chambres climatiques, dans le suivi des expérimentations et dans l'échantillonnage périlleux des populations de robinier en Angleterre.

Merci aux stagiaires qui sont venus décortiquer, trier, peser, compter, mesurer, extraire les innombrables robiniers qui ont été utilisés dans cette étude. Leur patience a été mise à rude épreuve : Emmi A., Florian M., Marion P., Alexandre R.

Merci à toutes les personnes qui m'ont encouragé lorsque je suis reparti sur les bancs de la fac. Tous ces encouragements m'ont particulièrement touché et m'ont aidé à conserver ma motivation toutes ces années. Je pense particulièrement à Barbara M. et à Philippe T. avec qui j'ai adoré enseigner à Blangy.

Merci aussi à Camille D., à nos nombreuses conversations sur l'épistémologie des sciences. Je cite pour toi Ernest Rutherford qui a si souvent alimenté nos débats : *"All science is either physics or stamp collecting"*.

Merci également à Richard C. dont l'énergie communicative fait qu'il n'y a pas de sommet trop haut.

Je pense également à Charles-Henri L., Raphaël B., Thibault D., Thomas L., Charles LMDKDLC, Bertrand H., Ronald S., Arnaud H., Alexandre C., Olivier H., Claire T., Lilian M. On était jeune, on était beau ; on sentait bon la Kanterbrau. Et si la jeunesse ne dure pas et que nous voilà papas, merci pour votre soutien, pour toutes ces années et celles à venir, comme dirait P.B. « il vaut mieux vivre avec des remords qu'avec des regrets ».

Je remercie aussi mes parents qui m'ont toujours dit que je devais faire de la recherche mais que je n'ai pas voulu écouter, ainsi que mes frères et sœurs, Éric, Sophie et Julie, qui m'ont toujours soutenu et encouragé dans ce projet.

Pour finir, je remercie Claire pour son soutien indéfectible, ses encouragements et sa motivation. Ses nombreuses relectures, ses conseils m'ont été indispensables tout au long de ces années. Je pense également à nos deux meilleurs projets que nous co-encadrons, Madeleine et Gustave, qui sont venus nous accompagner merveilleusement durant ces années et qui ont démontré que le sommeil n'est finalement pas indispensable.



# Table des matières

<b>Préface.....</b>	<b>1</b>
<b>Enseignement.....</b>	<b>4</b>
<b>Annexes de la partie Enseignement .....</b>	<b>6</b>
<b>Travail de recherche .....</b>	<b>7</b>
<b>1. Introduction générale.....</b>	<b>8</b>
1.1. Les invasions biologiques existent-t-elles ?.....	8
1.2. Au-delà de la polémique : comment une espèce devient-elle invasive ?.....	10
1.2.1. Les barrières de l'invasion .....	10
1.2.2. Les raisons du succès : l'interaction de plusieurs facteurs .....	13
1.3. A la lumière de l'évolution .....	18
1.3.1. Du génotype au phénotype.....	18
1.3.2. Comment modifie-t-on son phénotype ?.....	19
1.4. Mettre en évidence l'évolution : du laboratoire à l'expérimentation. ....	22
1.4.1. La génétique des populations pour déterminer la structure des populations .....	22
1.4.2. Comprendre la variabilité phénotypique grâce à la génétique quantitative.....	25
1.4.3. Génétique quantitative et plasticité phénotypique .....	30
1.5. L'adaptation locale, une des raisons du succès des espèces invasives ?.....	31
1.6. De la dissémination du robinier dans les forêts européennes .....	35
1.7. Objectifs de ce travail de recherche .....	40

<b>Chapitre 1. Few North-Appalachian populations are the sources of European black locust.....</b>	<b>43</b>
<b>Chapitre 2. Using automated sanding to homogeneously break seed dormancy in black locust (<i>Robinia pseudoacacia</i> L., Fabaceae).....</b>	<b>99</b>
<b>Chapitre 3. No evidence for genetic differentiation in juvenile traits between Belgian and French populations of the invasive tree <i>Robinia pseudoacacia</i>.....</b>	<b>114</b>
<b>Chapitre 4 The seeds of invasion: Enhanced germination among invasive European populations of black locust compared to native American populations. ....</b>	<b>138</b>
<b>Discussion générale .....</b>	<b>200</b>
<b>3. Discussion.....</b>	<b>201</b>
3.1. D'où vient le robinier européen ou comment retracer le scénario de l'introduction ?.....	201
3.1.1. Une introduction restreinte depuis les États-Unis suivie d'une multiplication en Europe .....	201
3.1.2. Considérations méthodologiques quant à l'inférence de l'histoire de l'invasion .....	204
3.2. Structure génétique des populations des Etats-Unis à l'Europe.....	206
3.2.1. Des processus naturels anciens ont façonné la diversité aux États-Unis tandis que son introduction en Europe est marquée par l'empreinte de l'Homme .....	206
3.2.2. Considérations méthodologiques sur l'estimation de la diversité à l'aide de marqueurs SNPs ....	210
3.3. Le caractère invasif du robinier peut-il être expliqué par l'évolution de certains traits phénotypiques ?	212
3.3.1. Un robinier qui se reproduit davantage de façon asexuée dans l'aire d'introduction .....	212
3.3.2. Une germination potentielle améliorée dans l'aire d'introduction .....	214
<b>4. Conclusion.....</b>	<b>218</b>
<b>Bibliographie .....</b>	<b>219</b>



# Préface

Cette thèse s'est appuyée sur une approche combinée de génétique des populations et de génétique quantitative afin d'analyser le rôle des processus évolutifs à l'œuvre dans le succès d'une espèce invasive ligneuse : le robinier faux-acacia (*Robinia pseudoacacia* L.)<sup>1</sup>. Ce travail de recherche a été réalisé au sein de l'UMR BIOGECO de l'Université de Bordeaux et du département d'écologie des forêts, des prairies et des milieux aquatiques de l'INRA. Il a été financé par une bourse nationale au mérite attribuée par le Ministère de la Recherche et de l'Enseignement Supérieur. De nombreux organismes et institutions ont permis de financer les expérimentations et les missions d'échantillonnage et ont ainsi rendu possible ce travail. Je remercie la Société Française d'Ecologie et d'Evolution, le labex COTE, l'Agence de l'eau Adour Garonne, l'ANR Xyloforest, le programme européen COST action NNEXT ainsi que les unités de recherche Gembloux AgroBioTech et BIOGECO.

Dans ce manuscrit, je présenterai tout d'abord les activités d'enseignement que j'ai menées de façon conjointe à mon travail de recherche, que j'exposerai dans une seconde partie.

La partie enseignement exposera brièvement les cours les différentes activités d'enseignement dans lesquelles je me suis investi. Elle sera suivie par une seconde partie concernant mon travail de recherche à proprement parler qui sera elle-même décomposée en trois volets : une introduction générale suivie par les chapitres ou articles de cette thèse et enfin une synthèse. L'introduction et la synthèse peuvent se lire pratiquement indépendamment des articles. J'y ai volontairement limité la description de la méthodologie expérimentale et si le lecteur souhaite en savoir davantage sur les éléments méthodologiques qui ont suscité ce travail, il peut se référer aux chapitres.

En premier lieu, j'ai donc rédigé une introduction générale, qui replace dans une perspective plus large les travaux entrepris. Ainsi, j'expose les concepts importants en écologie des invasions et fais le lien avec l'écologie évolutive. A mesure que je dévoile ces aspects, je présente brièvement quelques aspects méthodologiques et je conclus cette introduction en synthétisant les connaissances initiales sur le robinier et en

---

<sup>1</sup> Dans tout cet ouvrage, le terme « robinier » doit s'entendre comme une référence au robinier faux-acacia (*Robinia pseudoacacia* L.)

exposant mes questions de recherche. En second lieu, les chapitres de cette thèse sont présentés. Ils sont constitués par les quatre articles scientifiques rédigés durant ces 3 années. Deux articles ont été publiés, un troisième est en révision et il est prévu que le dernier soit soumis prochainement. Enfin, j'ai terminé ce rapport par une synthèse qui reprend les principaux résultats que je souhaite souligner et je reviens sur quelques considérations méthodologiques importantes.

# Enseignement

# Enseignement

Après six ans passés à enseigner les sciences physiques et chimiques dans les collèges et lycées de Seine-Maritime, il m'apparaissait indispensable de continuer à enseigner pendant mon contrat doctoral. Au cours de ces trois années, j'ai toujours eu à cœur d'allier recherche et enseignement afin de transmettre les connaissances scientifiques acquises ainsi que mes compétences notamment en analyse de données. L'interaction avec les étudiants est, à mon sens, particulièrement bénéfique aussi pour le chercheur, et lui permet de s'approprier ses questions de recherches tout en prenant du recul sur son travail.

J'ai donc accompli deux missions de monitorat au cours de mes deuxième et troisième années de doctorat à travers des cours magistraux, des TD et des TP. Je suis intervenu auprès des étudiants de licence et Master de l'université de Bordeaux et de Master de l'école d'ingénieur Bordeaux Sciences Agro. Le tableau 1 synthétise les différents enseignements dans lesquels je suis intervenu pour un total de 139h.

Table 1 Bilan des enseignements effectués au cours des deux missions de monitorat qui m'ont été confiées (2016 – 2018). CM : Cours Magistraux, TD : Travaux Dirigés, TDm : TD machine, TP: travaux pratiques

Intitulé	Type	Niveau	Nb d'heures
Génétique quantitative	CM	Master 1	20h
Ecophysiologie et assimilation carbonée	CM	Master 1	3h
Ecologie	CM	Master 1	4h
Biostatistiques avec R	TDm	Master 1	17h
Biostatistiques avec R	TDm	Licence 3	31h
Botanique et systématique végétale	TP	Licence 3	23h
Génétique des populations et analyse moléculaire des organismes	TD	Licence 3	14h
Ecologie	TD	Licence 2	27h

De façon complémentaire, je me suis investi dans la rédaction d'un livre de biostatistiques avec R<sup>2</sup> dans lequel je suis coauteur et qui a été publié chez Dunod cette année. J'ai notamment rédigé les fiches du chapitre analyses multivariées. Un chapitre est intégré à ce rapport de doctorat à titre d'exemple.

Pour finir, j'ai participé à l'encadrement de plusieurs étudiants en stage durant mon doctorat, participant ainsi à la définition de leur mission, à la supervision du déroulement du stage, à la relecture critique du rapport de stage et à la préparation à la soutenance. Le détail des stagiaires encadrés est indiqué ci-dessous :

- Encadrement de M. Pucheu, Licence 1, Univ. de Poitiers, printemps 2016 (2 mois). Génotypage des individus européens.
- Encadrement de F. Moret, Master 2, Univ. de Picardie, janvier – juin 2017 (6 mois). Suivi de l'expérimentation de génétique quantitative du robinier.
- Encadrement de J. Brian, Master 1, Univ. Bordeaux, printemps 2017 (4 mois). Evaluation de la dynamique du robinier en ripisylve.

---

<sup>2</sup> Nogués, X., Garenne, A., Bouteiller, X.P., Fievet, V. (2018). Biostatistiques, Tout le cours en fiche, Dunod supérieur.



# **Annexes de la partie Enseignement**

**Exemple d'une des fiches rédigées pour le livre Biostatistiques**

Des scientifiques se rendent sur différents sites pour faire des comptages d'oiseaux. À la fin de leur campagne de terrain, ils disposent d'un tableau de présence / absence avec sur chaque ligne, les différentes stations de mesure et sur chaque colonne les espèces identifiées. Les variables sont binaires, les valeurs peuvent être 0 (espèce non observée à la station de mesure considérée et 1 espèce observée).

Comment peut-on regrouper les individus ou les stations de mesures ? Est-ce que ce regroupement a un sens écologique ?

## 1. Les caractéristiques de la CAH

### ■ Autres noms

*Hierarchical Clustering* en anglais

### ■ Utilisation et conditions d'utilisation

L'objectif de la CAH est de regrouper des individus en fonction de leur similarité. La CAH peut ainsi être utilisée en écologie, en génétique, en sciences économiques ... Nous pouvons citer comme autres exemples :

- la réalisation d'une typologie de milieux écologiques : les milieux sont regroupés par types.
- l'identification des espèces qui partagent les mêmes habitats.
- le regroupement des villes selon leur ressemblance en termes de climat.

Au-delà des regroupements, la CAH permet d'évaluer des degrés de proximité ou d'appartenance à ces groupes.

La CAH est une méthode qui vise à subdiviser une population en groupes (classes) d'individus les plus semblables possibles.

La classification peut être **ascendante** lorsqu'on part des individus en les considérant chacun isolé dans une classe. On les regroupe successivement en classes de plus en plus grande jusqu'à former une seule classe, c'est une méthode agrégative.

La classification **descendante** hiérarchique part de l'ensemble des individus regroupés dans une seule classe et les sépare successivement jusqu'à ce que chaque individu soit isolé dans une classe. C'est une méthode séparative. Elle demande d'énormes ressources de calcul c'est pourquoi elle est peu utilisée. La réalisation d'une CAH se déroule en deux étapes :

- Calcul de la matrice de distance entre individus. Cette étape est cruciale, plusieurs distances existent, il est nécessaire de les adapter au type de données (présence ou absence, abondance, comptage, quantitatives ...).
- Regroupement en classes, c'est l'étape de la CAH à proprement parler. Plusieurs algorithmes existent pour regrouper les individus.

### ■ Le choix de la distance

En mathématiques, une distance vraie obéit à trois axiomes.

1.  $d_{ab} = d_{ba}$
2.  $d_{ab} \geq 0$  et  $d_{ab} = 0$  si et seulement si  $a = b$
3.  $d_{ab} \leq d_{ac} + d_{ca}$

Le 3<sup>e</sup> axiome est appelé inégalité triangulaire, il peut sembler évident, mais en pratique il n'est pas toujours vérifié. Dans les cas où il n'est pas vérifié on ne parle alors plus strictement de distance mais de dissimilarité.

Pour les calculs de la matrice de distance ou de dissimilarité, les plus communes en écologie sont :

- **La distance euclidienne :**
  - distance vraie ;
  - convient aux variables quantitatives, par exemple des mesures morphométriques ;
  - il peut être nécessaire de centrer réduire les variables si les unités de mesure ou les variances sont différentes (fonction *scale* ()).
- **La distance de Bray-Curtis :**
  - c'est une dissimilarité avec une échelle comprise entre 0 et 1 ;
  - convient aux relevés d'abondance ou de comptage ;
  - le principe est de faire la somme des différences d'abondance pour chacune des variables (par exemple la liste d'espèces) entre deux individus statistiques (par exemple les sites de relevés) rapportée à l'abondance totale.
- **La distance de Jaccard :**
  - c'est une dissimilarité avec une échelle comprise entre 0 et 1 ;
  - convient aux données de présence absence ;
  - le principe est de comparer les individus deux à deux en supprimant les variables pour lesquelles il y a co-absence puis de faire le rapport entre le nombre d'observations faites uniquement pour l'un des individus et le nombre d'observations totales.

Pour réaliser la CAH à partir de la matrice de distance ou de dissimilarité, de nombreux algorithmes existent. Le but n'est pas ici de les détailler tous. Les plus utilisées sont :

- **Le lien moyen (UPGMA) :**
  - consiste à calculer la distance moyenne entre toutes les paires de classe ;
  - elle est sensible aux valeurs extrêmes et suppose une évolution constante dans toutes les branches, ce qui peut être problématique par exemple pour construire un arbre phylogénétique.
- **La distance de Ward :**
  - c'est souvent la plus performante ;
  - elle consiste à maximiser l'inertie inter classe à chaque étape et donc à minimiser l'inertie intra classe.

### ■ Fonction R, package et organisation des données dans le fichier

`vegdist{vegan}` pour calculer la matrice de distance ou de dissimilarité

- Arguments : `method="euclidean"` , `"bray"` , `"jaccard"` `hclust stats` pour réaliser la CAH
- Arguments : `method="average"` , `"ward.D2"`

Le fichier de données est un tableau avec les individus statistiques en ligne et les variables en colonnes.

## 2. Résolution du cas d'étude

### ■ Importation et observation des données

```
OisPA <- read.table("Ois_PA.csv" , sep = ";", h=T)
str(OisPA)
'data.frame': 20 obs. of 16 variables:
 $ Site      : Factor w/ 4 levels "Champ","Foret",...: 3 3 3 3 3
              1 1 ...
 $ Aigrette  : int 0 0 0 0 0 0 0 0 0 0 ...
```

```

$ Buse      : int 0 0 0 0 1 1 0 1 1 0 ...
$ Canard    : int 0 0 0 0 0 0 1 1 0 0 ...
...
OisPA2 <- OisPA[,-1]#retirer la première colonne (site d'échan-
tillonnage)

```

### ■ Calcul de la matrice de distance

Les données étant de type présence / absence, nous allons utiliser la distance de Jaccard pour établir la matrice de dissimilarité.

```

library(vegan)# Charger la librairie vegan
DisJac <- vegdist(OisPA2, method="jaccard")

```

### ■ Classification ascendante hiérarchique

Nous utilisons la distance de Ward qui vise à maximiser l'inertie inter-classe.

```

ClustSite <- hclust(DisJac,method="ward.D2" )
plot(ClustSite,# 1er argument : matrice de distance
hang=-1 , # pour que les branches de l'arbre soient alignées en bas
labels =OisPA[,1]) # Nom des branches
rect.hclust(ClustSite, k=4, border="black") # Afficher sur le den-
drogramme

```

le découpage en classe

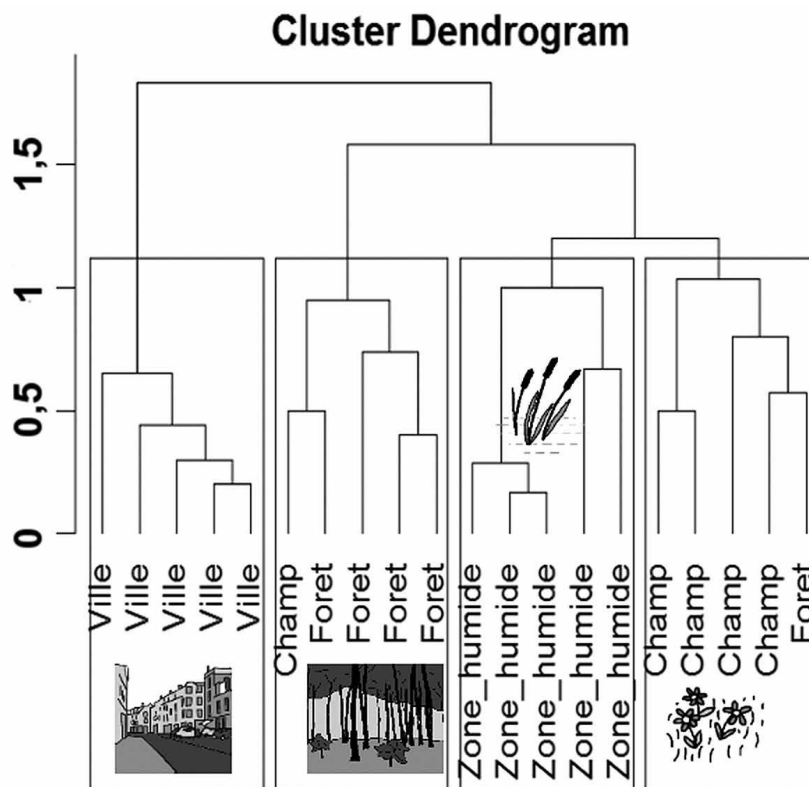


Figure 97.1 Dendrogramme des milieux.

# **Travail de recherche**

# Introduction générale

## 1. Introduction

### 1.1. Les invasions biologiques existent-t-elles ?

*Il fallait être Newton pour apercevoir que la lune tombe,  
quand tout le monde voit bien qu'elle ne tombe pas.  
Paul Valéry, Mélange, 1939*

L'acte fondateur de la science des invasions biologiques est souvent attribué à Charles Elton pour la publication du livre *the Ecology of animal and plant invasions* (Elton, 1958). Pour la première fois alors, Charles Elton formalisait le cadre conceptuel lié à l'introduction d'espèces en dehors de leur aire de distribution naturelle et à leur expansion dans un nouvel habitat. Dès lors, l'intérêt scientifique pour cette discipline n'a cessé de croître et le nombre de publications scientifiques sur ce thème a suivi une progression exponentielle (Richardson and Pyšek, 2008).

Toutefois de récentes critiques ont émergé à l'encontre de ce concept (Richardson and Ricciardi, 2013), allant jusqu'à un appel pour la disparition de cette discipline (Davis et al., 2011; Valéry et al., 2013). Selon ses détracteurs, les invasions biologiques pourraient même avoir des impacts positifs et contribuer à l'augmentation de la diversité spécifique locale (Brown and Sax, 2004; Thomas, 2013) ; de plus, toujours selon ces auteurs, elles seraient une persécution irrationnelle des espèces non natives voire une sorte de xénophobie à leur encontre (Thomas, 2013) ; enfin, la distinction entre espèce native et non native serait artificielle et sans intérêt scientifique (Davis et al., 2011; Valéry et al., 2013). Nous allons brièvement démontrer les limites de ces arguments et mettre en évidence qu'il convient de regarder au-delà de ce clivage les questions posées par les invasions biologiques qui dépassent ce seul cadre.

En premier lieu, de nombreux travaux ont démontré que les espèces invasives constituent une menace sérieuse sur la biodiversité (Lonsdale, 1999; Vitousek et al., 1996). Des études empiriques ont en effet mis en évidence que des espèces invasives sont responsables au moins en partie du déclin d'espèces locales par compétition, c'est notamment le cas de l'arbre invasif *Cinchona pubescens* introduit aux Galapagos (Jäger et al., 2009), ou par hybridation comme entre *Spartina alterniflora* et *S. foliosa* en Californie (Ayres et al., 2004; Simberloff, 1996). Il apparaît ainsi que les espèces invasives sont, à elles seules ou en interaction avec d'autres facteurs, une cause importante du déclin de la biodiversité (Clavero and García-Berthou, 2005; Gilbert and Levine, 2013; Ricciardi et al., 2004; Simberloff, 2011, 1996). Considérer que les espèces invasives favorisent la

biodiversité locale apparaît être un argument biaisé oubliant que la richesse spécifique ne saurait être l'indicateur le plus pertinent de la qualité d'un écosystème et que certaines espèces locales peuvent être impactées irrémédiablement par cette introduction. Que toutes les espèces natives ne soient pas impactées, que certaines espèces bénéficient même de la présence du taxon introduit ou encore que d'autres effets interagissent avec l'invasion comme l'augmentation de la pression anthropique sur l'habitat ne contredit pas nécessairement ce constat. Enfin, le temps de latence entre l'introduction et le début de l'invasion et la nécessité de suivi à long terme pour attester de l'impact sur la diversité locale (Boero, 2011) doivent inciter à la prudence avant d'affirmer qu'une espèce introduite a ou non un impact sur la biodiversité.

De plus, certains auteurs dénoncent la volonté d'éradiquer les espèces introduites davantage par principe que sur un fondement scientifique raisonné invoquant même un lien avec la xénophobie (Brown and Sax, 2004). Cela n'est pas sans rappeler la discussion de Simberloff (1996) qui, suggérant de limiter l'hybridation et l'introgession, anticipe les critiques qu'on pourrait lui faire de vouloir préserver la « pureté des races ». Il convient d'ailleurs de rappeler que la majorité des gestionnaires ne souhaite pas éradiquer totalement l'espèce invasive mais davantage contrôler l'invasion plutôt que de la supprimer. Richardson & Ricciardi (2013) soulignent avec raison que les gestionnaires ont souvent des moyens limités et qu'ils ciblent en priorité les espèces qui leur posent problèmes. Un autre argument des contradicteurs est que la lutte contre les espèces invasives est souvent coûteuse et peu efficace et donc inutile (Davis et al., 2011). Cela nous semble être un syllogisme contestable, que la lutte ne soit pas efficace n'empêche pas qu'elle soit nécessaire et pourrait même indiquer l'inverse. Le coût des espèces invasives tous taxons confondus a été estimé à plus de 380 milliards de dollars par an (Pimentel et al., 2000), en prenant en compte uniquement six pays. Pour donner un exemple plus précis, aux États-Unis, la plante *Lythrum salicaria* colonise plus de 100 000 ha chaque année et 45 millions de dollars sont dépensés chaque année pour tenter de contrôler sa dispersion et compenser les pertes fourragères (Pimentel et al. 2005). Pour les plantes, le coût annuel avoisinait 0.178 milliards de dollars pour les Etats-Unis et l'Afrique du Sud seuls (Pimentel et al., 2000). Ce montant inclue certes les coûts de gestion des invasions mais aussi les pertes causées par les espèces invasives notamment liés aux dégâts à l'agriculture. Les auteurs soulignent d'autre part qu'il est difficile de donner une valeur à la disparition d'espèces, induite par certaines invasions, biaisant ainsi l'estimation économique des impacts.

Enfin, certains auteurs contestent la distinction entre plantes natives et non natives la jugeant artificielle et sans fondement scientifique (Thomas, 2013; Valéry et al., 2013). Pourtant, l'impact des espèces exotiques dans leur milieu d'introduction semble supérieur à celui des espèces locales. Une méta-analyse démontre ainsi que les

espèces introduites causent davantage de dommages que les espèces locales (Paolucci et al., 2013). Il est vrai que des espèces natives peuvent également causer des dégâts, mais moins fréquemment que les espèces introduites (Simberloff et al., 2012). Une autre critique consiste à dire que les mouvements d'espèces ont toujours existé, mais c'est à notre sens confondre expansion d'aires et invasions qui méritent d'être distinguées (Blondel et al., 2013; van Kleunen et al., 2011a). Les mouvements observés lors des invasions sont plus rapides que ceux observés lors des colonisations post glaciaires (Ricciardi, 2007; Richardson and Pyšek, 2006). Les introductions d'espèces invasives liées aux activités humaines dépassent ainsi en vitesse, en étendue et en intensité les expansions naturelles et la distinction entre espèces natives et non natives mérite d'être reconnue (Richardson and Ricciardi, 2013) afin d'être mieux appréhendée.

Afin de dépasser ces polémiques, il convient de voir au-delà. Les espèces invasives existent, elles se dispersent, elles menacent par endroit la biodiversité locale, certes elles apportent également des services sans que cela ne soit contradictoire mais surtout elles questionnent et ouvrent ainsi de nombreux pans de recherche fondamentale et appliquée.

## **1.2. Au-delà de la polémique : comment une espèce devient-elle invasive ?**

*La clef de toutes les sciences est sans contredit le point d'interrogation,  
nous devons la plupart des grandes découvertes au : Comment ?  
et la sagesse dans la vie consiste peut-être à se demander à tout propos : Pourquoi ?  
Honoré de Balzac, La Peau de chagrin, 1831*

### **1.2.1. Les barrières de l'invasion**

Nous avons pu discerner que la science des invasions biologiques est une discipline relativement récente, en pleine expansion et, comme toute théorie scientifique, soumise à des remises en question. Il semble important de définir correctement les concepts et le lexique afin d'éviter les confusions et amalgames trompeurs. La notion d'invasion fait souvent recours à des concepts évoquant un certain anthropocentrisme (Richardson et al., 2000), le terme « invasion » est en lui-même frappant dans son appartenance au champ lexical guerrier et sa connotation négative, ce qui peut induire des réactions émotionnelles. Ces auteurs insistent donc sur la nécessité d'utiliser une terminologie rigoureuse lors de l'étude de l'écologie des invasions biologiques et recommandent ainsi plusieurs définitions (Table 1, Richardson et al., 2000).



Table 1 Terminologie recommandée par Richardson et al. (2000) en écologie des plantes invasives. Les termes anglais sont indiqués en italique.

Terminologie	Définition
Plantes allochtones ( <i>Alien plants</i> )	Taxons dont la présence dans une aire donnée est due à une introduction intentionnelle ou accidentelle liée à des activités humaines.
Plantes allochtones temporaires acclimatées ( <i>Casual alien plants</i> )	Plantes étrangères qui peuvent fleurir et éventuellement se reproduire occasionnellement dans une aire donnée, mais qui ne forment pas de populations pérennes dépendant ainsi d'introductions répétées pour leur persistance.
Plantes naturalisées ( <i>Naturalized plants</i> )	Plantes étrangères qui se reproduisent de façon régulière et peuvent entretenir des populations sur de nombreux cycles de vie, sans (ou malgré) une intervention humaine.
Plantes invasives ou exotiques envahissantes ( <i>Invasive plants</i> )	Plantes naturalisées qui ont la capacité de former des nouvelles pousses, souvent nombreuses, à des distances importantes de la plante mère (i.e. plus de 100 m en moins de 50 ans pour les taxons se dispersant par des graines ou des propagules et plus de 6 m en 3 ans dans le cas de taxons se propageant par des rhizomes, des stolons, des racines ou des tiges rampantes)
Plantes nuisibles ( <i>Weeds / Pests</i> )	Plantes, <b>non nécessairement étrangères</b> , qui croissent dans des zones où elles sont indésirables et qui produisent souvent des effets détectables sur l'environnement ou l'économie. Les plantes nuisibles à l'environnement ( <i>environmental weeds</i> ) sont des plantes étrangères qui envahissent la végétation naturelle en faisant habituellement diminuer la richesse spécifique des espèces natives et en altérant le fonctionnement de l'écosystème (Randall, 1997)
Transformateurs ( <i>Transformers</i> )	Ensemble des plantes invasives qui modifient une propriété, une condition ou la nature d'un écosystème sur une surface significative de celui-ci. Il s'agit d'une sous sélection des plantes invasives qui ont un fort impact négatif sur les écosystèmes.

Les recherches actuelles afin de théoriser l'écologie des invasions biologiques nécessitent d'éclaircir les processus de naturalisation et d'invasion. Ainsi, pour devenir invasif, un taxon donné doit franchir différentes barrières biotiques et abiotiques. Richardson et al. (2000) proposent un modèle conceptuel des étapes conduisant à l'invasion biologique. Ce modèle conceptuel a été remis à jour récemment avec une volonté d'unifier la discipline (Blackburn et al., 2011). Chacune des phases du processus correspond au franchissement d'une nouvelle barrière. Tout d'abord l'introduction signifie que la plante ou une propagule a réussi à franchir une barrière géographique importante par le biais d'une activité humaine (transport accidentel, introduction volontaire). Cette première étape correspond au processus de migration longue distance. L'espèce peut alors survivre dans son nouvel environnement et même se reproduire sexuellement ou végétativement, elle est dite acclimatée. Cependant la population peut ne pas se maintenir à long terme. Dès lors, les plantes étrangères temporaires doivent compter sur des introductions régulières pour leur persistance, la pression de propagules est alors fondamentale pour que les populations introduites se maintiennent (Richardson and Pyšek, 2006). La naturalisation commence uniquement quand la barrière environnementale n'empêche plus les individus de survivre dans l'aire d'introduction puis quand les différentes barrières qui limitent la reproduction sont dépassées. Par conséquent, la population est considérée comme naturalisée après avoir franchi ces trois premières étapes. Le nombre d'individus est alors suffisant pour

que la probabilité d'extinction en raison d'événements environnementaux stochastiques soit faible (Menges, 2000; Panetta & Randall, 1994).

Pour que la phase d'invasion commence, les plantes naturalisées doivent être capables de se disperser dans de nouvelles régions à distance de la zone d'introduction initiale en faisant face aux conditions biotiques et abiotiques de l'aire générale de répartition. Beaucoup vont commencer par se disséminer dans des habitats perturbés. Enfin, l'invasion de communautés non perturbées matures nécessite que le taxon surpasse la résistance opposée par différents moyens par les espèces natives comme le mutualisme, l'allélopathie ou la compétition (Richardson et al., 1999). Les espèces invasives sont celles qui sont capables de s'adapter dans de nouveaux habitats et de nouvelles conditions environnementales, qui se maintiennent sur un site depuis au moins dix ans en s'auto-remplaçant et qui ont le potentiel pour se disperser largement sur une grande surface ; cette définition se base sur la dynamique des espèces et n'inclue pas forcément de connotation d'impact (Blackburn et al., 2011; Richardson and Rejmánek, 2011). Richardson et al. (2000) estiment qu'entre 50% et 80% des espèces invasives peuvent être classées comme nuisibles, les autres, qualifiées d'espèces invasives bénignes, sont celles dont les dommages sont pratiquement indécélables la plupart du temps, il s'agit par exemple de *Arabidopsis thaliana*, *Cakile edentula* ou *Myosotis stricta* (Richardson et al., 2000). Les transformateurs représenteraient, elles, environ 10% des espèces invasives, c'est le cas notamment de *Acacia mearnsii*, *Bromus tectorum*, *Melaleuca quinquenervia* ou *Mimosa pigra* (Richardson et al., 2000).

De la capacité des espèces à se maintenir dans le nouvel habitat d'introduction face aux pressions de sélection dépend donc le succès de l'invasion. La « règle des dix » a été proposée afin d'estimer le nombre d'espèces introduites qui deviennent des invasives nuisibles (Williamson & Fitter, 1996). Il s'agit d'une estimation de la proportion d'espèces franchissant chaque stade de l'invasion et induisant des impacts négatifs. Elle prédit que 10% des espèces introduites deviennent des espèces acclimatées, 10% de celles-ci deviennent à leur tour naturalisées et enfin 10% des naturalisées deviennent des invasives nuisibles (Richardson & Pyšek, 2006). De façon empirique, une telle estimation suggère ainsi qu'entre 0,5% et 0,7% des espèces dans le groupe global des arbres et des buissons introduits en-dehors de leur aire de répartition naturelle seraient invasifs (Richardson & Rejmánek, 2011). Cependant des ratios très différents existent entre les espèces et les habitats, par exemple environ 0,01% des espèces introduites seraient devenues nuisibles en Angleterre, 1,3% des herbacées en Australie et 12% parmi les pins en Nouvelle-Zélande (P. A. Williams, 2003).

### 1.2.2. Les raisons du succès : l'interaction de plusieurs facteurs

Le succès d'une invasion est un processus complexe où plusieurs événements interagissent : (1) l'histoire de l'introduction, (2) l'invasibilité du milieu ou de l'écosystème récepteur (*invasibility*<sup>1</sup>) et (3) le potentiel invasif de l'espèce (*invasiveness*). Ainsi l'étude des facteurs d'invasion qui peuvent provenir des espèces ou de l'habitat peut permettre d'évaluer le potentiel invasif d'une espèce (Williams, 2003) et également de mieux comprendre la dynamique de l'invasion.

Tout d'abord, l'histoire de l'introduction joue un rôle important sur le succès d'une invasion. En premier lieu, plus une espèce a été introduite dans un nouvel environnement plus elle a de probabilité de devenir invasive, c'est ce qu'on appelle la pression de propagules (Colautti et al., 2006; Rejmánek, 2000; Richardson and Pyšek, 2006). La pression de propagules dépend à la fois du nombre d'individus introduits initialement et de la fréquence des nouvelles introductions. Une expérience a démontré que la taille d'une population de daphnie *Daphnia magna* après quarante jours de culture dans le milieu expérimental dépendait à la fois du nombre d'individus inoculés initialement dans le milieu mais aussi de la fréquence des nouvelles introductions et qu'il y avait une interaction entre ces deux facteurs (Drake et al., 2005). Ainsi en Afrique du Sud, il existe une corrélation significative entre le nombre d'occurrences spontanées et le nombre de plantations pour 57 espèces d'*Eucalyptus* introduits (Rejmánek, 2000). De même au sein du parc Kruger, le rôle de la contribution de la pression de propagules dans l'expansion de la population du cactus invasif *Opuntia stricta* a été souligné (Foxcroft et al., 2004). Finalement, Colautti et al. (2006) ont montré dans une méta-analyse que la pression de propagules est un important facteur explicatif de l'invasion : en effet parmi 64 études passées en revues, 55 ont montré une association significative entre la pression de propagules et le succès de l'invasion. D'autre part, un second aspect à prendre en compte dans l'histoire de l'introduction est qu'il existe un temps de latence entre la date d'introduction et la date à laquelle un taxon est déclaré invasif (Richardson and Pyšek, 2006). Ce temps de latence est variable selon les espèces. Dans une méta-analyse sur les espèces invasives du centre des États-Unis, Larkin (2011) a comparé les données historiques de 151 espèces ; il a mis en évidence que la durée du temps de latence s'échelonnait entre 3 et 140 ans avec une moyenne de  $47,3 \pm 34,6$  (SD) ans ce qui indique une forte variabilité (Larkin, 2011). Dans cette étude, Larkin a observé que le temps de latence était plus court pour les espèces introduites de façon accidentelle, mais

---

<sup>1</sup> Dans la mesure du possible, j'essaie d'éviter les termes anglophones, mais il peut arriver que je fasse référence à l'expression anglaise notamment dans le cas de concepts largement employés par la communauté scientifique

aussi qu'il n'y avait pas de différence significative dans la durée de la phase de latence entre les espèces ligneuses et herbacées. Ainsi la probabilité qu'une espèce devienne invasive augmenterait avec le temps de résidence dans la nouvelle aire. Un exemple est donné pour un buisson invasif, *Ligustrum robustum* (Oleaceae). Il a été introduit sur l'île Maurice à la fin du XIX<sup>ème</sup> siècle et sur la Réunion beaucoup plus récemment, en 1969. Il est maintenant un problème majeur sur l'île Maurice et a envahi toutes les forêts mauriciennes alors qu'à la Réunion, il commence tout juste à devenir problématique pour les forêts insulaires (Lavergne et al., 1999). Ainsi le temps écoulé depuis l'introduction devrait être pris en compte lorsqu'une espèce étrangère introduite est reportée ou non invasive et il est probable que le nombre d'espèces invasives identifiées va continuer d'augmenter, même dans le cas hypothétique où les nouvelles introductions cesseraient (Richardson and Pyšek, 2006). Ce décalage entre introduction et invasion a été conceptualisé sous la terminologie de dette d'invasion (Essl et al., 2011).

Deuxièmement, les caractéristiques du milieu envahi sont également susceptibles d'influencer le succès d'une invasion. Ainsi, l'invasibilité d'un milieu peut être défini comme sa susceptibilité à être envahi par des espèces introduites (Colautti et al., 2006; Richardson and Pyšek, 2006). Afin de déterminer si un habitat est davantage susceptible d'être envahi davantage qu'un autre, les études comparatives de milieux quantifient le nombre d'invasions par habitat (Lonsdale, 1999). Ainsi, par exemple, une idée répandue est que les îles seraient davantage sensibles aux invasions que les continents car les communautés y seraient moins diverses (Denslow, 2003; Lonsdale, 1999; Richardson and Pyšek, 2006), mais les éléments empiriques ne semblent pas le confirmer, et d'autres études infirment ce résultat (Sol, 2000; Vilà et al., 2010). Ainsi, des variations dans l'intensité de l'invasion entre deux milieux peuvent être dues simplement à des pressions de propagules différentes ou à des espèces introduites différentes (Hierro et al., 2005; Lonsdale, 1999). Cela a été montré notamment dans les milieux montagnards, qui seraient actuellement moins envahis, non pas du fait de leur différence de vulnérabilité aux invasions, mais du fait d'une plus récente et plus réduite introduction des espèces invasives (Pauchard et al., 2009)

Plusieurs théories tentent d'expliquer les caractéristiques qui influencent l'invasibilité du milieu, nous allons brièvement en présenter les principales.

L'hypothèse de la richesse spécifique (*species richness hypothesis*) a été énoncée par Elton (1958) qui suggéra qu'il existe une corrélation négative entre la diversité spécifique d'un habitat et son invasibilité. Une diversité plus faible conduirait à des relations interspécifiques plus faibles et davantage de niches vacantes, les espèces invasives pouvant ainsi profiter de ressources non utilisées (Hierro et al., 2005). Plusieurs études ont en effet montré que les communautés les plus riches en espèces étaient moins susceptibles d'être envahies (Kennedy et al., 2002; Levine,

2000). Cependant d'autres études ont trouvé que les communautés avec les diversités les plus élevées étaient les plus envahies (Hierro et al., 2005; Lonsdale, 1999). A une grande échelle, les mêmes conditions abiotiques qui promeuvent une grande diversité des espèces natives (climat, substrat, hétérogénéité de l'habitat) peuvent également fournir un environnement adapté à l'implantation des invasives (Richardson & Pyšek, 2006). La méta-analyse réalisée sur les arbres invasifs indique cependant qu'aucune étude empirique ayant testé cette hypothèse de niche vacante n'a démontré sa réalité (Lamarque et al., 2011). Ainsi le lien entre diversité et invasibilité n'est pas évident à démontrer et il est sans doute nécessaire d'utiliser un meilleur indicateur que la richesse spécifique seule (Richardson and Pyšek, 2006).

La théorie de la fluctuation des ressources postule qu'une communauté végétale devient plus susceptible d'être envahie quand il se produit une augmentation d'une quantité de ressources inutilisées. La fluctuation de la disponibilité des ressources serait ainsi un facteur clef d'invasibilité (Davis et al., 2000). L'augmentation de la disponibilité d'une ressource peut se faire selon deux voies : l'utilisation de cette ressource par la végétation locale peut diminuer ou le réservoir de cette ressource peut augmenter plus vite que la population ne peut le prélever. L'utilisation de la ressource peut décliner en raison de différents facteurs comme une perturbation qui détruirait une partie de la végétation locale réduisant ainsi les prélèvements en eau, en lumière ou en nutriments (Davis et al., 2000). Dans leur étude, Lavergne et al. (1999) ont remarqué que les sites les plus perturbés par l'Homme étaient ceux qui étaient les plus infestés par *Ligustrum robustum*. Chez les arbres, cette hypothèse a été testée dans différentes situations et dans l'intégralité des études, la perturbation testée induisait une augmentation de l'invasibilité des systèmes (Lamarque et al., 2011). Généralement, dans les environnements perturbés les facteurs abiotiques semblent être des déterminants d'invasibilité plus importants que les facteurs biotiques (Richardson & Pyšek, 2006). Enfin l'augmentation des ressources peut se produire lors des années pluvieuses, en conséquence de l'eutrophisation du milieu ou en réponse à l'ouverture de la canopée (Davis et al., 2000). Un corollaire important à cette théorie est que l'invasibilité d'une communauté n'est pas un paramètre constant, mais une condition qui fluctue au cours du temps (Grime, 2001).

Enfin, la théorie de l'absence d'ennemis naturels énonce que l'absence ou la réduction de prédateurs et de pathogènes dans le milieu d'introduction favoriserait le succès de l'invasion (Elton, 1958; Keane and Crawley, 2002). Cette théorie se base sur trois points essentiels : tout d'abord les pathogènes et prédateurs sont importants pour la régulation des populations, deuxièmement les pathogènes auraient un impact supérieur moindre sur les plantes introduites et enfin les plantes seraient capables de bénéficier de cette réduction dans la prédation pour se développer davantage (Keane and Crawley, 2002). Les résultats sont contrastés, mais il semble que dans certains

cas les dégâts des pathogènes seraient réduits sur les invasives comparativement aux plantes natives ce qui pourrait jouer un rôle dans l'invasion (Keane and Crawley, 2002). Par exemple, il a été montré que les populations invasives de *Silene latifolia* étaient significativement bien moins impactées par les pathogènes et les prédateurs que les populations de l'aire native (Wolfe, 2002). Cette hypothèse testée sur arbres invasifs a été observée empiriquement de nombreuses fois (Lamarque et al., 2011).

Plusieurs théories tentent d'expliquer les différences d'invasibilité entre habitats et aucune théorie n'offre cependant de réponse universelle : il semble plutôt que plusieurs théories peuvent cohabiter (Lamarque et al., 2011). De plus, leur importance relative peut varier selon le contexte et interagir avec la pression de propagules, mais également avec les caractéristiques des espèces introduites. Nous allons donc maintenant discuter des capacités intrinsèques des espèces à devenir invasives.

Le potentiel invasif d'une espèce est défini comme la capacité d'une espèce à s'établir et se disperser dans un nouvel habitat. Plusieurs théories ont été proposées pour tenter d'expliquer le succès d'une espèce donnée par rapport à une autre.

Parmi les différentes familles, tous les taxons n'ont pas le même potentiel pour envahir un nouveau milieu (Richardson & Pyšek, 2006). Ainsi à l'échelle de la famille, les *Fabaceae*, les *Asteraceae* ou les *Poaceae* sont souvent des invasives efficaces (Wu et al., 2004). Parmi les ligneux, les *Salicaceae* et les *Rosaceae* sont particulièrement bien représentés. Une étude a mis en évidence que le potentiel invasif était lié à la taxonomie, en montrant qu'en Nouvelle-Zélande, sur 36 espèces de gymnospermes invasives, 28 sont des *Pinaceae* et 21 appartiennent au genre *Pinus* (Richardson & Rejmánek, 2004).

Une meilleure aptitude à la compétition pourrait expliquer la capacité de certaines plantes exotiques à devenir invasives (Gioria & Osborne, 2014). Certains auteurs ont pourtant trouvé que les plantes invasives n'étaient pas significativement de meilleures compétitrices, mais plutôt que la performance relative dépendait des conditions de croissance (Daehler, 2003). Cependant les études sur les arbres invasifs semblent montrer que le trait commun à tous les arbres invasifs est une forte capacité de croissance (Lamarque et al., 2011). D'autres études ont apporté un support au moins partiel à la théorie de la réallocation des ressources. Les espèces exotiques libérées de leurs prédateurs naturels pourraient perdre certains traits leur conférant une résistance ou une protection face aux ravageurs et pathogènes et réallouer les ressources à d'autres fonctions comme la reproduction ou la croissance (Richardson & Pyšek, 2006). Ainsi par exemple, *Ligustrum robustum* produit cinq fois plus de fruits par mètre carré de feuillage dans son aire d'introduction à la Réunion que dans son aire naturelle au Sri Lanka. De plus, il

surpasse dans la compétition les espèces natives en termes de croissance et de taux de survie (Lavergne et al., 1999). Ces changements de comportement ou de valeurs de traits en réponse aux changements environnementaux rencontrés dans la nouvelle aire reposent largement sur la mise en œuvre d'une plasticité phénotypique ; je développerai ce point dans la partie 1.4.

Pour aller plus loin, la théorie des « nouvelles armes » (novel weapons) énonce que les espèces invasives pourraient également tirer parti de phénomènes allélopathiques pour surpasser les espèces locales (Callaway and Ridenour, 2004). Ainsi les composés extraits d'*Ailanthus altissima* affecteraient la végétation sur place par allélopathie en limitant le développement des plantules et diminueraient les attaques de ravageurs par leur effet insecticide (Kowarik & Säumel, 2007). *Centaurea diffusa* bénéficierait des effets allélopathiques de ses exsudats racinaires pour être invasif dans son nouvel environnement où la flore n'est pas adaptée à ces molécules, contrairement à celle de son aire native (Hierro et al., 2005). Il a été montré que *Centaurea maculosa* émet des exsudats racinaires allélopathiques qui impactent davantage les espèces voisines de la zone d'introduction que celles de l'aire native en inhibant leur croissance et leur germination (Bais, 2003).

Pour finir, les lois de Baker (Baker, 1967; Cheptou, 2012) énoncent que les espèces pratiquant davantage la reproduction asexuée seraient de meilleurs invasives. Il est probable que le mode de reproduction influence le potentiel invasif d'une espèce et des changements de système de reproduction ont été observés entre les aires natives et les aires d'introduction pour plusieurs espèces invasives (Barrett et al., 2008; Petanidou et al., 2012; Rambuda and Johnson, 2004). Des changements d'un mode de reproduction sexué vers l'asexué ont été observés parmi plusieurs espèces invasives telles *Eichhornia crassipes* (Barrett et al., 2008) ou *Fallopia japonica* (Hollingsworth and Bailey, 2000). Cependant ce n'est pas systématique et certaines invasives très fructueuses pratiquent exclusivement la reproduction sexuée comme *A. artemisiifolia* (Friedman and Barrett, 2008). De plus, des modèles théoriques ont mis en évidence que le potentiel invasif ne dépendrait pas de façon monotone du taux de reproduction asexuée et qu'il existerait un compromis entre le taux de reproduction sexuée et asexuée idéal. Les espèces pratiquant la reproduction asexuée et qui conserveraient une faible part de reproduction sexuée auraient le meilleur potentiel invasif (Bazin et al., 2013). En effet, la reproduction clonale assure la multiplication dans la nouvelle aire sans que les individus n'aient nécessairement besoin de trouver un partenaire ou un pollinisateur (Barrett et al., 2008) alors que la reproduction sexuée favorisait la dispersion longue distance ou la survie à l'hiver (Barrett et al., 2008; Bazin et al., 2013).

Pour conclure cette partie, plusieurs théories apportent des éléments pour mieux répondre à la question : comment une espèce devient-elle invasive ? Il est toutefois illusoire de vouloir trouver une réponse universelle à cette question car c'est de l'interaction de plusieurs facteurs que naissent les invasions ; certaines dépendant de l'histoire de l'introduction, d'autres de l'habitat envahi mais aussi des espèces introduites. Concernant les caractéristiques de l'espèce, je n'ai pas encore évoqué le rôle des mécanismes évolutifs au cœur de mon travail de recherche durant cette thèse et que je présenterai plus en détail dans les prochaines parties.

### 1.3. A la lumière de l'évolution

*Mr. McGuire: I want to say one word to you. Just one word.*  
*Benjamin: Yes, sir.*  
*Mr. McGuire: Are you listening?*  
*Benjamin: Yes, I am.*  
*Mr. McGuire: Plastics.*  
*Benjamin: Exactly how do you mean?*  
*Mr. McGuire: There's a great future in plastics. Think about it. Will you think about it?*  
*The Graduate, 1967*

#### 1.3.1. Du génotype au phénotype

D'un point de vue théorique, le phénotype d'un individu dépend à la fois de son génotype et de son environnement. Nous entendons l'environnement au sens large, c'est-à-dire tout ce qui n'est génétique : le climat, le sol, les agents biotiques, les traitements appliqués, la pente ... Pour aller plus loin, il faut également considérer que deux génotypes différents peuvent réagir différemment à la même variation environnementale, nous parlons alors d'interaction génotype environnement. Cela peut se résumer dans l'équation conceptuelle (Falconer and Mackay, 1996):

$$P = G + E + G \times E \quad (1)$$

Où P représente le phénotype, G le génotype et  $G \times E$  représente l'interaction génotype environnement.

Quels sont les conséquences de cette équation conceptuelle dans la nature?, Nous allons nous interroger dans le paragraphe suivant sur les forces à l'origine de la variabilité de génotypes que l'on trouve en populations naturelles ; nous retrouvons également une multitude d'environnements différents selon le fait que l'on se place à l'échelle globale, le biome par exemple, ou locale comme le site expérimental, voire le pot de culture. Il en résulte que les traits phénotypiques quantitatifs sont variables. Cette variabilité environnementale et génétique implique que dans une population donnée, pour un trait considéré, nous pouvons estimer une moyenne et une dispersion



autour de cette moyenne. De plus pour dénouer les effets G, E et  $\times$  E il va falloir concevoir des dispositifs expérimentaux spécifiques (Zenni et al., 2014).

### 1.3.2. Comment modifie-t-on son phénotype ?

Nous allons revenir sur la notion d'évolution avant de faire le lien entre ce concept et l'écologie des invasions biologiques.

En écologie évolutive, certains traits sont plus importants que d'autres : les traits liés à la *fitness*<sup>2</sup>. Ce sont les traits qui influencent directement les taux de natalité et de mortalité des populations d'organismes étudiés et donc leur succès reproductif, leur distribution et leur abondance. Ces taux sont eux-mêmes dépendants des interactions biotiques et abiotiques dans l'environnement. Les espèces du genre *Arabidopsis* sont des plantes modèles qui ont permis de démontrer relativement facilement ces relations, par ailleurs compliquées à mettre en évidence sur des organismes avec une durée de vie plus longue. Ainsi les populations norvégiennes d'*Arabidopsis lyrata* présentent un meilleur taux de croissance dans les environnements du Nord de l'Europe associé à la production d'un plus grand nombre de graines, assurant ainsi une fitness supérieure à celle d'autres populations de la même espèce dans ces environnements (Leinonen et al., 2009).

Toutefois, un trait favorable dans un environnement donné peut être défavorable dans un autre environnement. Par exemple, la couleur de la peau chez les humains est fortement corrélée à la latitude et au niveau d'irradiation aux rayonnements ultraviolets (UV). Aux basses latitudes, le phénotype couleur de peau plus foncée prévient la dégradation du folate, une molécule essentielle pour la formation du fœtus, par le rayonnement ultraviolet, tandis qu'aux hautes latitudes, où l'intensité du rayonnement UV est moindre, une peau plus claire permet un meilleur rendement pour la synthèse de vitamine D3, elle aussi indispensable pour la formation de l'embryon (Jablonski and Chaplin, 2000). Des phénotypes différents favorisent donc le succès reproductif des individus en fonction de la latitude où ils vivent en raison de la pression causée par les rayonnements UV, ce qui a induit une sélection divergente et a favorisé l'adaptation locale des populations. Ainsi, de nombreuses études ont démontré l'existence de différenciations phénotypiques entre populations d'arbres correspondant à une adaptation locale, notamment entre phénologie foliaire et température (Alberto et al., 2013). D'un point de vue quantitatif, les valeurs moyennes et les variances ont ainsi tendance à différer entre populations. D'autre part,

---

<sup>2</sup> Le concept de *fitness* sera ensuite traduit en « succès reproductif ». Il s'agit d'un concept fondamental en biologie évolutive et le terme anglais est empreint d'un sens riche difficile à traduire.

derrière le phénotype, il y a un génotype, ainsi l'évolution peut être vue comme un changement des fréquences alléliques entre populations (Conner and Hartl, 2004). Mais quels sont les mécanismes pouvant créer la variabilité existante et expliquer la différenciation des populations ?

Ces mécanismes sont au nombre de quatre : les forces évolutives recouvrent la mutation, la migration, la dérive génétique et la sélection. Ces quatre forces interagissent mais vont varier dans leur intensité, dans leurs effets et dans leur temporalité. Mutation et dérive génétiques sont deux processus aléatoires, qui vont avoir tendance à accroître les différenciations entre populations, par apparition ou disparition d'allèles. La migration correspond à un flux (d'individus ou de gènes) entre populations qui va donc avoir plutôt pour conséquence d'harmoniser les fréquences alléliques et donc de réduire les différenciations entre populations (Falconer and Mackay, 1996). Enfin, la dernière force évolutive est la sélection. C'est une force évolutive complexe car ses effets peuvent contribuer à différencier des populations ou au contraire à les rapprocher. La sélection repose sur le fait que dans un environnement donné certains traits phénotypiques vont être plus avantageux pour un individu et donc favoriser son succès reproductif. Les phénotypes dépendant en partie du génotype, certains génotypes vont donc être favorisés et la sélection va donc contribuer à modifier les fréquences alléliques au sein d'une population. Si les mêmes traits sont sélectionnés dans le même sens dans plusieurs populations, la sélection va être directionnelle et elle va avoir pour conséquence de rapprocher les populations en termes de fréquence allélique. Inversement si des traits différents sont sélectionnés dans différentes populations ou si le même trait est sélectionné dans des directions différentes entre populations, la sélection va avoir tendance à différencier les populations. La sélection évolutive est la seule force évolutive qui puisse conduire à de l'adaptation locale. L'adaptation locale est en quelque sorte « l'avantage de jouer à domicile ». On teste l'adaptation locale lorsqu'en comparant des individus provenant de plusieurs populations plantées (ou élevées) dans différents environnements, ce sont à chaque fois les populations locales qui ont la meilleure fitness dans leur environnement d'origine. Une méta-analyse portant sur l'utilisation de dispositifs de transplantations réciproques a permis de mettre en évidence que dans 70% des cas étudiés, les populations de plantes herbacées présentaient une adaptation locale (Leimu and Fischer, 2008).

Au sens strict, la plasticité phénotypique est définie comme la capacité d'un génotype à produire différents phénotypes en fonction de l'environnement (Ghalambor et al., 2007; Pigliucci, 2006). Elle dépend donc du trait étudié, de l'individu et de l'environnement. La définition de la plasticité phénotypique peut être étendue de l'échelle de l'individu à celle d'individus apparentés (frères, demi frères), de la population ou encore de l'espèce

(Davidson et al., 2011; Richards et al., 2006; Valladares et al., 2006). Le lien entre la plasticité phénotypique et l'évolution et *a fortiori* l'adaptation locale n'est pas évident. La plasticité peut être adaptative, si les individus les plus plastiques sont ceux qui ont le meilleur succès reproductif et sont ainsi sélectionnés ; il faut également qu'il y ait une base génétique c'est-à-dire l'existence d'une interaction  $G \times E$  ; enfin, la plasticité peut être adaptative mais elle peut également être mal-adaptative ou freiner l'évolution (Ghalambor et al., 2007). Si les phénotypes induits par la plasticité phénotypique sont optimaux dans chaque environnement alors la sélection ne contribuera pas à différencier les populations en fonction de leur habitat : au niveau de la population, nous n'observerions pas d'effet génétique ou d'interaction  $G \times E$  mais uniquement un effet environnemental. C'est par exemple le cas de l'herbacée *Pennisetum setaceum* à Hawaï dont les populations présentent une forte divergence phénotypique, tant morphologique que physiologique, mais la mise en place de test de transplantations réciproques a montré qu'il s'agissait uniquement d'un effet environnement de plasticité phénotypique et non de divergence génétique (Williams et al., 1995). D'un autre côté, si la plasticité modifie le phénotype dans le même sens que l'action de la sélection sans que le phénotype soit optimal, elle permet de faciliter l'évolution. Une étude sur les guppies de Trinidad, *Poecilia reticulata*, a ainsi mis en évidence que les descendants des individus situés en amont des cours d'eau étaient moins nombreux mais plus grands ; or il existe un gradient de prédation et de ressources décroissant de l'aval vers l'amont ; des tests de transplantation réciproque ont permis de démontrer qu'à la fois la plasticité phénotypique et la divergence génétique entre populations expliquaient ces caractéristiques, qui confèrent un avantage compétitif et adaptatif aux individus les plus grands en conditions environnementales limitantes (Ghalambor et al., 2007).

En conclusion, de récents débats proposent une voie vers une théorie synthétique de l'évolution expliquant notamment le lien entre plasticité phénotypique et évolution (Müller, 2017). La plasticité pourrait faciliter les transitions écologiques (Price, 2006). Davantage d'études empiriques sont nécessaires pour appréhender à la fois le rôle de l'évolution et de la plasticité phénotypique notamment dans le cas des espèces invasives. Mais quels sont les outils à notre disposition pour étudier ces phénomènes ?

## 1.4. Mettre en évidence l'évolution : du laboratoire à l'expérimentation

*It doesn't matter how beautiful your theory is,  
it doesn't matter how smart you are.  
If it doesn't agree with experiment, it's wrong.  
Richard P. Feynman*

Les forces évolutives modifient les fréquences alléliques entre populations et contribuent ainsi à homogénéiser ou au contraire à différencier génétiquement les populations entre elles. Ces modifications des fréquences alléliques peuvent se traduire également par des modifications du phénotype moyen qui résulte de l'action conjointe de l'environnement et du génotype. De quelles méthodes disposons-nous pour appréhender cette différenciation entre populations ? Nous allons développer le rôle de la génétique des populations et de la génétique quantitative afin d'étudier ces phénomènes.

### 1.4.1. La génétique des populations pour déterminer la structure des populations

L'essence de la génétique des populations est la variabilité génétique entre individus, c'est-à-dire le matériau brut sur lequel l'évolution peut agir (Conner and Hartl, 2004). Savoir si cette variabilité se répartit entre populations ou au sein des populations devient une question fondamentale. La première étape nécessaire est de pouvoir accéder à l'information moléculaire. Les progrès technologiques ont été considérables dans le domaine des marqueurs moléculaires au cours des dernières décennies avec comme tendance générale la diminution des coûts et l'augmentation du débit, c'est-à-dire du nombre de bases pouvant être lues. L'idée ici n'est pas de faire une revue complète des méthodes de génotypage mais plutôt de comprendre quelles ont été les avancées cruciales dans ce domaine et les motivations de ces progrès. Les premières méthodes dans les années 1960 employaient des migrations de protéines par électrophorèse sur un gel, ces protéines étaient souvent des enzymes couramment appelées isozymes ou allozymes. Ces techniques relativement peu coûteuses permettaient d'obtenir des informations importantes comme la distinction des individus homozygotes et des hétérozygotes, mais un inconvénient majeur était que les protéines résultant de l'expression d'un gène ne révèlent des différences que dans les régions codantes du génome qui sont pourtant minoritaires (Conner and Hartl, 2004; Falque, 2011). La mise au point de la méthode de PCR (*Polymerase Chain Reaction*) dans les années 1980 (Saiki et al., 1985) a révolutionné la génétique moléculaire moderne : il est alors devenu possible d'amplifier l'ADN, et donc malgré de faibles concentrations, d'étudier son polymorphisme. Schématiquement, le polymorphisme de l'ADN est dû à deux types de variation entre individus : soit un locus donné uniquement est substitué, on parle alors de *Single Nucleotide Polymorphism* (SNP) ; soit des séquences plus ou moins longues d'une à plusieurs mégabases sont

insérées ou supprimées, il s'agit d'*Insertion Deletion Polymorphism* (IDP) (Falque, 2011). Parmi les IDP, mentionnons le cas particulier des microsatellites qui consistent en une répétition d'un motif d'une à quatre bases et dont le nombre de répétitions peut varier fortement entre individus.

Depuis les années 2000, de nouvelles méthodes se développent rapidement appelées les *Next Generation Sequencing* (NGS). Toutefois pour utiliser les NGS à des fins de génotypage il est nécessaire d'obtenir une bonne couverture du génome (Falque, 2011), de réaliser un gros travail bioinformatique pour réassembler les séquences. Ce travail est facilité pour les espèces modèles dont le génome a été séquencé entièrement (Falque, 2011; Metzker, 2010). D'autre part, même si les NGS permettent d'avoir accès à une quantité d'information très importante, leur coût est plus élevé que pour un génotypage avec des marqueurs SNP ou microsatellites dont l'automatisation récente a permis de réduire les coûts.

Les microsatellites et les SNP sont des marqueurs fréquemment employés pour génotyper des espèces non modèles (Morin et al., 2004; Coates et al., 2009; Helyar et al., 2011). Les SNP présentent de nombreux avantages car ils sont répartis en grand nombre le long du génome et peuvent être relativement facilement développés, de plus le génotypage est facilement répliquable entre laboratoires (Morin et al., 2004; Coates et al., 2009; Helyar et al., 2011). Il est toutefois nécessaire de disposer de davantage de marqueurs SNP comparativement aux microsatellites pour obtenir le même degré de précision, principalement parce que les microsatellites sont multialléliques alors que les SNP sont en majorité bialléliques (Morin et al., 2004).

Toutefois un des problèmes des SNP, lorsqu'on les utilise pour caractériser la diversité génétique est le biais de recrutement (Morin et al., 2004; Coates et al., 2009; Helyar et al., 2011). En effet, les SNP avec une fréquence allélique de l'allèle rare élevée (ou *Minor Allelic Frequency* – MAF –) sont davantage susceptibles d'être échantillonnés alors qu'ils sont théoriquement moins abondants que ceux avec une MAF faible, ce qui biaise la distribution de la MAF dans les populations génotypées et altère l'estimation de la diversité génétique (Helyar et al., 2011). Pour pallier ce problème il est nécessaire d'augmenter le nombre de marqueurs utilisés pour éviter les conclusions trompeuses basées sur des faux positifs ou des faux négatifs (Morin et al., 2004; Helyar et al., 2011). De façon empirique, certaines études comparant ces deux types de marqueurs n'ont pas établi de différence entre les approches (van Inghelandt et al., 2010; Filippi et al., 2015). Cependant, il a aussi été montré que les SNP étaient plus performants que les microsatellites pour étudier la structure des populations alors que les microsatellites surpassaient les SNP dans l'estimation de la diversité génétique. (Singh et al., 2013).

Le génotypage des individus échantillonnés nous permet de déterminer si l'individu est homozygote dominant ou récessif ou encore hétérozygote. Tout l'enjeu va être ensuite d'analyser la variabilité génétique observée et de déterminer sa structuration parmi les populations, en analysant à quel niveau elle se situe : est-elle répartie entre individus au sein d'une population, entre populations ? Différentes méthodes d'analyse existent. D'une façon générale, elles peuvent être faites au niveau de l'individu ou de la population. A l'échelle de la population, nous allons pouvoir estimer plusieurs paramètres : la fréquence allélique, les taux d'hétérozygotie théoriques ou observés, la richesse allélique. Ces indices nous renseignent sur la diversité génétique de chaque population et leur comparaison entre populations va permettre de comprendre les événements qui ont contribué à la différenciation des populations. Par exemple, une diminution du nombre d'allèles ou du taux d'hétérozygotie peut indiquer un goulot d'étranglement (diminution importante de la taille de la population) ou un effet de fondation (formation d'une population à partir d'un petit nombre d'individus).

Pour aller plus loin que ces estimateurs de la diversité génétique, Sewall Wright, un des fondateurs de la génétique des populations, a développé les F-Statistiques (Wright, 1951). L'objectif de ces statistiques est de décrire la répartition de la variabilité génétique d'une espèce en trois niveaux hiérarchiques : variation au sein des populations, entre populations et variation totale. Trois indices sont ainsi calculés le  $F_{IS}$ , le  $F_{ST}$  et le  $F_{IT}$ . Nous allons décrire ici en quelques lignes le  $F_{ST}$ , indice qui permet de renseigner sur le degré de différenciation entre populations. Plusieurs façons de le calculer existent mais une façon couramment admise est celle de Weir et Cockerham (1984) qui se base sur un rapport de variance :

$$F_{ST} = \frac{\sigma_B^2}{(\sigma_B^2 + \sigma_W^2)} \quad (2)$$

ou  $\sigma_B^2$  (B pour between) représente la variance génétique entre populations et  $\sigma_W^2$  (W pour within) représente la variance génétique au sein des populations (Weir and Cockerham, 1984). La somme de  $\sigma_B^2$  et de  $\sigma_W^2$  représente la variance génétique totale. Ainsi ce ratio nous renseigne bien sur la part de la variabilité génétique totale due à la variabilité génétique entre populations. Comme le  $F_{ST}$  est calculé à partir de marqueurs moléculaires neutres, il renseigne sur la différenciation entre populations due aux forces évolutives neutres (migration, dérive génétique et mutation) et ne prend pas en compte l'effet de la sélection. D'après Wright, une valeur de  $F_{ST}$  comprise entre 0 et 0.05 suggère une différenciation entre populations faible, entre 0.05 et 0.15 le  $F_{ST}$  traduit une différenciation modérée, entre 0.15 et 0.25 il témoigne d'une différenciation importante et au-delà de 0.25, le  $F_{ST}$  illustre une différenciation très importante. Pour parachever les analyses de génétique au niveau des populations, différentes méthodes basées sur des analyses à l'échelle de l'individu existent, telles celles implémentées par le logiciel STRUCTURE (Pritchard et al., 2000), BAPS (Corander et al., 2008; Corander and Marttinen, 2006), TESS (Chen

et al., 2007; François et al., 2006), GENELAND (Guillot et al., 2005) ou une analyse discriminante en composante principale (DAPC) intégrée au package R ADEGENET (Jombart, 2008; Jombart et al., 2010). Selon le logiciel utilisé, il est possible d'inclure ou non l'information de la localisation géographique des populations. Nous avons employé le logiciel STRUCTURE pour nos analyses, cette méthode permet d'attribuer un coefficient d'appartenance de chaque individu aux K groupes génétiques définis dans l'analyse, K étant un paramètre que nous pouvons faire varier. Ainsi un individu totalement attribué à un groupe aura un coefficient d'appartenance de 100% pour ce groupe, ce coefficient pouvant quantifier une appartenance partielle à plusieurs groupes (individu dit « *admixed* »). Il est ainsi possible d'identifier le nombre de groupes génétiques parmi l'ensemble des individus et de comparer la structuration des individus sur critère génétique et leur structuration en population sur critère géographique. Ce type d'analyses permet de discerner des structures génétiques intéressantes entre populations ou de retrouver les populations sources. Dans le cas d'introduction d'espèces invasives, l'utilisation des marqueurs moléculaires pour déterminer l'histoire de l'introduction et la structure des populations est encore récente (Cristescu, 2015; Estoup and Guillemaud, 2010; Fitzpatrick et al., 2012). Toutefois, un nombre croissant d'études utilisent les apports des marqueurs moléculaires pour retracer l'histoire de l'invasion, chez les arbres nous pouvons citer les exemples de : *Acacia saligna* (Thompson et al., 2015), *Olea europaea* (Besnard et al., 2014), *Prunus serotina* (Pairon et al., 2010) et *Quercus rubra* (Merceron et al., 2017).

#### **1.4.2. Comprendre la variabilité phénotypique grâce à la génétique quantitative**

Pour étudier la variabilité phénotypique à l'aulne de l'évolution, nous disposons de la génétique quantitative. La génétique quantitative s'intéresse à la variation de caractères continus : taille, masse, taux de croissance, concentration en tanins etc. pour évaluer leurs bases génétiques et environnementales (Beilharz et al., 1993). La relation est souvent complexe entre le phénotype et le génotype et plusieurs gènes à plusieurs loci sont souvent impliqués tandis que l'environnement influence également le phénotype. L'hypothèse fondamentale de la génétique quantitative est que l'hérédité des différences quantitatives dépend de gènes qui sont soumis aux mêmes lois de transmission et ont les mêmes propriétés générales que les gènes qui régissent les différences qualitatives : la génétique quantitative est donc une extension de la génétique mendélienne. Cependant les méthodes d'étude diffèrent. L'unité d'étude est la « population » et non l'individu, on s'intéresse ainsi à la variabilité entre individus au sein de la population. Les objectifs de la génétique quantitative ainsi sont, de quantifier statistiquement la variation totale d'un trait et sa répartition due au génotype et due à l'environnement, afin de prédire la ressemblance

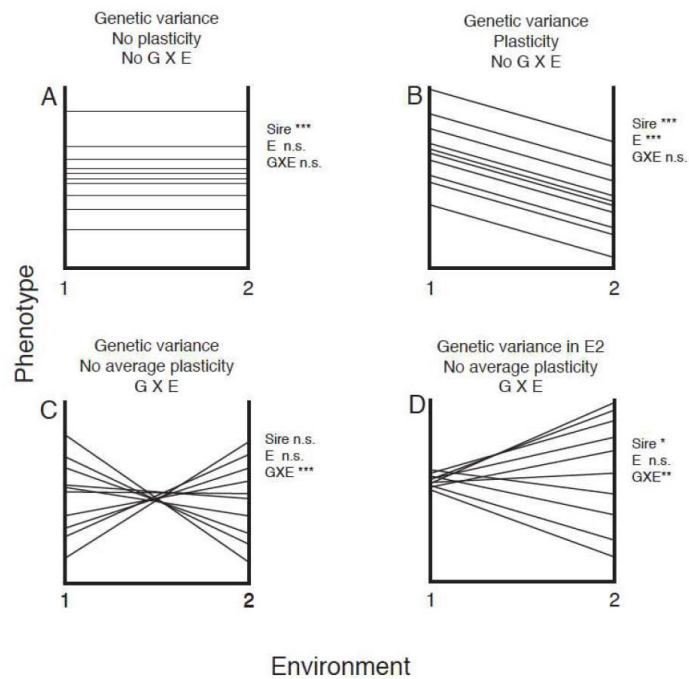
entre individus apparentés ou l'intensité de la réponse à la sélection ou d'estimer la différenciation entre populations à partir de l'analyse de traits phénotypiques continus (Conner and Hartl, 2004; Falconer and Mackay, 1996).

Le dispositif expérimental de base en génétique quantitative permettant d'évaluer ces parts de variation est l'utilisation de jardins communs, mais quel est son principe ?

Repartons de l'équation 1 :  $P = G + E + G \times E$ .

Si nous prenons deux clones différents que nous noterons 1 et 2 (mais cela peut également se faire à l'échelle de la famille si il est difficile en pratique d'obtenir des clones) et que nous les plantons dans un même environnement (Fig 1 A) alors les différences phénotypiques observées entre clones ne seront dues uniquement qu'aux effets des deux génotypes du fait que que l'environnement est commun aux deux clones: En utilisant un second environnement, il est possible de tester si l'environnement influence le phénotype. Si pour un clone donné, la valeur d'un trait est la même quel que soit l'environnement (Fig 1 A) alors l'effet de l'environnement n'est pas significatif. Si au contraire la valeur du trait est modifiée (Fig 1 B) il y a un effet significatif de l'environnement on parle alors de plasticité phénotypique. Si l'on mesure par exemple le taux de croissance d'un arbre dans deux environnements chaud et froid et que le taux de croissance est supérieur dans l'environnement chaud nous dirions que le trait phénotypique « taux de croissance » est plastique à la température. Enfin l'intérêt d'utiliser plusieurs génotypes dans plusieurs environnements est de pouvoir comparer les réponses des génotypes à la même variation environnementale, si les génotypes ne répondent pas de la même façon à la même variation environnementale, il existe alors une interaction génétique x environnement (Fig 1 C et D). Dans notre exemple précédent, ce serait le cas si nous plantions deux clones différents d'une même espèce d'arbre, et qu'un des clones voyait son taux de croissance augmenter avec l'augmentation de température alors que l'autre voyait son taux de croissance diminuer.





**Fig 1** Normes de réaction pour quatre situations théoriques. La valeur moyenne du trait mesuré pour chaque génotype (ou famille) est indiquée sur l'axe des ordonnées tandis que l'axe des abscisses représente les deux environnements utilisés. Il est ainsi possible de tester s'il existe un effet génétique dans la variabilité phénotypique observée (A, B, C, D), un effet de l'environnement (B, C, D) correspondant à de la plasticité phénotypique. Lorsque les pentes des normes de réaction diffèrent entre génotypes (C et D), une interaction génotype – environnement est mise en évidence. (D'après Conner & Hartl, 2004).

Ainsi le jardin commun est le dispositif expérimental de base en génétique quantitative et sa réplication sur plusieurs environnements permet de séparer effets génétiques, environnementaux et d'interactions. En pratique les plans expérimentaux sont plus complexes que ceux présentés ci-dessus pour illustrer le principe de fonctionnement.

L'analyse des données de jardins communs correspond à une analyse statistique de la variance du trait et l'équation 1 se transforme en :

$$V_P = V_G + V_E + V_{G \times E} + 2 \text{COV}_{GE} \quad (3)$$

Où  $V_P$  représente la variance Phénotypique,  $V_G$  la variance Génétique,  $V_E$  la variance Environnementale,  $V_{G \times E}$  : la variance liée à l'interaction Génétique x Environnement et  $\text{COV}_{GE}$  la Covariance Génétique Environnement. Ce terme de covariance correspond à une corrélation entre l'effet environnemental et l'effet génétique qui ne sont

alors plus indépendants de l'autre. C'est le cas par exemple si nous apportons plus d'eau aux plantes qui poussent le mieux. Dans un dispositif expérimental, il est possible d'annuler ce terme de l'expression en répartissant aléatoirement les individus dans le dispositif et en standardisant les conditions environnementales au sein de chaque environnement.

La variance environnementale inclut par définition toute la variabilité phénotypique d'origine non génétique, liée à la variabilité environnementale (nutriments, climat etc.) mais également les effets maternels, qui sont susceptibles de créer de la ressemblance entre individus apparentés. Dans le cas de graines prélevées sur des arbres en milieu naturel, certains arbres-mères sont ainsi susceptibles d'avoir bénéficié de conditions plus favorables (meilleure topographie ou exposition par exemple) et par conséquent d'avoir fourni davantage de ressources aux graines qu'ils ont produites. La ressemblance entre individus ne serait pas uniquement liée alors à un effet génétique, ce qui peut biaiser l'estimation de la variance génétique (Whitlock, 2008). Pour retirer les effets maternels, il faudrait récolter les graines sur arbres-mères cultivés en jardin commun mais un tel dispositif est difficilement envisageable pour des arbres, espèces à longue durée de vie. Un test réalisé sur l'herbacée invasive *Senecio inaequidens* comparant les traits d'individus issus de graines récoltées en populations naturelles ou produites en jardin commun n'a pas cependant mis en évidence d'effet maternel, malgré une variabilité observée du poids des graines (Monty et al., 2009).

Pour aller plus loin dans la compréhension de la répartition de la variabilité entre individus, il est nécessaire de décomposer également les termes de la variance. Celle-ci se décompose en :

$$V_G = V_A + V_D + V_I \quad (4)$$

Avec  $V_A$  la Variance génétique Additive c'est-à-dire la part héréditaire de la variance transmise directement par les allèles lors de la reproduction,  $V_D$  la Variance génétique de Dominance c'est-à-dire causée par l'interaction entre allèles d'un même locus et  $V_I$  la Variance génétique d'Epistasie causée par l'interaction entre allèles de différents locus.

En génétique quantitative, la source de variation la plus importante est la variance additive  $V_A$  car les individus transmettant une seule version de leurs allèles à leur descendance, lors de la reproduction sexuée, cela va déterminer le changement du phénotype moyen de la population entre les générations. Pour pouvoir estimer correctement  $V_A$ , on étudie la ressemblance de traits entre individus apparentés ; l'utilisation de familles

composées de demi-frères plutôt que de plein frères facilite l'estimation de cette variance additive, et permet de limiter le biais dans l'estimation du aux effets de dominance et d'épistasie (Leinonen et al., 2013; Whitlock, 2008).

En quantifiant la part de variance phénotypique due aux effets génétiques additifs, on calcule l'héritabilité au sens strict :

$$h^2 = \frac{V_A}{V_P} \quad (5)$$

L'héritabilité est adimensionnelle et comprise entre 0 et 1. Attention aux mauvaises compréhensions du concept : l'héritabilité fait référence à la variation dans la population (Visscher et al., 2008). Une héritabilité faible ou nulle ne signifie pas qu'un trait n'est pas affecté par le génotype. Cela signifie que par rapport à toute la variance phénotypique observée, une faible proportion est due à la variation génétique, que le phénotype n'est pas déterminé d'emblée par le génotype, mais modifié sous l'effet de l'environnement. De même, une héritabilité élevée ne signifie pas que l'environnement n'affecte pas un trait ou que toute la détermination du trait est purement génétique. Cela signifie que la plus grande partie de la variation observée dans la population est causée par la variation entre génotypes. Cela signifie également que dans ce type de population, le phénotype est un bon prédicteur du génotype. D'une façon générale, il semble que l'héritabilité a tendance à être plus faible dans un environnement stressant : par exemple chez des drosophiles, l'héritabilité de la largeur des ailes a été estimée à 0.09 dans un environnement stressant et à 0.69 dans un environnement non stressant (Conner et Hartl, 2004). Nous savons d'autre part, que les traits liés à la fitness ont tendance à avoir une héritabilité faible, une des explications possibles étant que la sélection réduit la variabilité génétique sur ce type de traits (Visscher et al., 2008).

Enfin un indice analogue au  $F_{ST}$  (équation 2) a été défini pour évaluer la différenciation génétique en se basant sur l'analyse phénotypique de traits quantitatifs : le  $Q_{ST}$  (Spitze, 1993) :

$$Q_{ST} = \frac{\sigma_B^2}{\sigma_B^2 + 2\sigma_W^2} \quad (6)$$

Où,  $\sigma_B^2$  représente la variance génétique du trait entre populations et  $\sigma_W^2$  représente la variance génétique au sein des populations.  $\sigma_W^2$  est dans le cadre du  $Q_{ST}$  équivalent à la variance additive  $V_A$ . Plus précisément, pour le  $Q_{ST}$ , sous l'hypothèse d'additivité des effets génétiques, il est possible d'estimer  $\sigma_W^2$  en fonction du degré d'apparentement entre individus d'une même famille (Le « ^ » indique une valeur estimée d'après des données) :

$$\hat{\sigma}_W^2 = \frac{1}{r} \hat{\sigma}_{family}^2 \quad (7)$$

Avec  $r$ , le coefficient d'apparement entre individus d'une même famille, et  $\hat{\sigma}_{family}^2$  la variance additive des familles dans une population (Gilbert and Whitlock, 2014; Leinonen et al., 2013). Le coefficient  $r$  vaut  $\frac{1}{4}$  pour des demi-frères et  $\frac{1}{2}$  pour des pleins frères (Gilbert and Whitlock, 2014; Santiso et al., 2015; Spitze, 1993). Ainsi, dans le cas de famille de demi-frères, le  $Q_{ST}$  est estimé comme suit : (Gilbert and Whitlock, 2014; Spitze, 1993)

$$\hat{Q}_{ST} = \frac{\hat{\sigma}_B^2}{\hat{\sigma}_B^2 + 8 \hat{\sigma}_{family}^2} \quad (8)$$

En supposant que tous les gènes agissant sur les traits phénotypiques étudiés ont des effets égaux, additifs et indépendants, la comparaison des deux estimateurs de différenciation génétique  $Q_{ST}$  et  $F_{ST}$  permet de faire des inférences évolutives sur le moteur des différences entre populations. Si le  $Q_{ST}$  est égal au  $F_{ST}$ , la dérive génétique neutre suffit à expliquer la divergence entre les populations sans l'intervention de la sélection naturelle (Edelaar et al., 2011; McKay and Latta, 2002; Merilä and Crnokrak, 2001). Par contre, si les traits quantitatifs sont soumis à une sélection divergente entre les populations, des phénotypes différents seraient donc favorisés entre les populations et le  $Q_{ST}$  devrait être supérieur au  $F_{ST}$ . Enfin une sélection stabilisante qui favoriserait les mêmes phénotypes dans les différentes populations se traduirait par un  $Q_{ST}$  inférieur au  $F_{ST}$  (Edelaar et al., 2011; Merilä and Crnokrak, 2001).

Une des difficultés dans la comparaison  $Q_{ST}$   $F_{ST}$  est d'obtenir un estimateur non biaisé et avec un intervalle de confiance pour le  $Q_{ST}$ . Un certain nombre de facteurs sont susceptibles de biaiser l'estimation du  $Q_{ST}$ : présence de variance de dominance ou d'épistasie, effets maternels. Il est difficile de maîtriser l'influence de ces sources de biais sur l'estimation, mais il semble qu'il soit plus probable qu'elles biaisent vers le bas le  $Q_{ST}$  rendant l'expérience plus conservative (Leinonen et al., 2013; Whitlock, 2008; Whitlock and Guillaume, 2009). Il a été montré que les approches bayésiennes d'estimation minimisent le biais comparativement aux méthodes fréquentistes (Leinonen et al., 2013; O'Hara and Merilä, 2005). D'autre part, afin de réduire l'incertitude de l'estimation, il est recommandé d'utiliser un minimum de vingt populations avec 10 familles par population et 5 individus par famille pour avoir une précision satisfaisante, ce qui correspond donc à des dispositifs expérimentaux lourds.

### 1.4.3. Génétique quantitative et plasticité phénotypique

Compte tenu des effets génétique, environnement et des interactions génétique x environnement, l'analyse de génétique quantitative requiert de travailler non dans un seul mais plusieurs jardins communs (Williams et al., 2008). L'analyse des traits entre environnements permet d'étudier la plasticité, et plusieurs indices permettent alors d'estimer quantitativement la plasticité. Le plus immédiat est la pente ou la forme de la courbe correspondant à la norme de réaction entre le trait et les environnements. Toutefois, il en existe de nombreux et Valladares et al (2006) en recensent 19. Ces indices sont pour la plupart corrélés entre eux (Valladares et al., 2006) et les auteurs proposent d'utiliser le *Relative Distance of Plasticity Index* (RDPI) car il est plus robuste pour réaliser des tests statistiques :

$$RDPI = \sum \frac{|X_{ij} - X_{i'j}|}{(X_{ij} + X_{i'j})} \times \frac{1}{n} \quad (9)$$

Où  $i$  fait référence à un environnement donné et  $i'$  à un autre environnement,  $j$  fait référence au génotype (ou à la famille),  $X_{ij}$  est donc la valeur du trait pour l'individu  $j$  dans l'environnement  $i$ .  $n$  correspond au nombre de distances entre environnement calculées (par exemple pour 2 environnements A et B, il y a 1 distance possible AB, mais pour 3 environnements A, B et C il y a 3 distances possibles AB, AC, BC). Ainsi on obtient une valeur de RDPI par individu et par trait quel que soit le nombre d'environnements (Godoy et al., 2012). Le RDPI est adimensionnel et compris entre 0 et 1, une valeur de 1 indiquant une plasticité maximale et celle de 0 l'absence de plasticité.

### 1.5. L'adaptation locale, une des raisons du succès des espèces invasives ?

*Ici, le calao se nourrit des fruits du strychnos, l'arbre à strychnine.  
Les noyaux sont l'un des plus forts poisons qui existent.  
L'oiseau mange la pulpe, jamais le noyau.  
L'adaptation est bien plus mystérieuse que la survie des plus aptes.  
André Malraux, Antimémoires, 1967*

Longtemps les études sur les invasions biologiques se sont concentrées sur les aspects écologiques et ce n'est que plus récemment que le lien entre invasion et évolution a commencé à émerger (Colautti and Lau, 2015). En effet, les invasions biologiques représentent des expériences évolutives naturelles, elles permettent ainsi d'étudier les processus évolutifs contemporains, quelques décennies ou siècles après leur introduction dans leur nouvelle aire (Colautti and Lau, 2015; Keller and Taylor, 2008; Lee, 2002; Prentis et al., 2008).

Quand une espèce est introduite dans une nouvelle aire, il est généralement attendu qu'elle subisse un goulot d'étranglement génétique conduisant à une réduction de la diversité génétique (richesse allélique ou taux d'hétérozygotie) entre l'aire native et d'introduction (Dlugosch et al., 2015). En cohérence avec cette hypothèse, la plante invasive *Heracleum mantegazzianum* présentait ainsi une diversité génétique plus faible dans son aire d'introduction ce qui confirmait l'effet de fondation (Henry et al., 2009). Toutefois, des méta analyses récentes (Dlugosch et al., 2015; Dlugosch and Parker, 2008) ont montré que la réduction de diversité dans l'aire d'introduction était souvent très modérée (de l'ordre de 15 – 20 % en moyenne). Une explication possible est que des introductions multiples de l'espèce invasive ont permis de limiter la réduction de diversité, comme par exemple pour *Phalaris arundinaceae* et *Prunus serotina* (Lavergne and Molofsky, 2007; Pairon et al., 2010). La diversité génétique peut même dans certain cas augmenter dans l'aire d'introduction en cas de brassage génétique entre populations géographiquement distantes dans leur aire d'origine et mises en contact dans l'aire d'introduction (Dlugosch et al., 2015; Dlugosch and Parker, 2008), bien que de fortes augmentations soient rares (Uller and Leimu, 2011). C'est le cas par exemple de l'invasion par *Phalaris arundinacea* qui a été favorisée par une augmentation de la variation génétique (Lavergne and Molofsky, 2007). De même, un brassage génétique entre populations a sûrement contribué au succès de l'invasive *Silene vulgaris* dans son aire d'introduction américaine (Keller et al., 2014; Keller and Taylor, 2010). Un mécanisme peut ainsi être l'apparition de nouveaux génotypes car des introductions multiples successives peuvent restaurer ou augmenter la diversité perdue par brassage génétique (Bock et al., 2014). La réduction de la variabilité génétique serait ainsi généralement limitée dans l'aire d'introduction ce qui est important évolutivement parlant. Une variance génétique additive suffisante est essentielle pour l'adaptation évolutive en réponse à un changement environnemental. Le succès d'une invasion pourrait donc être favorisé par la présence d'une variabilité génétique additive initiale dans la population source sur laquelle la sélection naturelle pourrait agir (Lee, 2002)

La sélection agirait au départ sur les capacités de dispersion ou sur la tolérance physiologique en réponse aux changements environnementaux rencontrés dans les habitats d'introduction. Suite à l'établissement de la population, l'adaptation consisterait dans la réponse aux pressions de sélection dans le nouvel habitat comme les gradients environnementaux (i.e. température, climat) ou les pressions biotiques exercées par les espèces rencontrées (Lee, 2002). Le succès d'une invasion biologique pourrait ainsi dépendre de la capacité des espèces invasives à s'adapter dans un sens évolutif à un nouvel environnement. D'une part, le processus d'adaptation repose sur la variabilité génétique préexistante dans l'aire d'origine de l'espèce ; d'autre part la variabilité dépendra

des mutations *de novo* c'est-à-dire depuis l'introduction. Le temps de latence précédant souvent le succès d'une invasion après l'introduction pourrait ainsi résulter du fait que ce temps serait nécessaire permettant à la population introduite d'atteindre un certain niveau de variance génétique plutôt qu'uniquement du nombre d'individus (Lee, 2002; Pappert et al., 2000). En effet, l'adaptation grâce à de nouvelles mutations est plus lente que l'adaptation liée à une variabilité génétique préexistante : le temps nécessaire pour que des mutations apparaissent et augmentent en fréquence dans la population dépend à la fois du système de reproduction et du nombre d'individus introduits (Bock et al., 2014). Le rôle des mutations pourrait être plus probable dans les populations d'espèces invasives, car les modèles d'adaptation de Fisher (1930) indiquent que de nouvelles mutations sont davantage susceptibles d'être bénéfiques dans des populations éloignées de leur optimum, ce qui est probablement le cas pour une nouvelle espèce envahissante (Bock et al., 2014). Cependant, dans le cas des arbres invasifs, l'âge de la maturité sexuelle est long comparativement aux espèces herbacées, aussi le nombre de générations écoulées depuis l'introduction est plus faible et cela pourrait diminuer le potentiel adaptatif et donc invasif (Zenni et al., 2016). Cela interroge sur la façon dont ces ligneux ont acquis leur potentiel invasif et pose plus globalement la question de la vitesse à laquelle l'évolution adaptative se fait en populations naturelles.

Plusieurs études ont ainsi souligné qu'une évolution récente dans l'aire d'introduction aurait favorisé l'invasion (Barrett et al., 2008; Blair and Wolfe, 2004; Colautti and Barrett, 2013; Dlugosch and Parker, 2008; Felker-Quinn et al., 2013; Keller and Taylor, 2008; Lee, 2002; Maron et al., 2004; Prentis et al., 2008; Whitney and Gabler, 2008). En effet, les différences environnementales entre les deux aires sont susceptibles de générer de fortes pressions de sélection. Dans leur étude sur *Hypericum perforatum*, plante invasive originaire d'Europe et introduite en Amérique du Nord à la fin du dix-huitième siècle, Maron et al. (2004) ont réalisé plusieurs expériences en jardin commun selon un gradient latitudinal en mêlant des populations de l'aire native et de l'aire d'introduction. À travers l'étude de traits phénotypiques, ils ont mis en évidence les mêmes clines latitudinaux parmi les populations originaires de l'aire native *versus* celles originaires de l'aire invasive. Cela peut être lié à deux processus : soit des introductions multiples à partir de sources génétiques diversifiées, dans ce cas les génotypes les plus efficaces persistent là où ils sont pré-adaptés, il n'y a pas d'évolution ; soit au contraire une adaptation évolutive. Grâce à l'examen de marqueurs moléculaires AFLP, il est apparu que beaucoup de populations avaient été fondées par des individus probablement non pré-adaptés à la latitude et aux conditions climatiques. Aussi les auteurs concluent que ces populations semblent avoir évolué rapidement, en 150 ans, en réponse aux conditions environnementales de leur nouvel habitat (Maron et al., 2004). Mais chez les plantes

invasives, une des démonstrations les plus abouties d'adaptation locale rapidement acquise dans la nouvelle aire provient de *Lythrum salicaria*. Lors d'une première étude portant sur des populations de l'aire d'introduction américaine, les auteurs ont mis en évidence un cline latitudinal de la période de floraison, ce cline étant similaire à celui observé entre les populations natives en Europe (Barrett et al., 2008). De plus, une étude utilisant des marqueurs moléculaires AFLP a montré que les populations natives et introduites étaient différenciées (Chun et al., 2009). Grâce à un test de transplantation réciproque consistant en trois jardins communs répartis le long d'un gradient latitudinal dans l'aire d'introduction, il a ensuite été montré que le trait, fortement relié à la fitness et au succès reproductif, « production de graines » avait une base génétique indiquant une différenciation entre populations. Cette différenciation a été attribuée à de l'adaptation locale, la population la plus performante étant toujours celle plantée dans son propre environnement. Cette adaptation locale serait survenue rapidement, en une période de l'ordre de la centaine d'années, dans l'aire d'introduction en réponse à une pression de sélection environnementale le long du gradient latitudinal (Barrett et al., 2008; Colautti and Barrett, 2013). Pour conclure, une méta-analyse récente a mis en évidence que l'adaptation locale dans l'aire d'introduction n'est pas rare parmi les plantes invasives (Oduor et al., 2016). Ainsi l'adaptation locale expliquerait une partie des divergences de traits phénotypiques observées entre populations de l'aire d'introduction et pourrait interagir avec la plasticité phénotypique qui est un autre mécanisme proposé pour expliquer le potentiel invasif d'une espèce.

Il a été souligné que la plasticité phénotypique pourrait favoriser le potentiel invasif des espèces (Agrawal, 2001; Davidson et al., 2011; Richards et al., 2006). De façon conceptuelle, elle semble un facteur clef pour permettre l'acclimatation des espèces exotiques introduites, lors du passage de l'aire native à l'aire d'introduction (Richardson and Pyšek, 2006). Aussi de nombreuses études ont porté sur la caractérisation de la plasticité des espèces invasives, soit en effectuant des comparaisons avec les espèces natives des communautés envahies, soit en comparant pour une espèce donnée la plasticité dans des populations issues de l'aire native et de l'aire invasive à l'aide de dispositifs de jardins communs. Ainsi les populations invasives de l'arbre *Melaleuca quinquenervia* présentent une plasticité supérieure aux populations natives en réponse à la variation de pH (Kaufman and Smouse, 2001). La méta-analyse réalisée par Davidson et al. (2011) sur la plasticité des plantes invasives conclut que l'augmentation des ressources disponibles induit une réponse plastique plus marquée chez les espèces invasives comparativement aux espèces natives. Ceci a été observé par exemple chez *Acer negundo*, un arbre américain invasif en Europe, dont la compétitivité par rapport aux arbres natifs de la ripisylve est liée à sa plus forte réponse plastique face à une augmentation simultanée de lumière et d'azote (Porté et al., 2011). Cependant, une plus grande



plasticité ne semble pas forcément associée à un gain de *fitness* (Davidson et al., 2011). De plus une méta-analyse et une comparaison multi espèces n'ont pas montré de différence de plasticité entre les espèces natives et invasives (Godoy et al., 2011; Palacio-López and Gianoli, 2011). Pour tester si la plasticité peut être adaptative ou non, des suivis multigénérationnels, pendant 11 ans, ont été réalisés sur la réponse de populations de l'herbacée invasive *Polygonum cespitosum* (Sultan et al., 2013). Les populations récentes sont plus plastiques en réponse à une augmentation de lumière ou de nutriments, mais elles ont aussi gagné en *fitness* (évaluée par les capacités reproductives), suggérant que la plasticité observée serait adaptative (Sultan et al., 2013). Récemment, plusieurs études ont cependant montré que la plasticité des espèces invasives pouvaient être non pas le résultat d'une adaptation post-introduction mais était un caractère pré-adaptatif, c'est-à-dire une caractéristique acquise dans l'aire d'origine avant l'introduction (Lamarque et al., 2013; van Kleunen et al., 2011b).

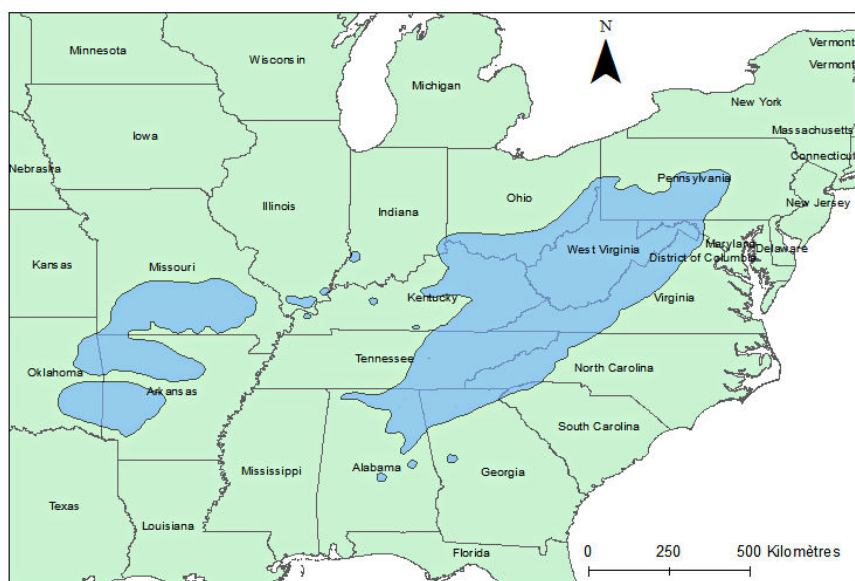
Pour conclure cette partie, l'étude des invasions biologiques sous un angle évolutif a permis de dévoiler que l'évolution rapide se produisait naturellement et pouvait conduire à de l'adaptation locale chez les espèces invasives expliquant leur succès.

## 1.6. De la dissémination du robinier dans les forêts européennes

*Et les bois toujours verts s'iront rajeunissant*  
Victor Hugo, *Soleils Couchants*, 1829

Mon travail de recherche portait sur l'histoire de l'introduction et le rôle de l'évolution lors de l'invasion par le robinier faux acacia : *Robinia pseudoacacia* en Europe.

Le robinier faux-acacia est un arbre pionnier de la famille des Fabacées, pouvant atteindre 30 mètres de haut, qui se reproduit à la fois de façon sexuée et asexuée par drageonnage (Cierjacks et al., 2013). Ses graines présentent une forte dormance physique due à leurs téguments résistants et imperméables. C'est un arbre originaire d'Amérique du Nord où il occupe une aire disjointe (Fig 2) dans l'est des États-Unis (Little, 1971). La plus grande partie de son aire de répartition, à l'est, correspond au plateau des Appalaches et chevauche onze États tandis que la plus petite à l'ouest s'étend sur les monts Ozark et englobe partiellement trois États.



**Fig 2** Aire de répartition naturelle du robinier faux acacia (*Robinia pseudoacacia*) en Amérique du nord (en bleu). *R. pseudoacacia* occupe une aire disjointe dans l'est des États-Unis. La plus grande partie à l'est correspond au plateau des Appalaches et chevauche onze États tandis que la plus petite à l'ouest correspond aux monts Ozark et englobe partiellement trois États (d'après Little 1971).

Il semble que cet arbre à croissance rapide et dont le bois possède des propriétés imputrescibles ait été remarqué par les populations natives qui l'auraient cultivé. Les colons se seraient également vite appropriés ces usages et l'auraient déplacé abondamment en Amérique comme l'attestent Michaux dans son livre « Histoire des arbres forestiers d'Amérique septentrionale tome III » (1813) et William Cobbett dans son ouvrage « The woodlands » (1825). Cela a même pu induire en erreur certains voyageurs, comme Miller qui certifie que le robinier est natif du Massachusetts (Michener, 1988). Actuellement on le retrouve dans chaque État des États-Unis ainsi qu'au Canada (Schütt 1994). Pourtant son bois n'est plus employé dans son aire native en raison d'un insecte pathogène dont les chenilles y creusent des galeries dans le bois le rendant impropre à la construction. De façon prophétique, Michaux relevait dans son livre<sup>3</sup> : « malheureusement, depuis 15 à 20 ans, il est survenu une sorte d'obstacle, qui sera un grand empêchement aux tentatives qu'on pourra faire, dans la suite pour propager l'Acacia dans toutes les parties des États-Unis, anciennement habitées ; cet obstacle est dû à un insecte ailé qui, depuis cette époque, attaque le tronc des Acacias vivans [...] ; il perce l'écorce, s'introduit dans le centre et le ronge en tous sens [...], les arbres

---

<sup>3</sup> J'ai volontairement conservé l'orthographe originelle dans cette citation.

ainsi affectés, cèdent facilement aux efforts des vents et se brisent. Cet inconvénient est tel que beaucoup de personnes renoncent à planter des Acacias. [...] Il doit donc résulter de ceci que les Acacias, venant successivement à disparaître dans les forêts Américaines par la consommation journalière qui s'en fait, et ne pouvant plus être reproduits dans les endroits anciennement habités, à cause de cet insecte, ils pourront devenir très rares dans les pays d'où ils sont originaires, pendant qu'en Europe, où on n'éprouve pas cet accident, on les propagera avec la plus grande facilité » (Michaux, 1813).

L'introduction du robinier en Europe est toujours enveloppée de mystère. Il a été introduit en Europe au début du XVII<sup>ème</sup> siècle et il est maintenant présent dans tous les pays européens. Les ouvrages populaires indiquent fréquemment que sa première introduction aurait eu lieu en France en 1601 et que ce serait Jean Robin (1550 – 1629) le botaniste du roi Henri IV qui en serait l'auteur. Un siècle plus tard, Carl Von Linné aurait doté le robinier de son nom en reconnaissance du travail accompli par Jean Robin et son fils Vespasien dans l'acclimatation et la dissémination du robinier en Europe. Il semble pourtant peu probable que cette date de 1601 soit exacte car le robinier est absent (Wein 1930 cité par Cierjacks et al., 2013) de la liste éditée par Jean et Vespasien Robin en 1601 (« Catalogue de son jardin ») et 1620 (« Histoire des plantes nouvellement trouvées en l'isle de Virginie »). D'après nos recherches, la première mention de cette espèce est faite en Angleterre par John Tradescant (1634) dans son catalogue « Plantarum in horto Iohannem Tradescanti » (cité dans « Early British botanists and their gardens » par Gunther, 1922) sous le nom *Locusta virginiana arbor* et presque simultanément en France dans l'ouvrage Plantarum canadensium historia (1635) par Jacques Philippe Cornuti, l'ami de Vespasien Robin, sous l'appellation *Acacia americana robini* (Fig 3).

Hamel (1854) rapporte que le fils de John Tradescant, a voyagé en Virginie, qui recouvrait alors les états actuels de Virginie, Virginie occidentale et du Kentucky, avant 1640 et il est attesté qu'il a ramené diverses espèces de plantes avec lui, il est ainsi probable que ce soit à lui que l'on doit la première importation du robinier en Europe depuis la Virginie. Il est de plus établi que John Tradescant père échangeait des graines avec Jean Robin depuis 1601. En effet John Tradescant avait été chargé du jardin de la propriété de Lord Salisbury à Hatfield en 1611 et ses registres indiquent qu'il a reçu des variétés de « Master Robyns » qui n'est autre que Jean Robin. Nous retrouvons plus tard, en 1629, des traces d'un échange de plantes ornementales (*Iris affracanis*, *Iris pertyca* ...) entre Robin et John Tradescant qu'il cultiva dans son jardin de Lambeth (Gunther 1922). Ainsi une hypothèse probable est que Tradescant échangeant des graines depuis 1601 avec les Robin, ait envoyé des graines de robinier

à Vespasien Robin qui les aurait cultivées dont l'arbre planté en 1634 au jardin des plantes, anciennement appelé le jardin royal, à Paris (Biographie Universelle, 1824).



**Fig 3** Gravure du robinier alors connu sous le nom *Acacia americana robini* parue dans le livre *Plantarum canadesium historia* (1635) de Jacques Philippe Cornuti, ami de Vespasien Robin.

En Europe, après une période d'intérêt pour son aspect ornemental, son feuillage gracile et ses fleurs odorantes, il passe de mode et tombe en désuétude. Liger dans son *Dictionnaire pratique du bon ménager* (1722) en parle en ces termes : « On a vu autrefois, et surtout dans la nouveauté, qu'on élevait beaucoup d'acacias ; mais aujourd'hui on s'en est pour ainsi dire dégoûté. On en plantait des allées, et on en formait des berceaux. Il croît fort vite et produit dans le printemps des fleurs dont l'odeur est fort agréable : mais comme il est fort sujet à se verser, qu'il a l'écorce raboteuse et le feuillage très petit, on ne s'en est plus soucié dans les jardins d'ornement ; et si l'on en voit encore en quelques endroits, c'est qu'on les y a bien voulu laisser. On étête cet arbre de temps en temps ; mais cette opération le rend difforme. On ne peut souffrir, aujourd'hui l'acacia tout au plus que dans quelques cours où il exhale une odeur fort agréable ».

Cependant, il semble être redécouvert pour les qualités de son bois à partir du milieu du 18<sup>ème</sup> siècle. Nicolas François de Neufchateau s'exprime ainsi dans son livre *Lettre sur le robinier* (1807) « Mais depuis cinquante ans,

cet arbre mieux connu a été vengé de l'oubli qui avait succédé à sa première vogue. On a senti le tort qu'on avait de l'abandonner, tandis qu'il réunit différentes propriétés qui le mettent au rang de nos arbres les plus utiles ». En effet, les explorateurs américains de retour en Europe assurèrent la promotion de cet arbre pour la foresterie. Michaux (1813) décrit l'abondance de cet arbre dans les mont Alleghany en Pennsylvanie et Virginie Occidentale et indiqua que depuis la fin du XVIII<sup>ème</sup> siècle, cet arbre était davantage considéré pour les propriétés de son bois que pour son aspect ornemental. A la même époque, l'homme politique anglais, exilé aux États-Unis, William Cobbett, fut dithyrambique sur les qualités de cette essence et assura la publicité de l'espèce en Europe. Dans son ouvrage, il témoigne ainsi de son rôle dans la dissémination du robinier en Europe : « I sold the plants; and, since that time, I have sold altogether more than a million of them » ajoutant que « My seed has always come from the neighborhood of Harrisburgh in Pennsylvania »<sup>4</sup> (Cobbett, 1825).

A partir de tous ces éléments historiques, nous pouvons conclure que la propagation du robinier en Europe a probablement connu un temps de latence entre sa première introduction au début du XVII<sup>ème</sup> siècle, probablement depuis la Virginie (états actuels de Virginie et Virginie occidentale), et sa redécouverte au cours du XVIII<sup>ème</sup> siècle qui a vraisemblablement entériné une nouvelle vague d'introduction, probablement en partie depuis la Pennsylvanie. Plus récemment, au cours du XX<sup>ème</sup> siècle, des programmes de sélection ont été entrepris par plusieurs pays d'Europe centrale (Keresztesi, 1983; Liesebach et al., 2004; Straker et al., 2015). Actuellement la Hongrie est le meneur européen dans la production de semis et leurs provenances sélectionnées sont distribuées largement à travers l'Europe (Keresztesi, 1983; Liesebach et al., 2004; Straker et al., 2015).

En dehors de la propagation par plantation pour l'ornementation ou la foresterie, le robinier faux-acacia s'est largement dispersé dans les milieux naturels (forêts, prairies, zones urbaines) d'Europe. Ainsi, en Hongrie, il est devenu la première essence forestière du pays, non seulement par ses plantations, mais aussi par sa dynamique naturelle : les plantations de robinier représentent 22% des surfaces forestières (Balogh et al., 2007), mais il est présent dans 30 à 50% des zones boisées feuillues et 40 à 65 % des prairies (Botta-Dukát, 2008). Sa propagation à courte distance des arbres adultes est principalement liée à la production de drageons (Radtke et al., 2013) et des suivis temporels ou des études moléculaires ont mis en évidence des distances de dispersion de 1 à 2 m par an (Cierjacks et al., 2013; Crosti et al., 2016). Des études menées en Allemagne et au Japon ont montré que ses graines étaient dispersées à longue distance, dans les gousses : par le vent ou les courants d'air des véhicules, elles peuvent

---

<sup>4</sup> J'ai volontairement laissé les citations de William Cobbett en version originale. Les traductions en français seraient : « J'ai vendu les plantes; et, depuis lors, j'en ai vendu plus d'un million » et « Mes graines sont toujours venues du voisinage de Harrisburg en Pennsylvanie ».

ainsi atteindre des distances de plusieurs dizaines de mètres (67 m) (Morimoto et al., 2010; von der Lippe and Kowarik, 2006) et par les rivières, elles parcourent facilement 400 m, 25% des graines suivies ayant atteint au moins 1200 m (Säumel and Kowarik, 2013). Ainsi dans la plupart des pays d'Europe, l'espèce naturalisée présente une forte dispersion, analysée par de nombreuses études (Cierjacks et al., 2013; Sitzia et al., 2016; Vítková et al., 2017) et le robinier faux-acacia est depuis classé parmi les 100 espèces les plus invasives d'Europe (DAISIE, 2009).

De cette analyse et de nos échanges avec les différents acteurs territoriaux dans le Sud-Ouest de la France, nous pouvons constater que le robinier faux-acacia se retrouve au centre d'un conflit d'usage : d'un côté les forestiers promeuvent la plantation de cette essence aux qualités indéniables et de l'autre les gestionnaires de milieux naturels tentent de limiter sa progression dans les zones qu'ils souhaitent protéger. Ainsi outre l'intérêt fondamental dans la compréhension des mécanismes évolutifs sous-jacents à l'invasion, il apparaît important d'appréhender l'origine du robinier européen et les raisons de son succès invasif.

## **1.7. Objectifs de ce travail de recherche**

Dans ce travail de recherche, j'ai voulu analyser l'histoire du robinier faux-acacia en Europe, d'un point de vue évolutif et écologique et ainsi apporter des éléments permettant de faire le lien entre la biologie évolutive et la compréhension de l'invasion européenne par le robinier. Les objectifs généraux étaient de comprendre l'histoire de l'introduction du robinier en Europe et ses conséquences sur la diversité et la structuration génétique des populations en Europe, d'évaluer l'existence d'une réponse adaptative des populations dans ce nouvel environnement, liée à une évolution rapide de caractères phénotypiques.

Dans un premier temps (chapitre 1), j'ai voulu éclaircir l'histoire de l'introduction du robinier européen, en recherchant l'origine de ses populations dans l'aire native à l'aide d'une étude comparative de génétique des populations au sein de l'aire invasive et de l'aire native. Pour cela nous avons entrepris un large échantillonnage de populations à travers de ces deux aires et nous avons génotypé les individus prélevés à l'aide de marqueurs SNP récemment développés dans notre laboratoire (Verdu et al., 2016). Nous nous sommes posés les questions suivantes, i) Quelles sont les populations sources du robinier introduit en Europe ? ii) Quel est le degré de différenciation génétique entre populations de l'aire native et de l'aire d'introduction ? iii) L'introduction s'est-elle accompagnée d'une perte de diversité génétique liée à un effet de fondation ?

D'autre part, nous avons souhaité comprendre quels traits favorisaient l'invasion, notamment les traits au stade juvénile, phase clé de l'installation. En effet, il apparaît que la germination et les stades juvéniles sont cruciaux pour le succès reproductif des individus et la sélection œuvrant sur ces stades juvéniles peut mener à de l'adaptation locale (Donohue et al., 2010; Postma and Ågren, 2016). Plus spécifiquement chez les espèces invasives, de nombreuses études ont montré l'importance de la germination et des stades juvéniles dans le succès de l'invasion (Beckmann et al., 2011; Blair and Wolfe, 2004; Erfmeier and Bruelheide, 2005; Gioria and Pyšek, 2016; Hirsch et al., 2017, 2012; Leiblein-Wild et al., 2014). Des changements phénotypiques dans cette phase précoce de la vie des individus ont été mis en évidence chez plusieurs espèces dans la vitesse de germination ou le taux de germination entre l'aire native et l'aire invasive. Par exemple les populations invasives de l'arbuste *Rhododendron ponticum* présentaient une phénologie des stades juvéniles significativement plus rapide que celle observée pour les populations de l'aire native (Erfmeier and Bruelheide, 2005) ; ou encore, l'arbuste *Hypericum perforatum* présentait un taux de germination plus élevé dans l'aire d'introduction (Beckmann et al., 2011). Nous avons ainsi mis en œuvre deux expérimentations de génétique quantitative, permettant d'évaluer la germination et le développement de jeunes robiniers soumis à des environnements thermiques différents.

Ces expérimentations ont nécessité de semer un grand nombre d'individus, en tout 5000 graines ont été mises à germer. Toutefois, les graines de robinier présentent une forte dormance physique qui nécessite d'être levée : la dormance est due au tégument dur et imperméable de la graine, il est nécessaire de les abîmer pour permettre à l'eau de traverser le tégument et d'assurer la germination. En effet sans traitement de scarification préalable, le taux de germination est naturellement très faible, inférieur à 10 % (Cierjacks et al., 2013). Dans la littérature scientifique, quelques méthodes sont proposées, mais sont souvent peu décrites et à l'efficacité parfois peu élevée. Les forestiers emploient fréquemment une bétonnière dans laquelle ils placent les graines avec du sable. Afin de disposer d'une méthode standardisée, répétable et gérable sur des échantillons de grande taille hiérarchisés (par populations / familles), préalablement à nos expérimentations, nous avons développé une méthode de levée de dormance mise au point sur graines de robinier et testée sur d'autres Fabacées invasives dont les graines présentent le même phénomène de dormance physique. Ce travail est présenté dans le chapitre 2.

Au sein de l'aire d'introduction d'une part (chapitre 3), j'ai cherché à tester l'existence d'une adaptation locale entre populations de la moitié sud (Aquitaine, France) et de la moitié nord (Wallonie, Belgique) de l'aire

européenne occidentale. Nous avons ainsi comparé vingt populations au sein de deux environnements caractérisés par 4°C de différence de température de l'air (différence moyenne au printemps entre Aquitaine et Wallonie). . Mes questions de recherche étaient : i) Observe-t-on de la plasticité phénotypique en réponse à la température sur les traits étudiés au niveau de la population ? ii) Existe-t-il des différences phénotypiques entre populations en relation avec la température et leur origine, suggérant une adaptation locale ?

Puis (chapitre 4), j'ai évalué le rôle de processus évolutifs (évolution rapide, plasticité) dans le succès de l'invasion par le robinier dans une approche de génétique quantitative évaluant les réponses à la température (3 niveaux) en jardins communs de populations natives et invasives de robinier. Mes questions de recherches étaient : i) Les populations ou les aires sont-elles génétiquement différenciées sur la base des traits juvéniles étudiés ? ii) Quelle est l'intensité de la plasticité en réponse à la variation de température ? iii) Observe-t-on un signal d'adaptation locale des populations de l'aire native et/ou de l'aire invasive ?



# **Chapitre 1**

**Few North-Appalachian populations are the sources of European black locust**

## TITLE

Few North-Appalachian populations are the sources of European black locust

## STATUS

Submitted to Evolutionary Applications

## AUTHORS

Bouteiller XP<sup>1</sup>, Verdu CF<sup>2</sup>, Aikio E<sup>3</sup>, Bloese P<sup>4</sup>, Dainou K<sup>2</sup>, Delcamp A<sup>1</sup>, De Thier O<sup>2</sup>, Guichoux E<sup>1</sup>, Mengal C<sup>2</sup>, Monty A<sup>5</sup>, Pucheu M<sup>1</sup>, van Loo M<sup>6</sup>, Porté AJ<sup>1</sup>, Lassois L<sup>2</sup>, Mariette S<sup>1</sup>

<sup>1</sup>BIOGECO, INRA, Univ. Bordeaux, FR-33610, Cestas, France

<sup>2</sup>Biodiversity and Landscape Unit, Gembloux Agro-Bio Tech, University of Liège, BE-5030, Gembloux, Belgium

<sup>3</sup>Department of Genetics and Physiology, University of Oulu, Oulu, Finland

<sup>4</sup>Department of Forestry, Michigan State University, East Lansing, Michigan, USA

<sup>5</sup>Forest Management Unit, Gembloux Agro-Bio Tech, University of Liège, BE-5030, Gembloux, Belgium

<sup>6</sup>Department of Botany and Biodiversity Research, University of Vienna, Vienna, Austria

## CORRESPONDENCE

bouteiller.xavier@gmail.com

stephanie.mariette@inra.fr

## Acknowledgments

We would like to thank all people who helped us for sampling black locust populations: Fred Huber from George Washington & Jefferson National Forests and all Rangers from the USDA Forest Service for their precious help in the sampling of black locust leaves in the USA (CV field work), Dr Gideon Bradburd from the Michigan State University for his support in America (XPB field work), Dr Santiago Gonzalez-Martinez (BIOGECO - France) who helped us to organize sampling through Spain and to all scientists who sampled for us or with us wild populations of black locust all through Europe, i.e. Dennis Born (Technische Universität- Germany), Marianne

Correard (INRA - Avignon -France), Christopher Galanos (Hellenic Open University - Greece), Dr Pablo Goikoetxea (Neiker - Spain), Dr Emilio Laguna(University of Valencia - Spain), Dr Roman Longaner and Antonin Martinik (Mendel University – Czech Republic), Dr Juan Pino (Universitat Autònoma de Barcelona - Spain), Alexandra Quenu and Sébastien Irola (SMABVC -France), Dr Volker Schneck (Thünen Institute of Forest Genetics -Germany), Dr Tomasz Wojda & Dr Martin Klisz (Forest Research Institute –Poland) and Dr Rafael Zas (Misión Biológica de Galicia - Spain) and Dr Kevin Walker and the Botanical Society of Britain and Ireland who kindly helped us to locate English populations of black locust. Special thanks to Régis Burtlett and Raphaël Segura for the help in sampling UK populations. We also thank Christophe Orazio (EFI – France), Nicolas Cheval and Dr Patrick Patuszka from (UE INRA – France) for helping us sampling European provenances within REINFFORCE project

This study was financially supported by the Forest and Nature Management Research Unit of Gembloux Agro-BioTech, The BioGeCo INRA-Univ Bordeaux Research Unit, the Special Research Fund of the University of Liège, the *Société Française d'Ecologie*, the French *Agence de l'Eau*, the ANR-10-EQPX-16 Xyloforest, the EU COST Action FP1403 "Non-Native Tree Species for European Forests - Experiences, Risks and Opportunities" and the Transnational Access to Research Infrastructures activity in the 7<sup>th</sup> Framework Program of the EC under the Trees4Future project (no.284181). Special thanks to Philippe Lejeune, for supporting the collaborative project.

## ABSTRACT

The role of evolution in biological invasion studies is often overlooked. In order to evaluate the evolutionary mechanisms behind invasiveness, it is crucial to identify the population sources of the introduction. Population genetics studies are underway on *Robinia pseudoacacia* L., a North American tree which is now one of the worst invasive tree species in Europe. We realized a large sampling both in invasive and native range, 63 populations were sampled and 818 individuals were genotyped using 113 SNPs. We then detected clonal genotypes in each population and analyzed population structure both between and within ranges and we compared genetic diversity among ranges, enlarging the number of SNPs to mitigate the ascertainment bias. First, we demonstrated that European black locust was introduced from only a limited number of populations located in Appalachians Mountains, this is in agreement with historical documents that we briefly reviewed in this study. Within America, population structure reflected long time process whereas in Europe it was largely impacted by human activities. Second, we evidenced a genetic bottleneck among ranges with a decrease in allelic richness and total number of alleles in Europe. Lastly, we found more clonality within European populations. Black locust became invasive in Europe although being introduced from a reduced part of its native distribution. Our results suggest that human activity was a major actor promoting invasion. Indeed human conducted breeding programs in Europe and traded seeds throughout the introduced range that likely promoted invasion. We suggest to not introduce the missing American genetic cluster to Europe.

## KEYWORDS

*Robinia pseudoacacia*, biological invasion, population genetics, introduction history, bottleneck, single nucleotide polymorphism

## 1. Introduction

Since their first definition in Charles Elton's book, biological invasions have been increasingly studied during the last decades. Compared to ecological impacts and management questions raised by many invasive species, the role of evolution in biological invasions has been long overlooked (Colautti & Lau, 2015). In order to fill this void of knowledge, it is crucial to identify the population sources of the introduction. Indeed, knowing the origin of the introduced populations allows a better understanding of evolutionary mechanisms behind invasiveness, such as the role of selection, local adaptation or admixture (Colautti & Lau, 2015; Dlugosch, Anderson, Braasch, Cang, & Gillette, 2015; Stephen R. Keller & Taylor, 2008). Practical applications of such studies are to identify source risk and predict invasive potential of a population (Chown et al., 2015).

When a species is introduced in a new range, it is generally expected to experience a genetic bottleneck leading to a loss of genetic diversity (i.e. allelic richness or heterozygosity) (Dlugosch et al., 2015). For example the invasive plant *Heracleum mantegazzianum* exhibited a lower diversity in the invasive range attesting a strong founder event (Henry et al., 2009). However, some studies (Dlugosch et al., 2015; Dlugosch & Parker, 2008) emphasized that the loss of genetic diversity among native and invasive range was generally weak (15 – 20 % in average). This can be explained by multiple introductions that limited the loss of diversity, as for *Phalaris arundinaceae* or *Prunus serotina* (Lavergne & Molofsky, 2007; Pairen et al., 2010). Genetic diversity is even bound to increase in the invasive range in the case of strong populations admixture (Dlugosch et al., 2015; Dlugosch & Parker, 2008), although large increase is rare (Uller & Leimu, 2011). For example, *Phalaris arundinacea* invasion was promoted by an increased genetic variation (Lavergne & Molofsky, 2007). Genomic admixture was likely to have favored the success of *Silene vulgaris* in its new American range (Keller, Fields, Berardi, & Taylor, 2014; Keller & Taylor, 2010).

Additionally to propagule pressure during introduction, the mating system can strongly impact the diversity and the genetic structure of populations. Clonal or self-fertilizing species are likely to experience a stronger loss of genetic diversity whereas a bottleneck effect may be reduced for outcrossing species (Baker, 1967; Pappert, Hamrick, & Donovan, 2000). For example, the loss of genetic diversity between native and invasive ranges was more important for purely clonal than for sexual populations of the invasive *Oxalis pes-caprae* (Ferrero et al., 2015).

Among all invasive species, trees and shrubs have been little studied (Richardson & Rejmánek, 2011) although they can strongly impact ecosystems. Contrary to many herbaceous species, invasive trees (357 tree species, i.e.

nearly 0.5% of all trees species; Richardson & Rejmánek, 2011) have often been voluntarily introduced in the new range for horticulture or forestry (Richardson & Rejmánek, 2011). Thus as a result of human interests, multiple repeated introductions could be expected and may have shaped their diversity in the introduced range (Hirsch, Richardson, & Le Roux, 2017). Furthermore, invasive trees are characterized by a longer generation time compared to invasive herbaceous species; this life-history trait may influence differentiation rate between ranges, and the few centuries and tree generations passed since the first introductions challenge the study of evolutionary processes in invasive trees (Hirsch et al., 2017). Population genetics studies of invasive trees especially including both invasive and native ranges have not commonly been conducted. To our knowledge, few studies evidenced mostly multiple introductions from native range to invasive range as for *Acacia saligna* (Thompson, Bellstedt, Richardson, Wilson, & Le Roux, 2015), *Pinus tadea* (Zenni, Bailey, & Simberloff, 2014) which have been widely sampled from native range and exhibited high admixture level within invasive range, or *Prunus serotina* (Pairen et al., 2010) which has been introduced several times but from a limited part of the native range.

Black locust *Robinia pseudoacacia* L. (Fabaceae) is native to North America and is considered an invasive tree at the world scale (eight regions among the 14 defined by Richardson & Rejmánek, 2011). In Europe, it is now recognized as one of the 100 worst invasive species (DAISIE, 2006, 2009). Although knowledge on its genetic diversity represents a key element to develop further ecological or evolutionary studies (Lawson Handley et al., 2011), little information exists on its genetic diversity and structure in introduced ranges, nor regarding its origin and differentiation from the population sources in North America. The only studies we know of in Europe compared four American populations to sixteen German and Hungarian populations (Lieseback & Schneck, 2012; Liesebach, Yang, & Schneck, 2004), but results remained mostly inconclusive, although suggesting a high genetic differentiation between American populations. Modern tools of molecular population genetics have proved to be useful to seek the geographic origins of invasive species, complementing or solving the lack of historical knowledge (Besnard et al., 2014; Chown et al., 2015; Cristescu, 2015; Hoos, Whitman Miller, Ruiz, Vrijenhoek, & Geller, 2010). Using SNP markers developed for black locust (Verdu et al., 2016), we investigated the introduction history and genetic diversity of black locust in its native range and European invasive range, in particular by answering the following questions: 1) Could we identify the native population sources of European black locust? 2) What is the genetic differentiation within and between ranges? 3) Can we detect a founder event associated with a loss of genetic diversity?

In order to put our results in perspective with the historiography of black locust introduction to Europe, we performed a bibliographic study that is presented hereafter.

## Short review of *Robinia pseudoacacia* introduction history in Europe

Black locust is native to North America where it naturally occurred in the eastern part of the United States, with a native range consisting of two disjoint areas (Little, 1971): the largest area corresponded to the Appalachian Mountains and partially encompasses several current States (i.e. Pennsylvania, Maryland, West Virginia, Virginia, North Carolina, South Carolina, Georgia, Alabama, Tennessee, Kentucky, and Ohio); the smallest area was located further west in the Ozark Mountains (i.e. Missouri, Arkansas and Oklahoma). In America, black locust was intensively displaced by the settlers due to the undeniable interest of its wood as attested by Michaux in "*Histoire des arbres d'Amérique septentrionales tome III*" (1813) or Cobbett in his book "the woodlands" (1825). These events sometimes lead to misinterpretation of its native distribution, for example Miller's in his *Gardeners Dictionary* (1756-1759) wrongly certified that black locust was native to Massachusetts (Michener, 1988). At the present, it is spread in every state of the contiguous USA and also in British Columbia, Québec, Newfoundland and Labrador in Canada (Schütt 1994). How and when its first introduction in Europe occurred is not exactly known and is still shrouded in mystery. It was introduced to Europe during the early 17<sup>th</sup> century and it is now present in all European countries. It is usually said in popular works that first European introduction was accomplished in France by Jean Robin (1550 – 1629), botanist of King Henri the Fourth in 1601. One century later, Carl Von Linné recognized the work realized by the Robins' father (Jean) and son (Vespasien) to acclimate black locust in Europe by giving it his current name. Actually, it seemed that the date of 1601 was very unlikely since black locust is absent (Wein 1930 cited by Cierjacks et al., 2013) from the lists edited by Jean and Vespasien Robin in 1601 ("*catalogue de son jardin*") and in 1620 ("*histoire des plantes nouvellement trouvées en l'isle de Virginie*"). To our knowledge, first citations of the species appeared in England in the John Tradescant list in "*plantarum in horto Iohannem Tradescanti*" (1634) (cited in "Early British botanists and their gardens" by Gunther, 1922) with the name *Locusta virginiana arbor* and quasi-simultaneously in France in the book *Plantarum canadensium historia* (1635) by Vespasien Robin friend's Jacques Philippe Cornuti with the name *Acacia americana robini*. Moreover, Tradescant the younger travelled in Virginia before 1640 (Hamel, 1854) and since it is attested that he brought back several plants with him, it is likely that he was the one who first introduced black locust seeds from Virginia to Europe. It is established that Tradescant the elder corresponded and exchanged seeds with the Robins since 1601 (Gunther, 1922); thus he would have sent some seeds to Vespasien Robin in France, who would have cultivated it (for example as the one planted in 1634 in the King's garden in Paris ("*jardin des plantes*"; in *Biographie Universelle*, 1824)).

In Europe, it seems that first introduction was followed by a period of interest for its ornamental aspect but it subsequently fell into disuse in the early 18<sup>th</sup> century as explained in a very critical dictionary about black locust of 1722 quoted by Nicolas François de Neufchateau in his book “*lettre sur le robinier*” (1807). Then starting from the middle of the 18<sup>th</sup> century, American explorers came back to Europe and promoted the use of black locust in forestry: for instance, Michaux (1813) described the abundance of this tree in the Allegheny mountains over Pennsylvania and West Virginia and indicated that since the end of the 18<sup>th</sup> century the tree was even more considered for the excellent qualities of its wood, than for the beauty of its foliage and flowers. At the same time, the English politician William Cobbett who emigrated to America in the late 18<sup>th</sup> century emphasized all the qualities of this tree, and promoted its planting in Europe: “I sold the plants; and, since that time, I have sold altogether more than a million of them” adding that “My seed has always come from the neighborhood of Harrisburgh in Pennsylvania” (Cobbett, 1825).

From these elements, we can conclude that black locust European dissemination seemed to have experienced a lag phase between its first introduction, possibly from Virginia during the early 17<sup>th</sup> century, and its rediscovery in the middle of the 18<sup>th</sup> century leading to a new wave of introductions probably from Pennsylvania and the Virginias. More recently starting at the beginning of the 20<sup>th</sup> century, breeding programs have been undertaken in central Europe (Keresztesi, 1983; Liesebach et al., 2004; Straker, Quinn, Voigt, Lee, & Kling, 2015). Currently, Hungary is the European leader in production of black locust seedlings and their selected provenances for wood production are now widely distributed in Europe for new forest plantations (Keresztesi, 1983; Liesebach et al., 2004; Straker et al., 2015).

## **2. Material and Methods**

### **2.1. Sampling**

Sixty-three populations of black locust were sampled both in the native range (29 populations) and in the European invasive range (34 populations). Sampling was conducted between spring 2014 and fall 2016 (Table 1 and Appendix 1) by different collaborators using a similar protocol: between 10 and 30 trees were sampled in each population. Samples were collected either in common gardens or in natural populations. The total sampling corresponded to 818 individuals, 402 from Europe and 416 from North America.



Table 1 General information on the sampled populations. Range corresponds either to Europe (EU) or the USA (US) with indication of either the country or the state. N is number of individuals genotyped per population. GPS coordinates are provided using the WGS84 geographic projection.

Range	Country/State	Population name	N genotyped	X Long WGS84	Y Lat WGS84
EU	France	Barthelasse Avignon	19	4.81781	43.96459
EU	Czech Republic	Brno	11	16.5178	49.0423
EU	Hungary	Budapest	20	19.1066667	47.6625
EU	Romania	Carei	11	22.44863	47.660695
EU	Belgium	Corphalie	10	5.2588	50.5389
EU	Poland	Drewnica	10	21.2508	52.2528
EU	Spain	Gafos Galicia	12	-8.61666667	42.3833333
EU	Bulgaria	Gorna Oryahovitsa-Kozarevec	12	25.694339	43.11872
EU	Netherland	Kelpen-Oler	12	5.824992	51.205358
EU	Germany	Klein	11	9.08888889	49.9911111
EU	France	La Flotte	6	-0.30478	44.38455
EU	France	La Gouaneyre	6	-0.27251	44.37598
EU	England	London Streat Ham	10	0.14327	51.4332
EU	England	London Wandsworth	10	0.16347	51.44574
EU	Macedonia	Macedonia	12	21.570794	41.506548
EU	Germany	Meppen	12	7.376964	52.703578
EU	Spain	Montseny	12	2.51236	41.8306
EU	Germany	Munchenberg	12	14.04553	52.55949
EU	Bulgaria	Novi Pazar-Kulevcha	12	27.195081	43.345317
EU	Hungary	Nyirsegi	12	19.041456	47.581206
EU	Germany	Obermeidenrich	12	6.815636	51.476075
EU	Poland	Pinczow	10	20.7022	50.2647
EU	Poland	Poznan	12	16.80831	52.31109
EU	Germany	Priesterweg Naturpark	12	13.3578	52.46072
EU	Hungary	Pusztavacs	10	19.505619	47.171762
EU	France	Remy	12	2.675194	49.460286
EU	Greece	Rhodos	12	27.9444444	36.2872222
EU	Slovakia	Slovakia	12	19.866761	48.719737
EU	Poland	Szczecin	12	14.548233	53.33744
EU	Turkey	Turkey	11	32.903719	40.159438
EU	Netherland	Uden	11	5.617717	51.685078
EU	Spain	Valencia	19	-0.784339	39.397275
EU	Spain	Vitoria	13	-1.94188	43.21603
EU	Austria	Wien	12	16.47286	48.2515
US	PA - Pennsylvania	ALTOONA	11	-78.3832	40.4885
US	VA - Virginia	Barbours Creek	21	-80.11	37.58
US	NC - North Carolina	Blue Ridge	22	-82.6722	35.4567
US	WV - West Virginia	CAMP CREEK	12	-81.1034	37.4876
US	TN - Tennessee	CHATTANOOGA	12	-85.783	35.1198
US	KY - Kentucky	DANIEL BOONE NF	12	-83.6452	37.7505
US	KY - Kentucky	Erline	22	-83.54	37.04
US	AR - Arkansas	FAYETTEVILLE	10	-94.205	36.0711
US	WV - West Virginia	FORT MILL RIDGE	12	-78.7971	39.3267
US	AR - Arkansas	FORT SMITH	12	-94.2895	35.3431
US	OH - Ohio	Ironton	22	-82.46	38.8
US	WV - West Virginia	LEWISBURG	7	-80.3812	37.7825
US	NC - North Carolina	Locust Cove	22	-83.71	35.36
US	KY - Kentucky	Morehead	22	-83.4663	38.0909
US	AR - Arkansas	OUACHITA	12	-93.8374	34.4492
US	WV - West Virginia	Perry	22	-78.66	39
US	AR - Arkansas	Pleasant Hill	4	-93.46	35.59
US	NC - North Carolina	SHOOTING CREEK	11	-83.6284	35.0546
US	WV - West Virginia	Slatyfork	17	-80	38.18
US	VA - Virginia	Stokesville	22	-79.3	38.28
US	MD - Maryland	US GRP 1	15	-78.75	39.65
US	WV - West Virginia	US GRP 2	11	-81.1	39.0666667
US	VA - Virginia	US GRP 3	6	-79.9333333	37.2666667
US	KY - Kentucky	US GRP 4	6	-84.5	38.0333333
US	KY - Kentucky	US GRP 5	8	-84.5333333	38.65
US	KY - Kentucky	US GRP 6	7	-83.6833333	36.75
US	AR - Arkansas	Victor	22	-93.05	35.65
US	OH - Ohio	WAYNE NF	12	-82.5935	38.6577
US	VA - Virginia	Whiteop	22	-81.6562	36.7694

Black locust propagates through sexual and asexual reproduction. In common gardens, since trees were grown from seeds of known origin, there was no risk of collecting clones, only to collect sibs. However, in natural populations, a minimal distance of 25 meters was kept between two sampled trees in order to minimize the risk of collecting clones.

Either leaves, cambium, buds or seeds were harvested depending on the season. For leaf sampling, a few leaflets on a green healthy leaf were collected using a manual tree pruner. For cambium sampling, external bark of the trunk was removed with a knife, then five rings of wood were collected using a 1 cm diameter punch. In the field, samples were stored into referenced tea bags and then placed into plastic boxes containing silica gel used for drying the samples. Silica gel was renewed after 24 h and 48 h and then until it kept its color. Plastic boxes were finally stored at ambient temperature in closed cupboards.

In natural population GPS coordinates either of the population or of each sampled tree were reported using a portable GPS (GPSMAP62, Garmin, Olathe, KS, USA). Concerning trees located in Michigan State University, geographic location of each mother tree was known and was used as coordinates of the sampled trees, and populations were defined by gathering trees from close geographic location (see Appendix 1).

## **2.2. DNA extraction and Genotyping**

For each individual, either a 1 cm<sup>2</sup> leaf sample was collected on a leaflet, or cambium was manually extracted from one ring of wood or five buds were collected; plant material was then crushed using an automated grinder (2010 Geno/Grinder, SPEX SamplePrep, Metuchen, NJ, USA). For four populations (Corphalie, Drewnica, Pinczow and Lewisburg), a few seeds from ten sampled mother trees were scarified and grown in the lab (Bouteiller, Porté, Mariette, & Monty, 2017) then the first fresh leaf on each specimen was used for genotyping. DNA was extracted and isolated using DNeasy 96 Plant Kit (Qiagen, Venlo, Netherlands) following the manufacturer's protocol. One negative control was set on each plate. DNA concentration was measured using an UV spectrophotometer NanoDrop 8000 (Thermo Fisher Scientific Inc., Wilmington, Delaware, USA) and confirmed using Quant-iT™ dsDNA Assay Kit (Thermo Fisher Scientific Inc., Wilmington, Delaware, USA). Besides DNA concentration, 260/280 and 260/230 absorbance ratio provided information about DNA purity. DNA concentrations were standardized to 10 ng/μL before SNP genotyping.

Two genotyping experiments were performed using SNPs recently developed on black locust (Verdu et al., 2016): (i) all collected samples (initial dataset) were genotyped using 113 SNPs selected following the procedure presented in Bouteiller et al., (2018) (ii) after clone removal (see results), 163 individuals were sub-sampled to maintain the sampling design between ranges (additional dataset: 69 individuals from the USA and 96 from Europe) and genotyped using a total of 251 SNPs. Four additional multiplexes of SNPs (138 SNPs in total) were designed with the MassArray Assay Editor 4.0.1.4 software (see Verdu et al., 2016 for more details regarding the procedure). SNP genotyping was performed with the Sequenom MassARRAY System (Agena Bioscience, San Diego, USA) at the Bordeaux Genome Transcriptome Facility (<https://pgtb.cgfb.u-bordeaux.fr/en>), using the iPLEX Gold chemistry genotyping kit according to the manufacturer's instructions. SNP data were visualized and validated using ViClust, a R program that we implemented for Galaxy (<https://usegalaxy.org/>); the program was also made available as a standalone R script for Linux or Windows (Bouteiller et al. 2018). All dataset will be accessible on Open Science Framework repository after acceptance

## **2.3. Data analysis**

### **2.3.1. Clone removal**

For the analysis of genetic diversity and structure within and among ranges, we chose to identify and remove clones from the analysis using R version 3.3.1 (R Development Core Team, 2016). Within populations, only markers without missing values were kept, and a pairwise comparison of each genotyped individual was done in order to detect putative clones.

For each population, the index of clonal diversity  $R$  (Arnaud-Haond, Duarte, Alberto, & Serrão, 2007) was calculated as:

$$R = \frac{(G-1)}{(N-1)} \quad (1)$$

where  $G$  is the number of unique genotypes in the considered population and  $N$  the sample size of the population. This index could vary from 0 in purely clonal population to 1 if all the individuals corresponded to different genotypes.

As some populations were sampled from trees in common garden or from lab-grown seedlings originated from seeds (Table 1), they were unlikely to contain clones; however, we checked that no clone was present in these populations and removed them from the next analysis. Testing difference of clonality among ranges was performed

using a Pearson  $\chi$  squared test with Yate's continuity correction using R version 3.3.1 (R Development Core Team, 2016).

### **2.3.2. Molecular genetic structure**

After removing clones from the dataset, molecular genetic differentiation was explored both among ranges and among populations within ranges using two approaches.

First, the typology of all sampled populations from both ranges was assessed on SNP data using a Correspondence Analysis (CA) at the population level, implemented using the R *adegenet* library (Jombart, 2008; Jombart, Devillard, & Balloux, 2010). As recommended by the authors, rare alleles, defined as allele recorded less than 10 % of the total allele number for each marker, were removed from the analysis.

Second, individual membership was analyzed using the Bayesian clustering approach *Markov Chain Monte Carlo* (MCMC) implemented in the STRUCTURE v2.3.4 software (Porrás-Hurtado et al., 2013; Pritchard, Stephens, & Donnelly, 2000), under an admixture model. Each model run assumed that the global diversity was structured into K clusters, and according to the SNP data, each individual was assigned proportionally to each cluster, thus we obtained for each individual membership coefficients to the K clusters associated with a 90% confidence interval by setting the ANCESTDIST parameter from 0 to 1.

Each run corresponded to a MCMC model with a burn-in period of 500,000 iterations followed by 500,000 iterations and it was repeated 10 times (Gilbert et al., 2012). The analysis was first performed using the initial dataset after clone removal to determine the structuration of populations from both ranges (K varying from 1 to 20), then it was performed for each range separately (K varying from 1 to 15). The most probable number of clusters according to each dataset was determined according to Evanno et al., (2005) as the peak in the  $\Delta K$  parameter calculated using the STRUCTURE HARVESTER software (Earl & vonHoldt, 2012). All runs were computed on the GenoToul bioinformatics cluster (<http://bioinfo.genotoul.fr/>) using the StrAuto script (Chhatre & Emerson, 2017) to produce STRUCTURE mainparams and extraparams and to automatize and parallelize the STRUCTURE analysis. Finally, the CLUMPAK software (Kopelman, Mayzel, Jakobsson, Rosenberg, & Mayrose, 2015) was used to synthesize STRUCTURE outputs and compute graphs of membership into the most likely K cluster number, as line charts at the individual level or as pie charts at the population level

We determined for each individual if it was significantly assigned to one of the K clusters or if it was admixed by using individual membership coefficient confidence intervals to the K clusters. When the confidence intervals

were not overlapping, individual was considered as significantly assigned to the  $K^{\text{th}}$  cluster for which it had the highest membership coefficient. Thus we were able to calculate for each population the ratio of individuals assigned for each  $K^{\text{th}}$  clusters or admixed. Finally, we assigned the population to the cluster for which the proportion of individuals assigned was the highest, in case of equality or if 100% of individuals were admixed, the population was declared as admixed.

To visualize within range genetic structure, we computed spatial interpolation using individual membership coefficients for the most likely  $K$  within each range. For interpolating data, mean individual membership averaged over 10 runs for the most likely  $K$  was used as input variable for the interpolation. An Inverse Distance Weighting (IDW) interpolation with a power value of 1 was realized using ArcGIS v10.2.2 geostatistical analyst tool (ESRI 2011). Neighborhood was searched using a 4 sectors circle with a maximal value of 25 neighbors and a minimum of 0.

### **2.3.3. Analysis of genetic differentiation and diversity**

The genetic differentiation between populations (with values between 0 - 1, none-full differentiation) was analyzed based on  $F_{ST}$  indexes (Wright, 1931) which quantify the.  $F_{ST}$  within and among ranges were calculated using the hierfstat v 0.04-28 R package (Jérôme Goudet, 2005) according to the Weir and Cockerham method (1984). 95% Confidence Intervals (CI) were estimated by performing 1000 bootstraps over loci.

Two datasets were analyzed for comparing genetic diversity among ranges, the initial dataset (i.e. 720 genotyped individuals using 113 SNPs) and the additional dataset (i.e. 163 genotyped individuals using 251 SNPs, see Methods). The aim of the second dataset was specifically to test for a potential bias in allelic frequency due to SNPs selection. First, allelic frequencies were evaluated by plotting the MAF distributions (Minor Allelic Frequency) per locus and per range. Diversity indices (ranging from 0 no to 1 maximum diversity) were calculated among ranges, among populations within ranges and within populations using the R package Hierfstat v 0.04-28 (Jérôme Goudet, 2005) and Fstat software v2.9.3 (J Goudet, 2013): observed heterozygosity ( $H_o$ ) which quantifies the proportion of heterozygous individuals, expected heterozygosity ( $H_e$ ), also known as genetic diversity that measures the expected proportion of heterozygous individuals under Hardy-Weinberg Equilibrium,, allelic richness (AR) which corresponds to the number of alleles weighted by the number of individuals in the smallest population and inbreeding coefficient  $F_{IS}$  within each population that measures the proportional deviation of

observed from expected heterozygosity within subpopulation, and Total Number of Alleles (TNA) per range was calculated by summing within each range the number of allele over all loci.

Differences in AR were tested by performing a non-parametric Wilcoxon paired test among loci between ranges. In order to evaluate differences in total number of alleles between ranges, a bootstrap over all loci and individuals was computed using 1000 simulations and differences were tested using a non-parametric Mann Whitney test using R version 3.3.1 (R Development Core Team, 2016).

#### **2.3.4. Isolation by distance analysis**

In natural populations, it is expected that genetic similarity was higher between spatially close populations and then decreased among populations with geographic distance, this pattern is named Isolation By Distance (IBD). IBD was tested within each range, using two approaches (i) the genetic distances between populations was calculated as the ratio  $F_{ST}/(1-F_{ST})$  using pairwise  $F_{ST}$  plotted against the logarithm of the pairwise geographic distances among populations (Rousset, 1997) and the correlation was tested using a Pearson coefficient test. (ii) A Mantel test between the matrix of pairwise geographic distances and the matrix of pairwise genetic distances was performed using R ade4 library based on 9999 replicates. For both methods, pairwise geographic distances among populations were calculated based on the GPS coordinates of each population using the R CalcDists function provided by Scott Chamberlain on GitHub (<https://gist.github.com/sckott/931445>). The matrix of pairwise genetic distances was estimated using the Cavalli-Sforza and Edwards Chord distance with hierfstat v 0.04-28 (Takezaki & Nei, 1996).

### **3. Results**

#### **3.1. More asexual reproduction in European populations**

Overall, a higher clonality was detected in the European populations compared to the American ones, with a significant range effect ( $\chi^2 = 29.04$ ,  $df = 1$ ,  $p = 7.10 \cdot 10^{-8}$ ). As expected, no clone was found within the common garden's populations nor in the populations obtained from seedlings germinated in the lab. When removing these populations from the analysis (thus leaving 280 European and 356 American individuals), 98 genotypes were found with a least one duplicated version (i.e. clones): 68 clones out of the 280 European samples and 30 clones

out of the 356 American samples. Keeping only one sample per genotype resulted in a dataset after clone removal that contained 720 genotypes out of the 818 sampled individuals, distributed as 334 genotypes in 34 European populations and 386 genotypes in 29 American populations.

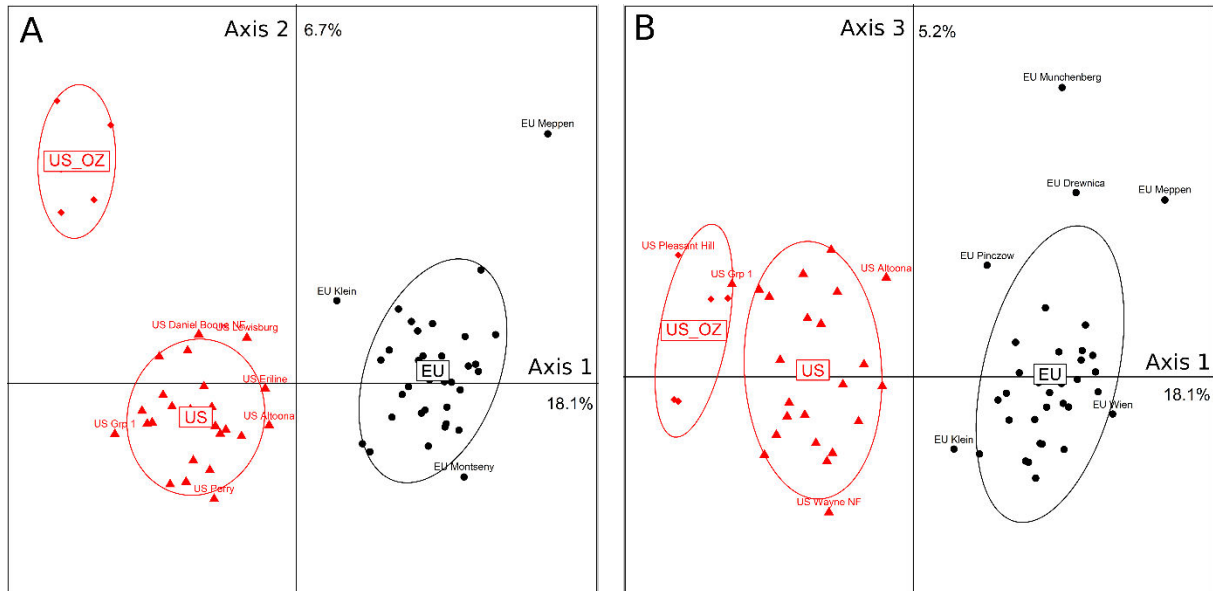
Among European populations (Appendix 2), the index of clonal diversity  $R$  ranged from 0.27 to 1 (mean = 0.82,  $sd = 0.26$ ), with significant differences in clonality between populations ( $\chi^2 = 88.33$ ,  $df = 60$ ,  $p = 0.01$ ). The most clonal populations (i.e.  $R < 0.5$ ) were Brno Czech republic ( $R = 0.4$ ), Gafos Galicia Spain ( $R = 0.36$ ), La Flotte France ( $R = 0.4$ ), Munchenberg Germany ( $R = 0.27$ ), Uden the Netherlands ( $R = 0.4$ ), Valencia Spain ( $R = 0.28$ ). Among American populations, the index of clonal diversity  $R$  ranged from 0.67 to 1 (mean = 0.93,  $sd = 0.10$ ) and there was no significant effect of population on the clonality ( $\chi^2 = 76.77$ ,  $df = 60$ ,  $p = 0.07$ ). The lowest  $R$  value was observed for the Pleasant Hill, in Arkansas ( $R = 0.67$ ).

### **3.2. Introduced populations are genetically close to a few Northeastern native populations**

Introduced European individuals of black locust are genetically close to only a few native individuals located in four populations from the Northeastern part of its native range. The correspondence analysis (CA) conducted at the population level demonstrated that American and European populations are overall strongly separated, with only three American populations (Altoona, Eriline, Lewisburg) slightly closer to the European cluster. At the individual level, the STRUCTURE analysis gave congruent results: among the 29 sampled American populations, only four (Altoona, Eriline, US Grp 3 and Wayne National Forest) exhibited individuals with a strong genetic relatedness to the European individuals.

First we used a CA to analyze the relatedness between populations. On the CA first axis (18.1 % of inertia), two clusters were clearly separated opposing European and American populations (Fig 1-A & B). Among the few outlying populations, the three American populations the closest to the European cluster were Altoona (PA), Eriline (KY) and Lewisburg (WV). The second axis (6.7 % of inertia) splitted the American cluster into two groups: the upper cluster corresponded to the Ozark western populations and the lower one to the Appalachian eastern populations (Fig 1-A). Within Europe, some outlying populations such as Drewnica (Poland), Klein

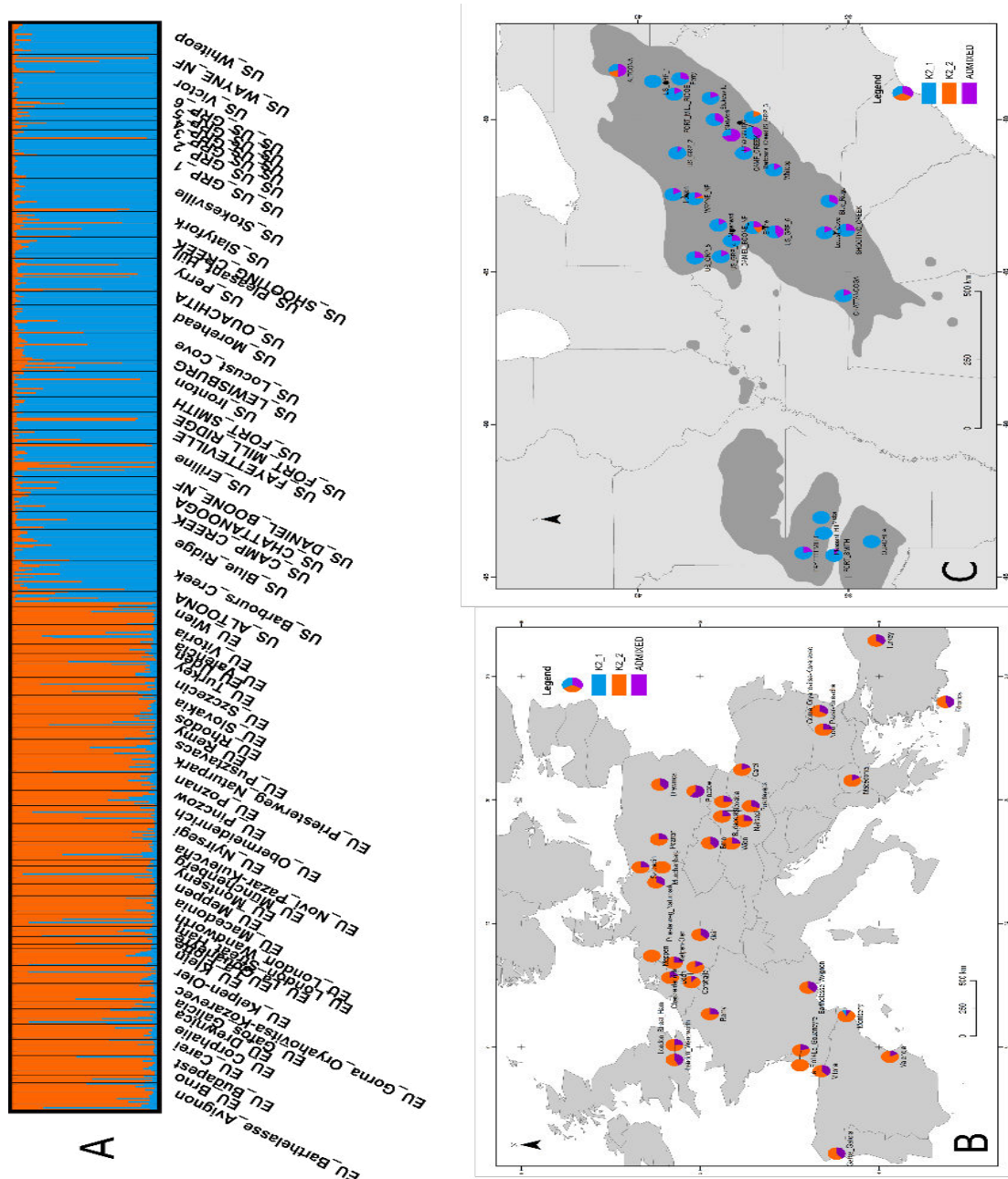
(Germany), Meppen (Germany), Munchenberg (Germany) or Pinczow (Poland) were detected on the third axis (5.2 % of inertia -Fig 1- B).



**Fig 1** Correspondence Analysis performed at the population level. A) Axis 2 plotted against axis 1 B) Axis 3 plotted against axis 1. Ellipses were plotted to illustrate the identified genetic clusters. They encompassed roughly 80 % of the populations. Populations within the ellipse were printed as red symbols (American populations) either triangle for Appalachians populations (US) or diamond for Ozark populations (US\_OZ) or black circles for European populations (EU). Name of the populations outside of the ellipse were printed above the point of the population name label.

Second, we used a STRUCTURE analysis that allowed to cluster the individuals based on their molecular genetic similarity. An optimal number of  $K = 2$  clusters (Appendix 3- A) was identified on the STRUCTURE analysis output using the initial dataset after clone removal (i.e. 720 individuals from the 63 populations from both ranges). The first cluster (Fig 2, in blue; K2\_1; Appendix 4) gathered a majority of the individuals from America (76.8 %), and only one European population (Montseny: 8.3 %) contained samples significantly attributed to this cluster based on individuals ancestry confidence intervals (0.245 % of the total number of European individuals).





**Fig 2** A) Individual assignment for the most likely number of clusters (i.e.  $K = 2$ ) as a result of the STRUCTURE analysis among ranges. Each colored vertical line represented one individual ancestry membership between the 2 clusters (orange, cluster K2\_1 and blue cluster K2\_2). Black vertical lines separated different populations. Both analyses were computed on the initial dataset after clone removal (i.e. 720 individuals from 63 populations genotyped using 113 SNPs). B) & C) Pie charts of the population assignment in Europe and the USA for the most likely number of clusters (i.e.  $K = 2$ ) as a result of the STRUCTURE analysis among ranges. In blue, proportion of individuals significantly assigned to cluster K2\_1, in orange proportion of individuals significantly assigned to cluster K2\_2, and in Purple proportion of individuals admixed in each population.

The second cluster (Fig 2 - C in orange; K2\_2; Appendix 4) gathered most of the European individuals (73.6 % of the total number of European samples) and a few samples from North America (2.18 %). The American samples significantly assigned to cluster K2\_2 belonged to only four American populations (Appendix 4), and only a small proportion of individuals were significantly attributed to cluster K2\_2 (Altoona PA: 20%, Eriline KY: 18.2%, US Grp 3 VA: 16.7%, Wayne National Forest OH: 8.33%). These 4 populations were all located in the northern part of the Appalachian Mountains (Fig 2C). Finally, admixed individuals (i.e. not attributed significantly to one cluster) were found in both ranges with relatively similar proportion (Europe: 26.1%, America: 21.0 %, Global: 23.8 %). In America, the proportion of admixed individuals per population ranged from 0 % (Ouachita, Pleasant Hill, US Grp 1, US Grp 3, Victor) to 71% (Lewisburg), whereas in Europe it is ranging from 0 % (Meppen, Munchenberg) to 60 % (Pinczow).

### **3.3. Significant genetic differentiation among populations in the native range contrary to the introduced range**

There was a significant genetic differentiation among all populations, estimated  $F_{ST}$  among all populations was 5.23 % (95% CI: 4.77 % - 5.70 %). Overall, within the native range, native black locust populations were clearly genetically differentiated, matching with a geographic structure, whereas in the introduced European range, the differentiation between populations was low and presented no structure across the continent.

Genetic differentiation among American populations was significant with an estimated  $F_{ST}$  of 4.46 % (95% CI: 3.94 % - 5.06 %). Clear signals of Isolation By Distance (IBD) were observed between populations in the native range, as both correlation tests were significant (Table 2, Appendix 5 A). IBD remained significant among the Appalachian populations, but not among the Ozark populations.

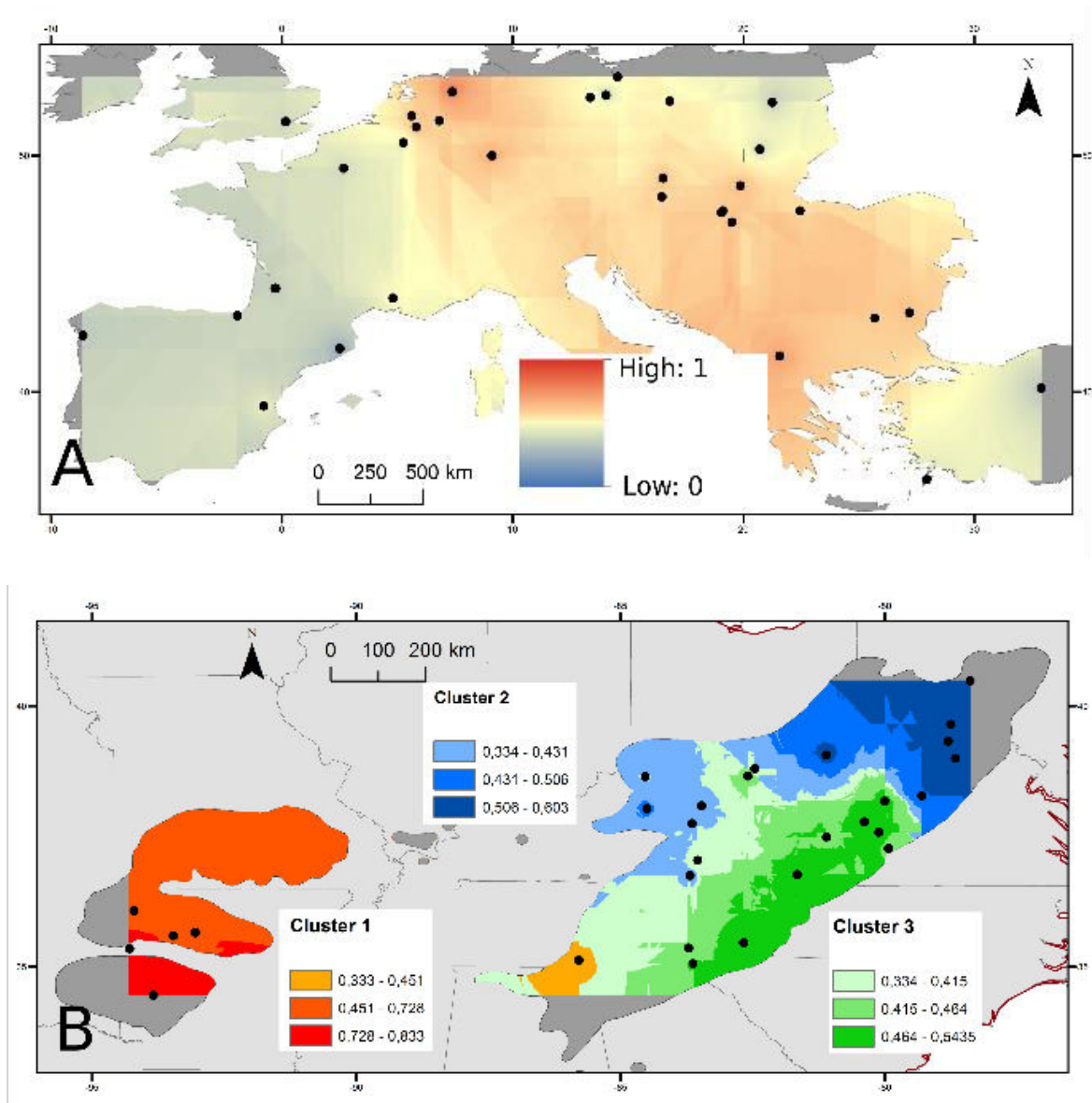
Table 2 Isolation By Distance correlation tests. Both regression of pairwise  $F_{ST}/(1-F_{ST})$  against logarithm of pairwise geographic distance (Rousset 1997) and Mantel test were performed within each range or within sub-selection of population in each range. Significant results were bolded.

Range	Pearson test		Mantel test	
	r	p	r	p
USA	<b>0.53</b>	<b>2.9 10<sup>-30</sup></b>	<b>0.479</b>	<b>3 10<sup>-4</sup></b>
Appalachians	<b>0.193</b>	<b>1.28 10<sup>-3</sup></b>		
Ozarks	0.386	0.27		
Europe	r	0.75	-0.028	0.562
K2_1:	0.105	0.223		
K2_2:	0.0692	0.483		

In the native range, the STRUCTURE analysis indicated that populations were organized in an optimal number of  $K = 3$  clusters (Appendix 3- B). The first cluster (K3\_1\_US, Fig 3 B, Appendix 6) corresponded to the Ozark western populations with high membership coefficients at the population level, all populations from Ozark had their highest proportion of individuals assigned to this cluster with a minimum for Fayetteville (22.2% of individuals assigned to K3\_1\_US and 77.8% of admixed individuals) and a maximum for Fort Smith (80.0% of individuals assigned to K3\_1\_US and 20% of admixed individuals). Mean admixture among populations assigned to this cluster was 51.2%.

The second cluster (Fig 3B, K3\_2\_US, Appendix 6) was located in the North of the Appalachian Mountains over Kentucky, Maryland, Ohio, Pennsylvania, Virginia and West Virginia: proportions of individual assigned were the highest for North Appalachians populations (Altoona, Perry, US Grp 3 and US Grp 4) and ancestry membership gradually decreased going to the South and the East of the Appalachian Mountains. Populations assigned to this cluster had a proportion of individuals significantly assigned ranging from 4.5% (Stokesville) to 50% (Us Grp 4). Four populations were also partially comprised of individuals assigned to other clusters, 6.7% of US Grp 1 individuals and 16.7% of US Grp 4 individuals were assigned to cluster K3\_1\_US and 5.6% of Locust Cove individuals were assigned to the cluster K3\_3\_US. Mean admixture among populations assigned to this cluster was 77.5%.

Heart of the third cluster (Fig3- B, K3\_3\_US) was based on the Eastern part of the Appalachians Mountains in North Carolina (Blue Ridge), Virginia (Barbours Creek, Whiteop) and Southern part of West Virginia (Slatyfork). The other populations were all made by a fraction of individuals significantly assigned to this cluster with values ranging from 4.5% (Eriline) to 30% (Barbours Creek). Mean admixture among populations assigned to this cluster with the exception of the 2 fully admixed populations was 83.9%



**Fig 3** . A) & B) Graphical IDW interpolation computed on the STRUCTURE individuals ancestry membership on the within range analysis for the most likely K in Europe (K =2) and in the USA (K = 3), respectively. The two European clusters are represented by a continuous color scale from blue (K2\_1\_EU) to red (K2\_2\_Eu) the three clusters in the USA are represented by a continuous red color scale (K3\_1\_US) through blue color scale (K3\_2\_US) to green color scale (K3\_3\_US).

Genetic differentiation among European populations was significant with an estimated  $F_{ST}$  of 3.08 % (95% CI: 2.60 % - 3.60 %); it was significantly lower than the  $F_{ST}$  observed in the native range. No significant signal of IBD was found either by testing the correlation among pairwise  $F_{ST}/(1-F_{ST})$  ratio and the logarithm of pairwise geographic distances (Pearson correlation test,  $r = -0.013$   $p = 0.75$ ) or by realizing a Mantel test ( $r = -0.028$ ,  $p =$

0.562). The correlation was not either significant within each cluster (K2\_1: Pearson correlation test,  $r = 0.105$ ,  $p = 0.223$ , K2\_2: Pearson correlation test,  $r = 0.0692$ ,  $p = 0.483$ ).

In the introduced range, the STRUCTURE analysis indicated that populations were organized with an optimal number of  $K = 2$  clusters (Appendix 3- C).

The first cluster (Fig 3 A in red; K2\_1\_EU Appendix 6) was composed of 14 populations mainly from Central Europe. All the populations of this cluster included a fraction of individuals significantly assigned with values ranging from 8.33% (Gorna) to 62.5% (Meppen). Moreover 2 populations (Macedonia, Nyirsegi) had 8.33% of their individuals significantly attributed to the second European cluster (K2\_2\_EU). Globally, mean individual admixture among populations assigned to this cluster, with the exception of the fully admixed population, was 75.6%, 19% of individuals were significantly assigned to this cluster and 5.3% were assigned to the other cluster.

The second cluster (Fig 3 A in blue; K2\_2\_EU Appendix 6) was mainly represented by 15 populations from Western and Eastern Europe. All populations included a fraction of individuals significantly assigned to this cluster with values ranging from 9.1% (Turkey) to 100% (Munchenberg). Moreover four populations had some individuals significantly attributed to the first European cluster (K2\_1\_EU): Vitoria (7.7%), Wien and Szczecin (8.33%) and Valencia (16.7%). Globally, mean individual admixture among populations assigned to this cluster was 70.9%.

### **3.4. MAF distribution and detection of a bottleneck in the introduced populations**

The MAF analysis performed on the initial and additional datasets pointed out a deficit in low frequency alleles when using 113 SNPs (MAF mode: 0.05 – 0.15), both in the native and invasive range, confirming that this initial set oversampled high MAF SNPs that would bias the evaluation of diversity indexes (Appendix 7 A). Consequently, the number of SNPs was increased to 251 SNPs by genotyping low MAF SNPs from our initial set of SNPs (Verdu et al., 2016) on a reduced number of samples.

When analyzing the 251 SNPs on 163 samples, no difference in heterozygosity was observed between both ranges but a lower AR was however showed in Europe compared to North America (1.91 vs 1.96, Wilcoxon,  $v = 1375.5$ ,  $p = 0.00453$ ) as well as a lower Total Number of Alleles in Europe compared to North America (485 vs 494) (Appendix 7 B).

## **4. Discussion**

Based on an extensive population genetics analysis in both native and European invasive ranges, we showed that black locust was introduced to Europe from a limited part of its northeastern native distribution, in the Appalachian Mountains. This founding effect brought about a bottleneck, detected only when we increased the number of SNPs with low MAF markers. A strong genetic structure was observed in the USA whereas a much weaker one was detected in Europe. Moreover, asexual propagation was bound to be more prevalent in the invasive range compared to the native one.

### **4.1. Populations genetics and introduction history: European black locust populations close to Northern Appalachian populations**

Genetic results clearly demonstrated that black locust has been introduced to Europe from a restricted number of native populations located in the northeastern Appalachian Mountains. The robustness of this result is based on a large sampling, at three levels: number of populations (29 in Europe and 34 in USA), number of individuals (334 in Europe and 386 in USA without clones, with a mean of 11.4 samples per population), and number of markers (113 SNPs). Besides, two types of analyses (CA, Structure) led to the same conclusion concerning the limited origin of European populations.

Our results regarding the geographic origin of European black locust using a molecular approach are congruent with historical records which pointed to the original sources of black locust in the Northeastern part of its native range in the Appalachian Mountains. Reviewing historical studies, we could indeed conclude that the first European black locust introduction during the early 17<sup>th</sup> century was likely to have occurred from Virginia and further introductions during the 17<sup>th</sup> and 18<sup>th</sup> centuries from Pennsylvania and West Virginia (Gunther, 1922, Michaux, 1813, Cobbett, 1825). Consequently, considering both the historical indications and the genetic proximity of all European black locust populations to few native ones, a sensible hypothesis would be that no subsequent introductions followed and that the expansion of the species in Europe through asexual reproduction or seeds were produced from the original black locust raised in Europe.

## 4.2. Evidencing the bottleneck depends on the set of genotyped SNPs

Given the few American populations close to European ones, a bottleneck is expected in European populations. The decrease in genetic diversity was showed only when using a larger number of SNPs. This is due to particular properties of SNPs.

SSRs and SNPs are two widely used markers for genotyping non-model species (Morin et al., 2004; Coates et al., 2009; Helyar et al., 2011). SNPs present many advantages: they can be easily developed based on NGS, genotyping is easy replicable among laboratories and SNPs are widely distributed through genome (Morin et al., 2004; Coates et al., 2009; Helyar et al., 2011). More SNPs are however needed than SSRs for reaching the same level of precision, essentially because SSRs are multiallelic whereas SNPs are mainly biallelic (Morin et al., 2004). One major problem in using SNPs is the ascertainment bias (Morin et al., 2004; Coates et al., 2009; Helyar et al., 2011). In particular, SNPs with high minor allele frequency are more susceptible to be sampled for genotyping populations and consequently this can alter diversity estimates (Helyar et al., 2011). Then, genetic diversity analysis conducted with SNP data may lead to false negative or false positive conclusions (Morin et al., 2004; Helyar et al., 2011). Some empirical studies indeed found that SNPs performed better than SSRs for studying population structure, whereas SSRs were more efficient for estimating diversity (Singh et al., 2013). Nevertheless, other empirical studies reached the same conclusion whatever the type of marker used (van Inghelandt et al., 2010; Filippi et al., 2015).

Our study allows emphasizing the importance of taking the SNP ascertainment bias into account when comparing genetic diversity among several groups. Using the initial dataset we indeed observed more frequent minor allele (modal class 0.05 – 0.15, Online Resource 3) confirming the bias. We partially corrected this bias by using the additional dataset (251 SNPs), where actual MAF distribution is closer to the expected MAF distribution (modal class 0 – 0.05, Appendix 7 A). If we only used the initial dataset for studying among ranges genetic diversity, we would have missed detecting the bottleneck (no difference in heterozygosity, allelic richness and total number of alleles among ranges, Appendix 7 B), however the analysis conducted with the additional dataset detected a

bottleneck (decrease in allelic richness and total number of alleles in the introduced range, Appendix 7 B). Indeed, both a loss in allelic richness and in total number of alleles are asserting a bottleneck whereas heterozygosity is expected to respond less strongly than allelic richness to founding event (Dlugosch et al., 2015) as we observed in our study.

Uller & Leimu (2011) demonstrated that genetic variation among native and invasive ranges was influenced by taxonomy. Indeed invasive animals often suffered an among ranges loss of genetic diversity whereas invasive plants often exhibited higher genetic diversity in the invasive range (Uller & Leimu, 2011). According to these authors, one factor contributing to this pattern is that invasive animal populations were often founded by single introduction events whereas for plants multiple introductions associated with admixture is more common (Uller & Leimu, 2011). Thus our results appear like an original pattern for plants of introductions from few populations coming from a restricted area of the native range combined with a loss of allelic richness. Many studies investigating the origins of herbaceous invasive plants documented multiple introductions from wide parts of their native range. For the European weed *Ambrosia artemisiifolia*, multiple introductions from two distinct genetic clusters of the native American range have been evidenced (van Boheemen et al., 2017). Similarly, two clusters in western and eastern Europe were identified as the sources of the American weed *Centaurea solstitialis* (Barker, Andonian, Swope, Luster, & Dlugosch, 2017). Thirdly, a wide scale study of the invasive weed *Mikania micrantha* in South-East Asia demonstrated the existence of two distinct genetics clusters that were resulting from separate introductions from the native American range (Yang et al., 2017). Studies on invasive trees are less numerous, but generally concluded to be multiple introductions (Besnard et al., 2014; Merceron et al., 2017; Pairon et al., 2010; Thompson et al., 2015) without the ability to clearly identify the population sources. A recent study on the genetic structuration of invasive tree populations of *Acacia saligna* within several invasive ranges suggested multiple introductions from populations distributed all through its Australian native range (Thompson et al., 2015).

#### **4.3. Long term evolutionary processes shaped structure in the native range whereas invasive range structure reflected anthropic action.**

Natural evolutionary processes shaped the genetic diversity and structure of black locust populations in the native range.



Indeed, three genetic clusters were identified within the native range with the strongest differentiation between this first cluster in the Ozark Mountains and the two clusters in the Northern and Southern Appalachian Mountains; together with the isolation by distance pattern it suggested the action of natural and long term evolutionary processes. On the contrary in Europe, the weak structure in two clusters with a few outlying populations and no isolation by distance would suggest a recent evolutionary history marked by human actions.

Within North America, glacial refugia have been identified on both sides of the Mississippi River (Hewitt, 2000; Swenson & Howard, 2005) over geographic areas closely related to the genetic clusters identified in this study. A strong differentiation on each side of the Mississippi River has been recorded for at least one other tree species, loblolly pine (*Pinus taeda* L.), which exhibited two distinct genetic clusters (Lu et al., 2016); the Mississippi River have likely acted as a physical barrier during post glaciation recolonization since the last glacial maximum during Wisconsinan 21,000 years ago (Pessino, Chabot, Giordano, & DeWalt, 2014).

Within the Appalachian mountains, similarly to our finding, distinct genetic clusters have been identified along a North South axis for several woody species such as *Scirpus ancistrochaetus* (Cipollini, Lavretsky, Cipollini, & Peters, 2017), *Tsuga caroliniana* (Potter, Campbell, Josserand, Nelson, & Jetton, 2017) and *Pinus strobus* (Nadeau et al., 2015). As emphasized by Swenson and Howard (2005), historical suture zone has been identified between the Northern Appalachian Mountains and Southern Appalachian Mountains. Consistently with this suture zone, two glacial refugia were detected for *P. strobus* in the Northern Appalachian Mountains and Southern Appalachian Mountains (Nadeau et al., 2015). Thus, the most probable assumption is that Appalachian black locust genetic structure was driven by the same historical processes and reflected postglacial colonization routes from glacial refugia on each side of the Appalachian Mountains.

On the contrary within the European range, two distinct genetic clusters were detected (over central Europe i.e. Austria, Bulgaria, Czech Republic, central Germany, Hungary, Macedonia, Netherland, Slovakia and Romania and from western and Eastern Europe i.e. England, France, Eastern Germany, Poland, Spain and Turkey) but with a very weak signal. Indeed most of the individuals were admixed between these two clusters (approx. 75%) and within each cluster, one outlying population was driving the signal of differentiation: Meppen (62.5% of individuals assigned to K2\_1) and Munchenberg (100% of individuals assigned to K2\_2). Although weak, three signals confirmed that the observed genetic structure was significant: (i) first, within each cluster some individuals were significantly assigned to the cluster (ii) STRUCTURE admixture parameter (i.e.  $\alpha$  parameter) was checked for convergence (i.e.  $\alpha$  should be relatively constant - with a range 0.2 or less - ) as recommended in the

STRUCTURE manual (Pritchard, 2010) (iii) finally, when removing the two outlying populations similar genetic structure results were observed within Europe (data not shown).

As discussed previously, black locust was most probably introduced to Europe from only a few populations located in a limited area of its native range, in the Northern part of the Appalachian Mountains. No support was found for an introduction from two different American genetic clusters that would have founded the two European clusters. Historical writings indicated that in the 17<sup>th</sup> century, Vespasien Robin disseminated through Europe, seeds collected from trees raised in Paris and that from 1634, it was largely planted through Europe as an ornamental species and seed orchards were developed in Europe by trading European raised seeds (Cierjacks et al., 2013, Cobbett, 1825, de Neufchateau, 1807, Jouannet, 1828). Doing so they propagated the same genetic material all through Europe. The origin of the central European cluster would then be related to human selection. Indeed, since the second half of the 18<sup>th</sup> century, extensive afforestation programs have been conducted in central Europe (in particular Germany, Hungary and the Czech republic) in order to promote the use of black locust in forestry (Cierjacks et al., 2013; Vítková, Müllerová, Sádlo, Pergl, & Pyšek, 2017). Moreover a genetic breeding program has been conducted since the beginning of the 20<sup>th</sup> century in Hungary (Keresztesi, 1983). One hypothesis could be that the genetic clustering within Europe would be a signal of evolution caused by artificial selection due to human oriented selection and tree breeding initiated in Central Europe since the 18<sup>th</sup> century. Thus we could talk of European black locust as partially domesticated, and wonder which traits could influence invasiveness. Further investigations including common garden surveys would be needed to assess whether genetic differences resulted in phenotypic differences. .

#### **4.1. Role of clonality in shaping genetic diversity in Europe**

Overall, we detected that European populations of black locust were more clonal than American populations. This was demonstrated by the higher index of clonal diversity  $R$ , but also by the analysis of the inbreeding coefficient  $F_{IS}$ . Since  $F_{IS}$  was lower in European populations than in American populations, this indicated an excess of heterozygosity within the European populations and clonality is expected to produce this pattern (Arnaud-Haond et al., 2007; Halkett, Simon, & Balloux, 2005; Stoeckel & Masson, 2014) as clonal reproduction can maintain heterozygosity over generations (Stoeckel et al., 2006).

Moreover, within the European range, two outlying populations (i.e. Meppen and Munchenberg) were clearly differentiated whatever the analysis. The Munchenberg population was already studied (code N° 7 – Hasenholz

in Liesebach et al., 2004) and results emphasized that this population was clearly differentiated from all others and two major clones were identified covering 80% of the plot (Liesebach et al., 2004). This is indeed consistent with our finding, since the Munchenberg population exhibited the lowest negative  $F_{IS}$  (-0.180) which indicated an excess of heterozygosity potentially due to a high level of clonality.

In Japan, both clonal and sexual reproduction have been evidenced to promote spread and invasion of black locust (Kurokochi & Hogetsu, 2014). Sexual regime is likely to influence invasiveness and shift in mating system have already been observed among ranges for several invasive species (Barrett, Colautti, & Eckert, 2008; Petanidou et al., 2012; Rambuda & Johnson, 2004). It has been demonstrated that essentially clonal populations didn't automatically exhibit a weak level of genetic diversity, conversely they can even maintain a high level of genetic diversity (Barrett, 2010). However, clonal populations are sensitive to founder event (Barrett, 2010). Effect of clonality is likely to strongly decrease  $F_{ST}$  and to slowly decrease genotypic diversity in purely clonal populations (Balloux, Lehmann, & De Meeûs, 2003), but partially clonal populations would be hard to differentiate from strictly sexual populations (Balloux et al., 2003). Theoretically, species invasiveness is non homogeneously dependent of clonal rate, it would exist as a tradeoff between clonal rate and sexual rate, and species with high clonal rate but a weak part of sexual reproduction would have the higher invasiveness (Bazin, Mathé-Hubert, Facon, Carlier, & Ravigné, 2013). Clonal reproduction provides reproductive assurance to invasive plant species (Barrett et al., 2008). Shift in mating system from outcrossing to clonality due to a strong founder event have been observed for invasive *Eichhornia crassipes* (Barrett et al., 2008) and *Fallopia japonica* (Hollingsworth & Bailey, 2000). However this is not systematic and some pure outcrossing species are successful invaders such as *A. artemisiifolia* (Friedman & Barrett, 2008). Anyway, it is possible that a shift towards more clonal reproduction occurred in the mating system of black locust between the native and the invasive range. It could have been produced by the founder event or by artificial selection since one traditional way for managing black locust plantations in Europe, was to stimulate clonal reproduction by damaging tree roots (Saint Jean de Crève Coeur, 1786, Bergeron, 1802, de Neufchateau, 1807).

## **5. Concluding remarks and recommendations on the invasive black locust management**

We found a remarkable congruence between genetic analysis and historical records regarding the geographic origin of European black locust, restricted to individuals of few founding populations from the Northern Appalachian Mountains. The introduction history of black locust is thus a unique pattern among invasive trees, which are

commonly characterized by multiple introduction events. As a consequence, only a small part of the genetic diversity of black locust from its native American range was introduced to Europe. Since this tree is highly invasive in Europe and since genetic diversity is assumed to be positively correlated with invasiveness (Dlugosch et al., 2015), we would recommend to avoid importing trees from the native range anymore, in particular from the missing clusters in the Southern Appalachian and the Ozark mountains.

Second, despite the fact that black locust suffered a genetic bottleneck and a loss of diversity at introduction, this did not prevent its successful colonization over its European range. Moreover, we found some support for a shift in mating systems between ranges with an increase of clonality in Europe, either resulting from natural or artificial selection. However, our sampling design was not specifically constructed to investigate changes in clonality during and post introduction; thus more studies including extensive populations and plots samplings would be needed to better understand the role of clonality in the success of this species. Still, this shift could promote the invasiveness of black locust at the local scale and reduce the invasion risk over longer distances away from artificial plantations; as a consequence, recommendations to forest plantation managers would be to avoid short rotation coppicing of the species and to set up large “safety belts” around their plantations to minimize local spreading of black locust in the vicinity.

## 6. Bibliography

- Arnaud-Haond, S., Duarte, C. M., Alberto, F., & Serrão, E. A. (2007). Standardizing methods to address clonality in population studies. *Molecular Ecology*, *16*(24), 5115–5139. <http://doi.org/10.1111/j.1365-294X.2007.03535.x>
- Baker, H. G. (1967). Support for Baker’s Law as a rule. *Evolution*, *21*(4), 853–856.
- Balloux, F., Lehmann, L., & De Meeûs, T. (2003). The population genetics of clonal and partially clonal diploids. *Genetics*, *164*(4), 1635–1644. <http://doi.org/10.1016/j.tree.2005.01.001>
- Barker, B. S., Andonian, K., Swope, S. M., Luster, D. G., & Dlugosch, K. M. (2017). Population genomic analyses reveal a history of range expansion and trait evolution across the native and invaded range of yellow starthistle (*Centaurea solstitialis*). *Molecular Ecology*, *26*(4), 1131–1147. <http://doi.org/10.1111/mec.13998>
- Barrett, S. C. H. (2010). *Why Reproductive Systems Matter for the Invasion Biology of Plants. Fifty Years of Invasion Ecology: The Legacy of Charles Elton*. <http://doi.org/10.1002/9781444329988.ch15>

- Barrett, S. C. H., Colautti, R. I., & Eckert, C. G. (2008). Plant reproductive systems and evolution during biological invasion. *Molecular Ecology*, *17*(1), 373–83. <http://doi.org/10.1111/j.1365-294X.2007.03503.x>
- Bazin, É., Mathé-Hubert, H., Facon, B., Carlier, J., & Ravigné, V. (2013). The effect of mating system on invasiveness: some genetic load may be advantageous when invading new environments. *Biological Invasions*, *16*(4), 875–886. <http://doi.org/10.1007/s10530-013-0544-6>
- Besnard, G., Dupuy, J., Larter, M., Cuneo, P., Cooke, D., & Chikhi, L. (2014). History of the invasive African olive tree in Australia and Hawaii: Evidence for sequential bottlenecks and hybridization with the Mediterranean olive. *Evolutionary Applications*, *7*(2), 195–211. <http://doi.org/10.1111/eva.12110>
- Bouteiller, X. P., Porté, A. J., Mariette, S., & Monty, A. (2017). Using automated sanding to homogeneously break seed dormancy in black locust (*Robinia pseudoacacia* L., Fabaceae). *Seed Science Research*, 1–8. <http://doi.org/10.1017/S0960258517000150>
- Bouteiller X.P., Barraquand F., Garnier-Géré P., Harmand N., Laizet Y., Raimbault A., Segura R., Lassois L., Monty A., Verdu C., Mariette S. & Porté A.J. (2018). No evidence for genetic differentiation in juvenile traits between Belgian and French populations of the invasive tree *Robinia pseudoacacia*. *Plant Ecology and Evolution*.
- Chhatre, V. E., & Emerson, K. J. (2017). StrAuto: automation and parallelization of STRUCTURE analysis. *BMC Bioinformatics*, *18*(1), 192. <http://doi.org/10.1186/s12859-017-1593-0>
- Chown, S. L., Hodgins, K. A., Griffin, P. C., Oakeshott, J. G., Byrne, M., & Hoffmann, A. A. (2015). Biological invasions, climate change and genomics. *Evolutionary Applications*, *8*(1), 23–46. <http://doi.org/10.1111/eva.12234>
- Cierjacks, A., Kowarik, I., Joshi, J., Hempel, S., Ristow, M., Von Der Lippe, M., & Weber, E. (2013). Biological Flora of the British Isles: *Robinia pseudoacacia*. *Journal of Ecology*, *101*(6), 1623–1640. <http://doi.org/10.1111/1365-2745.12162>
- Cipollini, K., Lavretsky, P., Cipollini, D., & Peters, J. L. (2017). Genetic Population Structure across the Range of Endangered Northeastern Bulrush, *Scirpus ancistrochaetus*. *International Journal of Plant Sciences*, *178*(1), 67–78. <http://doi.org/10.1086/688960>
- Colautti, R. I., & Lau, J. A. (2015). Contemporary evolution during invasion : evidence for differentiation, natural selection, and local adaptation. *Molecular Ecology*, *24*, 1999–2017. <http://doi.org/10.1111/mec.13162>
- Cristescu, M. E. (2015). Genetic reconstructions of invasion history. *Molecular Ecology*, *24*(9), 2212–2225. <http://doi.org/10.1111/mec.13117>

- Cobbett, W. (1828). *A Year's Residence in the United States of America* (3<sup>rd</sup> ed.). London, England.
- DAISIE. (2006). *Delivering Alien Invasive Species In Europe*. Retrieved February 26, 2015, from <http://www.europe-aliens.org>
- DAISIE. (2009). *Handbook of alien species in Europe*. (P. Hulme, W. Nentwig, P. Pyšek, & M. Vilà, Eds.). Dordrecht, Netherlands: Springer.
- Dlugosch, K. M., Anderson, S. R., Braasch, J., Cang, F. A., & Gillette, H. D. (2015). The devil is in the details: genetic variation in introduced populations and its contributions to invasion. *Molecular Ecology*, *24*, 2095–2111. <http://doi.org/10.1111/mec.13183>
- Dlugosch, K. M., & Parker, I. M. (2008). Founding events in species invasions: Genetic variation, adaptive evolution, and the role of multiple introductions. *Molecular Ecology*, *17*(1), 431–449. <http://doi.org/10.1111/j.1365-294X.2007.03538.x>
- Earl, D. a., & vonHoldt, B. M. (2012). STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, *4*(2), 359–361. <http://doi.org/10.1007/s12686-011-9548-7>
- ESRI. (2011). *ArcGIS Desktop: Release 10.2.2* Redlands, CA: Environmental Systems Research Institute. *Redlands CA*.
- Ferrero, V., Barrett, S. C. H., Castro, S., Caldeirinha, P., Navarro, L., Loureiro, J., & Rodríguez-Echeverría, S. (2015). Invasion genetics of the Bermuda buttercup (*Oxalis pes-caprae*): complex intercontinental patterns of genetic diversity, polyploidy and heterostyly characterize both native and introduced populations. *Molecular Ecology*, *24*(9), 2143–2155. <http://doi.org/10.1111/mec.13056>
- François de Neufchateau, N. (1807). *Lettre sur le robinier*, Galland, Paris.
- Friedman, J., & Barrett, S. C. H. (2008). High outcrossing in the annual colonizing species *Ambrosia artemisiifolia* (Asteraceae). *Annals of Botany*, *101*(9), 1303–1309. <http://doi.org/10.1093/aob/mcn039>
- Gilbert, K. J., Andrew, R. L., Bock, D. G., Franklin, M. T., Kane, N. C., Moore, J. S., ... Vines, T. H. (2012). Recommendations for utilizing and reporting population genetic analyses: The reproducibility of genetic clustering using the program structure. *Molecular Ecology*, *21*(20), 4925–4930. <http://doi.org/10.1111/j.1365-294X.2012.05754.x>
- Goudet, J. (2005). HIERFSTAT, a package for R to compute and test hierarchical F-statistics. *Molecular Ecology Notes*, *5*(1), 184–186. <http://doi.org/10.1111/j.1471-8278.2004.00828.x>
- Goudet, J. (2013). FSTAT: a computer program to calculate F-Statistics. *Journal of Heredity*, *104*, 586–590.

<http://doi.org/10.1093/jhered/est020>

Gunther, R.T. (1922). *Early British botanists and their gardens*, Oxford, England.

Halkett, F., Simon, J.-C., & Balloux, F. (2005). Tackling the population genetics of clonal and partially clonal organisms. *Trends in Ecology & Evolution*, *20*(4), 194–201. <http://doi.org/10.1016/j.tree.2005.01.001>

Henry, P., Le Lay, G., Goudet, J., Guisan, A., Jahodová, S., & Besnard, G. (2009). Reduced genetic diversity, increased isolation and multiple introductions of invasive giant hogweed in the western Swiss Alps. *Molecular Ecology*, *18*(13), 2819–2831. <http://doi.org/10.1111/j.1365-294X.2009.04237.x>

Hewitt, G. (2000). The genetic legacy of the quaternary ice ages. *Nature*, *405*(6789), 907–913. <http://doi.org/10.1038/35016000>

Hirsch, H., Richardson, D. M., & Le Roux, J. J. (2017). Introduction to the special issue: Tree invasions: Towards a better understanding of their complex evolutionary dynamics. *AoB PLANTS*, *9*(3). <http://doi.org/10.1093/aobpla/plx014>

Hollingsworth, M. L., & Bailey, J. P. (2000). Evidence for massive clonal growth in the invasive weed *Fallopia japonica* (Japanese Knotweed). *Botanical Journal of the Linnean Society*, *133*(4), 463–472. <http://doi.org/10.1111/j.1095-8339.2000.tb01589.x>

Hoos, P. M., Whitman Miller, A., Ruiz, G. M., Vrijenhoek, R. C., & Geller, J. B. (2010). Genetic and historical evidence disagree on likely sources of the Atlantic amethyst gem clam *Gemma gemma* (Totten, 1834) in California. *Diversity and Distributions*, *16*(4), 582–592. <http://doi.org/10.1111/j.1472-4642.2010.00672.x>

Jombart, T. (2008). adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics (Oxford, England)*, *24*(11), 1403–5. <http://doi.org/10.1093/bioinformatics/btn129>

Jombart, T., Devillard, S., & Balloux, F. (2010). Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genetics*, *11*(1), 94. <http://doi.org/10.1186/1471-2156-11-94>

Keller, S. R., Fields, P. D., Berardi, A. E., & Taylor, D. R. (2014). Recent admixture generates heterozygosity-fitness correlations during the range expansion of an invading species. *Journal of Evolutionary Biology*, *27*(3), 616–627. <http://doi.org/10.1111/jeb.12330>

Keller, S. R., & Taylor, D. R. (2008). History, chance and adaptation during biological invasion: Separating stochastic phenotypic evolution from response to selection. *Ecology Letters*, *11*, 852–866. <http://doi.org/10.1111/j.1461-0248.2008.01188.x>

Keller, S. R., & Taylor, D. R. (2010). Genomic admixture increases fitness during a biological invasion. *Journal*

- of Evolutionary Biology*, 23(8), 1720–1731. <http://doi.org/10.1111/j.1420-9101.2010.02037.x>
- Keresztesi, B. (1983). Breeding and cultivation of black locust, *Robinia pseudoacacia*, in Hungary. *Forest Ecology and Management*, 6, 217–244.
- Kopelman, N. M., Mayzel, J., Jakobsson, M., Rosenberg, N. A., & Mayrose, I. (2015). <scp>Clumpak</scp> : a program for identifying clustering modes and packaging population structure inferences across *K*. *Molecular Ecology Resources*, 15(5), 1179–1191. <http://doi.org/10.1111/1755-0998.12387>
- Kurokochi, H., & Hogetsu, T. (2014). Fine-scale initiation of non-native *Robinia pseudoacacia* riparian forests along the Chikumagawa River in central Japan. *Journal of Ecology and Environment*, 37(1), 21–29.
- Lavergne, S., & Molofsky, J. (2007). Increased genetic variation and evolutionary potential drive the success of an invasive grass. *Proceedings of the National Academy of Sciences of the United States of America*, 104(10), 3883–8. <http://doi.org/10.1073/pnas.0607324104>
- Lawson Handley, L. J., Estoup, A., Evans, D. M., Thomas, C. E., Lombaert, E., Facon, B., ... Roy, H. E. (2011). Ecological genetics of invasive alien species. *BioControl*, 56(4), 409–428. <http://doi.org/10.1007/s10526-011-9386-2>
- Lieseback, H., Yang, M., & Schneck, V. (2004). Genetic diversity and differentiation in a black locust (*Robinia pseudoacacia* L.) progeny test. *Forest Genetics*, 11(2), 151–161.
- Lieseback, H., & Schneck, V. (2012): Chloroplast DNA variation in planted and natural regenerated stands of black locust (*Robinia pseudoacacia* L.). *Silvae Genetica*, 61(1-2): 27-35.
- Little, E.L. (1971). Atlas of United States trees, vol. 1. Conifers and important hardwoods. U.S. Department of Agriculture, Miscellaneous Publication 1146. Washington, DC. 9 p., 313 maps.
- Lu, M., Krutovsky, K. V., Nelson, C. D., Koralewski, T. E., Byram, T. D., & Loopstra, C. A. (2016). Exome genotyping, linkage disequilibrium and population structure in loblolly pine (*Pinus taeda* L.). *BMC Genomics*, 17(1), 1–11. <http://doi.org/10.1186/s12864-016-3081-8>
- Merceron, N. R., Leroy, T., Chancerel, E., Romero-Severson, J., Borkowski, D., Ducouso, A., ... Kremer, A. (2017). Back to America: tracking the origin of European introduced populations of *Quercus rubra* L. *Genome*, 790(July), 1–54. <http://doi.org/10.1139/gen-2016-0187>
- Michaux, A. (1810). Histoire des arbres forestiers de l'Amérique septentrionale. Paris, France.
- Michener, D. C. (1988). The Introduction of Black Locust (*Robinia pseudoacacia* L.) to Massachusetts. *Arnoldia*, 48(4), 52–57.
- Miller, P. (1756-1759). The Gardeners Dictionary (7<sup>th</sup> ed.). London, England.
- Nadeau, S., Godbout, J., Lamothe, M., Gros-Louis, M.-C., Isabel, N., & Ritland, K. (2015). Contrasting patterns



- of genetic diversity across the ranges of *Pinus monticola* and *P. strobus*: A comparison between eastern and western North American postglacial colonization histories. *American Journal of Botany*, *102*(8), 1342–1355. <http://doi.org/10.3732/ajb.1500160>
- Pairon, M., Petitpierre, B., Campbell, M., Guisan, A., Broennimann, O., Baret, P. V., ... Besnard, G. (2010). Multiple introductions boosted genetic diversity in the invasive range of black cherry (*Prunus serotina*; Rosaceae). *Annals of Botany*, *105*(6), 881–890. <http://doi.org/10.1093/aob/mcq065>
- Pappert, R. A., Hamrick, J. L., & Donovan, L. A. (2000). Genetic variation in *Pueraria lobata* (Fabaceae), an introduced, clonal, invasive plant of the southeastern United States. *American Journal of Botany*, *87*(9), 1240–1245.
- Pessino, M., Chabot, E. T., Giordano, R., & DeWalt, R. E. (2014). Refugia and postglacial expansion of *Acroneuria frisoni* Stark & Brown (Plecoptera:Perlidae) in North America. *Freshwater Science*, *33*(1), 232–249. <http://doi.org/10.1086/675306>
- Petanidou, T., Godfree, R. C., Song, D. S., Kantsa, A., Dupont, Y. L., & Waser, N. M. (2012). Self-compatibility and plant invasiveness: Comparing species in native and invasive ranges. *Perspectives in Plant Ecology, Evolution and Systematics*, *14*(1), 3–12. <http://doi.org/10.1016/j.ppees.2011.08.003>
- Porras-Hurtado, L., Ruiz, Y., Santos, C., Phillips, C., Carracedo, J., & Lareu, M. V. (2013). An overview of STRUCTURE: Applications, parameter settings, and supporting software. *Frontiers in Genetics*, *4*(MAY), 1–13. <http://doi.org/10.3389/fgene.2013.00098>
- Potter, K. M., Campbell, A. R., Josserand, S. A., Nelson, C. D., & Jetton, R. M. (2017). Population isolation results in unexpectedly high differentiation in Carolina hemlock (*Tsuga caroliniana*), an imperiled southern Appalachian endemic conifer. *Tree Genetics and Genomes*, *13*(5). <http://doi.org/10.1007/s11295-017-1189-x>
- Pritchard, J. K. (2010). Documentation for structure software: Version 2.3.6(3), 321–326. <http://doi.org/10.1002/spe.4380060305>
- Pritchard, J. K., Stephens, M., & Donnelly, P. J. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*, 945–949.
- R Development Core Team. (2016). R: A Language and Environment for Statistical Computing. Version 3.3.1. *R Foundation for Statistical Computing Vienna Austria*, 0, {{ISBN}} 3--900051--07--0. <http://doi.org/10.1038/sj.hdy.6800737>
- Rambuda, T. D., & Johnson, S. D. (2004). Breeding systems of invasive alien plants in South Africa: does Baker's

- rule apply? *Diversity and Distributions*, 10(5–6), 409–416. <http://doi.org/10.1111/j.1366-9516.2004.00100.x>
- Richardson, D. M., & Rejmánek, M. (2011). Trees and shrubs as invasive alien species - a global review. *Diversity and Distributions*, 17(5), 788–809. <http://doi.org/10.1111/j.1472-4642.2011.00782.x>
- Rousset, F. (1997). Genetic Differentiation and Estimation of Gene Flow from FStatistics Under Isolation by Distance. *Genetics*, 145, 1219–1228.
- Saint-Jean de Crève Coeur, J.H. (1786). Mémoire sur la culture et les usages du faux Acacia dans les États-Unis de l'Amérique septentrionale. Paris, France.
- Schütt, P. (1994). *Robinia pseudoacacia*. In: Enzyklopädie der Holzgewächse: Handbuch und Atlas der Dendrologie. Wiley-VCH.
- Stoeckel, S., Grange, J., Fernández-Manjarres, J. F., Bilger, I., Frascaria-Lacoste, N., & Mariette, S. (2006). Heterozygote excess in a self-incompatible and partially clonal forest tree species -- *Prunus avium* L. *Molecular Ecology*, 15(8), 2109–18. <http://doi.org/10.1111/j.1365-294X.2006.02926.x>
- Stoeckel, S., & Masson, J. P. (2014). The exact distributions of  $F_{IS}$  under partial asexuality in small finite populations with mutation. *PLoS ONE*, 9(1), 1–15. <http://doi.org/10.1371/journal.pone.0085228>
- Straker, K. C., Quinn, L. D., Voigt, T. B., Lee, D. K., & Kling, G. J. (2015). Black Locust as a Bioenergy Feedstock: a Review. *Bioenergy Research*, 8(3), 1117–1135. <http://doi.org/10.1007/s12155-015-9597-y>
- Swenson, N. G., & Howard, D. J. (2005). Clustering of Contact Zones, Hybrid Zones, and Phylogeographic Breaks in North America. *The American Naturalist*, 166(5), 581–591. <http://doi.org/10.1086/491688>
- Takezaki, N., & Nei, M. (1996). Genetic distances and the setting of conservation priorities. *Biological Conservation*, 75(3), 311. [http://doi.org/10.1016/0006-3207\(96\)85988-X](http://doi.org/10.1016/0006-3207(96)85988-X)
- Thompson, G. D., Bellstedt, D. U., Richardson, D. M., Wilson, J. R. U., & Le Roux, J. J. (2015). A tree well travelled: Global genetic structure of the invasive tree *Acacia saligna*. *Journal of Biogeography*, 42(2), 305–314. <http://doi.org/10.1111/jbi.12436>
- Uller, T., & Leimu, R. (2011). Founder events predict changes in genetic diversity during human-mediated range expansions. *Global Change Biology*, 17(11), 3478–3485. <http://doi.org/10.1111/j.1365-2486.2011.02509.x>
- van Boheemen, L. A., Lombaert, E., Nurkowski, K. A., Gauffre, B., Rieseberg, L. H., & Hodgins, K. A. (2017). Multiple introductions, admixture and bridgehead invasion characterize the introduction history of *Ambrosia artemisiifolia* in Europe and Australia. *Molecular Ecology*, 26(20), 5421–5434. <http://doi.org/10.1111/mec.14293>
- Vítková, M., Müllerová, J., Sádlo, J., Pergl, J., & Pyšek, P. (2017). Black locust (*Robinia pseudoacacia*) beloved

and despised: A story of an invasive tree in Central Europe. *Forest Ecology and Management*, 384, 287–302. <http://doi.org/10.1016/j.foreco.2016.10.057>

Yang, M., He, Z., Huang, Y., Lu, L., Yan, Y., Hong, L., ... Shi, S. (2017). The emergence of the hyperinvasive vine, *Mikania micrantha* (Asteraceae), via admixture and founder events inferred from population transcriptomics. *Molecular Ecology*, 26(13), 3405–3423. <http://doi.org/10.1111/mec.14124>

Zenni, R. D., Bailey, J. K., & Simberloff, D. (2014). Rapid evolution and range expansion of an invasive plant are driven by provenance-environment interactions. *Ecology Letters*, 17(6), 727–735. <http://doi.org/10.1111/ele.12278>

## Supplementary information

Appendix 1. A) Complete information on the sampled populations. Range corresponds either to Europe (EU) or the USA (US) with indication of either the country or the state. N is number of individuals genotyped per population. GPS coordinates were provided using the WGS84 geographic projection. Altitude was extracted from the Worldclim database (REF) using the geographic coordinates. Information regarding harvesting are provided as the date of harvesting, names of the harvesters, origin of the funding, the type of stand (natural or Common Garden) and supplementary details regarding the location forest. B) Populations defined using the trees sampled at Michigan State University common garden. X Long. Used and Y Lat. used refer to the geographic coordinates of the created population (population name) provided using the WGS84 geographic projection. "Original" columns refer to the original sampling location and coordinates of the seeds. Trees from close original geographic location were gathered into the same created population, in order to get a sufficient number of individuals within each population.

A)	N°	Range	Country	Population name	Number X Longitude - Y Latitude W Altitude (m)	Sample used	Date of harw	Harvester 1	Harvester 2	Harvester 3	Harvester 4	Funding	Type of population
1	EU	France	Bartheleasse Avignon	19	4.81781	43.96459	19 LEAVES	02/06/2016	J. Thevenet	M. Corraud			Natural population
2	EU	Czech Republic	Brno	11	16.5178	49.0423	223 CAMBIUM	06/04/2016	X. Bouteiller	A. Martink	R. Longaner		Natural population
3	EU	Hungary	Budapest	20	19.106667	47.6625	109 LEAVES	May 2016	A. Monty				Natural population
4	EU	Romania	Carei	11	22.44863	47.660695	LEAVES	21/06/16	X. Bouteiller	S. Mariette	M. Pucheu	N. Cheval	Common Garden
5	EU	Belgium	Corphalie	10	5.2588	50.5389	100 LEAVES/SEED	Fall 2015	A. Monty				Natural population
6	EU	Poland	Drewnica	10	21.2508	52.2528	105 LEAVES	20/06/16	T. Woodja	M. Klisz			Natural population
7	EU	Spain	Gatos Galicia	12	-8.616667	42.383333	255 LEAVES	May 2016	R. Zas				Natural population
8	EU	Bulgaria	Gorna Oryahovitsa-K	12	25.694339	43.11872	LEAVES	21/06/16	X. Bouteiller	S. Mariette	M. Pucheu	N. Cheval	Common Garden
9	EU	Netherlands	Kelpen-Oler	12	5.824992	51.205358	28 LEAVES	Fall 2015	A. Monty				Natural population
10	EU	Germany	Klein	11	9.088889	49.991111	117 LEAVES	01/08/2016	A. Monty				Natural population
11	EU	France	La Flotte	6	-0.30478	44.38455	96 LEAVES	09/05/2016	X. Bouteiller	R. Segura			Natural population
12	EU	France	La Gouaneyre	6	-0.27251	44.37598	95 LEAVES	09/05/2016	X. Bouteiller	R. Segura			Natural population
13	EU	England	London Street Ham	10	0.14327	51.4332	39 LEAVES	14/02/2016	R. Segura	R. Burlett			Natural population
14	EU	England	London Wandsworth	12	0.16347	51.44574	12 LEAVES	15/02/2016	R. Segura	R. Burlett			Natural population
15	EU	Macedonia	Macedonia	10	21.570794	41.506548	LEAVES	21/06/16	X. Bouteiller	S. Mariette	M. Pucheu	N. Cheval	Common Garden
16	EU	Germany	Meppen	12	7.376964	52.703578	18 LEAVES	Fall 2015	A. Monty				Natural population
17	EU	Spain	Montseny	12	2.51236	41.8306	403 CAMBIUM	09/03/2016	X. Bouteiller	S. Mariette	J. Pino		Natural population
18	EU	Germany	Muntenberg	12	14.04553	52.55949	85 CAMBIUM	08/04/2016	X. Bouteiller	S. Mariette	V. Schneek		Natural population
19	EU	Bulgaria	Novi Pazar-Kulevcha	12	27.195081	43.3465317	LEAVES	21/06/16	X. Bouteiller	S. Mariette	M. Pucheu	N. Cheval	Common Garden
20	EU	Hungary	Nyirseg	12	19.041456	47.581206	LEAVES	21/06/16	X. Bouteiller	S. Mariette	M. Pucheu	N. Cheval	Common Garden
21	EU	Germany	Oberneldrich	12	6.815636	51.476075	30 LEAVES	Fall 2015	A. Monty				Natural population
22	EU	Poland	Pinczow	10	20.7022	50.2647	203 LEAVES	20/06/16	T. Woodja	M. Klisz			Natural population
23	EU	Poland	Poznan	12	16.80831	52.31109	128 CAMBIUM	10/04/2016	X. Bouteiller	S. Mariette			Natural population
24	EU	Germany	Priestsweg Naturpar	12	13.3578	52.46072	41 CAMBIUM	11/04/2016	X. Bouteiller	S. Mariette	D. Born		Natural population
25	EU	Hungary	Pusztavacs	10	19.505619	47.171762	LEAVES	21/06/16	X. Bouteiller	S. Mariette	M. Pucheu	N. Cheval	Common Garden
26	EU	France	Remy	12	2.675194	49.460286	21 LEAVES	Fall 2015	A. Monty				Natural population
27	EU	Greece	Rhodos	12	27.944444	36.287222	83 LEAVES	04/04/2016	C. Galanos				Natural population
28	EU	Slovakia	Slovakia	12	19.866761	48.719737	LEAVES	21/06/16	X. Bouteiller	S. Mariette	M. Pucheu	N. Cheval	Common Garden
29	EU	Poland	Szczecin	12	14.548233	53.33744	-3 CAMBIUM	09/04/2016	X. Bouteiller	S. Mariette			Natural population
30	EU	Turkey	Turkey	11	32.903719	40.159438	LEAVES	21/06/16	X. Bouteiller	S. Mariette	M. Pucheu	N. Cheval	Common Garden
31	EU	Netherlands	Uden	11	5.617717	51.685078	26 LEAVES	Fall 2015	A. Monty				Natural population
32	EU	Spain	Valencia	19	-0.784339	39.397275	349 LEAVES	May 2016	E. Laguna				Natural population
33	EU	Spain	Vitoria	13	-1.94188	43.21603	250 LEAVES	10/03/2016	X. Bouteiller	S. Mariette	P. Galkoebwa		Natural population
34	EU	Austria	Wien	12	16.47286	48.2515	158 BUDS	05/04/2016	X. Bouteiller	M. Van Loo			Natural population
35	US	PA - Pennsylvania	ALTOONA	11	-78.3832	40.4885	465 LEAVES	Fall 2016	X. Bouteiller				Natural population
36	US	VA - Virginia	Barbours Creek	21	-80.11	37.58	540 LEAVES	June 2014	C. Verdú	C. Mengal			Natural population
37	US	NC - North Carol	Blue Ridge	22	-82.6722	35.4567	1033 LEAVES	June 2014	C. Verdú	C. Mengal			Natural population
38	US	WV - West Virgir	CAMP CREEK	12	-81.1034	37.4876	655 LEAVES	Fall 2016	X. Bouteiller				Natural population
39	US	TN - Tennessee	CHATTANOOGA	12	-85.783	35.1198	309 LEAVES	Fall 2016	X. Bouteiller				Natural population
40	US	KY - Kentucky	DANIEL BOONE NF	12	-83.6452	37.7505	365 LEAVES	Fall 2016	X. Bouteiller	C. Mengal			Natural population
41	US	KY - Kentucky	Erlinne	22	-83.54	37.04	292 LEAVES	June 2014	C. Verdú	C. Mengal			Natural population
42	US	AR - Arkansas	FAYETTEVILLE	10	-94.205	36.0711	387 LEAVES	Fall 2016	X. Bouteiller				Natural population
43	US	WV - West Virgir	FORT MILL RIDGE	12	-78.7971	39.3267	249 LEAVES	Fall 2016	X. Bouteiller				Natural population
44	US	AR - Arkansas	FORT SMITH	12	-94.2895	35.3431	121 LEAVES	Fall 2016	X. Bouteiller				Natural population
45	US	OH - Ohio	Ironton	22	-82.46	38.8	261 LEAVES	June 2014	C. Verdú	C. Mengal			Natural population
46	US	WV - West Virgir	LEWISBURG	7	-80.3812	37.7825	608 LEAVES/SEED	Fall 2016	X. Bouteiller				Natural population
47	US	NC - North Carol	Locust Cove	22	-83.71	35.36	890 LEAVES	June 2014	C. Verdú	C. Mengal			Natural population
48	US	KY - Kentucky	Morehead	22	-83.4663	38.0909	237 LEAVES	June 2014	C. Verdú	C. Mengal			Natural population
49	US	AR - Arkansas	OUACHITA	12	-93.8374	34.4492	304 LEAVES	Fall 2016	X. Bouteiller				Natural population
50	US	WV - West Virgir	Perry	22	-78.66	39	504 LEAVES	June 2014	C. Verdú	C. Mengal			Natural population
51	US	AR - Arkansas	Pleasant Hill	4	-93.46	35.59	324 LEAVES	June 2014	C. Verdú	C. Mengal			Natural population
52	US	NC - North Carol	SHOOTING CREEK	11	-83.6284	35.0946	761 LEAVES	Fall 2016	X. Bouteiller				Natural population
53	US	WV - West Virgir	Slayfork	17	-80	38.18	704 LEAVES	June 2014	C. Verdú	C. Mengal			Natural population
54	US	VA - Virginia	Stokesville	22	-79.3	38.28	607 LEAVES	June 2014	C. Verdú	C. Mengal			Natural population
55	US	MD - Maryland	US GRP 1	11	-78.75	39.65	247 CAMBIUM	Fall 2016	X. Bouteiller	P. Bloese			Common Garden
56	US	WV - West Virgir	US GRP 2	11	-81.1	39.06667	235 CAMBIUM	Fall 2016	X. Bouteiller	P. Bloese			Common Garden
57	US	VA - Virginia	US GRP 3	6	-79.933333	37.266667	286 CAMBIUM	Fall 2016	X. Bouteiller	P. Bloese			Common Garden
58	US	KY - Kentucky	US GRP 4	6	-84.5	38.033333	303 CAMBIUM	Fall 2016	X. Bouteiller	P. Bloese			Common Garden
59	US	KY - Kentucky	US GRP 5	8	-84.533333	38.65	277 CAMBIUM	Fall 2016	X. Bouteiller	P. Bloese			Common Garden
60	US	KY - Kentucky	US GRP 6	7	-83.683333	36.75	584 CAMBIUM	Fall 2016	X. Bouteiller	P. Bloese			Common Garden
61	US	AR - Arkansas	Victor	22	-93.05	35.65	513 LEAVES	June 2014	C. Verdú	C. Mengal			Natural population
62	US	OH - Ohio	WAYNE NF	12	-82.5935	38.6577	267 LEAVES	Fall 2016	X. Bouteiller				Natural population
63	US	VA - Virginia	Whiteop	22	-81.6562	36.7694	679 LEAVES	June 2014	C. Verdú	C. Mengal			Natural population

B)

Range	Country/State	Pop name	Number of gens	X Long. Used	Y Lat. used	Original State	Original Town	Original X Long.	Original Y Lat.
US	Maryland	US_GRP_1	15	-78.75	39.65	Maryland	LaVale	-78.8	39.63333333
						Maryland	Cumberland	-78.75833333	39.66333333
						Maryland	Lumberland	-78.75	39.65
US	West Virginia	US_GRP_2	11	-81.1	39.06666667	West Virginia	St. Marys	-81.21666667	39.4
						West Virginia	Woodsfield	-81.1	39.06666667
						Ohio	Fly	-81.08333333	39.55
US	Virginia	US_GRP_3	6	-79.93333333	37.26666667	Virginia	Blacksburg	-80.4	37.21666667
						Virginia	Christiansburg	-80.4	37.11666667
US	Kentucky	US_GRP_4	6	-84.5	38.03333333	Kentucky	Lexington	-84.5	38.03333333
US	Kentucky	US_GRP_5	8	-84.53333333	38.65	Kentucky	Dry Ridge	-84.61666667	38.63333333
						Kentucky	Williamston	-84.53333333	38.65
						Kentucky	Falmouth	-84.33333333	38.66666667
US	Kentucky	US_GRP_6	7	-83.68333333	36.75	Kentucky	Pineville	-83.68333333	36.75
						Tennessee	Maynardville	-83.83333333	36.28333333
						Tennessee	Lancing	-84.63333333	36.13333333

Appendix 2. Clonality results among the sampled populations. N: sample size of the population, G: number of unique genotypes, R: clonal diversity index. FIS Mean: mean value of the FIS for each population and FIS Lower and Higher Interval are the boundaries of the FIS 95% confidence interval estimated by 1000 bootstrap over loci.

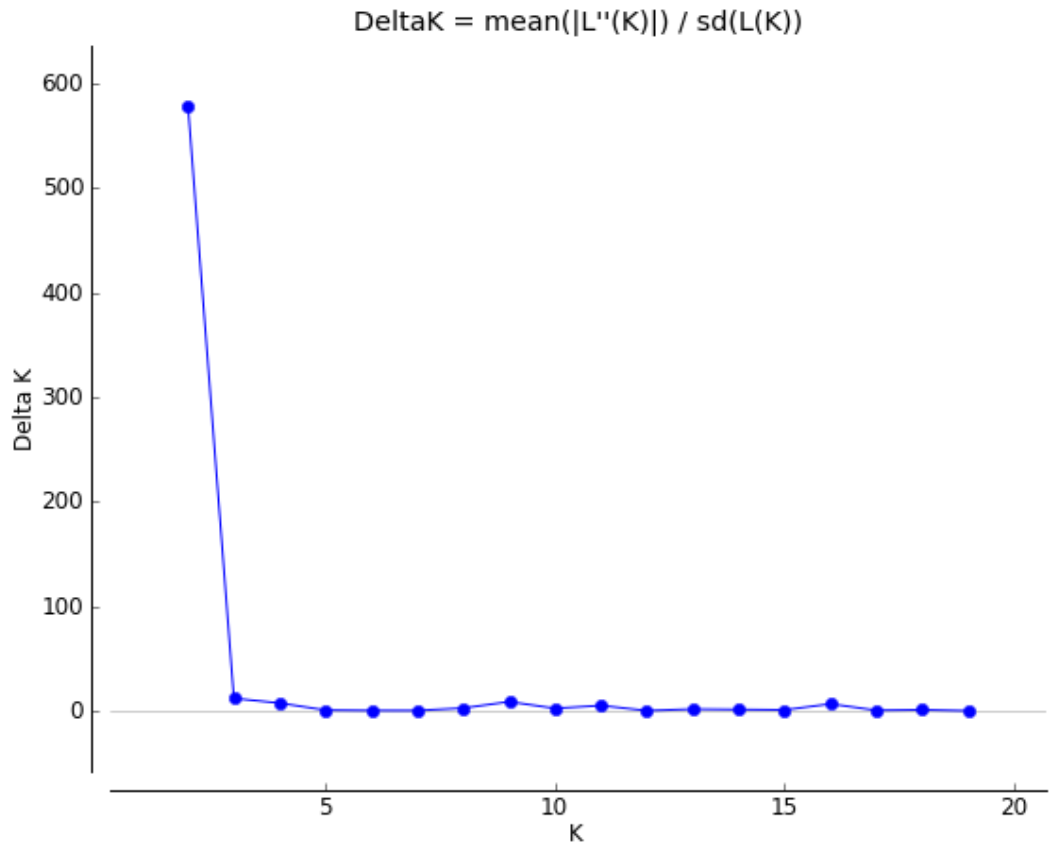
Range	Pop	N	G	R	Fis Mean	FIS Lower	Inte	Fis Higher	Interval
1 EU	Barthelasse_Avignon	19	18	0.944444444	0.1119	0.0477	0.1789	0.1789	
2 EU	Brno	11	5	0.4	-0.0256	-0.1311	0.0843	0.0843	
3 EU	Budapest	20	13	0.63157895	0.115	0.0466	0.1836	0.1836	
4 EU	Carei	11	11	1	0.0961	0.0259	0.1654	0.1654	
5 EU	Corphalie	10	10	1	0.1201	0.0424	0.1953	0.1953	
6 EU	Drewnica	10	10	1	0.0602	-0.0151	0.1384	0.1384	
7 EU	Gafos_Galicia	12	5	0.36363636	0.0562	-0.0466	0.1662	0.1662	
8 EU	Gorna_Oryahovitsa-Kozarevec	12	12	1	0.0996	0.0261	0.1759	0.1759	
9 EU	Kelpen-Oler	12	12	1	0.1188	0.0391	0.1973	0.1973	
10 EU	Klein	11	11	1	0.0101	-0.0986	0.1237	0.1237	
11 EU	La_Flotte	6	3	0.4	0.0617	-0.0874	0.2047	0.2047	
12 EU	La_Gouaneyre	6	5	0.8	0.1524	0.0199	0.2813	0.2813	
13 EU	London_Streat_Ham	10	7	0.66666667	0.1443	0.0234	0.2646	0.2646	
14 EU	London_Wandsworth	10	8	0.77777778	0.0849	0.0022	0.1655	0.1655	
15 EU	Macedonia	12	12	1	0.0861	0.0171	0.1566	0.1566	
16 EU	Meppen	12	8	0.63636364	-0.0974	-0.2105	0.0261	0.0261	
17 EU	Montseny	12	12	1	0.0679	-0.0092	0.1468	0.1468	
18 EU	Munchenberg	12	4	0.27272727	-0.1803	-0.3194	-0.0298	-0.0298	
19 EU	Novi_Pazar-Kulevcha	12	12	1	0.0906	0.0237	0.1605	0.1605	
20 EU	Nyirsegi	12	12	1	0.0866	0.0037	0.1728	0.1728	
21 EU	Obermeidenrich	12	12	1	0.0659	1.00E-04	0.1344	0.1344	
22 EU	Pinczow	10	10	1	0.1542	0.0667	0.2487	0.2487	
23 EU	Poznan	12	12	1	0.1188	0.0367	0.2032	0.2032	
24 EU	Priesterweg_Naturpark	12	12	1	0.1045	0.0349	0.1769	0.1769	
25 EU	Pusztavacs	10	10	1	0.0588	-0.0207	0.1467	0.1467	
26 EU	Remy	12	8	0.63636364	0.1041	0.0142	0.1932	0.1932	
27 EU	Rhodos	12	9	0.72727273	0.1039	0.0143	0.1971	0.1971	
28 EU	Slovakia	12	12	1	0.161	0.0751	0.2512	0.2512	
29 EU	Szczecin	12	12	1	0.0761	0.0033	0.151	0.151	
30 EU	Turkey	11	11	1	0.0809	-0.0047	0.1718	0.1718	
31 EU	Uden	11	5	0.4	0.1329	0.0153	0.2465	0.2465	
32 EU	Valencia	19	6	0.27777778	0.0656	-0.0475	0.1788	0.1788	



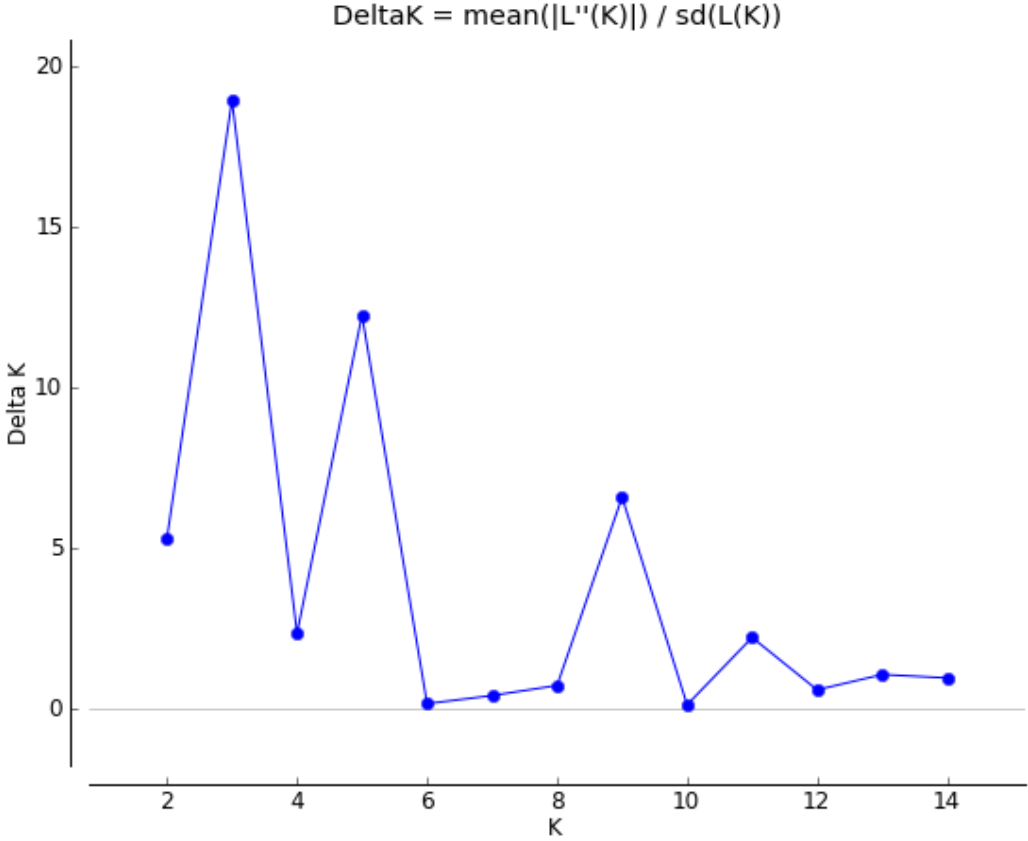
33 EU	Vitoria	13	13	1	0.1317	0.0543	0.2104
34 EU	Wien	12	12	1	0.0691	-0.013	0.1545
35 US	ALTOONA	11	10	0.9	0.1145	0.0358	0.1923
36 US	Barbours_Creek	21	20	0.95	0.1256	0.0642	0.192
37 US	Blue_Ridge	22	21	0.95238095	0.1182	0.0596	0.1763
38 US	CAMP_CREEK	12	12	1	0.0726	-0.008	0.1536
39 US	CHATTANOOGA	12	11	0.90909091	0.0651	-0.02	0.1509
40 US	DANIEL_BOONE_NF	12	12	1	0.147	0.067	0.2301
41 US	Eriline	22	22	1	0.082	0.0267	0.1376
42 US	FAYETTEVILLE	10	9	0.88888889	0.1855	0.0945	0.2768
43 US	FORT_MILL_RIDGE	12	12	1	0.0555	-0.0118	0.1244
44 US	FORT_SMITH	12	10	0.81818182	0.1167	0.0269	0.2149
45 US	Ironton	22	17	0.76190476	0.1047	0.0327	0.1814
46 US	LEWISBURG	7	7	1	0.0372	-0.061	0.1367
47 US	Locust_Cove	22	18	0.80952381	0.1148	0.0559	0.179
48 US	Morehead	22	19	0.85714286	0.1345	0.0644	0.2087
49 US	OUACHITA	12	9	0.72727273	0.0972	0.0154	0.178
50 US	Perry	22	22	1	0.0972	0.0256	0.1682
51 US	Pleasant_Hill	4	3	0.66666667	-0.0582	-0.2	0.0993
52 US	SHOOTING_CREEK	11	11	1	0.1112	0.0392	0.1857
53 US	Slatyfork	17	17	1	0.1079	0.0401	0.1749
54 US	Stokesville	22	22	1	0.1314	0.0599	0.2058
55 US	US_GRP_1	15	15	1	0.095	0.0297	0.1592
56 US	US_GRP_2	11	11	1	0.133	0.0521	0.2163
57 US	US_GRP_3	6	6	1	0.1228	0.0183	0.2284
58 US	US_GRP_4	6	6	1	0.0952	-0.0013	0.1958
59 US	US_GRP_5	8	8	1	0.1745	0.0841	0.2678
60 US	US_GRP_6	7	7	1	0.1565	0.047	0.2597
61 US	Victor	22	17	0.76190476	0.1512	0.0848	0.2236
62 US	WAYNE_NF	12	12	1	0.132	0.0489	0.2158
63 US	Whiteop	22	20	0.9047619	0.1137	0.049	0.1829

Appendix 3.  $\Delta K$  calculated according to the Evanno et al., (2005) method on the STRUCTURE output using STRUCTURE HARVESTER. A)  $\Delta K$  for the initial dataset after clone removal over both ranges, B)  $\Delta K$  for the American range dataset and C)  $\Delta K$  for the European range dataset.

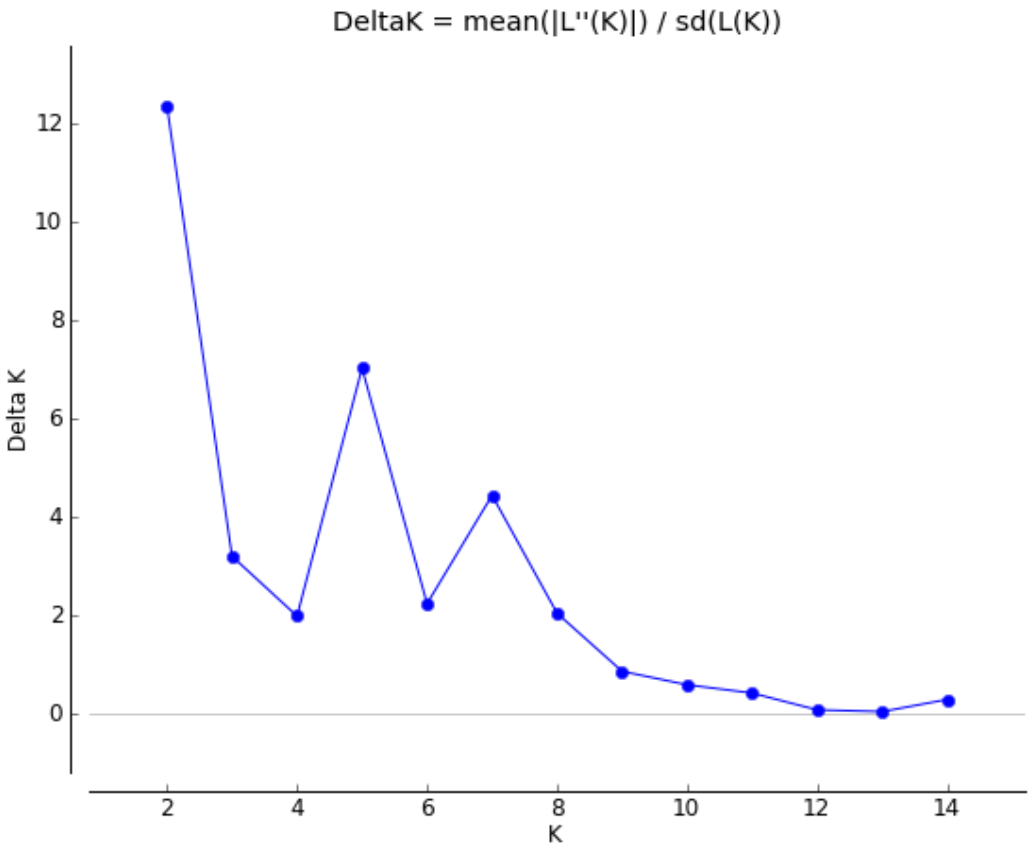
A –  $\Delta K$  plot produced by STRUCTURE HARVESTER for the 10 runs from  $K = 1$  to 20 of the whole data set (US + European populations. Most likely  $K$  is for  $K = 2$ .



B – ΔK plot produced by STRUCTURE HARVESTER for the 10 runs from K = 1 to 20 of the US data (US populations). Most likely K is for K = 3.



C – ΔK plot produced by STRUCTURE HARVESTER for the 10 runs from K = 1 to 20 of the European data (European populations). Most likely K is for K = 2.



Appendix 4. Proportion of individuals in each population assigned to the K clusters of the most likely K based on the STRUCTURE analysis conducted on the among ranges dataset. K2\_1: Proportion of individuals significantly assigned to cluster K2\_1; K2\_2: Proportion of individuals assigned to cluster K2\_2; Admixed: Proportion of admixed individuals and Assignment\_cluster: assignment cluster of the population based on the highest proportion of individuals significantly assigned.

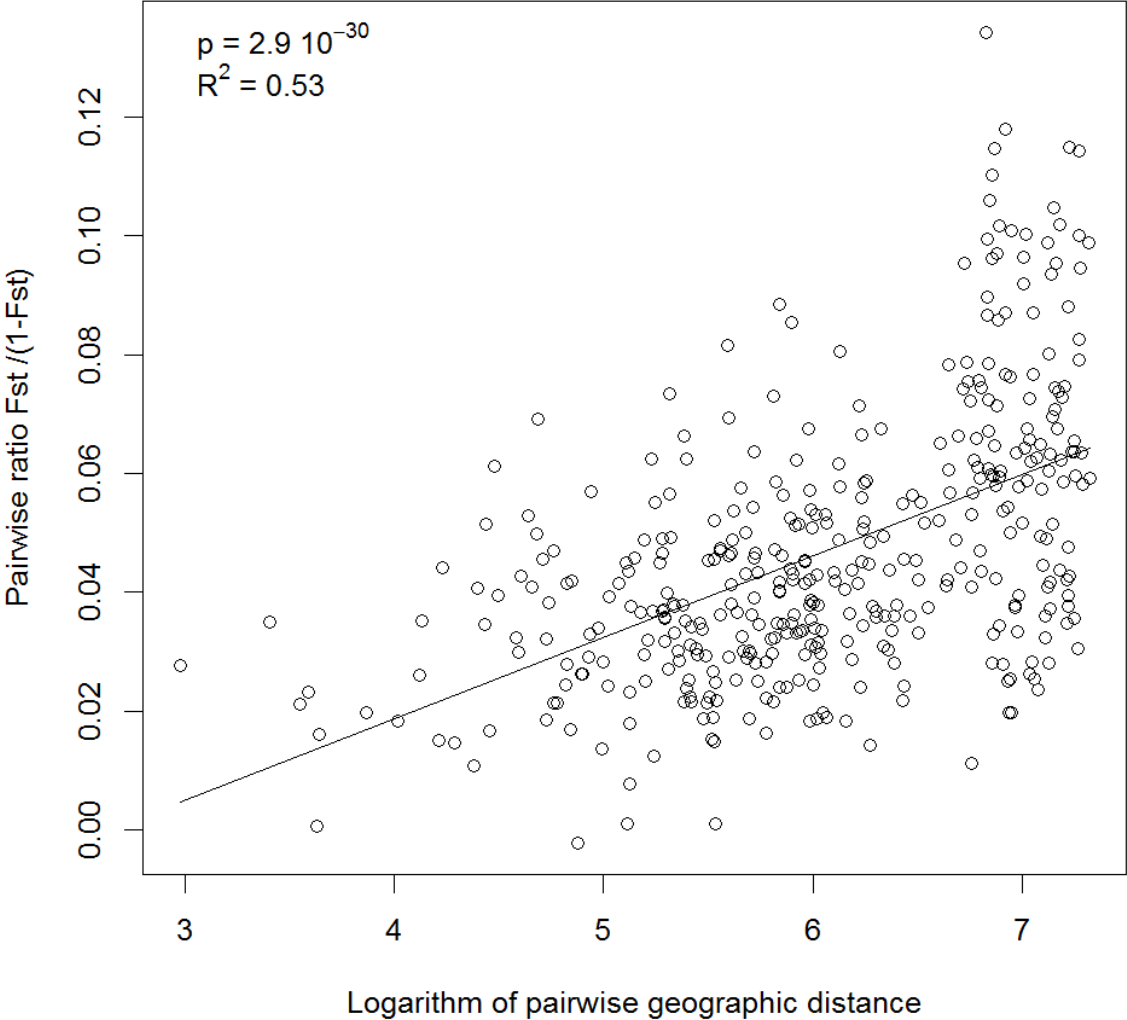
Pop	Range	Country/State	Pop_name	K2_1	K2_2	Admixed	Assign. Cluster
1	EU	France	Barthelasse_Avignon	0	0.61111	0.388889	K2_2
2	EU	Czech Republic	Brno	0	0.6	0.4	K2_2
3	EU	Hungary	Budapest	0	0.76923	0.230769	K2_2
4	EU	Romania	Carei	0	0.81818	0.181818	K2_2
5	EU	Belgium	Corphalie	0	0.9	0.1	K2_2
6	EU	Poland	Drewnica	0	0.6	0.4	K2_2
7	EU	Spain	Gafos_Galicia	0	0.6	0.4	K2_2
8	EU	Bulgaria	Ornina_Oryahovitsa-Kozarev	0	0.75	0.25	K2_2
9	EU	Netherland	Kelpen-Oler	0	0.83333	0.166667	K2_2
10	EU	Germany	Klein	0	0.63636	0.363636	K2_2
11	EU	France	La_Flotte	0	1	0	K2_2
12	EU	France	La_Gouaneyre	0	0.8	0.2	K2_2
13	EU	England	London_Streat_Ham	0	0.57143	0.428571	K2_2
14	EU	England	London_Wandworth	0	0.75	0.25	K2_2
15	EU	Macedonia	Macedonia	0	0.83333	0.166667	K2_2
16	EU	Germany	Meppen	0	1	0	K2_2
17	EU	Spain	Montseny	0.083	0.83333	0.083333	K2_2
18	EU	Germany	Munchenberg	0	1	0	K2_2
19	EU	Bulgaria	Novi_Pazar-Kulevcha	0	0.66667	0.333333	K2_2
20	EU	Hungary	Nyirsegi	0	0.75	0.25	K2_2
21	EU	Germany	Obermeidenrich	0	0.75	0.25	K2_2
22	EU	Poland	Pinczow	0	0.4	0.6	K2_2
23	EU	Poland	Poznan	0	0.75	0.25	K2_2
24	EU	Germany	Priesterweg_Naturpark	0	0.66667	0.333333	K2_2
25	EU	Hungary	Pusztavacs	0	0.7	0.3	K2_2
26	EU	France	Remy	0	0.75	0.25	K2_2
27	EU	Greece	Rhodos	0	0.55556	0.444444	K2_2
28	EU	Slovakia	Slovakia	0	0.75	0.25	K2_2
29	EU	Poland	Szczecin	0	0.75	0.25	K2_2
30	EU	Turkey	Turkey	0	0.63636	0.363636	K2_2
31	EU	Netherland	Uden	0	0.8	0.2	K2_2
32	EU	Spain	Valencia	0	0.83333	0.166667	K2_2
33	EU	Spain	Vitoria	0	0.61538	0.384615	K2_2
34	EU	Austria	Wien	0	0.75	0.25	K2_2
35	US	PA - Pennsylvania	ALTOONA	0.3	0.2	0.5	K2_1
36	US	VA - Virginia	Barbours_Creek	0.6	0	0.4	K2_1
37	US	NC - North Carolina	Blue_Ridge	0.667	0	0.333333	K2_1
38	US	WV - West Virginia	CAMP_CREEK	0.833	0	0.166667	K2_1
39	US	TN - Tennessee	CHATTANOOGA	0.818	0	0.181818	K2_1
40	US	KY - Kentucky	DANIEL_BOONE_NF	0.75	0	0.25	K2_1
41	US	KY - Kentucky	Eriline	0.591	0.18182	0.227273	K2_1
42	US/OZARK	AR - Arkansas	FAYETTEVILLE	0.778	0	0.222222	K2_1
43	US	WV - West Virginia	FORT_MILL_RIDGE	0.833	0	0.166667	K2_1
44	US/OZARK	AR - Arkansas	FORT_SMITH	0.9	0	0.1	K2_1
45	US	OH - Ohio	Ironton	0.824	0	0.176471	K2_1
46	US	WV - West Virginia	LEWISBURG	0.286	0	0.714286	K2_1
47	US	NC - North Carolina	Locust_Cove	0.833	0	0.166667	K2_1
48	US	KY - Kentucky	Morehead	0.842	0	0.157895	K2_1
49	US/OZARK	AR - Arkansas	OUACHITA	1	0	0	K2_1

50	US	WV - West Virginia	Perry	0.727	0	0.272727	K2_1
51	US/OZARK	AR - Arkansas	Pleasant_Hill	1	0	0	K2_1
52	US	NC - North Carolina	SHOOTING_CREEK	0.727	0	0.272727	K2_1
53	US	WV - West Virginia	Slatyfork	0.647	0	0.352941	K2_1
54	US	VA - Virginia	Stokesville	0.818	0	0.181818	K2_1
55	US	MD - Maryland	US_GRP_1	1	0	0	K2_1
56	US	WV - West Virginia	US_GRP_2	0.909	0	0.090909	K2_1
57	US	VA - Virginia	US_GRP_3	0.833	0.16667	0	K2_1
58	US	KY - Kentucky	US_GRP_4	0.833	0	0.166667	K2_1
59	US	KY - Kentucky	US_GRP_5	0.75	0	0.25	K2_1
60	US	KY - Kentucky	US_GRP_6	0.571	0	0.428571	K2_1
61	US/OZARK	AR - Arkansas	Victor	1	0	0	K2_1
62	US	OH - Ohio	WAYNE_NF	0.75	0.08333	0.166667	K2_1
63	US	VA - Virginia	Whiteop	0.85	0	0.15	K2_1

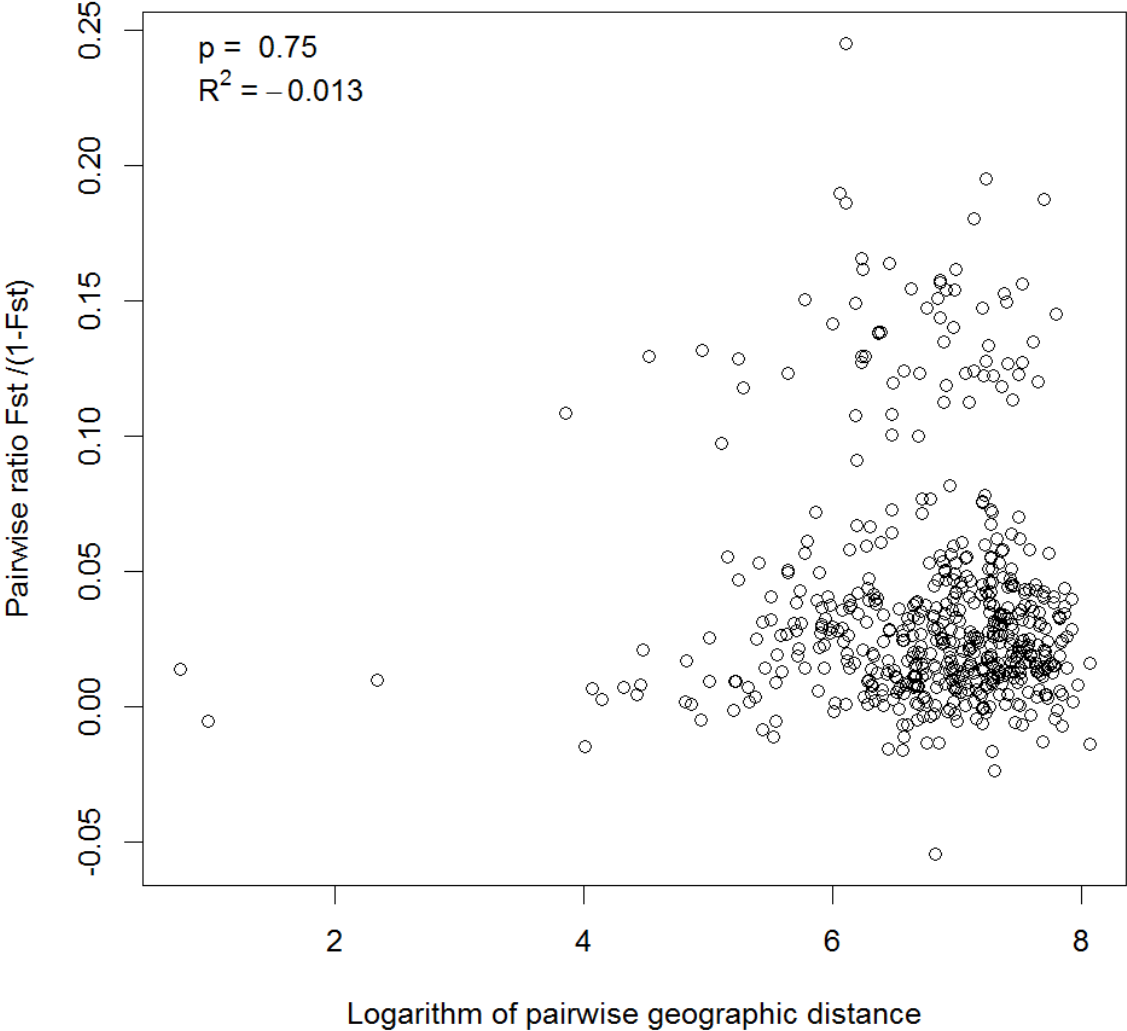


Appendix 5. IBD within America (A) and within Europe (B). ratio  $F_{ST} / (1 - F_{ST})$  from pairwise  $F_{ST}$  was calculated and plotted against the logarithm of the pairwise geographic distances among populations (Rousset 1997)

### Isolation by distance in America



### Isolation by distance in Europe



Appendix 6. Proportion of individuals in each population assigned to the K clusters of the most likely K based on the STRUCTURE analysis conducted the within range datasets. Within America: K3\_1\_US: Proportion of individuals significantly assigned to cluster K3\_1\_US; K3\_2\_US: Proportion of individuals assigned to cluster K3\_2\_US; K3\_3\_US: Proportion of individuals assigned to cluster K3\_3\_US; Admixed\_US: Proportion of admixed individuals and Assignment\_cluster\_US: assignment cluster of the population based on the highest proportion of individuals significantly assigned. Within Europe: K2\_1\_EU: Proportion of individuals significantly assigned to cluster K2\_1\_EU; K2\_2\_EU: Proportion of individuals assigned to cluster K2\_2\_EU; Admixed\_EU: Proportion of admixed individuals and Assignment\_cluster\_EU: assignment cluster of the population based on the highest proportion of individuals significantly assigned.

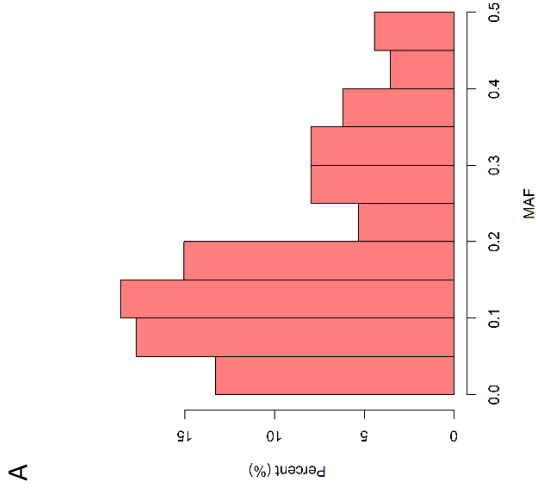
Range	Country/State	Pop_name	K3_1_US	K3_2_US	K3_3_US	Admixed_US	Assignment_clus	K2_1_EU	K2_2_EU	Admixed_EU	Assignment_cluster_EU
EU	France	Barthelesse_Avignon						0.1111	0.0000	0.8889	K2_1_EU
EU	Czech Republic	Brno						0.4000	0.0000	0.6000	K2_1_EU
EU	Hungary	Budapest						0.2308	0.0000	0.7692	K2_1_EU
EU	Romania	Carei						0.0909	0.0000	0.9091	K2_1_EU
EU	Belgium	Corphalie						0.0000	0.2000	0.8000	K2_2_EU
EU	Poland	Drewnica						0.0000	0.2000	0.8000	K2_2_EU
EU	Spain	Gafos_Galicia						0.0000	0.0000	1.0000	Admixed_EU
EU	Bulgaria	Gorna_Oryahovitsa-Kozarevec						0.0833	0.0000	0.9167	K2_1_EU
EU	Netherlands	Kelpen-Oler						0.2500	0.0000	0.7500	K2_1_EU
EU	Germany	Klein						0.2727	0.0000	0.7273	K2_1_EU
EU	France	La_Flotte						0.0000	0.0000	1.0000	Admixed_EU
EU	France	La_Gouaneyre						0.0000	0.4000	0.6000	K2_2_EU
EU	England	London_Streat_Ham						0.0000	0.1429	0.8571	K2_2_EU
EU	England	London_Wandsworth						0.1250	0.2500	0.6250	K2_2_EU
EU	Macedonia	Macedonia						0.3333	0.0833	0.5833	K2_1_EU
EU	Germany	Meppen						0.6250	0.0000	0.3750	K2_1_EU
EU	Spain	Montseny						0.0000	0.2500	0.7500	K2_2_EU
EU	Germany	Munchenber						0.0000	1.0000	0.0000	K2_2_EU
EU	Bulgaria	Novi_Pazar-Kulevcha						0.1667	0.0000	0.8333	K2_1_EU
EU	Hungary	Nyirsegi						0.1667	0.0833	0.7500	K2_1_EU
EU	Germany	Obermeidenrich						0.0000	0.0000	1.0000	Admixed_EU
EU	Poland	Pinczow						0.0000	0.1000	0.9000	K2_2_EU
EU	Poland	Poznan						0.0833	0.0833	0.8333	Admixed_EU
EU	Germany	Priesterweg_Naturpark						0.0833	0.0833	0.8333	Admixed_EU
EU	Hungary	Pusztavacs						0.1000	0.0000	0.9000	K2_1_EU
EU	France	Remy						0.0000	0.2500	0.7500	K2_2_EU
EU	Greece	Rhodos						0.0000	0.3333	0.6667	K2_2_EU
EU	Slovakia	Slovakia						0.2500	0.0000	0.7500	K2_1_EU
EU	Poland	Szczecin						0.0833	0.2500	0.6667	K2_2_EU
EU	Turkey	Turkey						0.0000	0.0909	0.9091	K2_2_EU
EU	Netherlands	Uden						0.0000	0.2000	0.8000	K2_2_EU
EU	Spain	Valencia						0.1667	0.0000	0.8333	K2_1_EU

EU	Spain	Vitoria									
EU	Austria	Wien									
US	PA - Pennsylvania	ALTOONA	0.0000	0.2000	0.0000	0.8000	K3_2_US	0.0769	0.1538	0.7692	K2_2_EU
US	VA - Virginia	Barbours_Creek	0.0000	0.0000	0.3000	0.7000	K3_3_US	0.0833	0.1667	0.7500	K2_2_EU
US	NC - North Carolina	Blue_Ridge	0.0000	0.0000	0.0952	0.9048	K3_3_US				
US	WV - West Virginia	CAMP_CREEK	0.0000	0.0000	0.0833	0.9167	K3_3_US				
US	TN - Tennessee	CHATTANOOGA	0.0000	0.0000	0.0000	1.0000	Admixed_US				
US	KY - Kentucky	DANIEL_BOONE_NF	0.0000	0.0000	0.0000	1.0000	Admixed_US				
US	KY - Kentucky	Erline	0.0000	0.0000	0.0455	0.9545	K3_3_US				
US/OZAR	AR - Arkansas	FAYETTEVILLE	0.2222	0.0000	0.0000	0.7778	K3_1_US				
US	WV - West Virginia	FORT_MILL_RIDGE	0.0000	0.0833	0.0000	0.9167	K3_2_US				
US/OZAR	AR - Arkansas	FORT_SMITH	0.8000	0.0000	0.0000	0.2000	K3_1_US				
US	OH - Ohio	Ironton	0.0000	0.1176	0.1176	0.7647	Admixed_US				
US	WV - West Virginia	LEWISBURG	0.1429	0.0000	0.1429	0.7143	Admixed_US				
US	NC - North Carolina	Locust_Cove	0.0000	0.1111	0.0556	0.8333	K3_2_US				
US	KY - Kentucky	Morehead	0.0000	0.0526	0.0000	0.9474	K3_2_US				
US/OZAR	AR - Arkansas	OJACHITA	0.5556	0.0000	0.0000	0.4444	K3_1_US				
US	WV - West Virginia	Perry	0.0000	0.2273	0.0000	0.7727	K3_2_US				
US/OZAR	AR - Arkansas	Pleasant_Hill	0.3333	0.0000	0.0000	0.6667	K3_1_US				
US	NC - North Carolina	SHOOTING_CREEK	0.0000	0.0000	0.0000	1.0000	Admixed_US				
US	WV - West Virginia	Slatyfork	0.0000	0.0000	0.2353	0.7647	K3_3_US				
US	VA - Virginia	Stokesville	0.0000	0.0455	0.0455	0.9091	Admixed_US				
US	MD - Maryland	US_GRP_1	0.0667	0.1333	0.0000	0.8000	K3_2_US				
US	WV - West Virginia	US_GRP_2	0.0000	0.1818	0.0000	0.8182	K3_2_US				
US	VA - Virginia	US_GRP_3	0.0000	0.3333	0.0000	0.6667	K3_2_US				
US	KY - Kentucky	US_GRP_4	0.1667	0.5000	0.0000	0.3333	K3_2_US				
US	KY - Kentucky	US_GRP_5	0.0000	0.0000	0.0000	1.0000	Admixed_US				
US	KY - Kentucky	US_GRP_6	0.0000	0.1429	0.0000	0.8571	K3_2_US				
US/OZAR	AR - Arkansas	Victor	0.5294	0.0000	0.0000	0.4706	K3_1_US				
US	OH - Ohio	WAYNE_NF	0.0000	0.0000	0.1667	0.8333	K3_3_US				

Appendix 7. A) MAF distributions assessed using the initial dataset after clone removal (i.e. 720 genotyped individuals using 113 SNPs, A B and C) and using the additional dataset (i.e. 163 genotyped individuals using 251 SNPs, D E and F). In blue for the American range, in red for the European range, in COLOR for both ranges together. G -Distribution of the Total number of alleles estimated by 1000 bootstrap replications over all loci and individuals; in blue the American range and in red the European range. B) Genetic diversity indices per range estimated using Fstat (Goudet 2013) using the additional dataset (i.e. 163 genotyped individuals using 251 SNPs). Ho: Observed Heterozygosity, He: Expected Heterozygosity, AR: Allelic Richness, TNA: Total Number of Alleles. Bold characters indicates significantly differences ( $p < 0.05$ ).

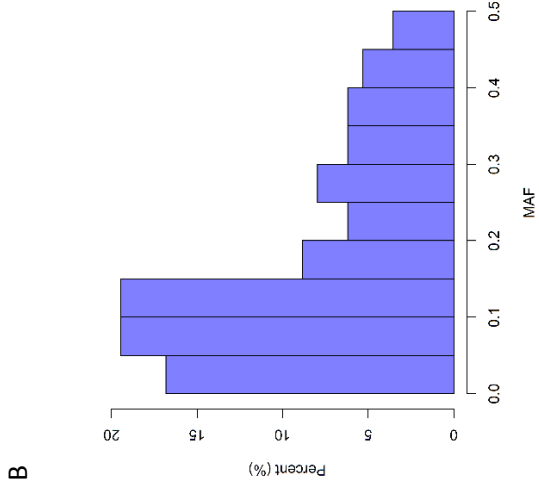
A)

European range



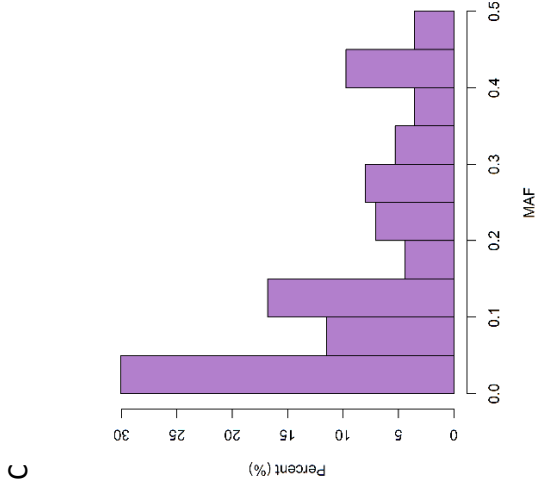
A

American range



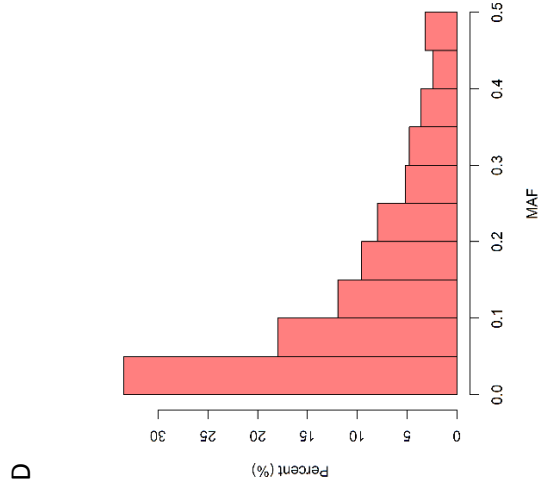
B

Both ranges

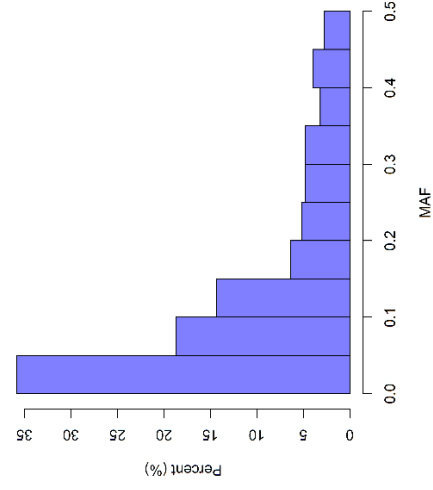


C

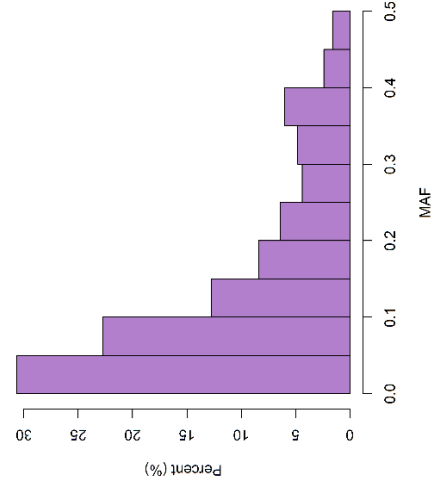
Additional dataset



D



E



F



B)

	Range	Ho	He	TNA	AR
Initial Dataset (113 SNPs)	EU	0.24	0.27	226	2.00
	US	0.22	0.26	226	2.00
Additional Dataset (251 SNPs)	EU	0.19	0.21	<b>485</b>	<b>1.91</b>
	US	0.17	0.20	<b>494</b>	<b>1.96</b>

## **Chapitre 2**

**Using automated sanding to homogeneously break seed dormancy in black locust  
(*Robinia pseudoacacia* L., Fabaceae)**

# Using automated sanding to homogeneously break seed dormancy in black locust (*Robinia pseudoacacia* L., Fabaceae)

Xavier P. Bouteiller<sup>1§</sup>, Annabel J. Porté<sup>1§\*</sup>, Stéphanie Mariette<sup>1</sup> and Arnaud Monty<sup>2</sup>

<sup>1</sup>Biogeco, INRA, Université de Bordeaux, 33615 Pessac, France; <sup>2</sup>Gembloux Agro-bio Tech, Université de Liège, Gembloux, Belgique

(Received 2 June 2016; accepted after revision 4 April 2017)

## Abstract

Physical dormancy of *Robinia pseudoacacia* seeds makes it a challenge for scientists and forest managers to obtain a homogeneous germination for larger seed samples. Water imbibition of the seeds can be achieved through manual piercing of the seed coat, but this method remains time consuming and heterogeneous. We tested several ecologically friendly methods to break seed dormancy, including manual pin puncture, water soaking, oven dry-heating (two temperatures) and sanding. Sanding was performed using an automatic grinder to control shaking duration (three durations) and get a homogeneous scraping of the coat. All methods, except dry-heating, resulted in successful dormancy breaking; water soaking was the least efficient method, attaining 57% germination. Sanding proved to be as efficient as puncturing (97%) but long duration sanding (10 or 15 min) could damage cotyledons, which would impede further development of the plant. Short-time sanding (5 min) proved to be the best method to reach high total germination and healthy (undamaged cotyledon) seedlings, and was successfully applied to 500 seeds. The reference puncture method and the automatic sanding were also tested on seeds of nine Fabaceae species and proved to be efficient for some species. Automated sanding can thus be used as a standard to break physical dormancy of black locust or other Fabaceae seeds to allow further comparative studies of plant populations or genotypes.

**Keywords:** automated grinder, dormancy breaking, Fabaceae, germination, sanding, scarification, standard protocol

## Introduction

Black locust (*Robinia pseudoacacia* L., Fabaceae) is a North American tree species introduced in Europe in the early seventeenth century (Cierjacks *et al.*, 2013) and used for ornamental and forestry purposes. It is currently considered of conflicting values in Europe (Dickie *et al.*, 2014): in forestry its mature wood is a rot-proof wood reaching the same durability class as teak (*Tectona grandis*) (Dünisch *et al.*, 2010) and young stems from short rotation coppice are a costly but efficient energy source (Manzone *et al.*, 2015); but in natural areas it is considered one of the most invasive tree species worldwide (Basnou, 2006; Richardson and Rejmánek, 2011) impacting local diversity and ecosystem functioning in particular as a consequence of its nitrogen-fixing capability (Cierjacks *et al.*, 2013). Black locust is capable of sexual reproduction and it also spreads locally by asexual root-sucker production (Kurokuchi *et al.*, 2010; Cierjacks *et al.*, 2013). However, its seeds present a physical dormancy that prevents germination when improperly broken (Cierjacks *et al.*, 2013).

The physical dormancy of Fabaceae seeds is due to the impermeable layers of the seed coat that prevent water absorption and, thus, embryo germination (Baskin *et al.*, 2000; Baskin and Baskin, 2004). Under natural conditions, the disruption of the lens allows permeability and further germination. A simple pin puncture or razor cut in the seed coat would allow water absorption and often lead to nearly 100% germination rate in Fabaceae (Funes and Venier, 2006; Bonner and Karrfalt, 2008; Abudureheman *et al.*, 2014; Giuliani *et al.*, 2015). However, this methodology cannot be considered when numerous seeds are required and/or when dormancy must be homogeneously broken, as for forestry purposes or for research. Previous studies tested various methods to break dormancy of Fabaceae seeds (Bonner and Karrfalt, 2008; Abudureheman *et al.*, 2014; Giuliani *et al.*, 2015). These mechanical methods were based on: (1) dry-heat treatments (testing various

\* Correspondence

Email: [annabel.porte@inra.fr](mailto:annabel.porte@inra.fr)

§ These authors contributed equally to this work.

temperature and duration) or on manual abrasion using sandpaper; (2) water imbibition based on seed soaking in hot water following several variants regarding water temperature (40–100°C), soaking duration (20 min to 72 h) and number of heating/cooling water cycles (1–7); and (3) concentrated sulphuric acid treatment, testing variable concentrations or duration (10–120 min). The latter method often resulted in high germination rates but was risky for the environment or could damage the seeds (Bonner and Karrfalt, 2008). Thus we decided not to consider it in our search for eco-friendly methods. Specific tests on black locust seeds were proven contradictory regarding their efficiency, for example, heat being selected by some authors whereas soaking in water or acid being indicated as the most efficient by others (Singh *et al.*, 1991; Masaka and Yamada, 2009; Basbag *et al.*, 2010; Giuliani *et al.*, 2015).

Considering this discrepancy in the various previous studies, our objective was to compare seed treatments of black locust to break physical dormancy reaching high germination and to find a method that allows treatment of large numbers of seeds in a short time and in a homogeneous way. The best methods were then validated on seeds from various Fabaceae species.

## Material and methods

### Dormancy break

Pods were collected in early February 2015 from ten trees in each of 10 populations located in South-

Western France (Gironde, Lot et Garonne), using a manual pruner. A visual sorting of the seeds was performed to remove dried, aborted or rotten seeds. Seeds were stored in paper bags in a cold room between 0 and 5°C (Bonner and Karrfalt, 2008; Cierjacks *et al.*, 2013). Two hundred and ten seeds from one tree of the Gabarnac population (Gironde, France) were used to test seven dormancy-breaking treatments, with 30 seeds per treatment: puncturing (P), consisting of making one pin puncture in the seed coat, using a needle; dry-heat (Basbag *et al.*, 2010) was applied by placing seeds in a Petri dish in an oven at 90°C for 20 and 30 min (H20, H30); imbibition by soaking in water (W) was performed in one cycle, placing seeds in nearly boiling water at 95°C (Turner and Dixon, 2009) and letting them cool for 15 h; finally we tested a novel sanding treatment (Fig. 1) using an automated grinder (2010 Geno/Grinder, SPEX SamplePrep, Metuchen, NJ, USA) that allows shaking of tubes controlling both intensity and duration. Five seeds were positioned in an Eppendorf tube half-filled with sand (sand for pool filtering, 5% free silica and grain size 0.8–1.4 mm; Aqua Espace Accessoires, Cadaujac, France) and agitated at 1500 r.p.m. for 5, 10 and 15 min, respectively (S5, S10, S15). Up to 96 tubes at a time can then be agitated, allowing homogeneous treatment of 480 seeds each time when using five seeds per tube.

After treatment, five seeds scarified with the same treatment were placed in a pot (QuickPot 35RW, HerkuPlast Kubern GmbH, Ering, Germany) filled with 25 g of substrate (substrate 307, Peltracom, Gent, Belgium) and covered with 6 g of substrate. The pots were then watered with 50 ml of water. During



**Figure 1.** The grinder used to break seed dormancy. The seeds and sand in the tubes are shaken homogeneously controlling both speed and duration to quickly achieve a homogeneous sanding of the seed coat. On the right, close up of the plates allowing shaking of 96 tubes at a time, with a view of the pods and seeds as collected from the pods and in the tube ready for sanding (5 seeds per tube). Photograph credits: X.P. Bouteiller and A.J. Porté.

germination, each pot was watered every other day with 50 ml of water, so that the plants did not suffer from water stress. Pots were positioned in a climate chamber (Fitoclima D1200, Aralab, S. Domingos de Rana, Portugal) at 60% air relative humidity, a day/night temperature of 22/20°C and light intensity of  $185 \pm 45 \mu\text{mol PAR m}^{-2} \text{s}^{-1}$  (quantum sensor Licor Li190, Lincoln, USA) with a photoperiod of 12 h day-light/10 h darkness and two 1 h periods of gradual transitions in-between. Phenological stages of seed germination were monitored for two weeks (4–6 measurements), following five stages adapted from *Glycine max* (Minost, 1997) and older *R. pseudoacacia* (Cierjacks *et al.*, 2013): stage 1, radical emergence; stage 2, cotyledon emergence; stage 3, cotyledon deployment; stage 4, first whole leaf; stage 5, first compound leaf (Supplementary information Fig. S1). Sometimes the cotyledons rapidly detached from the stems of young seedlings. In this case, seeds were indicated as germinated (stage 1) but since plant development would be further limited, they were recorded as germination with non-intact cotyledons. Germination rates were calculated separately based on all germination events (*total germination*, including germination with non-intact cotyledons) and based on *healthy germination*, defined as germination with intact cotyledons. Individual germination data (total and healthy germination) were analysed using a generalized linear mixed model with a binary distribution (1 germinated/0 not germinated at the end of the monitoring) and logit link function (procedure GLIMMIX, SAS9.4, SAS Institute, Cary, NC, USA) with treatment as a fixed effect, pot nested in treatment as a random effect.

We selected nine additional species with orthodox seeds (Table 1; Supplementary information Fig. S2) to evaluate seed dormancy breaking methods. Species were selected from invasive trees and shrubs from

the Fabaceae family, using the global database of invasive trees and shrubs from Rejmánek and Richardson (2013). Orthodoxy and seed weight information were obtained from the Royal Botanic Gardens Kew Seed Information Database (SID; version 7.1, January 2017; <http://data.kew.org/sid/>); species 1000-seed weight varied from 12.7 to 48 g, as we avoided big seeds that could get stuck in the tubes during sanding. Seeds were bought (Sunshine Seeds, Peter Krebs, Ahlen, Germany) and seeded in autumn 2016, following the same growing conditions as for *Robinia pseudoacacia*. Twenty seeds per species and treatment were treated following the control (C), puncture (P) and 5 min sanding (S5) protocols. Germination was evaluated 30 days after sowing, and individual germination data were analysed using a generalized linear mixed model with a binary distribution (1 germinated/0 not germinated at the end of the monitoring) and logit link function (procedure GLIMMIX, SAS9.4, SAS Institute) with treatment as a fixed effect, pot nested in treatment as a random effect.

### Germination curve

To evaluate the germination rate of *R. pseudoacacia* seeds using the best scarification method, 500 seeds (five seeds per 10 trees per 10 populations) were treated and monitored over 20 days. We then constructed a germination curve of the fraction of healthy germinated seeds [ $G(t)$ ]. Several non-linear models are available from the literature [logistic, Gompertz, Weibull, log-logistic (Torres and Frutos, 1989; Ritz *et al.*, 2013; Archontoulis and Miguez, 2015)] but only the Gompertz model matched the shape of the black locust germination time curve. The model was fitted using a non-linear procedure (NLIN, SAS9.4, SAS

**Table 1.** Main characteristics of the species used to test the automated sanding dormancy-breaking method

Genus	Species	Life form	Origin	Nb zones where invasive	Status in Europe	Seed weight (1000) in g
<i>Acacia</i>	<i>dealbata</i> Link	T	Australia	7	I	12.7
<i>Acacia</i>	<i>mearnsii</i> De Wild.	T	Australia	13	I	13.2
<i>Acacia</i>	<i>saligna</i> (Labill.) H.L.Wendl.	T	Australia	5	I	16.9
<i>Albizia</i>	<i>julibrissin</i> Durazz.	T	Asia, Middle East	2	E	36
<i>Dalbergia</i>	<i>sissoo</i> Roxb. Ex DC.	T	Asia	2	–	41.8
<i>Leucaena</i>	<i>leucocephala</i> (Lam.) de Wit (L.) Gillis	T	Central America	14	E	48
<i>Mimosa</i>	<i>pigra</i> L.	T/S	South America	8	–	18
<i>Robinia</i>	<i>pseudoacacia</i> L.	T	North America	8	I	19.2
<i>Senna</i>	<i>alata</i> (L.) Roxb.	S	Central America	6	–	38.5
<i>Senna</i>	<i>occidentalis</i> (L.) Link	S	Central, South America	8	–	17

Life form: T, tree; S, shrub. Number (Nb) of zones where invasive: Rejmanek and Richardson (2013) surveyed the literature to construct a database on invasive woody species, indicating their invasive status throughout the world which was divided into 15 geographical zones. Status in Europe: I, invasive; E, established; –, planted or not present. The status was indicated following Rejmanek and Richardson (2013), the DAISIE database of European alien species (DAISIE European Invasive Alien Species Gateway: <http://www.europe-alien.org/>) and the CABI database (<http://www.cabi.org/isc/search/?q=&types=7,19&sort=DateDesc>). Seed weight (mean of 1000 seeds, g) was extracted from the Royal Botanic Gardens Kew Seed Information Database [January 2017, Seed Information Database (SID), version 7.1, available at: <http://data.kew.org/sid/>].

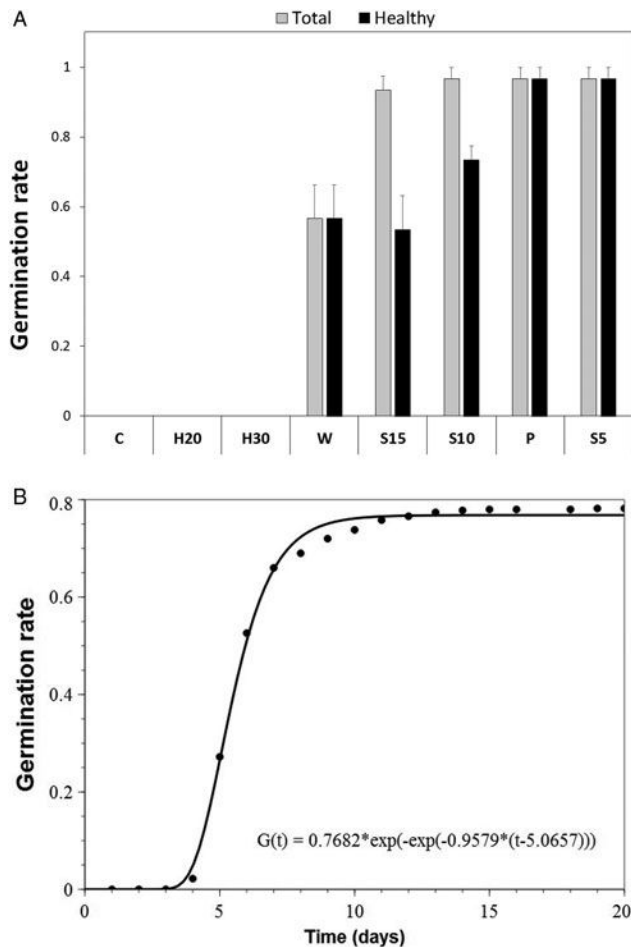
Institute:

$$G(t) = D * \exp(-\exp(-b(t - t_m))), \quad (1)$$

where parameter  $D$  is the maximum germination rate,  $b$  is the slope of the germination curve and  $t_m$  is the time at the inflexion point.

## Results

All treatments except dry-heat (H20, H30) permitted the breaking of seed dormancy of black locust (Fig. 2A). The highest germination of 97% (SE = 3.3) was reached with treatments puncture (P) and sanding



**Figure 2.** (A) Maximum germination (total and healthy) of black locust seeds according to the dormancy-breaking treatment (C, W, S15, S10, S5, P); as no seed germinated in the dry-heat treatments (H20, H30), bars were equal to zero (no variance). For statistics, see Table 2. (B) Germination-time curve (full line, Gompertz model,  $P < 0.0001$ ) fitted using healthy germination data on 500 seeds treated with the S5 sanding treatment (black circles). C, control; W, water soaking; S15, S10 and S5, automatic sanding for 15, 10 or 5 min; P, puncture.

of 5-min duration (S5); moreover, no damage was observed on cotyledons, thus no reduction in healthy germination was observed. Longer sanding treatments (S10, S15) induced no significant reduction *per se* in total germination success (97 and 93%, SE 3.3 and 4.2) but resulted in significantly lower healthy germination (maximum 73 and 53%, SE 4.2 and 9.9, Fig. 2A) due to stem emergence without cotyledons (Table 2). Water soaking (W) only resulted in moderate germination, with a total germination of 57% (SE 9.5); it was significantly lower than all effective treatments considering total germination but equivalent to long duration sanding treatments S10 and S15 when considering healthy germination.

When we applied the S5 treatment to 500 seeds collected from 100 different black locust trees (Fig. 2B), the germination success was lower than during the comparative test, but still reached a healthy total germination of 78.2%; 12.8% of the seeds germinated with broken cotyledons and 9% did not germinate. Overall, the Gompertz model was accurate to fit the data although slightly underestimating maximum germination (76.8%) as a result of reaching a plateau on still slightly increasing data.

Overall, puncturing (P) proved to break the dormancy of the Fabaceae seeds we tested, while they did not germinate in the absence of the treatment (control C). Some variation was observed in the degree of efficiency of sanding (S5) according to the species, which could be related to the seed characteristics. As observed for *R. pseudoacacia*, sanding S5 was as efficient as the reference puncture technique (P) to break dormancy of *A. dealbata*, *A. julibrissin* and *M. pigra* (Fig. 3, Table 3). For a second group of species, *A. saligna*, *L. leucocephala* and *S. alata*, sanding was not efficient, with a low level of germination, not significantly different from the control (C). Seeds of *S. occidentalis* demonstrated poor germination whatever the treatment. Finally, for *A. mearnsii* and *D. sissoo*, sanding (S5) resulted in a negative effect on germination, whereas puncture was as efficient as the control, suggesting that dormancy was already broken when seeds were received.

## Discussion

Together with manual puncturing (P), a short-time sanding using an automated grinder for 5 min (S5) was the most efficient treatment to break physical seed dormancy in black locust. Under natural conditions or without dormancy-breaking treatment, germination of these seeds is low, ranging from 3 to 22% (Singh *et al.*, 1991; Cierjacks *et al.*, 2013). All treatments, except dry-heat, at least doubled natural germination. The needle puncture or blade cut is frequently cited in the literature as effective in breaking physical

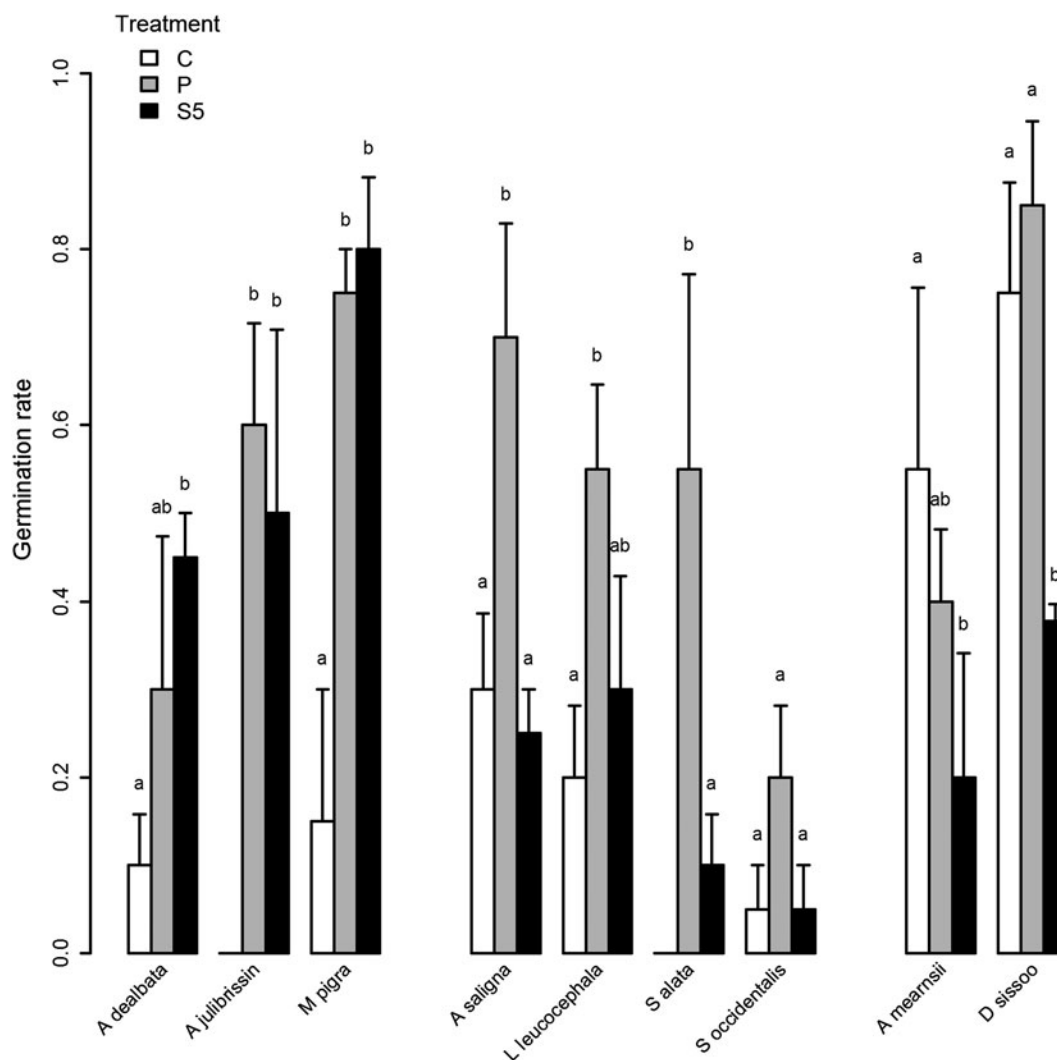
**Table 2.** Mixed model analysis testing for germination success (total and healthy) of black locust according to treatment (W, S15, S10, S5, P)

	Treatment			Pot (treatment)		Mean differences
	d.f.	F	P	ChiSq	P	
Germination	4	5.66	<b>0.0022</b>	0.02	0.45	W <sup>a</sup> /S15 <sup>b</sup> , S10 <sup>b</sup> , S5 <sup>b</sup> , P <sup>b</sup>
Healthy germination	4	4.52	<b>0.0018</b>	–	–	S15 <sup>a</sup> , W <sup>a</sup> , S10 <sup>a</sup> /S5 <sup>b</sup> , P <sup>b</sup>

Analyses were performed without the dry-heat treatments which presented no germination and, thus, no variance. Treatment was considered as a fixed effect (*F*-test value for fixed effects) and pot (treatment) as a random effect (ChiSq in proc glimmix). d.f. represents degree of freedom. Bold font indicates statistically significant values (*P* < 0.05). Random effect was removed from the model when variance was estimated to be 0 (Kiernan *et al.*, 2012). Means with the same letter are considered not to be significantly different.

dormancy of black locust, Fabaceae or other physically dormant seeds (Singh *et al.*, 1991; Funes and Venier, 2006; Masaka and Yamada, 2009; Cierjacks *et al.*, 2013; Abudureheman *et al.*, 2014; Giuliani *et al.*, 2015), so that it can be used as a reference method. The

absence of germination following dry-heat treatments contradicted previous studies that reported 90–94% germination (Masaka and Yamada, 2009; Basbag *et al.*, 2010). Hot water soaking is proposed by forest managers to break physical dormancy (e.g. in France



**Figure 3.** Germination of nine Fabaceae species seeds according to the dormancy-breaking treatment (C, P, S5): *Acacia dealbata*, *Acacia mearnsii*, *Acacia saligna*, *Albizia julibrissin*, *Dalbergia sissoo*, *Leucaena leucocephala*, *Mimosa pigra*, *Senna alata* and *Senna occidentalis*. C, control; S5, automatic sanding for 5 min; P, puncture. Different letters indicate significantly different germination rates among the dormancy breaking treatments for each species.

**Table 3.** Mixed model analysis testing for the germination success (total) of nine species of Fabaceae according to treatment (C, P, S5)

	Treatment			Pot (treatment)		Mean differences
	d.f.	F	P	ChiSq	P	
<i>Acacia dealbata</i>	2	2.56	0.09	0.07	0.39	C <sup>a</sup> /P <sup>ab</sup> /S <sup>b</sup>
<i>Acacia mearnsii</i>	2	2.34	0.11	0.08	0.39	C <sup>a</sup> /P <sup>ab</sup> /S5 <sup>b</sup>
<i>Acacia saligna</i>	2	4.42	<b>0.02</b>	0.10	0.38	C <sup>a</sup> , S5 <sup>a</sup> /P <sup>b</sup>
<i>Albizia julibrissin</i>	1	0.15	0.70	0.06	0.40	S5 <sup>a</sup> , P <sup>a</sup>
<i>Dalbergia sissoo</i>	2	3.53	<b>0.04</b>	0.10	0.38	C <sup>a</sup> , P <sup>a</sup> /S5 <sup>b</sup>
<i>Leucaena leucocephala</i>	2	2.57	0.09	0.08	0.39	C <sup>a</sup> /S5 <sup>ab</sup> /P <sup>b</sup>
<i>Mimosa pigra</i>	2	7.96	<b>0.0009</b>	0.08	0.39	C <sup>a</sup> /S5 <sup>b</sup> , P <sup>b</sup>
<i>Senna alata</i>	1	7.34	<b>0.01</b>	0.05	0.41	S5 <sup>a</sup> /P <sup>b</sup>
<i>Senna occidentalis</i>	2	1.39	0.26	0.06	0.40	C <sup>a</sup> , S5 <sup>a</sup> , P <sup>a</sup>

Treatment was considered as a fixed effect (*F*-test value for fixed effects) and pot (treatment) as a random effect (ChiSq in proc glimmix). d.f. represents degree of freedom (either 2 or 1 when a treatment presented no germination thus no variance). Bold font indicates statistically significant values ( $P < 0.05$ ). Means with the same letter are considered not to be significantly different.

(CRPF, 2007). However, although it was indeed proved highly efficient sometimes to break physical dormancy of black locust (Singh *et al.*, 1991) or other Fabaceae species (Ventura de Souza *et al.*, 2012; Abudurehman *et al.*, 2014), our results are in accordance with results indicating only moderate (20–50%) efficiency of breaking physical dormancy (Turner and Dixon, 2009; Abudurehman *et al.*, 2014; Giuliani *et al.*, 2015).

To our knowledge, automatic sanding has never been tested on physically dormant seeds. A more rough sanding in a cement mixer or similar machine is generally proposed by forest managers in both the native and invasive ranges to treat tonnes of physically dormant seeds at a time (Argel and Paton, 1999; CRPF, 2007; Kimura and Islam, 2012); sandpaper scraping of black locust seeds has been mentioned but without indicating its efficiency (Morimoto *et al.*, 2010; Dini-Papanastasi *et al.*, 2012). However, since Fabaceae species are cropped for cattle or human food (Graham and Vance, 2003) with many species of this family presenting seeds with physical dormancy (Baskin *et al.*, 2000; Baskin and Baskin, 2004; Jayasuriya *et al.*, 2013), dormancy-breaking treatments of other Fabaceae species have been reported. Mechanical scarification has largely been applied using sandpaper; most of the time it resulted in the highest germination success ranging from 75 to 100%, compared with control treatments of 0 to 30% (González-Castañeda *et al.*, 2004; Alves de Oliveira *et al.*, 2008; Nogueira Camargos *et al.*, 2008; Ali *et al.*, 2011; Barcelo Gomes *et al.*, 2013; Martins de Mesquita Matos *et al.*, 2014; Campos Dayrell *et al.*, 2015; Ferreras *et al.*, 2015; Statwick, 2016). The only indication of automation was developed in the 1970s by air pulsing seeds on a rotating sandpaper-covered drum (Townsend and McGinnies, 1972, cited by Kimura and Islam, 2012). Using this

device to break seed dormancy of two different Fabaceae species, it was, however, not efficient with 7 and 29% germination, respectively (Patanè and Gresta, 2006). Using 10 Fabaceae species, we demonstrated that automated sanding can be more generally efficient in breaking physical dormancy: in four of the species a 5-min sanding was highly efficient, whereas for three of them we hypothesize that a longer period is necessary. Indeed, for *A. saligna*, *L. leucocephala* and *S. alata*, seed dormancy was not broken; however, seeds were still viable since they germinated with a 75–100% success when punctured and seeded again after completion of the experiment (data not shown). For these species, or other species of interest that we did not test, we suggest calibrating the automatic sanding duration until dormancy is broken. For *S. occidentalis*, a previous study tested a mechanical cut in the coat which resulted in 100% germination, compared with the control (Teketay, 1996). Our opposite result where even mechanical puncture did not allow high germination rates to be reached thus suggests that the seeds we used had lost their germination capacity. Finally, the negative impact of sanding for *A. mearnsii* and *D. sissoo* was probably related to the use of non-dormant seeds. Seed coats of *A. mearnsii* were already partly broken when received, with apparent white soft tissues that could have been too much damaged by sanding to allow proper development. Regarding *D. sissoo*, some publications indicated that seeds display no or only weak physical dormancy, allowing immediate and high levels of germination (Khan and Khan, 1992; Murali, 1997; Jøker, 2002). Thus sanding the seeds probably resulted in degrading them and preventing germination. Still, the method we propose using automated sanding allows us to treat large samples in a short time and assures homogeneity in the scarification process, as demonstrated by the dormancy-breaking



treatments. The lower germination that was observed when sanding was applied to 500 seeds of *R. pseudoacacia* can be related to the heterogeneity of the seeds, coming from 10 different populations, compared with the test performed on one homogeneous seed sample; indeed physical dormancy breaking can present some heterogeneity related to seed origins (Masaka and Yamada, 2009). When using a standard protocol, the discrepancy between total and healthy germination can then be interpreted as an indication of the proportion of unhealthy seeds with a damaged or thinner coat. Comparing studies using different protocols to break seed dormancy could be misleading, considering the high variability of their efficiency. Therefore, we propose that automated sanding be used as a standard to break dormancy of black locust seeds, and in general for Fabaceae physically dormant seeds, to allow further comparative studies of plant populations or genotypes. This standard protocol would only require to be calibrated for other species with seeds presenting a physical dormancy.

## Acknowledgements

This study has been carried out with financial support from the French National Research Agency (ANR) in the frame of the Investments for the future Program (ANR-10- LABX-45) and from the University of Liège (Fonds Spéciaux pour la Recherche).

## Supplementary material

**S1:** The five seed germination phenological stages; **S2:** Photographs of seeds of all species used.

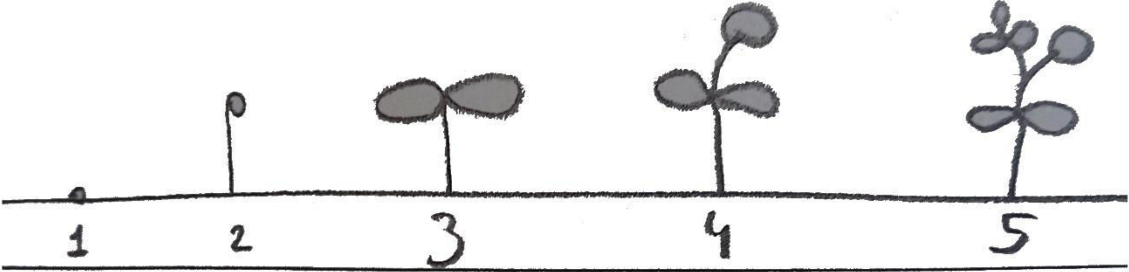
To view supplementary material for this article, please visit <https://doi.org/10.1017/S0960258517000150>

## References

- Abudurehman, B., Liu, H., Zhang, D. and Guan, K. (2014) Identification of physical dormancy and dormancy release patterns in several species (Fabaceae) of the cold desert. *Seed Science Research* **24**, 133–145.
- Ali, H.H., Tanveer, A., Nadeem, M.A. and Asghar, H.N. (2011) Methods to break seed dormancy of *Rhynchosia capitata*, a summer annual weed. *Chilean Journal of Agricultural Research* **71**, 483–487.
- Alves de Oliveira, D., Ferreira Nunes, Y.R., Almeida Rocha, E., Fagundes Braga, R., Silva Pimenta, M.A. and Magalhães Veloso, M.D. (2008) Potencial germinativo de sementes de fava d'anta (*Dimorphandra mollis* Benth. – Fabaceae: Mimosoideae) sob diferentes procedencias, datas de coleta e tratamentos de escarificação. *Revista Arvore* **32**, 1001–1009.
- Archontoulis, S.V. and Miguez, F.E. (2015) Non-linear regression models and applications in agricultural research. *Agronomy Journal* **107**, 786–798.
- Argel, P.J. and Paton, C.J. (1999) Overcoming legume hard-seedness, pp. 247–259 in Loch, D.S. and Ferguson, J.E. (eds), *Forage Seed Production*, volume 2: *Tropical and Subtropical Species*, Wallingford, UK, CABI International.
- Barcelo Gomes, M., Assis de Faria, A., Santos Cerqueira, D. and Lima Bailão, L. (2013) Avaliação de métodos para a superação de dormência de sementes de jatobá (*Hymenaea courbaril* L.). *Revista Eletrônica Interdisciplinar* **1**, 6–9.
- Basbag, M., Aydin, A. and Ayzit, D. (2010) The effect of different temperatures and durations on the dormancy breaking of black locust (*Robinia pseudoacacia* L.) and honey locust (*Gleditsia triacanthos* L.) seeds. *Notulae Scientia Biologicae* **2**, 125–128.
- Baskin, J.M. and Baskin, C.C. (2004) A classification system for seed dormancy. *Seed Science Research* **14**, 1–16.
- Baskin, J.M., Baskin, C.C. and Li, X. (2000) Taxonomy, anatomy and evolution of physical dormancy in seeds. *Plant Species Biology* **15**, 139–152.
- Basnou, C. (2006) *Robinia pseudoacacia* factsheet. Retrieved from: <http://www.europe-aliens.org/speciesFactsheet.do?speciesId=11942>
- Bonner, F.T. and Karrfalt, R.P. (2008) The woody plant seed manual. *USDA Agricultural Handbook* **727**, 1–1223.
- Campos Dayrell, R.L., Gonçalves-Alvim, S.D.J., Negreiros, D., Fernandes, G.W. and Oliveira Silveira, F.A. (2015) Environmental control of seed dormancy and germination of *Mimosa calodendron* (Fabaceae): implications for ecological restoration of a highly threatened environment. *Brazilian Journal of Botany* **38**, 395–399.
- Cierjacks, A., Kowarik, I., Joshi, J., Hempel, S., Ristow, M., von der Lippe, M. and Weber, E. (2013) Biological flora of the british isles: *Robinia pseudoacacia*. *Journal of Ecology* **101**, 1623–1640.
- CRPF (2007) Le robinier faux-acacia. Retrieved from: <http://www.crfp-poitou-charentes.fr/IMG/pdf/robinier.pdf>
- Dickie, I.A., Bennett, B.M.B., Burrow, L.E., Nuñez, M.A., Peltzer, D.A., Porté, A.J., Richardson, D.M., Rejmánek, M., Rundel, P.W. and van Wilgen, B.W. (2014) Conflicting values: ecosystem services and invasive tree management. *Biological Invasions* **16**, 705–719.
- Dini-Papanastasi, O., Kostopoulou, P. and Radoglou, K. (2012) Effects of seed origin, growing medium and mini-plug density on early growth and quality of black locust (*Robinia pseudoacacia* [L.] seedlings. *Journal of Forest Science* **58**, 8–20.
- Dünisch, O., Richter, H.-G. and Koch, G. (2010) Wood properties of juvenile and mature heartwood. *Wood Science and Technology* **44**, 301–313.
- Ferreras, A.E., Funes, G. and Galetto, L. (2015) The role of seed germination in the invasion process of Honey locust (*Gleditsia triacanthos* L., Fabaceae): comparison with a native confamilial. *Plant Species Biology* **30**, 126–136.
- Funes, G. and Venier, P. (2006) Dormancy and germination in three *Acacia* (Fabaceae) species in central Argentina. *Seed Science Research* **16**, 77–82.
- Giuliani, C., Lazzaro, L., Mariotti Lippi, M. and Foggi, B. (2015) Temperature-related effects on the germination capacity of black locust (*Robinia pseudoacacia* L., Fabaceae) seeds. *Folia Geobotanica* **50**, 275–282.
- González-Castañeda, J., Angoa-Pérez, M.V., Frías-Hernández, J.T., Olalde-Portugal, V., Flores-Ancira, E.,

- Terrones-Rincón, T.R.L., Van Cleemput, O. and Dendooven, L.** (2004) Germination of seeds of huisache (*Acacia schaffneri*) and catclaw (*Mimosa monanctris*) as affected by sulphuric acid and mechanical scarification and subsequent growth and survival in a greenhouse and field experiment. *Seed Science and Technology* **32**, 727–738.
- Graham, P.H. and Vance, C.P.** (2003) Legumes: importance and constraints to greater use. *Plant Physiology* **131**, 872–877.
- Jayasuriya, K.M.G.G., Wijetunga, A.S.T.B., Baskin, J.M. and Baskin, C.C.** (2013) Seed dormancy and storage behaviour in tropical Fabaceae: a study of 100 species from Sri Lanka. *Seed Science Research* **23**, 257–269.
- Jøker, D.** (2002) *Dalbergia sissoo* Roxb. ex DC. Retrieved from: [http://dfsc.dk/pdf/Seedleaflets/Dalbergia%20sissoo\\_65\\_int.pdf](http://dfsc.dk/pdf/Seedleaflets/Dalbergia%20sissoo_65_int.pdf)
- Khan, M. and Khan, R.** (1992) Woody Plant Seed Manual (Pakistan). Retrieved from: [http://pdf.usaid.gov/pdf\\_docs/PNABW321.pdf](http://pdf.usaid.gov/pdf_docs/PNABW321.pdf)
- Kiernan, K., Tao, J. and Gibbs, P.** (2012) Tips and strategies for mixed modeling with SAS/STAT<sup>®</sup> procedures. *SAS Global Forum* **2012**, 332–2012.
- Kimura, E. and Islam, M.A.** (2012) Seed scarification methods and their use in forage legumes. *Research Journal of Seed Science* **5**, 38–50.
- Kurokochi, H., Toyama, K. and Hogetsu, T.** (2010) Regeneration of *Robinia pseudoacacia* riparian forests after clear-cutting along the Chikumagawa river in Japan. *Plant Ecology* **210**, 31–41.
- Manzone, M., Bergante, S. and Facciotto, G.** (2015) Energy and economic sustainability of woodchip production by black locust (*Robinia pseudoacacia* L.) plantations in Italy. *Fuel* **140**, 555–560.
- Martins de Mesquita Matos, J., Oliveira Ramos, K.M., de Carvalho Cristo Martins, R., Vasconcelos de Oliveira, D. and Pereira Martins, A.** (2014) Estudo das técnicas de superação da dormência das sementes de *Enterolobium contortisiliquum* (vell.) Morong. *Heringeriana* **4**, 60–64.
- Masaka, K. and Yamada, K.** (2009) Variation in germination character of *Robinia pseudoacacia* L. (Leguminosae) seeds at individual tree level. *Journal of Forest Research* **14**, 167–177.
- Minost, C.** (1997) Soybean. Retrieved from <https://www7.inra.fr/hyppz/CULTURES/6c--000.htm>
- Morimoto, J., Kominami, R. and Koike, T.** (2010) Distribution and characteristics of the soil seed bank of the black locust (*Robinia pseudoacacia*) in a headwater basin in northern Japan. *Landscape and Ecological Engineering* **6**, 193–199.
- Murali, K.S.** (1997) Patterns of seed size, germination and seed viability of tropical tree species in Southern India. *Biotropica* **29**, 271–279.
- Nogueira Camargos, V., Moreira de Carvalho, M.L., Vieira de Araújo, D. and Linhares Magalhães, F.H.** (2008) Superação da dormência e avaliação da qualidade fisiológica de sementes de *Sesbania virgata*. *Ciência e Agrotecnologia, Lavras* **32**, 1858–1865.
- Patanè, C. and Gresta, F.** (2006) Germination of *Astragalus hamosus* and *Medicago orbicularis* as affected by seed-coat dormancy breaking techniques. *Journal of Arid Environments* **67**, 165–173.
- Rejmánek, M. and Richardson, D.M.** (2013) Trees and shrubs as invasive alien species – 2013 update of the global database. *Diversity and Distributions* **19**, 1093–1094.
- Richardson, D.M. and Rejmánek, M.** (2011) Trees and shrubs as invasive alien species – a global review. *Diversity and Distributions* **17**, 788–809.
- Ritz, C., Phipper, C.B. and Streibig, J.C.** (2013) Analysis of germination data from agricultural experiments. *European Journal of Agronomy* **45**, 1–6.
- Singh, D.P., Hooda, M.S. and Bonner, F.T.** (1991) An evaluation of scarification methods for seeds of two leguminous trees. *New Forests* **5**, 139–145.
- Statwick, J.M.** (2016) Germination pretreatments to break hard-seed dormancy in *Astragalus cicer* L. (Fabaceae). *PeerJ* **4**, e2621.
- Teketay, D.** (1996) The effect of different pre-sowing seed treatments, temperature and light on the germination of five *Senna* species from Ethiopia. *New Forest* **11**, 155–171.
- Torres, M. and Frutos, G.** (1989) Analysis of germination curves of aged fennel seeds by mathematical models. *Environmental and Experimental Botany* **29**, 409–415.
- Townsend, C.E. and McGinnies, W.J.** (1972) Establishment of nine forage legumes in the Central Great Plains. *Agronomy Journal* **64**, 699–702.
- Turner, S.R. and Dixon, K.W.** (2009) Seed dormancy and germination in the Australian baobab, *Adansonia gregorii* F. Muell. *Seed Science Research* **19**, 261–266.
- Ventura de Souza, T., Heinig Voltolini, C., Santos, M. and Silveira Paulilo, M.T.** (2012) Water absorption and dormancy-breaking requirements of physically dormant seeds of *Schizolobium parahyba* (Fabaceae – Caesalpinioideae). *Seed Science Research* **22**, 169–176.

Supplementary information S1: drawing and photographs of the five seed germination phenological stages used for *Robinia pseudoacacia*, adapted from Minost (1997) and Cierjacks et al. (2013). © XP Bouteiller, AJ Porté



Supplementary information S2: photographs of the seeds for each species



*Acacia dealbata*, Steve Hurst, hosted by the USDA-NRCS PLANTS Database



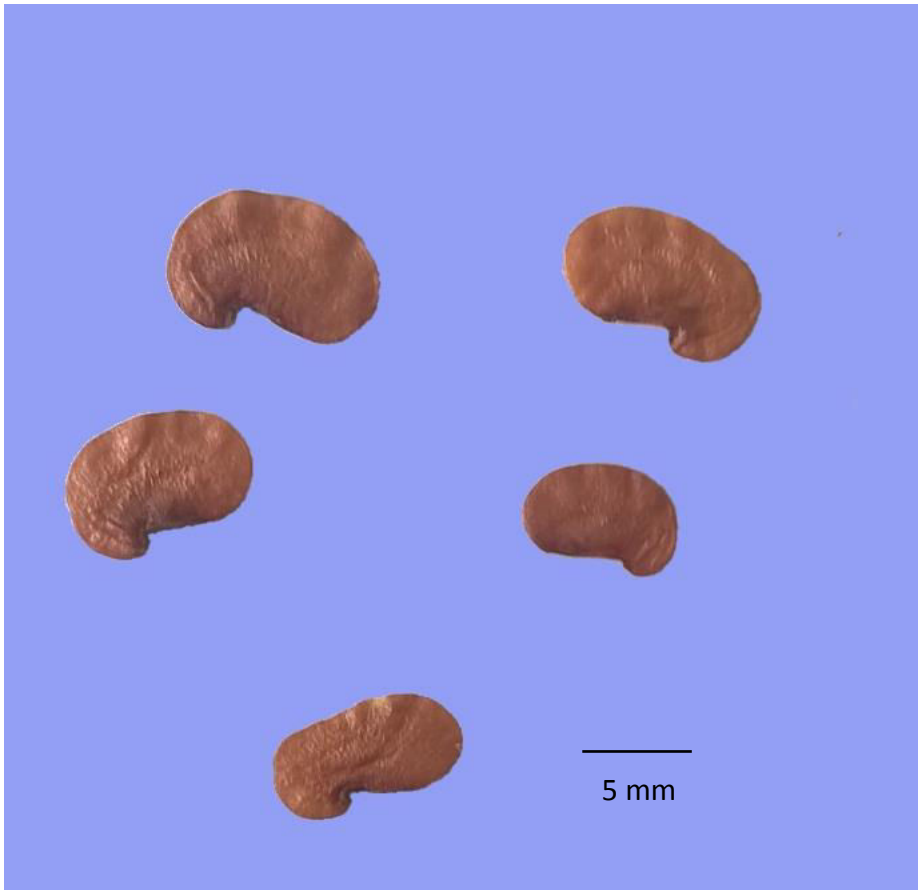
*Acacia mearnsii*, Steve Hurst, hosted by the USDA-NRCS PLANTS Database



*Acacia saligna*, Tracey Slotta, hosted by the USDA-NRCS PLANTS Database



*Albizia julibrissin*, Steve Hurst, hosted by the USDA-NRCS PLANTS Database



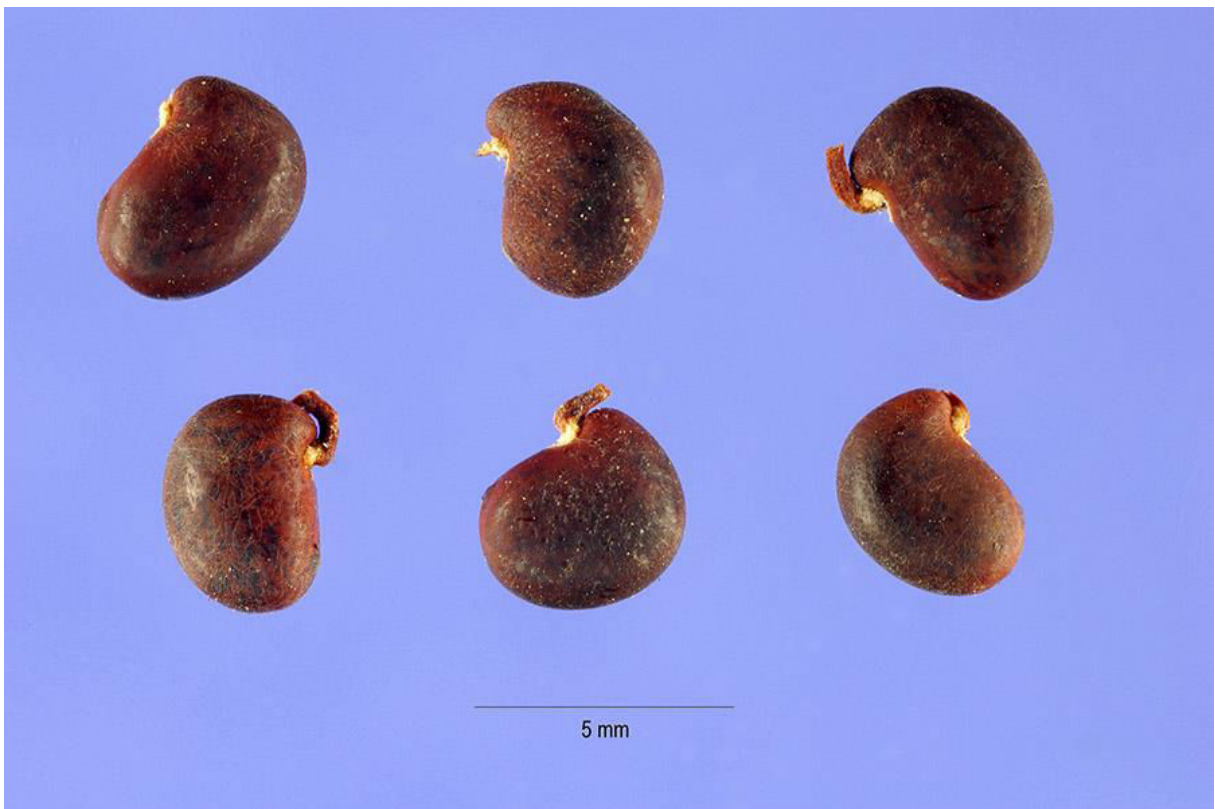
*Dalbergia sissoo*, Annabel J. Porté



*Leucaena leucocephala*, Steve Hurst, hosted by the USDA-NRCS PLANTS Database



*Mimosa pigra*, Steve Hurst, hosted by the USDA-NRCS PLANTS Database



*Robinia pseudoacacia*, Steve Hurst, hosted by the USDA-NRCS PLANTS Database



*Senna alata*, Tracey Slotta, hosted by the USDA-NRCS PLANTS Database



*Senna occidentalis*, Steve Hurst, hosted by the USDA-NRCS PLANTS Database



**Reference:**

All photographs are free of use. Except for *Dalbergia sissoo*, all photographs were extracted from the Image database of the Plant database from the USDA.

USDA, NRCS. 2017. The PLANTS Database (<http://plants.usda.gov>, 11 January 2017). National Plant Data Team, Greensboro, NC 27401-4901 USA.

## **Chapitre 3**

**No evidence for genetic differentiation in juvenile traits between Belgian and French populations of the invasive tree *Robinia pseudoacacia***

# No evidence for genetic differentiation in juvenile traits between Belgian and French populations of the invasive tree *Robinia pseudoacacia*

Xavier P. Bouteiller<sup>1,\*</sup>, Frédéric Barraquand<sup>2</sup>, Pauline Garnier-Géré<sup>1</sup>, Noémie Harmand<sup>1</sup>, Yec'han Laizet<sup>1</sup>, Alexandre Raimbault<sup>1</sup>, Raphaël Segura<sup>1</sup>, Ludivine Lassois<sup>3</sup>, Arnaud Monty<sup>4</sup>, Cindy Verdu<sup>3</sup>, Stéphanie Mariette<sup>1</sup> & Annabel J. Porté<sup>1</sup>

<sup>1</sup>BIOGECO, INRA, Univ. Bordeaux, FR-33610, Cestas, France

<sup>2</sup>IMB, CNRS & LabEx COTE, Univ. Bordeaux, FR-33600, Pessac, France

<sup>3</sup>Biodiversity and Landscape Unit, Gembloux Agro-Bio Tech, University of Liège, BE-5030, Gembloux, Belgium

<sup>4</sup>Forest Management Unit, Gembloux Agro-Bio Tech, University of Liège, BE-5030, Gembloux, Belgium

\*Author for correspondence: [bouteiller.xavier@gmail.com](mailto:bouteiller.xavier@gmail.com)

**Background** – The role of evolution in biological invasion studies is often overlooked. In order to evaluate the evolutionary mechanisms behind invasiveness, both quantitative and population genetics studies are underway on *Robinia pseudoacacia* L., one of the worst invasive tree species in Europe.

**Methods** – A controlled experiment was set up using 2000 seeds from ten populations in Southern France and ten populations in Belgium. Seedlings were cultivated in two climatic chambers set at 18°C and 22°C. Early development life history traits (e.g. seedling phenology) and functional traits (e.g. growth rates) were monitored. Genotyping using SNP markers was used to evaluate the genetic differentiation among the populations and a  $Q_{ST} - F_{ST}$  comparison was done in order to test for the role of selection.

**Results** – Populations exhibited a strong plasticity to temperature for all measured traits, the warmer environment being generally more suitable, irrespective of their origin. No significant departure from neutral evolution was evidenced by the  $Q_{ST} - F_{ST}$  comparisons, although we found a slightly significant differentiation at the molecular level.

**Conclusion** – Plasticity for the functional and life history traits was evidenced but no genetic interaction suggesting no possible evolution of plasticity at those traits. Moreover, no support for genetic differentiation and local adaptation was found among studied populations within invasive range, raising two main questions: first, what is the role of selection on functional and life-history traits; and second, is the elapsed time since first introduction sufficient to allow evolution and local adaptation?

**Key words** – Biological invasion,  $Q_{ST} - F_{ST}$  comparisons, local adaptation, selection, temperature, *Robinia pseudoacacia*.

## INTRODUCTION

Biological invasions are recognized as a major threat on biodiversity (Heywood 1989, Vitousek et al. 1996, 1997, Pimentel et al. 2001). Most studies have been concentrating on ecological aspects, and the role of Darwinian evolution was often overlooked in the study of invasive species (Colautti & Lau 2015). However, one interesting aspect of biological invasions is that they represent a natural experiment allowing to study evolutionary processes (Lee 2002, Prentis et al. 2008, Colautti & Lau 2015). Invasive species are indeed an opportunity to study contemporary evolution, several decades or hundreds of years after their introduction in a new range (Keller & Taylor 2008, Colautti & Lau 2015).

Following introductions, species may have experienced modifications in their selection regime (Keller & Taylor 2008). Genetically-based changes in phenotypic traits have already been observed (Lee 2002, Blair & Wolfe 2004, Maron et al. 2004, Keller & Taylor 2008, Barrett et al. 2008, Colautti & Barrett 2013, Monty et al. 2013); for instance, a latitudinal cline in flowering period similar in both the native and invasive ranges was evidenced for populations of *Lythrum salicaria* L. (Barrett et al. 2008). Using reciprocal transplant experiments with three common gardens to evaluate phenotypic changes (Colautti & Barrett 2013) and genetic structure analysis (Chun et al. 2009), Colautti & Barrett (2013) demonstrated that local adaptation in the new range

was the most likely process explaining the observed differentiation between invasive populations of this species.

Phenotypic plasticity is another mechanism that could enhance introduced populations' establishment (Agrawal 2001, Richards et al. 2006). Indeed, thanks to plasticity, organisms could express favourable phenotypes under a wider range of environments and maintain a higher fitness (Richards et al. 2006). Consequently, plasticity could be an important factor promoting invasions success (Davidson et al. 2011, Richards et al. 2006). In this context, phenotypic plasticity is defined at the population level (Valladares et al. 2006), as the ability of genetically-related individuals to express differences in phenotypes under different environments; this level of genetic diversity is particularly relevant for understanding the ecological role of plasticity in natural populations (Monty et al. 2013, Richards et al. 2006).

Coutts et al. (2011) underlined, using a modelling approach, the importance of understanding seedling development and survival for woody invasive relatively to herbaceous ones. Germination, although the earliest stage in plant development, influences survival, fitness and adaptive capacity of plants (Donohue et al. 2010). It has been demonstrated that natural selection acted on germination timing in response to the season of seeds dispersal in the model plant *Arabidopsis thaliana* (Donohue et al. 2005). On the same species, a reciprocal transplant experiment using recombinant inbred lines demonstrated that selection occurring during germination and juvenile stages promoted local genotypes and favoured local adaptation (Postma & Ågren 2016). Therefore, understanding germination is critical for understanding local adaptation of populations (Postma & Ågren 2016), and seems particularly relevant to evaluate on newly introduced populations.

At large spatial scales, temperature is a major factor controlling species distributions (Woodward 1987, De Frenne et al. 2010) and alteration in temperature can dramatically modify tree species distributions and impact their fitness (Prentice et al. 1992, Iverson & Prasad 1998). Moreover, temperature interacts strongly with germination and seedlings recruitment by driving dormancy (i.e. initiation and break), germination phenology or seedlings vigour (Walck et al. 2011).

Among the 100 worst invasive species in Europe, seven trees and shrubs were identified (DAISIE 2009) among which *Robinia pseudoacacia* L. (Fabaceae). Black locust is a deciduous tree species, native to North America. The first European introduction in 1601 in Paris was often attributed Jean Robin, the botanist of the French King Henri IV. However, this date is discussed and it seems that the tree was probably introduced later, around 1634 (Cierjacks et al. 2013, Vítková et al. 2017), in England and France simultaneously. It is now widely distributed in natural lands throughout the continent (DAISIE 2009). In Western Europe, the species grows under a wide range of temperatures; adaptive divergences among European populations in response to this abiotic constraint can thus be hypothesized.

Our objective was to test for genetic differentiation and local adaptation to temperature in the European invasive range of *R. pseudoacacia* by comparing juvenile growth of

20 populations sampled in Wallonia (Belgium) and Aquitaine (southern France). Specifically, we will address the following questions:

- (i) Can we detect plasticity on the studied life history traits and functional traits at the population level?
- (ii) Are there phenotypic differences among populations in relation to temperature at their location of origin?
- (iii) Can we detect any departure from neutral evolution on the life history traits and functional traits?

In order to test for local adaptation within ranges, a multiple common gardens experiment and a quantitative genetic analysis of  $Q_{ST} - F_{ST}$  comparisons were combined (Keller & Taylor 2008). In common garden experiments, local adaptation is revealed when populations grown in their original environment outperform other populations. Moreover, the comparison of phenotypic  $Q_{ST}$  and genotypic  $F_{ST}$  differentiation indexes allows to propose some evolutionary inferences: under neutrality, the expectation is that  $Q_{ST}$  would be similar to  $F_{ST}$ , a  $Q_{ST}$  value significantly higher than a  $F_{ST}$  value would indicate a divergent selection acting among populations; on the contrary, a  $Q_{ST}$  value inferior to a  $F_{ST}$  value would signal a stabilizing selection (Whitlock & Guillaume 2009, Leinonen et al. 2013, Colautti & Lau 2015). The two methodological approaches are complementary to understand the evolutionary forces acting on traits.

## MATERIALS AND METHODS

### Sampling

Between January and March 2015, ten populations of *R. pseudoacacia* were sampled per range, in the Belgian region of Wallonia and in the French region of Aquitaine (fig. 1, electronic appendix 1). We sampled unplanted populations derived from natural regeneration. Mean, Minimal and Maximal temperatures of May were extracted from WorldClim version 2 (<http://worldclim.org/version2>) using the 30 seconds resolution raster. WorldClim version 2 (Fick & Hijman 2017) provides average monthly climate data for minimum, mean, and maximum temperature over the 1970–2000 period. The different populations were spread over geographic areas of the same dimensions (i.e. Gironde and Lot et Garonne, French departments in Aquitaine: 15 000 km<sup>2</sup> and Belgian Wallonia: 16 000 km<sup>2</sup>). Overall, French and Belgian populations were distant almost 800 km from one another. In each population, 10 to 100 pods were collected on ten trees with a tree pruner. Given that the species is known to spread by root-suckering (Cierjacks et al. 2013), a minimal distance of twenty meters was kept between two sampled trees in order to minimize the risk of collecting the same genotype.

Seeds were manually extracted from pods and placed into a tea paper filter to be stored in a cold room at 0–5°C (Forest Research 2015, Royal Botanic Gardens Kew 2015). Only well conserved seeds (i.e. seeds without mould stains or without damaged tegument) were kept and seeds with moist or empty appearance were systemically eliminated. Seeds were counted for each maternal tree and weighted with a 0.1mg accuracy (Practum 224-1S, Sartorius, Goettingen, Germany).

## Experimental design for phenotypic measurements

A controlled experiment was set up in two climatic chambers (Fitoclima D1200, Aralab, S Domingos de Rana, Portugal) with a total of 2000 individuals: five seeds corresponding to half-sibs  $\times$  10 maternal trees  $\times$  10 populations  $\times$  2 geographic range  $\times$  2 climatic chambers. Prior to seeding, seeds were mechanically scarified using an automated sand blasting technique (Bouteiller et al. 2017) to ensure a controlled dormancy breaking. Seeds were sowed into plastic trays (QuickPot 35RW, HerkuPlast Kubern GmbH, Ering, Germany); each pot was filled with 25 g of substrate (Substrate 307, Peltracom, Gent, Belgium) plus 6 g after sowing to cover the seeds (Bonner & Karrfalt 2008).

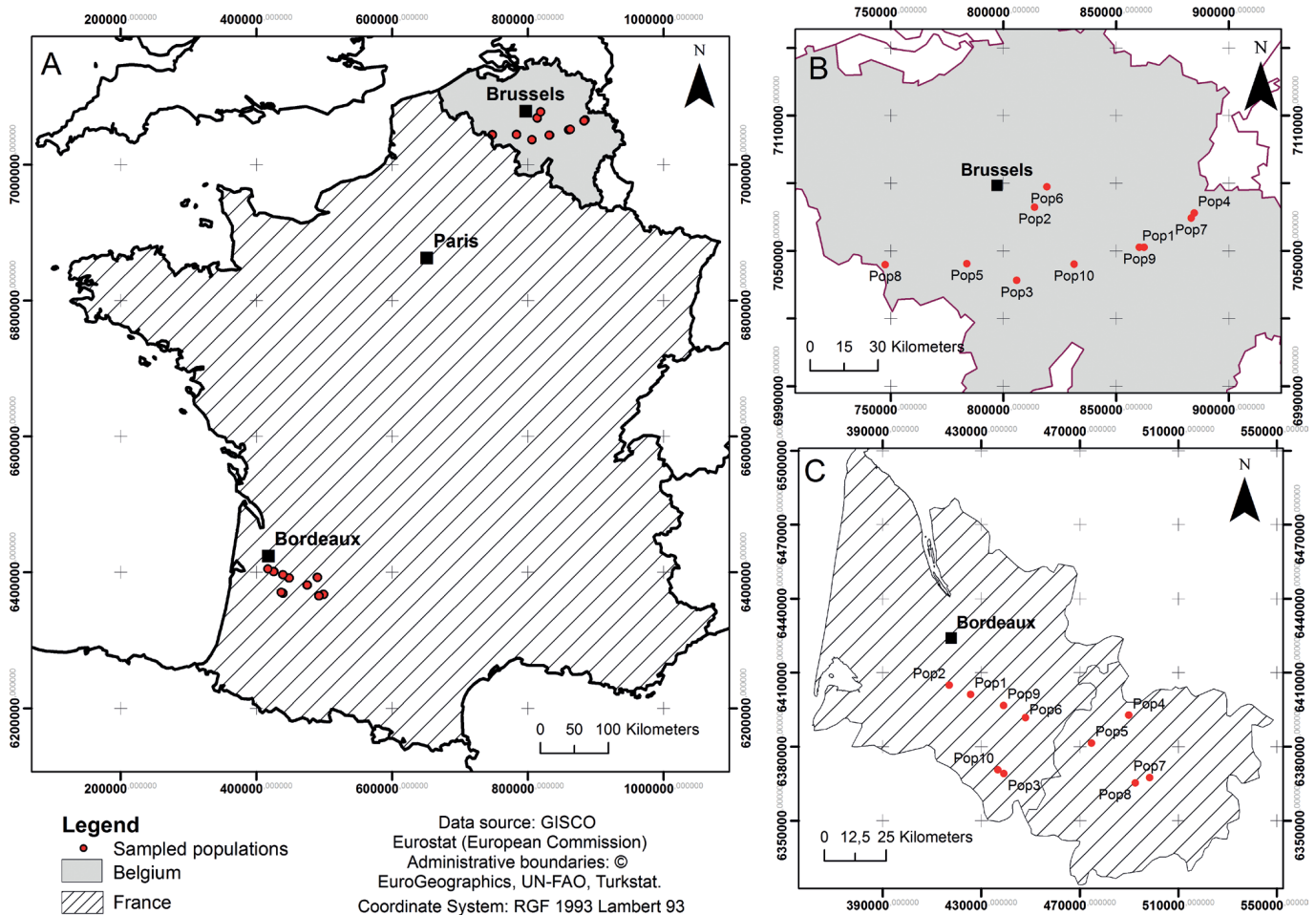
Environmental conditions within both chambers were set equal except for air temperature: a day/night 12/10 h photoperiod with progressive day/night transitions of one hour,  $60 \pm 5\%$  of air relative humidity,  $185 \pm 45 \mu\text{mol}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$  photosynthetic photon flux density (LiCOR Li190, Lincoln, USA).  $\text{CO}_2$  concentration was equal to ambient atmospheric concentration. Temperature conditions were corresponding to the average maximal temperature of May as measured at the weather station of INRA in Villenave d'Ornon (Aquitaine, France, 1987–2007) and at the weather station of Uccle (Brussels-Capital, Belgium, 1981–2010):  $22^\circ\text{C}/20^\circ\text{C}$  day/

night temperature in the first climatic chamber (French temperature conditions) and  $18^\circ\text{C}/16^\circ\text{C}$  day/night temperature in the second climatic chamber (Belgian temperature conditions).

Watering (50 mL) was provided every two days in order to supply a non-limiting resource. After 31 days, seedlings were fertilized with a liquid fertilizer (NPK 7.5.6, Florendi Jardin SAS, Dinard, France), renewed twice every 10 days.

## Phenotypic life history traits and functional traits

**Life history traits** – Germination and seedling phenology were monitored daily for each individual over a period of 440 growing degree days (GDD). GDD were calculated by multiplying the climatic chamber day temperature condition (in  $^\circ\text{C}$ ) and the number of days spent in the climatic chamber, base temperature was considered equal to 0 since no reference value was found for black locust. Five phenological stages have been defined as follows: 1, emergence; 2, straight stem; 3, open cotyledons; 4, first leaf; 5, second leaf (see Bouteiller et al. 2017 for details and photographs of the phenological stages). According to air temperature, the phenological survey was stopped after 20 days in the warmer climatic chamber and after 25 days in the colder cli-



**Figure 1** – A, map of France and Belgium presenting the sampled populations with a highlight of: B, Wallonia's Belgian populations; and C, of Aquitaine's French populations. © EuroGeographics for the administrative boundaries.

matic chamber. For each phenological stage and each individual, the minimum number of days required to reach the stage was noted. When a stage was missing, the number of days was calculated as the average between the numbers of days required to reach the previous and the stage. When an individual died after germinating, it was further recorded as missing value.

**Functional traits** – At the end of the phenological survey (440 GDD), total height was measured from collar to the apical bud using a ruler (cm ~1mm). Seedlings were thinned from 5 to 3 plants per pot in order to avoid competition, preferentially removing damaged individuals. After 1342 GDD corresponding to 61 and 75 days in the 22°C and 18°C climatic chamber, respectively, total eight (cm ~1mm), stem collar diameter (mm ~ 0.01mm) and the effective quantum yield of photosystem II (PSII yield) were measured using a ruler, an electronic calliper and a portable chlorophyll fluorometer (PAM 2100, Heinz Walz GmbH, Effeltrich, Germany) respectively.

The PSII yield was measured on the terminal leaflet of the youngest fully-developed. This trait provides information about the proportion of light absorbed by the PSII that is used in photochemistry (Genty et al. 1989, Santiso et al. 2015).

Then, seedlings were separated into leaves and stem, and oven dried at 65°C (Universal oven, Memmert, Swabach, Germany). After minimum one drying week to reach constant weight, leaves and stem were weighted with a 0.1 mg accuracy (Practum 224-1S, Sartorius, Goettingen, Germany). Growth rates in height, diameter and leaf and stem biomass were calculated by dividing the trait value by the number of days spent in each climatic chamber. Data are accessible on Open Science Framework repository (<https://osf.io/q3ebp/>).

### DNA extraction and SNP genotyping

In order to evaluate molecular genetic diversity, a 1 cm<sup>2</sup> leaf sample was collected on one individual per family sampled randomly among the thinned individuals. DNA was extracted and isolated using DNAeasy 96 Plant Kit (Qiagen, Venlo, Netherlands) following the manufacturer's protocol. One negative control was set on each plate. DNA concentration was measured using an UV spectrophotometer NanoDrop 8000 (Thermo Fisher Scientific Inc., Wilmington, Delaware, USA) and confirmed using QuantiFluor dsDNA System (Promega, Fitchburg, Wisconsin USA). Besides DNA concentration, 260/280 and 260/230 absorbance ratio provide information about DNA purity. Since the concentration was about 10ng/μL, DNA was further used for SNP amplification without dilution. SNPs were recently developed on black locust using the RADseq approach (Verdu et al. 2016). Four multiplexes of 36 SNPs each were further designed with the MassArray assay design 4.1 software and genotyped with the Sequenom MassARRAY iPLEX platform at the Bordeaux Genome Transcriptome facility (<https://pgtb.cgfb.u-bordeaux.fr/>), using the iPLEX Gold genotyping kit according to the manufacturer's instructions. Products were then analysed using the MassArray mass spectrophotometer (Sequenom) and data were acquired in real time using the MassArray RT software.

SNP data were then visualized and validated using ViClust, a R program that we implemented for Galaxy (<https://usegalaxy.org/>) and that is also available as a standalone R script for Linux or Windows at <https://github.com/garniergere/ViClust/>. Briefly, ViClust imports the unmodified xml file to export from the Typer-4\_0\_20 in-house Sequenom software. It then allows to visualize the clusters proposed by the Sequenom method based on Gaussian mixtures (Johansen et al. 2013), proposes additional and more or less stringent filters on the different allele signal magnitudes, and performs an alternative hierarchical clustering method using the Ward algorithm (Ward 1963). This method relaxes the constraining assumption of normal distribution for allele signals. The program gives SNPs plots across the different alternatives in batch, allows comparing and validating each SNPs more easily, and exports automatically the proposed genotype assignments, making the whole process much less time consuming than using Typer-4\_0\_20. Out of the initial 144 SNPs genotyped, 132 SNPs were finally, kept for analyses. Data are accessible on Open Science Framework repository (<https://osf.io/q3ebp/>).

### Data analyses

Molecular genetic differentiation was explored using two approaches. First, the typology of populations was assessed using a Correspondence Analysis (CA) on SNP data; it was implemented using the adegenet library (Jombart 2008) in R version 3.2.2 (R Development Core Team 2015). As recommended by Jombart et al. (2009), rare alleles, defined as those recorded less than 5% of the total allele number for each marker, were removed from the analysis.

Second, population structure was analysed using STRUCTURE v2.3.4 (Pritchard et al. 2000). Using the admixture model with no a priori hypothesis, 20 independent Markov chains, each with K (number of assumed clusters) values allowed to range from 1 to 20, were run with a burn-in period of 500,000 iterations followed by 1 500 000 iterations. STRUCTURE runs were computed on the GenoToul bioinformatics facility (<http://bioinfo.genotoul.fr/>). The script StrAuto (Chhatre and Emerson 2017) was used to produce STRUCTURE mainparams and extraparams and to automatize and parallelise STRUCTURE analysis. STRUCTURE HARVESTER (Earl & vonHoldt 2012) was employed to calculate ΔK as described in Evanno et al. (2005) in order to determine the most likely K. Finally, CLUMPAK (Kopelman et al. 2015) was utilised to synthetize STRUCTURE outputs and to compute graphical STRUCTURE outputs.

To evaluate the germination rate of *R. pseudoacacia* seeds among ranges and temperatures, we constructed a germination curve on the fraction of germinated seeds,  $G(t)$  (for more details see Bouteiller et al. 2017). A Gompertz model was fitted using a Bayesian procedure (equation 1). The code used in this study is provided on the open source GitHub platform (<https://github.com/xbouteiller/GompertzFit>).

$$G(t) = D * \exp(-\exp(-b(t - t_m))) \quad (1)$$

where parameter  $D$  equals the maximum germination rate,  $b$  reflects the slope of the germination curve and  $t_m$  is the time at the inflexion point.

To estimate differentiation between populations in life history traits and functional traits between climatic chambers (i.e. the temperature effect), a Bayesian mixed model of ANCOVA was fitted for each trait, using the mothers' average seed weight as covariate. There is indeed a significant difference (electronic appendix 3A) among ranges for seed weight. French families' average seed weight ( $23.84 \pm 3.33$  mg ranging from 13.79 mg to 31.65 mg) is significantly higher than Belgium families' average seed weight ( $18.99 \pm 3.34$  mg ranging from 11.47 mg to 26.02 mg). Moreover, there are significant differences among populations (electronic appendix 3B) with some populations presenting heavier (e.g. Fr.Pop4, Fr.Pop1) and lighter (e.g. Bel.Pop8) seed weights.

The between-chambers model was defined as:

$$Y_{ijklm} = b_o + \underbrace{b_{1,i}}_{temp} + \underbrace{b_{2,j}}_{range} + \underbrace{B_{3,jk}}_{population} + \underbrace{B_{4,jkl}}_{tree} + \underbrace{B_{5,ijk}}_{temp \times pop} + \underbrace{c(x_{ijkl} - \bar{x})}_{maternal\ effect} + \varepsilon_{ijklm} \quad (2)$$

With the residual as:

$$\varepsilon_{ijklm} \sim N(0, \sigma^2) \quad (3)$$

Lowercases indicate the fixed effects and uppercases indicate random effects, with

$$i \in 1, 2$$

for the 18 and 22°C temperature chamber conditions;

$$j \in 1, 2$$

for the two ranges: Belgium, France;

$$k \in [1, 10]$$

for the 10 populations in each range;

$$l \in [1, 10]$$

for the 10 trees per population; and

$$m \in [1, 5]$$

the 5 individuals (seeds) per tree.

$x_{ijkl}$  represents the seed weight at the tree (i.e. family level) and  $\bar{x}$  the global mean seed weight.

We used JAGS 3.4.0 (Plummer 2005), the R2jags package of R (Su & Yajima 2012) and R version 3.2.2 (R Development Core Team 2015) to compute the Bayesian model. For each life history traits and functional traits, 150,000 iterations were run with a burn-in of 125 000 iterations and a thinning interval of 10, using 4 different chains. Autocorrelation and convergence were assessed using the "autocorr.plot" and "gelman.plot" native JAGS functions and Rhat convergence criterion was inferior to 1.01 as recommended (Kruschke 2014).

Codes provided by O'Hara & Merilä (2005) and Kruschke (2014) were adapted to our design using uninformative inverse gamma conjugate priors for random effects and normal priors for fixed effects.

Genetic differentiation among populations can be defined from molecular markers analyses using  $F_{ST}$  indices calculated as follows (Wright 1949):

$$F_{ST} = \frac{\sigma_B^2}{(\sigma_B^2 + \sigma_W^2)} \quad (4)$$

The  $F_{ST}$  value for the whole dataset was estimated using the HICKORY 1.1 software (Holsinger 1999).  $F_{ST}$  was estimated by running the software native's full model (i.e. assuming non-null inbreeding coefficient and significant genetic structure among populations) and compared to the non-inbreeding and no population structure models, as recommended by Holsinger (1999). The full model outperformed the other models, based on the DIC criterion values (see HICKORY documentation for more details on the models). Pairwise  $F_{ST}$  were calculated between all pairs of the 20 populations using Genepop v4.3.3 (Raymond & Rousset 1995, Rousset 2008).

For each quantitative life history traits and functional traits, the  $Q_{ST}$  index of quantitative genetic differentiation among populations was calculated as follows (Spitze 1993):

$$Q_{ST} = \frac{\sigma_B^2}{(\sigma_B^2 + 2\sigma_W^2)} \quad (5)$$

where

$$\sigma_W^2$$

is the within population genetic variance and

$$\sigma_B^2$$

is the between population genetic variance estimated using

$$\hat{\sigma}_{pop}^2$$

Assuming that individuals obtained from seeds of the same tree were half-sibs, within population variance can be estimated as

$$\hat{\sigma}_W^2 = 4\hat{\sigma}_{tree}^2$$

$Q_{ST}$  values were computed per temperature chamber condition using the same model that presented above (equation 1) but without the temperature and temperature x population interaction effects.

The code used in this study is provided on the open source GitHub platform (<https://github.com/xbouteiller/BayesMix>). It was tested by comparison to the same model computed with SAS software using the frequentist PROC MIXED procedure (SAS, version 9.1, SAS Institute, Cary, NC, USA); Bayesian  $Q_{ST}$  estimations (mean + CI) were compared to estimates using first the delta method computed with SAS (O'Hara & Merilä 2005, Isik 2009) and second using the R package QstFstComp (Gilbert & Whitlock 2015). Comparisons are presented in electronic appendix 2. Results were similar in almost all traits. Frequentist methods can estimate weak negative variances, when those are close to zero at the population level, which we found in preliminary explorations. Given that our choice of priors in the Bayesian models (with positive support) made such negative estimates of variances impossible, and that the models showed overall a good performance, we adopted a Bayesian approach to estimate  $Q_{ST}$  (O'Hara & Merilä 2005).

RESULTS

**Molecular genetic differentiation: a weak structure and few outlying populations**

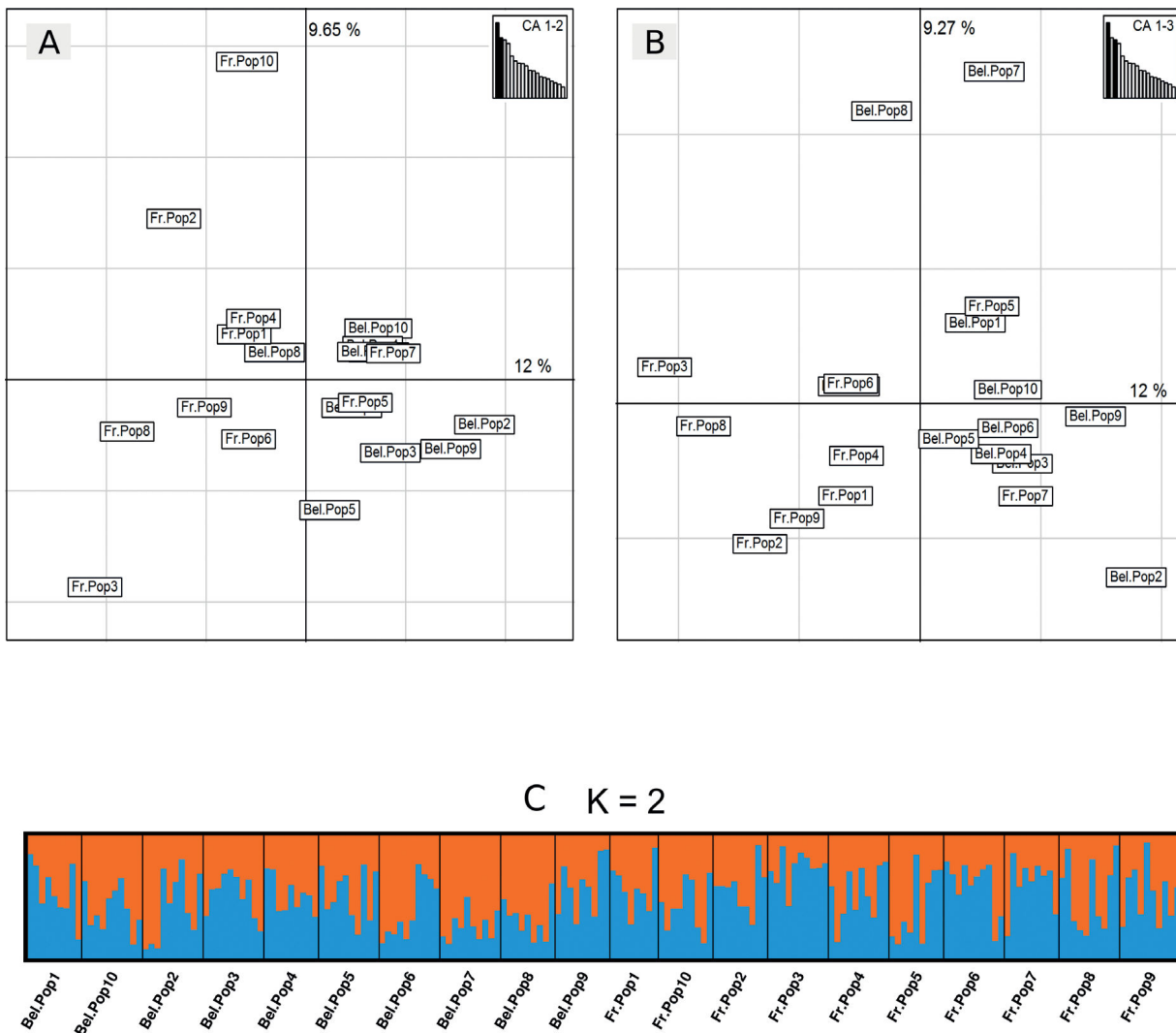
CA on SNP data was first characterized by the grouping of most of the populations, both French and Belgian (fig. 2A & B). Nevertheless, respectively three populations (Fr.Pop2, 3, 10) and five populations (Fr.Pop3, 8 and Bel.Pop2, 7, 8) were separated from this melted group of populations when considering either the first two axes or axes 1 and 3. Fig. 2 shows the major mode identified by CLUMPAK on 20 STRUCTURE runs for K = 2 (fig. 2C). Higher  $\Delta K$  was obtained for K = 2. Thus optimal clustering was most likely for K = 2 clusters (Earl & vonHoldt 2012, Evanno et al. 2005). Structure analysis exhibited a similar pattern than the one observed in the CA: a weak population structure has been revealed. Indeed the STRUCTURE analysis revealed a very

high level of admixture in most populations: less than 3.5% of the individuals had an inferred ancestry higher than 90%.

**Genetic vs environment effect on life history traits and functional traits: a strong response to temperature without genetic interaction**

Germination curves per range and temperature as well as the fitted Gompertz model are presented on fig. 3. Germination was faster under the warmest temperature condition but maximum germination rates, D, were similar whatever the temperature and the range: total germination rate was established at  $84.9 \pm 7.4\%$ .

*Robinia pseudoacacia* seedlings demonstrated a strong positive response to temperature for all considered traits, whatever the populations and range of origin in Europe. Fig. 4A–L show the distributions of trait values within each



**Figure 2** – Correspondence analysis (A & B) and STRUCTURE analysis (C) and on SNP genotyping data. For the Correspondence Analysis, first (A) axis 1 was plotted against axis 2, second (B) axis 1 was plotted against axis 3. First axis contains 13% of variation, second axis contains 10.3% of variation and third axis contains 9.54% of variation. STRUCTURE plots represent major mode for K = 2 (C) identified by CLUMPAK on 20 STRUCTURE runs for 190 individuals from 20 populations using 132 SNP markers. K = 2 corresponds to the optimal K number based on  $\Delta K$  calculated using STRUCTURE HARVESTER. Labels are indicating sampled populations. Fr: French, Bel: Belgian.



**Table 1 – Mean values ( $\pm$  standard deviations) of all the surveyed traits within each climatic chamber (i.e. 18°C and 22°C).**

Bold characters indicate significant differences between chambers in response to temperature. The temperature effect was always significant except for the biomass traits. GR: Growth Rate

Climatic chamber	18°C	22°C
<b>Stage 1 (d)</b>	8.2 $\pm$ 2.5	6.0 $\pm$ 1.7
<b>Stage 2 (d)</b>	10.0 $\pm$ 2.8	7.5 $\pm$ 1.7
<b>Stage 3 (d)</b>	13.0 $\pm$ 3.0	9.6 $\pm$ 2.1
<b>Stage 4 (d)</b>	20.0 $\pm$ 2.7	14.0 $\pm$ 1.9
<b>Stage 5 (d)</b>	25.0 $\pm$ 1.5	18.0 $\pm$ 1.5
<b>Height GR 1 (cm.d<sup>-1</sup>)</b>	0.16 $\pm$ 0.034	0.2 $\pm$ 0.045
<b>Height GR 2 (cm.d<sup>-1</sup>)</b>	0.15 $\pm$ 0.055	0.18 $\pm$ 0.059
<b>PSII Yield</b>	0.0014 $\pm$ 0.0011	0.003 $\pm$ 0.0012
Leaf weight GR (g.d <sup>-1</sup> )	0.0032 $\pm$ 0.002	0.0034 $\pm$ 0.0023
Stem weight GR (g.d <sup>-1</sup> )	0.0010 $\pm$ 0.00073	0.0011 $\pm$ 0.00099
Total weight GR (g.d <sup>-1</sup> )	0.0042 $\pm$ 0.0027	0.0045 $\pm$ 0.003
<b>Diameter GR (mm.d<sup>-1</sup>)</b>	0.026 $\pm$ 0.0066	0.030 $\pm$ 0.0088

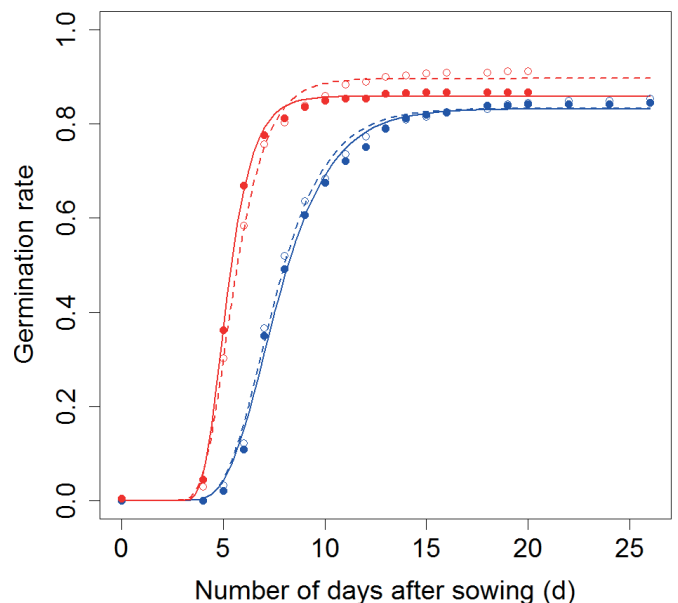
climatic chamber and among populations: a strong response to temperature could be evidenced, but no clear population effect or range effect was easily identifiable. Seedling development was faster under warmer temperature conditions (fig. 5, electronic appendix 4A–La). For example, (table 1), number of days for seeds to germinate (to reach phenological stage 1) was 6.0  $\pm$  1.7 days at 22°C versus 8.2  $\pm$  2.5 days at 18°C; similarly, time to reach phenological stage 5 with a fully developed composed leaf was shorter at 22°C compared to 18°C (18  $\pm$  1.5 days and 25  $\pm$  1.5 days, respectively). For functional and physiological traits, increases in trait values with increases in temperature were observed for most of the traits: height and diameter growth rates were significantly higher at 22°C as well as PSII yield (table 1). However, biomass showed no significant increase with increased temperature.

Neither a genetic nor a genetic  $\times$  environment interaction could be found, irrespective of the trait (fig. 4, electronic appendix 4A–L b & c). The genetic variability was observed at the tree level for all traits (electronic appendix 4A–L d), variances between trees being 18 to 37 times higher than variances between populations. For instance, for phenological stage 3 and diameter growth rate, variance between trees equalled and , whereas variance between populations only represented and . Finally, as observed on fig. 5 and in electronic appendix 4d, no temperature  $\times$  population interaction was observed. The temperature  $\times$  population interaction variance was of the same magnitude than the between populations variance, and always lower than the between trees variance, for all traits.

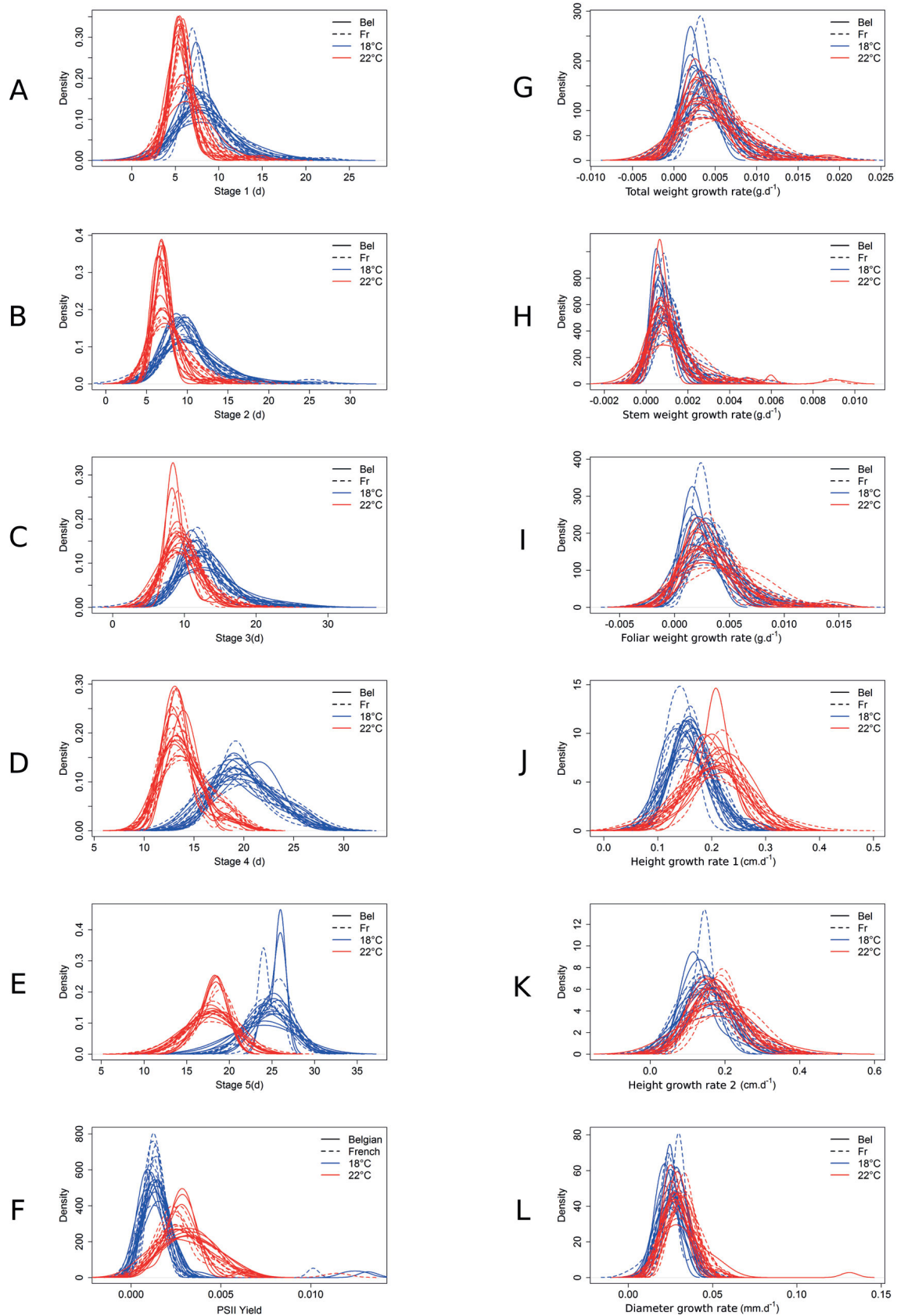
To sum up, populations exhibited a strong plasticity to temperature for all measured traits, the warmer environment being generally more suitable. Thus, at the population and range levels, similar responses have been observed in response to an increase in temperature. Belgian populations did not outperform French populations when in their natural temperature range and reciprocally.

### $Q_{ST} - F_{ST}$ comparisons: no evidence for deviation from neutrality

A moderate but significant genetic differentiation among populations was evidenced at the molecular level, with a  $F_{ST}$  value of 0.026 (fig. 5, 95%CrI = 0.021 – 0.030). Using all measured life history traits and functional traits, no genetic differentiation was evidenced, with  $Q_{ST}$  values not significantly different from 0 (fig. 5), whatever the temperature conditions. Beware that  $Q_{ST}$  values for phenological stage 5



**Figure 3 – Germination curves of *Robinia pseudoacacia* seeds per temperature condition (blue: 18°C, red: 22°C) and per range (solid lines, full circles: Belgian range, dashed lines, open circles: French range). Data were reported with circles and fitted curves using a Gompertz function with lines.**



**Figure 4** – Density distribution functions for each life history trait and functional trait according to the temperature conditions (red line 22°C, blue line 18°C) and the range (full line Belgium, dotted line France). A–E, phenological stages 1 to 5; F, PSII Yield, G, total weight GR, H, stem weight GR, I: Foliar weight GR, J–K, height GR 1 and 2, L, diameter GR. GR: Growth Rate.

at 18°C and stem weight increment at 22°C are not robust estimates due to a large number of missing values.

Moreover, no deviation from neutrality was evidenced for all traits among the twenty populations. Indeed, all  $Q_{ST}$  values were not significantly different from the  $F_{ST}$  value except for the PSII yield at 22°C that was significantly lower (0.0026, 95%CrI = 0 – 0.013).

## DISCUSSION

### A strong response to temperature for all traits

A strong response to temperature for most traits was evidenced inducing a faster phenological development and superior trait values in the warmest environment.

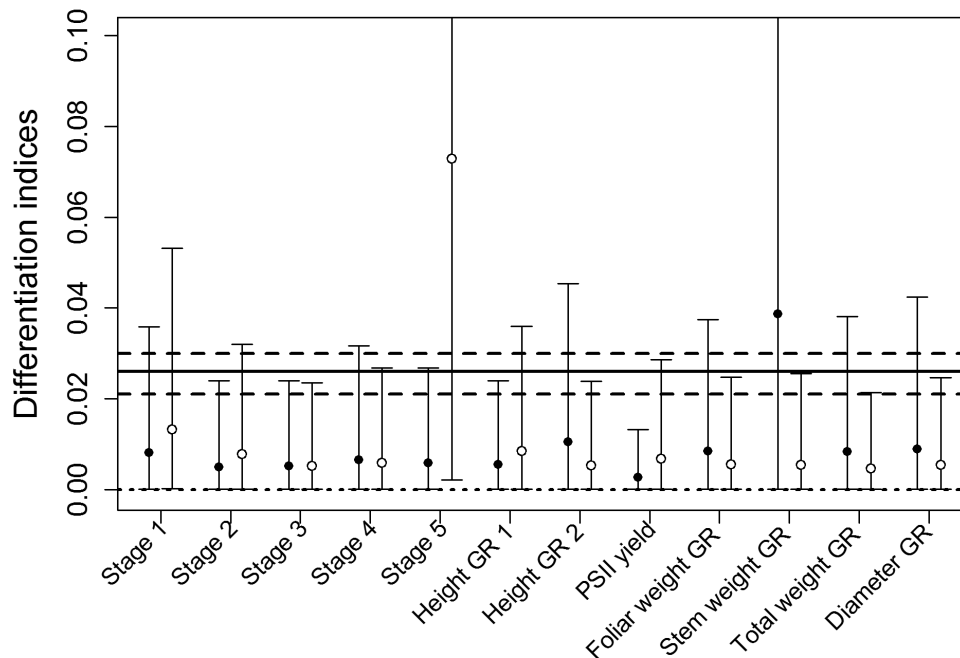
The strong plasticity at the population level in response to temperature observed on juvenile *R. pseudoacacia* in this study could help understand its large invasion range in Europe. Indeed, broad environmental tolerance species are more susceptible to encounter favourable conditions when facing a new environment (Davidson et al. 2011). Invasive plants are commonly more phenotypically plastic than co-occurring native species for a wide variety of traits such as growth or fitness-related traits (Maron et al. 2007, Hyldgaard & Brix 2012, Molina-Montenegro & Naya 2012, Lamarque et al. 2015). For instance, a strong plasticity in response to temperature has been demonstrated for five invasive populations of *Taraxacum officinale* F.H.Wigg. in several ecophysiological and fitness related traits (Molina-Montenegro & Naya 2012); this plasticity could be an important promoting factor of in-

vasion success. Moreover, plasticity in germination traits could be an important factor for successful establishment and persistence of invasive species impacting their spreading and persisting capacities in new spaces (Gioria & Pyšek 2017). A study comparing germination plasticity among twelve native and twelve invasive species using three environmental regimes based on temperature, day length, and soil moisture demonstrated that invasive species had higher performances (germination speed and rate) and plasticity than native species (Wainwright & Cleland 2013).

Plasticity of phenotypic traits is adaptive when it improves survival and reproduction (Richards et al. 2006). In the context of biological invasions, phenotypic plasticity could evolve rapidly during the lag phase after introduction for ecologically important traits and thus contribute to the invasion success (Agrawal 2001, Richards et al. 2006). However, it seems that the majority of phenotypic plasticity is selectively neutral (Davidson et al. 2011), although it could be particularly beneficial during the early stages of the invasion. In order to assess the adaptive role of plasticity of juvenile traits of *R. pseudoacacia* during the invasion further comparisons need to be made using both native and invasive populations (Richards et al. 2006, Davidson et al. 2011).

### No signal of local adaptation among European populations of *R. pseudoacacia*

In the present study, local adaptation was tested by growing populations in climatic chambers at two different temperatures; we also compared genetic structure using molecular markers.



**Figure 5** – Differentiation indices calculated using phenotypic life history traits and functional traits ( $Q_{ST}$ ) and using SNP markers ( $F_{ST}$ ).  $Q_{ST}$  values (mean and 95% confidence interval) were calculated on the 20 populations per trait and growing chamber (closed circle: 22°C, open circle: 18°C). The full line represents the mean  $F_{ST}$  and the dashed lines its 95% confidence interval based on allelic variation in SNP's loci. Estimations were calculated using Bayesian methods. GR: Growth Rate.

A weak structure was detected between French and Belgian populations since only few individuals had an inferred ancestry higher than 90%. Moreover, Belgian populations have not outperformed French populations under the coolest temperature conditions resembling the Belgian conditions, nor reciprocally for the French populations under the warmer environment. Thus, based on the measures made on very early stage life history traits and functional traits, no signal of local adaptation was evidenced among the twenty European populations studied. On the contrary, signs of local adaptation have already been observed in invasive populations of the herbaceous species *Hypericum perforatum* L. (Maron et al. 2004), *Eschscholzia californica* Cham. (Leger & Rice 2007) and *Lythrum salicaria* (Colautti & Barrett 2013). However, the opposite result was also evidenced among invasive populations of *Polygonum cespitosum* Blume (Mate-sanz et al. 2012) or when testing introduced populations of invasive shrubs as *Buddleja davidii* Franch. whose populations responded similarly to environmental changes but exhibited high phenotypic plasticity (Ebeling et al. 2011).

In addition, our results showed no differences between  $Q_{ST}$  and  $F_{ST}$  values: no deviation from neutrality can be inferred for the measured traits. They would only be influenced by genetic drift, mutation and migration (Keller & Taylor 2008, Whitlock 2008). However, as presented in the introduction, early life stages are likely to be under selection so we can wonder if methodological biases could impede detection of local adaptation in  $Q_{ST} - F_{ST}$  comparisons. This could happen if any maternal, environmental, dominance or epistasis effects are present (Whitlock 2008, Leinonen et al. 2013). To avoid these biases, a robust 20 populations of half sibs experiment was designed, which was suitable to estimate additive variance and  $Q_{ST}$  (O'Hara & Merilä 2005, Goudet & Büchi 2006, Whitlock 2008, Leinonen et al. 2013). Similarly to our study, Merilä & Crnokrak (2001) found that life history traits could exhibit no significant divergence between  $Q_{ST}$  and  $F_{ST}$  whereas functional traits showed significant differences between  $Q_{ST}$  and  $F_{ST}$ ; they suggested that life history traits contain more non additive effects, and are consequently more subject to biases downward or upward the  $Q_{ST}$ . Still, in our study, both functional and life history traits were surveyed, and we used the mean seed weight per family as a covariate to control for maternal effects (Whitlock 2008, Chun et al. 2011). As a conclusion, we can consider that methodological biases are unlikely to explain our results that are consistent across all the traits considered.

A handful studies used  $Q_{ST} - F_{ST}$  comparisons to infer the role of directional selection occurring during biological invasion, and their results could therefore be compared to ours, but their results are so far contrasted. For instance, Chun et al. (2011) found a  $Q_{ST}$  for reproductive allocation superior to the  $F_{ST}$  by comparing ten invasive populations of *Ambrosia artemisiifolia* L., introduced to Europe in the second part of the 19<sup>th</sup> century, suggesting that divergent selection is acting among populations (Chun et al. 2011). Similarly, populations of the invasive *Phalaris arundinacea* L. exhibited  $Q_{ST}$  superior to  $F_{ST}$  in the invasive range (Lavergne & Molofsky 2007). Lastly, a similar pattern was observed for *Centaurea solstitialis* L. (Eriksen et al. 2012). In contrast, a biological invasion study on *Lythrum salicaria* failed to demonstrate

$Q_{ST}$  superior to  $F_{ST}$  even though recent local adaptation in the new range was demonstrated (Chun et al. 2009, Colautti & Barrett 2013). This highlights that in the case of invasive species,  $Q_{ST} - F_{ST}$  comparisons can be difficult to interpret, especially if  $F_{ST}$  is small or if highly differentiated native populations are introduced into narrower environmental range or because of the history of introduction (Colautti & Lau 2015). Other methods exist to estimate  $Q_{ST}$  such as multivariate approaches that can strengthen precision when population number is low, or could allow to disentangle genetic drift and selection with a better accuracy when neutral differentiation is high (Ovaskainen et al. 2011, Leinonen et al. 2013). Considering that we used a robust experimental design and that we observed a very weak genetic differentiation both with neutral molecular markers and with quantitative traits, we kept a trait by trait approach (O'Hara & Merilä 2005, Leinonen et al. 2013).

In our study, the most probable hypothesis is thus that evolution in the local range did not play a role yet in the genetic differentiation and local adaptation of the studied European populations of *R. pseudoacacia*. Indeed, ability to detect a  $Q_{ST} - F_{ST}$  difference evolves with time since the introduction (Hendry 2002) and in young systems in which selection would not have enough time to act and to produce divergent populations, similar results with  $Q_{ST}$  not exceeding  $F_{ST}$  would be expected (Whitlock 1999). In invasion studies, it has been shown that selection and local adaptation can act fast over a few generations (Koskinen et al. 2002, Dlugosch & Parker 2008). Since its first introduction in the early 17<sup>th</sup> century in Europe, *R. pseudoacacia* probably produced a maximum of 70 generations, considering a first flowering age of approximately six years (Burns & Honkala 1990). Although that could appear to be a sufficient number of generations when compared to results observed in *Hypericum canariense* L. (Dlugosch & Parker 2008), the number of generations required to evidence local adaptation through adaptive evolution may be increased in *R. pseudoacacia* due to the contribution of its clonal reproduction. To fully conclude on the role of evolution in shaping the diversity of *R. pseudoacacia* in Europe, a broader sampling including both the native and invasive range would be required (Keller & Taylor 2008, Colautti & Lau 2015).

## CONCLUSION

In this study we studied the phenotypic differentiation of the invasive tree *Robinia pseudoacacia* among 20 populations from two parts of the invasive range (i.e. Aquitaine, France and Wallonia, Belgium). We followed germination and juvenile stages development. We demonstrated that most variability was within populations at the family (i.e. tree) level without significant differentiation among populations. Moreover, a  $Q_{ST} - F_{ST}$  comparison highlighted no deviation from neutrality. Studies including a broader sampling both in the native and invasive ranges are needed to investigate the role of evolution during the invasion by *R. pseudoacacia*.

## SUPPLEMENTARY DATA

Supplementary data are available in pdf at *Plant Ecology and Evolution*, Supplementary Data Site (<http://www.ingenta-connect.com/content/botbel/plecevo/supp-data>) and consist of: (1) summary of the sampled populations with their geographic coordinates; (2) results of the mixed model computed using frequentist delta method with SAS and comparisons of the estimated  $Q_{ST}$  computed with 3 different methods; (3) results of the Bayesian analysis for the seed weight trait; and (4) results of the Bayesian analyses for all phenotypic traits computed using the full between chambers mixed model.

## ACKNOWLEDGEMENTS

We are very grateful to Alexandra Quénu, Benjamin Dencausse, Nastasia Merceron, Patrick Reynet for their help in choosing populations, and to Adline Delcamp and Myriam Paillet for their help in realizing the lab work. This study has been carried out with financial support from the French National Research Agency (ANR) in the frame of the Investments for the future Program, within the Cluster of Excellence COTE (ANR-10-LABX-45). This study was funded by the ANR-10-EQPX-16 Xyloforest.

## AUTHORS CONTRIBUTIONS

L. Lassois, S. Mariette, A. Monty, A.J. Porté and C. Verdu conceived the experiments; X.P. Bouteiller conducted the experiments with the contribution of C. Verdu, S. Mariette, A.J. Porté, A. Raimbault and R. Segura; P. Garnier-Géré, N. Harmand, Y. Laizet and S. Mariette developed the program ViClust for SNP validation; X.P. Bouteiller analysed the SNP and phenotypic data with the contribution of A. Raimbault, S. Mariette, F. Barraquand and A.J. Porté; X.P. Bouteiller, S. Mariette and A.J. Porté wrote the paper, S. Mariette and A.J. Porté are both senior authors of the study. All authors read and approved the final version of the manuscript. They declare no conflict of interest.

## REFERENCES

- Agrawal A.A. (2001) Phenotypic plasticity in the interactions and evolution of species. *Science* 294: 321–326. <https://doi.org/10.1126/science.1060701>
- Barrett S.C.H., Colautti R.I., Eckert C.G. (2008) Plant reproductive systems and evolution during biological invasion. *Molecular Ecology* 17: 373–383. <https://doi.org/10.1111/j.1365-294X.2007.03503.x>
- Blair A.C., Wolfe L.M. (2004) The evolution of an invasive plant: an experimental study with *Silene latifolia*. *Ecology* 85: 3035–3042. <https://doi.org/10.1890/04-0341>
- Bonner F.T., Karrfalt R.P. (2008) The woody plant seed manual. Agriculture Handbook 727. Washington, D.C., U.S. Department of Agriculture, Forest Service.
- Bouteiller X.P., Porté A.J., Mariette S., Monty A. (2017) Using automated sanding to homogeneously break seed dormancy in black locust (*Robinia pseudoacacia* L., Fabaceae). *Seed Science Research* 27: 243–250. <https://doi.org/10.1017/S0960258517000150>
- Burns R.M., Honkala B. (1990) Silvics of North America: 2. Hardwoods. Agriculture Handbook 654. Washington, D.C., U.S. Department of Agriculture, Forest Service.
- Chhatre V.E., Emerson K.J. (2017) StrAuto: automation and parallelization of STRUCTURE analysis. *BMC Bioinformatics* 18: 192. <https://doi.org/10.1186/s12859-017-1593-0>
- Chun Y.J., Nason J.D., Moloney K.A. (2009) Comparison of quantitative and molecular genetic variation of native vs. invasive populations of purple loosestrife (*Lythrum salicaria* L., Lythraceae). *Molecular Ecology* 18: 3020–3035. <https://doi.org/10.1111/j.1365-294X.2009.04254.x>
- Chun Y.J., 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. <https://doi.org/10.1111/j.1365-294X.2011.05013.x>
- Cierjacks A., Kowarik I., Joshi J., Hempel S., Ristow M., von der Lippe M., Weber E. (2013) Biological Flora of the British Isles: *Robinia pseudoacacia*. *Journal of Ecology* 101: 1623–1640. <https://doi.org/10.1111/1365-2745.12162>
- Colautti R.I., Barrett S.C.H. (2013) Rapid adaptation to climate facilitates range expansion of an invasive plant. *Science* 342: 364–366. <https://doi.org/10.1126/science.1242121>
- Colautti R.I., Lau J.A. (2015) Contemporary evolution during invasion: evidence for differentiation, natural selection, and local adaptation. *Molecular Ecology* 24: 1999–2017. <https://doi.org/10.1111/mec.13162>
- Coutts S.R., van Klinken R.D., Yokomizo H., Buckley Y.M. (2011) What are the key drivers of spread in invasive plants: dispersal, demography or landscape: and how can we use this knowledge to aid management? *Biological Invasions* 13: 1649–1661. <https://doi.org/10.1007/s10530-010-9922-5>
- DAISIE (2009) Handbook of alien species in Europe. Dordrecht, Springer series in invasion ecology. <https://doi.org/10.1007/978-1-4020-8280-1>
- Davidson A.M., Jennions M., Nicotra A.B. (2011) Do invasive species show higher phenotypic plasticity than native species and, if so, is it adaptive? A meta-analysis. *Ecology Letters* 14: 419–431. <https://doi.org/10.1111/j.1461-0248.2011.01596.x>
- De Frenne P., Graae B.J., Kolb A., Brunet J., Chabrierie O., Cousins S.A.O., Decocq G., Dhondt R., Diekmann M., Eriksson O., Heinken T., Hermy M., Jögar Ü., Saguez R., Shevtsova A., Stanton S., Zindel R., Zobel M., Verheyen K. (2010) Significant effects of temperature on the reproductive output of the forest herb *Anemone nemorosa* L. *Forest Ecology and Management* 259: 809–817. <https://doi.org/10.1016/j.foreco.2009.04.038>
- Dlugosch K.M., Parker I.M. (2008) Invading populations of an ornamental shrub show rapid life history evolution despite genetic bottlenecks. *Ecology Letters* 11: 701–709. <https://doi.org/10.1111/j.1461-0248.2008.01181.x>
- Donohue K., Dorn L., Griffith C., Kim E., Aguilera A., Polisetty C.R., Schmitt J. (2005) The evolutionary ecology of seed germination of *Arabidopsis thaliana*: variable natural selection on germination timing. *Evolution* 59: 758–770. <https://doi.org/10.1111/j.0014-3820.2005.tb01751.x>
- Donohue K., Rubio de Casas R., Burghardt L., Kovach K., Willis C.G. (2010) Germination, postgermination adaptation, and species ecological ranges. *Annual Review of Ecology, Evolution, and Systematics* 41: 293–319. <https://doi.org/10.1146/annurev-ecolsys-102209-144715>
- Earl D.A., vonHoldt B.M. (2012) STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* 4: 359–361. <https://doi.org/10.1007/s12686-011-9548-7>

- Ebeling S.K., Stöcklin J., Hensen I., Auge H. (2011) Multiple common garden experiments suggest lack of local adaptation in an invasive ornamental plant. *Journal of Plant Ecology* 4: 209–220. <https://doi.org/10.1093/jpe/rtr007>
- Eriksen R.L., Desronvil T., Hierro J.L., Kesseli R. (2012) Morphological differentiation in a common garden experiment among native and non-native specimens of the invasive weed yellow starthistle (*Centaurea solstitialis*). *Biological Invasions* 14: 1459–1467. <https://doi.org/10.1007/s10530-012-0172-6>
- Evanno G., Regnaut S., Goudet J. (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* 14: 2611–2620. <https://doi.org/10.1111/j.1365-294X.2005.02553.x>
- Fick S.E., Hijmans R.J. (2017) Worldclim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*. Available from <http://worldclim.org/version2> [accessed 6 Dec. 2017].
- Forest Research (2015) Seed storage and pretreatment for Robinia pseudoacacia [online]. Available from <https://www.forestry.gov.uk/fr/infnd-7fachj> [accessed 10 Feb. 2015].
- Genty B., Briantais J.-M., Baker N.R. (1989) The relationship between the quantum yield of photosynthetic electron transport and quenching of chlorophyll fluorescence. *Biochimica et Biophysica Acta* 990: 87–92. [https://doi.org/10.1016/S0304-4165\(89\)80016-9](https://doi.org/10.1016/S0304-4165(89)80016-9)
- Gilbert K.J., Whitlock M.C. (2015)  $Q_{ST}$ – $F_{ST}$  comparisons with unbalanced half-sib designs. *Molecular Ecology Resources* 15: 262–267. <https://doi.org/10.1111/1755-0998.12303>
- Gioria M., Pyšek P. (2017) Early bird catches the worm: germination as a critical step in plant invasion. *Biological Invasions* 19: 1055–1080. <https://doi.org/10.1007/s10530-016-1349-1>
- Goudet J., Büchi L. (2006) The effects of dominance, regular inbreeding and sampling design on  $Q_{ST}$ , an estimator of population differentiation for quantitative traits. *Genetics* 172: 1337–1347. <https://doi.org/10.1534/genetics.105.050583>
- Hendry A.P. (2002)  $Q_{ST} > = \neq < F_{ST}$ ? Trends in Ecology & Evolution 17: 502. [https://doi.org/10.1016/S0169-5347\(02\)02603-4](https://doi.org/10.1016/S0169-5347(02)02603-4)
- Heywood V.H. (1989) Patterns, extents and modes of invasions by terrestrial plants. In: Drake J.A. et al. (eds) *Biology invasions: a global perspective*: 31–60. Chichester, Wiley.
- Holsinger K.E. (1999) Analysis of genetic diversity in geographically structured populations: a Bayesian perspective. *Heredity* 130: 245–255. <https://doi.org/10.1111/j.1601-5223.1999.00245.x>
- Hyltdgaard B., Brix H. (2012) Intraspecific differences in phenotypic plasticity: invasive versus non-invasive populations of *Ceratophyllum demersum*. *Aquatic Botany* 97: 49–56. <https://doi.org/10.1016/j.aquabot.2011.11.004>
- Isik F. (2009) Analysis of half-sib progeny tests in forestry. SAS Paper 1–45. Available from <http://www4.ncsu.edu/~fisik/Analysis%20of%20Half-sib%20Progeny%20Test%20Data%20in%20Forestry.pdf> [accessed 9 Jan. 2018].
- Iverson L.R., Prasad A.M. (1998) Predicting abundance of 80 tree species following climate change in the eastern United States. *Ecological Monographs* 68: 465–485. [https://doi.org/10.1890/0012-9615\(1998\)068\[0465:PAOTSF\]2.0.CO;2](https://doi.org/10.1890/0012-9615(1998)068[0465:PAOTSF]2.0.CO;2)
- Johansen P.P., Andersen J.D., Børsting C., Morling N. (2013) Evaluation of the iPLEX® Sample ID Plus Panel designed for the Sequenom MassARRAY® system. A SNP typing assay developed for human identification and sample tracking based on the SNP for ID panel. *Forensic Science International: Genetics* 7: 482–487. <https://doi.org/10.1016/j.fsigen.2013.04.009>
- Jombart T. (2008) adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* 24: 1403–1405. <https://doi.org/10.1093/bioinformatics/btn129>
- Jombart T., Pontier D., Dufour A.-B. (2009) Genetic markers in the playground of multivariate analysis. *Heredity* 102: 330–341. <https://doi.org/10.1038/hdy.2008.130>
- Keller S.R., Taylor D.R. (2008) History, chance and adaptation during biological invasion: separating stochastic phenotypic evolution from response to selection. *Ecology Letters* 11: 852–866. <https://doi.org/10.1111/j.1461-0248.2008.01188.x>
- Kopelman N.M., Mayzel J., Jakobsson M., Rosenberg N.A., Mayrose I. (2015) CLUMPAK: a program for identifying clustering modes and packaging population structure inferences across K. *Molecular Ecology Resources* 15: 1179–1191. <https://doi.org/10.1111/1755-0998.12387>
- Koskinen M.T., Haugen T.O., Primmer C.R. (2002) Contemporary fisherian life-history evolution in small salmonid populations. *Nature* 419: 826–830. <https://doi.org/10.1038/nature01029>
- Kruschke J.K. (2014) *Doing Bayesian data analysis: a tutorial with R, JAGS, and Stan*. 2nd Ed. Boston, Academic Press.
- Lamarque L.J., Lortie C.J., Porté A.J., Delzon S. (2015) Genetic differentiation and phenotypic plasticity in life-history traits between native and introduced populations of invasive maple trees. *Biological Invasions* 17: 1109–1122. <https://doi.org/10.1007/s10530-014-0781-3>
- Lavergne S., Molofsky J. (2007) Increased genetic variation and evolutionary potential drive the success of an invasive grass. *Proceedings of the National Academy of Sciences of the United States of America* 104: 3883–3888. <https://doi.org/10.1073/pnas.0607324104>
- Lee C.E. (2002) Evolutionary genetics of invasive species. *Trends in Ecology & Evolution* 17: 386–391. [https://doi.org/10.1016/S0169-5347\(02\)02554-5](https://doi.org/10.1016/S0169-5347(02)02554-5)
- Leger E.A., Rice K.J. (2007) Assessing the speed and predictability of local adaptation in invasive California poppies (*Eschscholzia californica*). *Journal of Evolutionary Biology* 20: 1090–1103. <https://doi.org/10.1111/j.1420-9101.2006.01292.x>
- Leinonen T., McCairns R.J.S., O’Hara R.B., Merilä J. (2013)  $Q_{ST}$ – $F_{ST}$  comparisons: evolutionary and ecological insights from genomic heterogeneity. *Nature Reviews Genetics* 14: 179–190. <https://doi.org/10.1038/nrg3395>
- Maron J.L., Vilà M., Bommarco R., Elmendorf S., Beardsley P. (2004) Rapid evolution of an invasive plant. *Ecological Monographs* 74: 261–280. <https://doi.org/10.1890/03-4027>
- Maron J.L., Elmendorf S.C., Vilà M. (2007) Contrasting plant physiological adaptation to climate in the native and introduced range of *Hypericum perforatum*. *Evolution* 61: 1912–1924. <https://doi.org/10.1111/j.1558-5646.2007.00153.x>
- Matesanz S., Horgan-Kobelski T., Sultan S.E. (2012) Phenotypic plasticity and population differentiation in an ongoing species invasion. *PLoS One* 7: e44955. <https://doi.org/10.1371/journal.pone.0044955>
- Merilä J., Crnokrak P. (2001) Comparison of genetic differentiation at marker loci and quantitative traits. *Journal of Evolutionary Biology* 14: 892–903. <https://doi.org/10.1046/j.1420-9101.2001.00348.x>
- Molina-Montenegro M.A., Naya D.E. (2012) Latitudinal patterns in phenotypic plasticity and fitness-related traits: assessing the Climatic Variability Hypothesis (CVH) with an invasive plant species. *PLoS One* 7: e47620. <https://doi.org/10.1371/journal.pone.0047620>
- Monty A., Bizoux J.-P., Escarré J., Mahy G. (2013) Rapid plant invasion in distinct climates involves different sources of pheno-

- typic variation. *PLoS ONE* 8: e55627. <https://doi.org/10.1371/journal.pone.0055627>
- O'Hara R.B., Merilä J. (2005) Bias and precision in  $Q_{ST}$  estimates: problems and some solutions. *Genetics* 171: 1331–1339. <https://doi.org/10.1534/genetics.105.044545>
- Ovaskainen O., Karhunen M., Zheng C., Arias J.M.C., Merilä J. (2011) A new method to uncover signatures of divergent and stabilizing selection in quantitative traits. *Genetics* 189: 621–632. <https://doi.org/10.1534/genetics.111.129387>
- Pimentel D., McNair S., Janecka J., Wightman J., Simmonds C., O'Connell C., Wong E., Russel L., Zern J., Aquino T., Tsomondo T. (2001) Economic and environmental threats of alien plant, animal, and microbe invasions. *Agriculture, Ecosystems & Environment* 84: 1–20. [https://doi.org/10.1016/S0167-8809\(00\)00178-X](https://doi.org/10.1016/S0167-8809(00)00178-X)
- Plummer M. (2005) JAGS: a program for analysis of Bayesian graphical models using Gibbs sampling. In: Hornik K., Leisch F., Zeileis A. (eds) *Proceedings of the 3rd International Workshop on Distributed Statistical Computing (DSC 2003)*, March 20–22, 2003, Technische Universität Wien, Vienna, Austria. Available from <https://www.r-project.org/conferences/DSC-2003/Proceedings/Plummer.pdf> [accessed 9 Jan. 2018].
- Postma F.M., Ågren J. (2016) Early life stages contribute strongly to local adaptation in *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences of the United States of America* 113: 7590–7595. <https://doi.org/10.1073/pnas.1606303113>
- Prentice I.C., Cramer W., Harrison S.P., Leemans R., Monserud R.A., Solomon A.M. (1992) A global biome model based on plant physiology and dominance, soil properties and climate. *Journal of Biogeography* 19: 117–134. <https://doi.org/10.2307/2845499>
- Prentis P.J., Wilson J.R.U., Dormontt E.E., Richardson D.M., Lowe A.J. (2008) Adaptive evolution in invasive species. *Trends in Plant Science* 13: 288–294. <https://doi.org/10.1016/j.tplants.2008.03.004>
- Pritchard J.K., Stephens M., Donnelly P.J. (2000) Inference of population structure using multilocus genotype data. *Genetics* 155: 945–949.
- R Development Core Team (2015) R: a language and environment for statistical computing. Available from <https://www.r-project.org/> [accessed 2 Jan. 2018].
- Raymond M., Rousset F. (1995) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity* 86: 248–249. <https://doi.org/10.1093/oxfordjournals.jhered.a111573>
- Richards C.L., Bossdorf O., Muth N.Z., Gurevitch J., Pigliucci M. (2006) Jack of all trades, master of some? On the role of phenotypic plasticity in plant invasions. *Ecology Letters* 9: 981–993. <https://doi.org/10.1111/j.1461-0248.2006.00950.x>
- Rousset F. (2008) GENEPOP'007: a complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources* 8: 103–106. <https://doi.org/10.1111/j.1471-8286.2007.01931.x>
- Royal Botanic Gardens Kew (2015) Seed Information Database. Available from <http://data.kew.org/sid/> [accessed 10 Feb. 2015].
- Santiso X., López L., Gilbert K.J., Barreiro R., Whitlock M.C., Retuerto R. (2015) Patterns of genetic variation within and among populations in *Arbutus unedo* and its relation with selection and evolvability. *Perspectives in Plant Ecology, Evolution and Systematics* 17: 185–192. <https://doi.org/10.1016/j.ppees.2015.02.006>
- Spitze K. (1993) Population structure in *Daphnia obtusa*: quantitative genetic and allozymic variation. *Genetics* 135: 367–374.
- Su Y., Yajima M. (2012) R2jags: a package for running jags from R. Available from <https://cran.r-project.org/web/packages/R2jags/index.html> [accessed 2 Jan. 2018].
- Valladares F., Sanchez-Gomez D., Zavala M.A. (2006) Quantitative estimation of phenotypic plasticity: bridging the gap between the evolutionary concept and its ecological applications. *Journal of Ecology* 94: 1103–1116. <https://doi.org/10.1111/j.1365-2745.2006.01176.x>
- Verdu C.F., Guichoux E., Quevauvillers S., De Thier O., Laizet Y., Delcamp A., Gévaudan F., Monty A., Porté A.J., Lejeune P., Lassois L., Mariette S. (2016) Dealing with paralogy in RAD-seq data: in silico detection and single nucleotide polymorphism validation in *Robinia pseudoacacia* L. *Ecology and Evolution* 6: 7323–7333. <https://doi.org/10.1002/ece3.2466>
- Vítková M., Müllerová J., Sádlo J., Pergl J., Pyšek P. (2017) Black locust (*Robinia pseudoacacia*) beloved and despised: a story of an invasive tree in Central Europe. *Forest Ecology and Management* 384: 287–302. <https://doi.org/10.1016/j.foreco.2016.10.057>
- Vitousek P.M., D'Antonio C.M., Loope L.L., Westbrooks R. (1996) Biological invasions as global environmental changes. *American Scientist* 84: 468–478.
- Vitousek P.M., D'Antonio C.M., Loope L.L., Rejmánek M., Westbrooks R. (1997) Introduced species: a significant component of human-caused global change. *New Zealand Journal of Ecology* 21: 1–16.
- Wainwright C.E., Cleland E.E. (2013) Exotic species display greater germination plasticity and higher germination rates than native species across multiple cues. *Biological Invasions* 15: 2253–2264. <https://doi.org/10.1007/s10530-013-0449-4>
- Walck J.L., Hidayati S.N., Dixon K.W., Thompson K., Poschold P. (2011) Climate change and plant regeneration from seed. *Global Change Biology* 17: 2145–2161. <https://doi.org/10.1111/j.1365-2486.2010.02368.x>
- Ward J.H. (1963) Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*. 58: 236–244. <https://doi.org/10.1080/01621459.1963.10500845>
- Whitlock M.C. (1999) Neutral additive genetic variance in a metapopulation. *Genetical Research* 74: 215–221. <https://doi.org/10.1017/S0016672399004127>
- Whitlock M.C. (2008) Evolutionary inference from  $Q_{ST}$ . *Molecular Ecology* 17: 1885–1896. <https://doi.org/10.1111/j.1365-294X.2008.03712.x>
- Whitlock M.C., Guillaume F. (2009) Testing for spatially divergent selection: Comparing  $Q_{ST}$  to  $F_{ST}$ . *Genetics* 183: 1055–1063. <https://doi.org/10.1534/genetics.108.099812>
- Woodward F.I. (1987) *Climate and plant distribution*. Cambridge, Cambridge University Press.
- Wright S. (1949) The genetical structure of populations. *Annals of Eugenics* 15: 323–354. <https://doi.org/10.1111/j.1469-1809.1949.tb02451.x>

Manuscript received 28 Sep. 2017; accepted in revised version 3 Jan. 2018.

Communicating Editor: Pierre Meerts.

Electronic appendix 1 to:

Xavier P. Bouteiller, Frédéric Barraquand, Pauline Garnier-Géré, Noémie Harmand, Yec'han Laizet, Alexandre Raimbault, Raphaël Segura, Ludivine Lassois, Arnaud Monty, Cindy Verdu, Stéphanie Mariette & Annabel J. Porté (2018)

No evidence for genetic differentiation in juvenile traits between Belgian and French populations of the invasive tree *Robinia pseudoacacia*

Plant Ecology and Evolution 151(1)

**Appendix 1 – Populations of *Robinia pseudoacacia* sampled in France (Aquitaine) and Belgium (Wallonia).**

GPS coordinates are provided in the WGS84 geographic projection. Mean, Minimal and Maximal temperatures of May were extracted from WorldClim version 2 (<http://worldclim.org/version2>). WorldClim version 2 provides average monthly climate data for minimum, mean, and maximum temperature over the 1970–2000 period.

Range	Population name	Population identifier	Latitude (°)	Longitude(°)	Mean T° of May (°C)	Minimal T° of May (°C)	Maximal T° of May (°C)
France	Garingail	Fr.Pop1	44.65486	-0.45972	15.00	10.50	19.50
France	Le Brousteyrot	Fr.Pop2	44.68538	-0.57156	14.90	10.50	19.30
France	La Gouaneyre	Fr.Pop3	44.37143	-0.2729	14.90	9.90	20.00
France	Miramont de Guyenne	Fr.Pop4	44.60188	0.35302	15.10	9.70	20.60
France	Marmande	Fr.Pop5	44.49577	0.1674	15.20	9.50	21.00
France	La Magdeleine	Fr.Pop6	44.5795	-0.1744	15.1	9.9	20.3
France	Granges Sur Lot	Fr.Pop7	44.37632	0.4701	15.30	9.70	21.00
France	Lafitte Sur Lot	Fr.Pop8	44.35574	0.39729	15.30	9.50	21.10
France	Gabarnac	Fr.Pop9	44.6201	-0.28791	15.10	9.80	20.50
France	La Flotte	Fr.Pop10	44.38445	-0.30404	14.90	9.90	19.80
Belgium	Carrière de l'Amspin	Bel.Pop1	50.5433	5.28715	13.30	8.50	18.00
Belgium	Wavre	Bel.Pop2	50.7052	4.60953	13.30	8.60	18.00
Belgium	Charleroi	Bel.Pop3	50.4185	4.49233	13.20	8.50	17.90
Belgium	Herstol	Bel.Pop4	50.6651	5.60891	13.50	9.00	18.00
Belgium	Glaverbel	Bel.Pop5	50.492	4.17391	13.00	8.20	17.70
Belgium	Nethen	Bel.Pop6	50.7851	4.68551	12.90	8.30	17.50
Belgium	Citadelle	Bel.Pop7	50.653	5.5845	13.80	9.30	18.20
Belgium	Blaton	Bel.Pop8	50.492	3.6691	12.90	8.20	17.70
Belgium	Falaise d'Ampsin	Bel.Pop9	50.5391	5.26143	13.30	8.50	18.10
Belgium	Carrière-de Asty Moulin	Bel.Pop10	50.4772	4.85306	13.10	8.40	17.80



Electronic appendix 2 to:

Xavier P. Bouteiller, Frédéric Barraquand, Pauline Garnier-Géré, Noémie Harmand, Yec'han Laizet, Alexandre Raimbault, Raphaël Segura, Ludivine Lassois, Arnaud Monty, Cindy Verdu, Stéphanie Mariette & Annabel J. Porté (2018)  
No evidence for genetic differentiation in juvenile traits between Belgian and French populations of the invasive tree *Robinia pseudoacacia*

Plant Ecology and Evolution 151(1)

## **Appendix 2 – Comparison of the results of the Bayesian and frequentist approaches to compute mixed model and to estimate $Q_{ST}$ .**

### **A – Results for the ANCOVA mixed model testing for the effect of Range of origin (France vs. Belgium) and Temperature (22°C vs. 18°C) as fixed effects with average families' seed weight as covariates and Population nested within Range, Tree nested within population and interaction between Temperature and Population as random effects.**

df: degree of freedom, F: F-value, p: p-value, SW: Seed Weight, R: Range, T: Temperature, GR: Growth Rate. Significant values are in bold, not consistent results among frequentist and Bayesian methods are highlighted in grey.

Consistent results among analysis were always observed, except for (i) Range effect of Height 1 GR, however p-value near to the boundary ( $p = 0.03$ ) give low support to the significance of the effect (ii) Temperature effect for the Leaf Weight GR which is significant in Bayesian contrary to the frequentist result, however confidence intervals of the Bayesian estimated parameter are very close to 0 (Appendix 4 I a) and (iii) Population effect of the Leaf Weight GR which is significant with the frequentist method but not with Bayesian method, however once again p-value is close to the boundary (0.03). Thus, the few times where Bayesian and frequentist methods gave different results, the signal is weak and could be attributed to the limit of accuracy of each method.

### **B – $Q_{ST}$ mean values and 95% confidence limits (CL).**

$Q_{ST}$  were calculated using different approaches: a Bayesian method implemented with JAGS (Plummer 2005) from an adaptation of the code provided by O'Hara & Merilä (2005) using a uniform prior and two frequentist methods.

The first one used a mixed model of analysis of variance in SAS and CL were calculated following the delta method (Isik 2009) and the second one used the QstFstComp package (Gilbert & Whitlock 2014).

Consistent results among the three methods were observed in all cases. However, note that due to weak variance at the population level, the estimated  $Q_{ST}$  could be negative in the frequentist methods whereas using the Bayesian method estimated  $Q_{ST}$  and their confidence intervals are strictly positive due to prior definition which is more coherent because  $Q_{ST}$  is a ratio of variance.

A

Traits		<i>Fixed effects</i>			<i>Random effects</i>		
		Range	Temperature	Seed weight	Population(range)	Tree(population)	Temp X Pop
Stage 1 (d)	df	1	<b>1</b>	1	18	<b>1515</b>	1515
	F or Z	1.01	<b>337.19</b>	0.28	-0.13	<b>4.24</b>	0.93
	p	0.3280	<b>&lt; 0.0001</b>	0.5950	0.8961	<b>&lt; 0.0001</b>	0.3528
Stage 2 (d)	df	1	<b>1</b>	1	18	<b>1343</b>	1343
	F or Z	2.30	<b>323.85</b>	0.48	-1.30	<b>5.49</b>	1.45
	p	0.1466	<b>&lt; 0.0001</b>	0.4877	0.1926	<b>&lt; 0.0001</b>	0.1466
Stage 3 (d)	df	1	<b>1</b>	1	18	<b>1296</b>	1296
	F or Z	1.12	<b>615.79</b>	0.04	-0.61	<b>4.52</b>	-0.05
	p	0.3048	<b>&lt; 0.0001</b>	0.8386	0.5397	<b>&lt; 0.0001</b>	0.9602
Stage 4 (d)	df	1	<b>1</b>	1	18	<b>1234</b>	1234
	F or Z	3.59	<b>1475.64</b>	3.58	-1.90	<b>5.91</b>	1.52
	p	0.0745	<b>&lt; 0.0001</b>	0.0589	0.0574	<b>&lt; 0.0001</b>	0.1274
Stage 5 (d)	df	1	<b>1</b>	1	18	<b>747</b>	747
	F or Z	0.13	<b>4756.19</b>	1.42	0.09	<b>4.65</b>	0.14
	p	0.7214	<b>&lt; 0.0001</b>	0.2339	0.9303	<b>&lt; 0.0001</b>	0.8903
Height GR 1 (cm.d <sup>-1</sup> )	df	<b>1</b>	<b>1</b>	<b>1</b>	18	<b>853</b>	853
	F or Z	<b>5.45</b>	<b>421.28</b>	<b>16.21</b>	-0.36	<b>5.56</b>	0.02
	p	<b>0.0313</b>	<b>&lt; 0.0001</b>	<b>&lt; 0.0001</b>	0.7220	<b>&lt; 0.0001</b>	0.9841
Height GR 2 (cm.d <sup>-1</sup> )	df	1	<b>1</b>	<b>1</b>	18	<b>787</b>	787
	F or Z	0.10	<b>53.27</b>	<b>11.59</b>	-0.62	<b>5.96</b>	1.08
	p	0.7520	<b>&lt; 0.0001</b>	<b>&lt; 0.0001</b>	0.5376	<b>&lt; 0.0001</b>	0.2813
PSII Yield	df	1	<b>1</b>	1	18	<b>778</b>	778
	F or Z	2.31	<b>210.32</b>	0.05	-0.57	<b>3.77</b>	0.49
	p	0.1457	<b>&lt; 0.0001</b>	0.8177	0.5694	<b>0.0002</b>	0.6267
Leaf weight GR (g.d <sup>-1</sup> )	df	1	<b>1</b>	<b>1</b>	<b>18</b>	<b>787</b>	787
	F or Z	0.07	2.08	<b>23.40</b>	<b>-2.14</b>	<b>5.50</b>	1.52
	p	0.1654	0.7979	<b>&lt; 0.0001</b>	<b>0.0327</b>	<b>&lt; 0.0001</b>	0.1296
Stem weight GR (g.d <sup>-1</sup> )	df	1	1	<b>1</b>	18	<b>785</b>	785
	F or Z	0.05	0.73	<b>32.71</b>	-1.12	<b>3.25</b>	0.13
	p	0.8331	0.4037	<b>&lt; 0.0001</b>	0.2646	<b>0.0011</b>	0.8986
Total weight GR (g.d <sup>-1</sup> )	df	1	1	<b>1</b>	18	<b>785</b>	785
	F or Z	0.02	2	<b>26.30</b>	-1.84	<b>5.35</b>	1.22
	p	0.8809	0.1734	<b>&lt; 0.0001</b>	0.0662	<b>&lt; 0.0001</b>	0.2241
Diameter GR (mm.d <sup>-1</sup> )	df	1	<b>1</b>	<b>1</b>	18	<b>786</b>	786
	F or Z	1.72	<b>93.14</b>	<b>45.23</b>	-0.26	<b>4.16</b>	-0.55
	p	0.2068	<b>&lt; 0.0001</b>	<b>&lt; 0.0001</b>	0.7946	<b>&lt; 0.0001</b>	0.5840

B

Temp.	Traits	Q <sub>ST</sub> Bayesian			Q <sub>ST</sub> Delta method			Q <sub>ST</sub> QstFstComp		
		Mean	2.5% CI	97.5% CI	Mean	2.5% CI	97.5% CI	Mean	2.5% CI	97.5% CI
22°C	Stage 1	0.006	0.000	0.032	-0.004	-0.025	0.018	0.001	-0.014	0.026
	Stage 2	0.004	0.000	0.023	-0.009	-0.034	0.016	0.002	-0.013	0.029
	Stage 3	0.004	0.000	0.020	-0.015	-0.048	0.018	-0.007	-0.019	0.013
	Stage 4	0.005	0.000	0.027	-0.008	-0.032	0.017	0.001	-0.015	0.026
	Stage 5	0.005	0.000	0.025	-0.008	-0.033	0.016	-0.007	-0.018	0.012
	Height GR	0.004	0.000	0.020	-0.005	-0.023	0.012	0.000	-0.014	0.010
	Height GR	0.009	0.000	0.041	0.001	-0.018	0.019	0.000	-0.015	0.024
	PSI Yield	0.002	0.000	0.011	-0.010	-0.025	0.006	-0.009	-0.015	0.001
	Leaf weight GR	0.007	0.000	0.036	-0.003	-0.025	0.020	0.002	-0.013	0.029
	Stem weight GR	0.061	0.000	0.642	-0.021	-0.130	0.089	0.003	-0.186	0.197
	Total weight GR	0.007	0.000	0.033	-0.005	-0.030	0.021	0.002	-0.013	0.027
Diameter GR	0.007	0.000	0.037	-0.006	-0.034	0.022	-0.001	-0.023	0.036	
18°C	Stage 1	0.012	0.000	0.050	0.006	-0.008	0.020	0.004	-0.013	0.033
	Stage 2	0.005	0.000	0.025	-0.001	-0.016	0.014	-0.002	-0.013	0.016
	Stage 3	0.004	0.000	0.019	-0.007	-0.027	0.013	-0.008	-0.017	0.006
	Stage 4	0.005	0.000	0.023	-0.002	-0.016	0.013	0.000	-0.011	0.017
	Stage 5	0.092	0.001	0.424	0.053	NA	NA	-0.051	-0.146	-0.020
	Height GR	0.008	0.000	0.037	0.003	-0.011	0.017	0.003	-0.012	0.027
	Height GR	0.009	0.000	0.035	0.006	-0.006	0.017	0.007	-0.007	0.028
	PSI Yield	0.005	0.000	0.025	-0.004	-0.022	0.015	-0.005	-0.016	0.013
	Leaf weight GR	0.009	0.000	0.034	0.006	-0.005	0.017	0.015	-0.003	0.041
	Stem weight GR	0.015	0.000	0.056	0.010	0.002	0.019	0.022	-0.003	0.065
	Total weight GR	0.009	0.000	0.037	0.006	-0.005	0.017	0.016	-0.003	0.046
Diameter GR	0.008	0.000	0.035	0.005	-0.008	0.017	0.013	-0.006	0.037	

Electronic appendix 3 to:

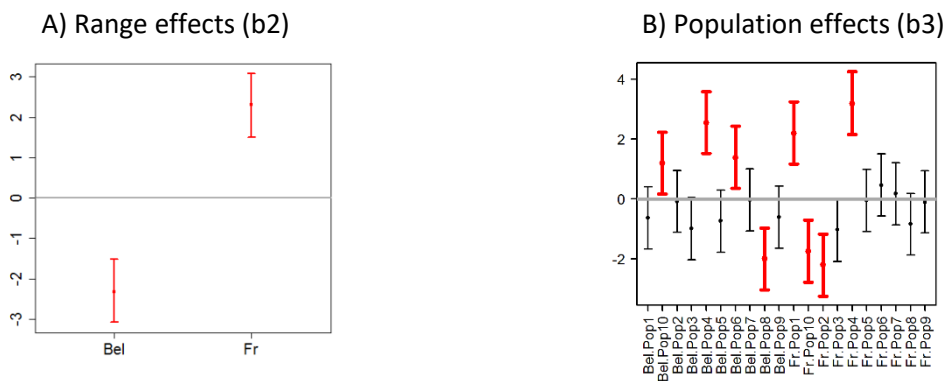
Xavier P. Bouteiller, Frédéric Barraquand, Pauline Garnier-Géré, Noémie Harmand, Yec'han Laizet, Alexandre Raimbault, Raphaël Segura, Ludivine Lassois, Arnaud Monty, Cindy Verdu, Stéphanie Mariette & Annabel J. Porté (2018)

No evidence for genetic differentiation in juvenile traits between Belgian and French populations of the invasive tree *Robinia pseudoacacia*

Plant Ecology and Evolution 151(1)

**Appendix 3** – Results of the seed weight analysis using the Bayesian model testing for A: range effects, B: population effects. Each dot represents the mean value of the estimate, and bars represent the Bayesian 95% credible interval. Effects significantly different from 0 are drawn in red. Both range and population effects are significant.

To test if there were significant differences in seed weight at the family level, data were analyzed using a similar Bayesian model than the main model (equation 1) including only the range and population effects.



Electronic appendix 4 to:

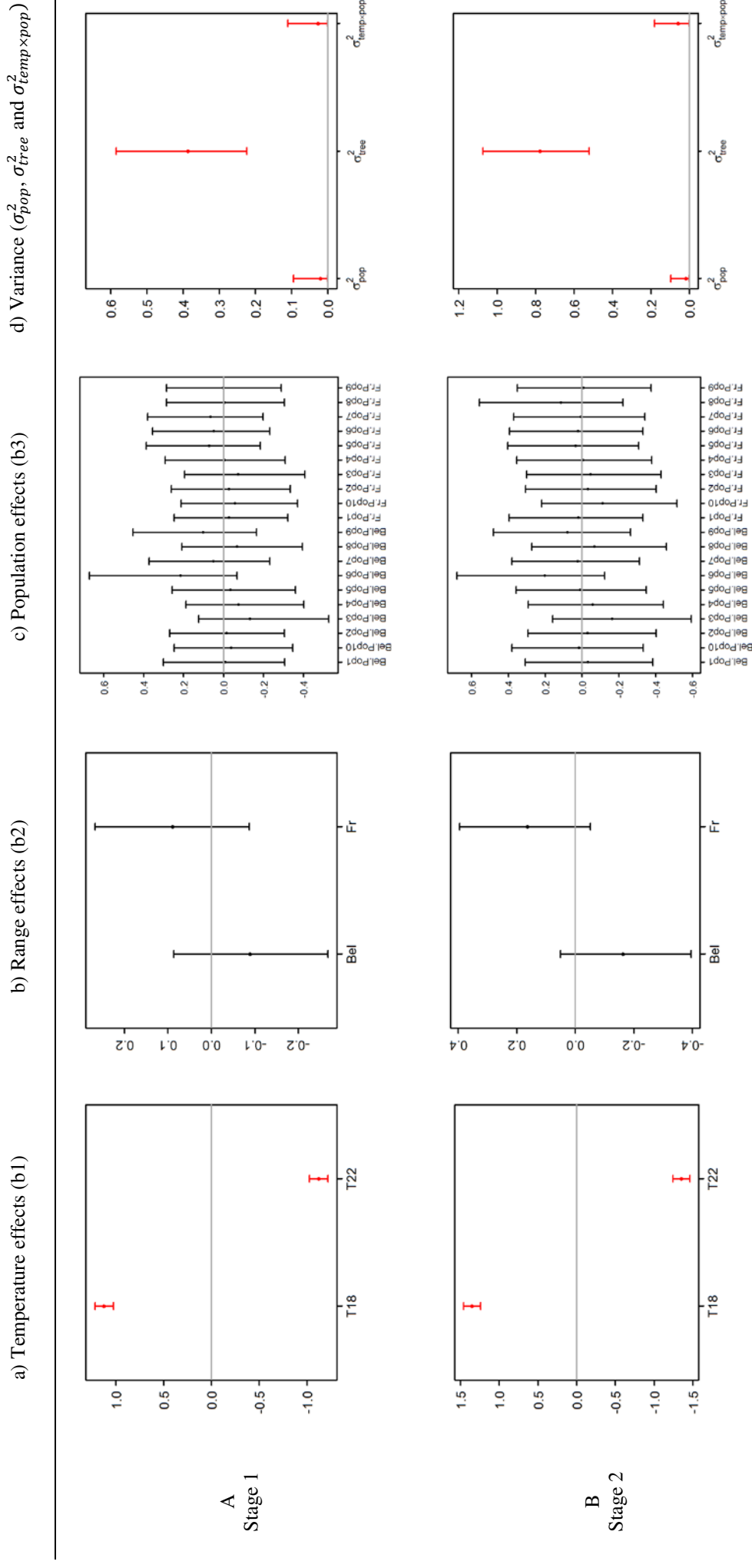
Xavier P. Bouteiller, Frédéric Barraquand, Pauline Garnier-Géré, Noémie Harmand, Yec'han Laizet, Alexandre Raimbault, Raphaël Segura, Ludivine Lassois, Arnaud Monty, Cindy Verdu, Stéphanie Mariette & Annabel J. Porté (2018)

No evidence for genetic differentiation in juvenile traits between Belgian and French populations of the invasive tree *Robinia pseudoacacia*

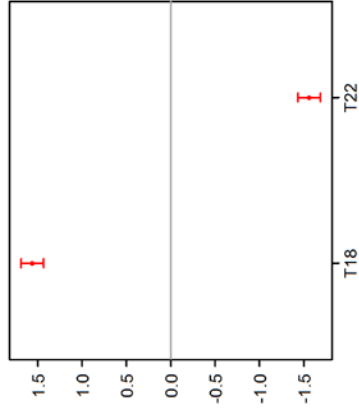
Plant Ecology and Evolution 151(1)

**Appendix 4 – Results of the Bayesian analyses for all phenotypic traits using the full between chamber model presented in the main document (equation 1).**

A–E, phenological stages 1 to 5; F–G, height Growth Rate (GR) 1 and 2; H, PSII yield; I–K, leaf, stem and total weight GR; L, diameter GR; with a: temperature effects (b1), b: range effect (b2), c: population effects (b3), d: estimated population, tree and temperature  $\times$  population interaction variance. Each dot represents the mean value of the estimated effect, and bars represent the 95% credible interval over the mean value. Effects significantly different from 0 are drawn in red.

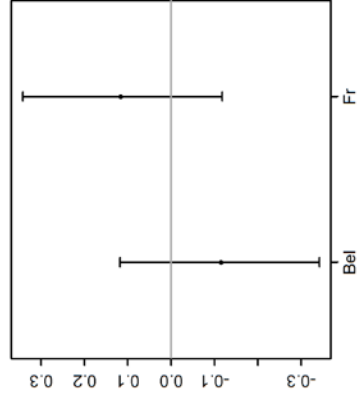


a) Temperature effects (b1)

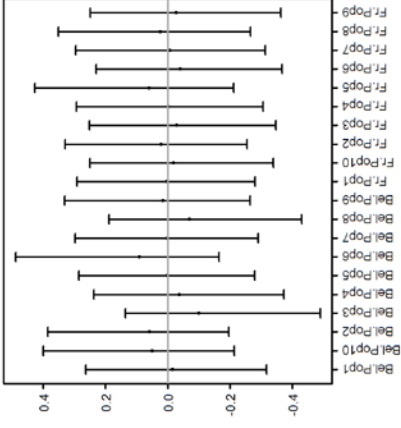


C  
Stage 3

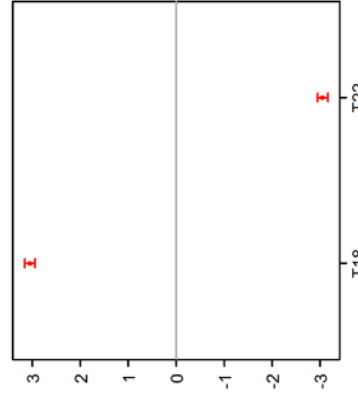
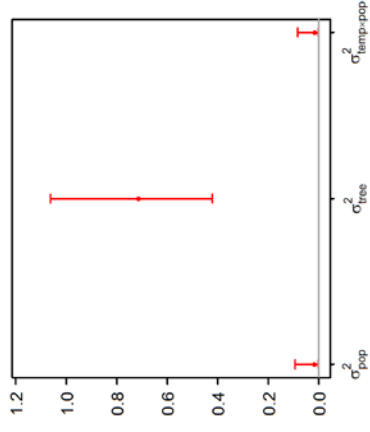
b) Range effects (b2)



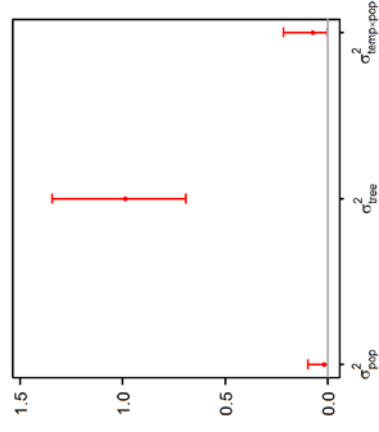
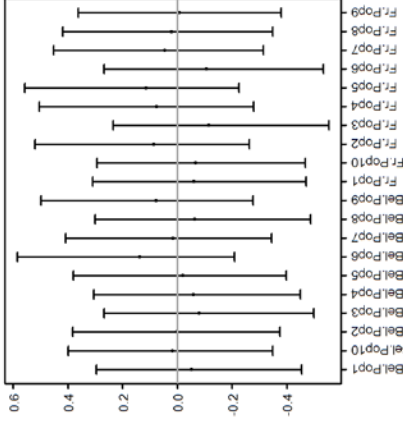
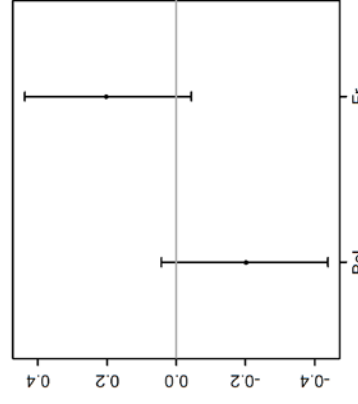
c) Population effects (b3)



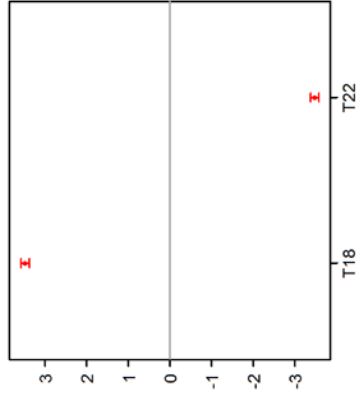
d) Variance ( $\sigma_{pop}^2$ ,  $\sigma_{tree}^2$  and  $\sigma_{Temp \times pop}^2$ )



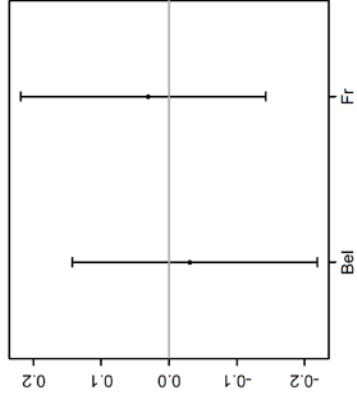
D  
Stage 4



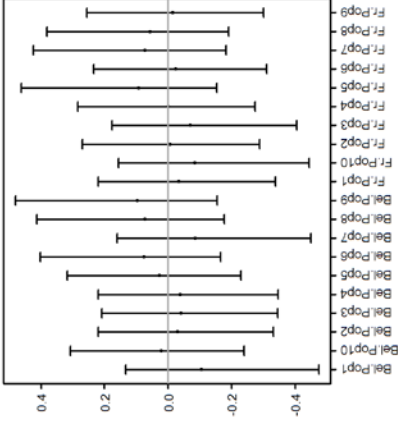
a) Temperature effects (b1)



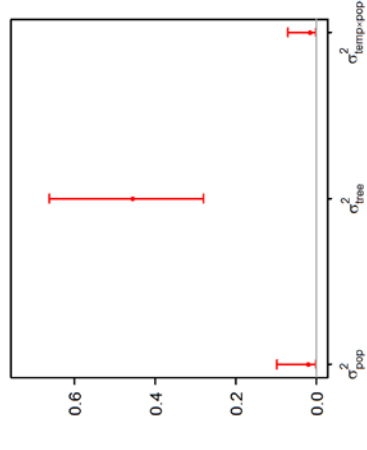
b) Range effects (b2)



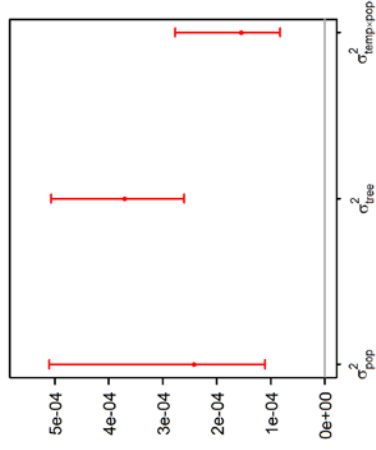
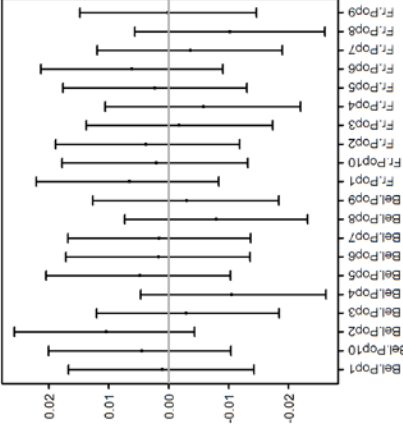
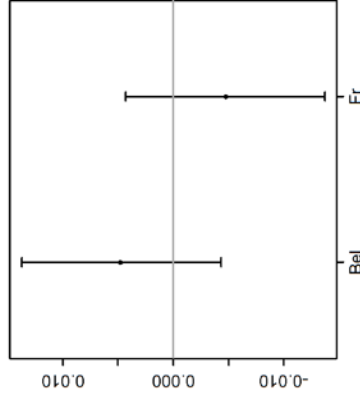
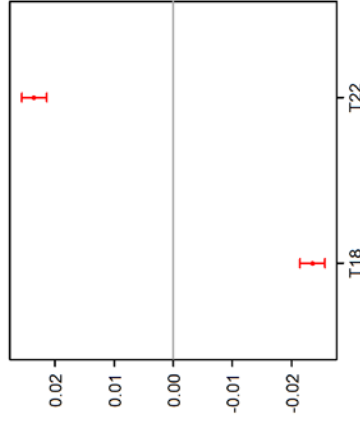
c) Population effects (b3)



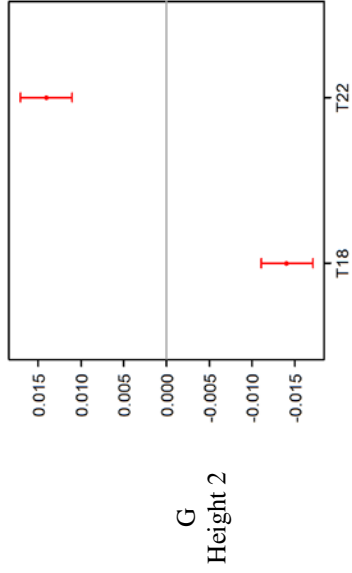
d) Variance ( $\sigma_{pop}^2$ ,  $\sigma_{tree}^2$  and  $\sigma_{temp \times pop}^2$ )



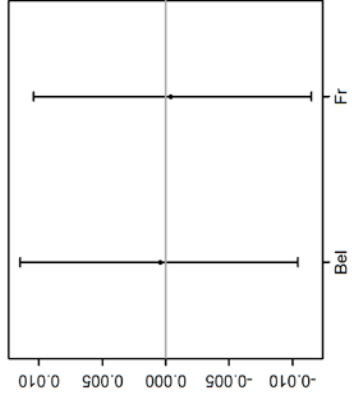
F  
Height 1



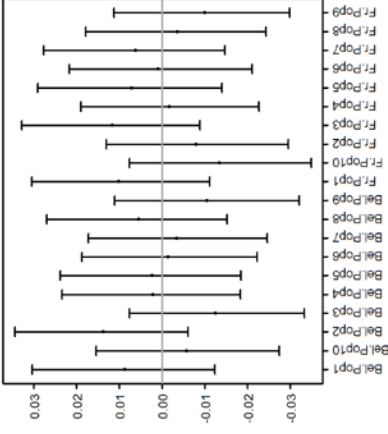
a) Temperature effects (b1)



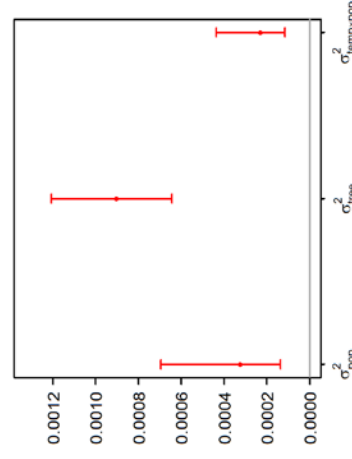
b) Range effects (b2)



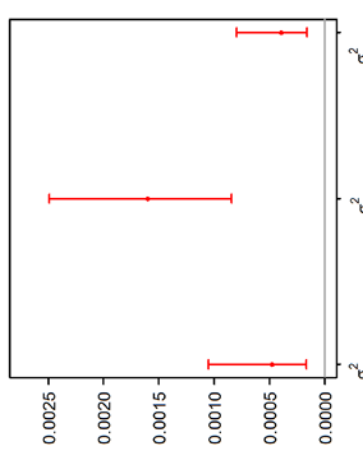
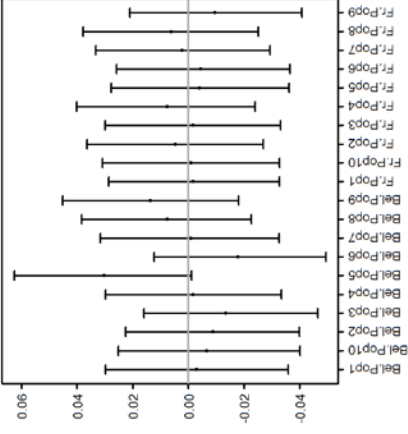
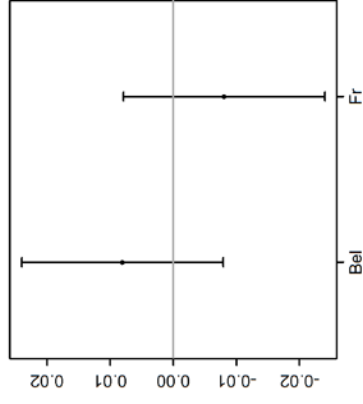
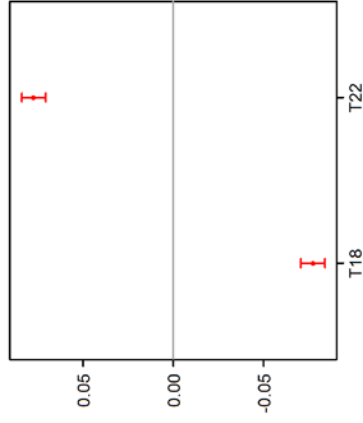
c) Population effects (b3)



d) Variance ( $\sigma_{pop}^2$ ,  $\sigma_{tree}^2$  and  $\sigma_{Temp \times pop}^2$ )

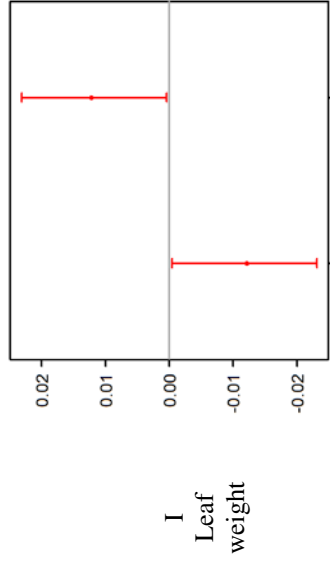


H  
PSII Yield

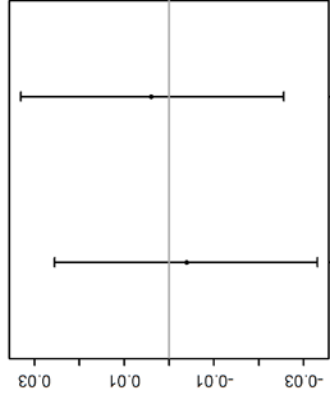




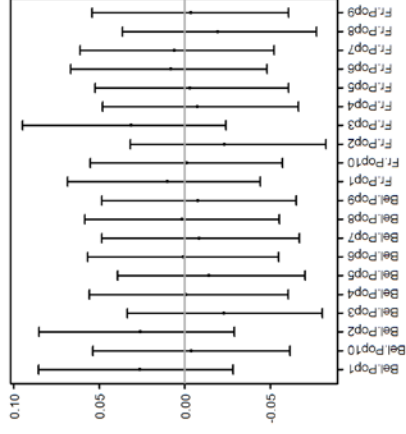
a) Temperature effects (b1)



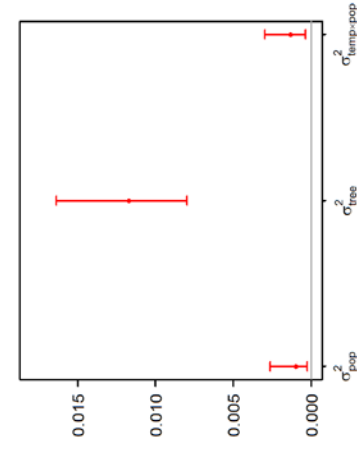
b) Range effects (b2)



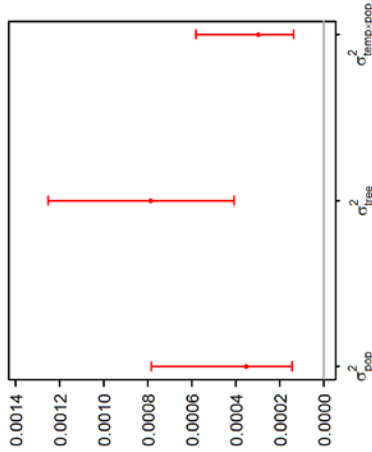
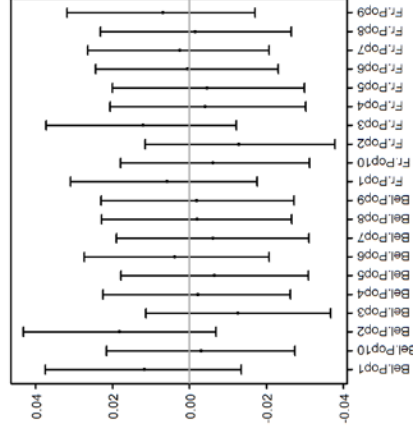
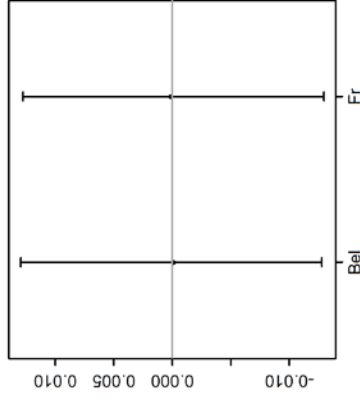
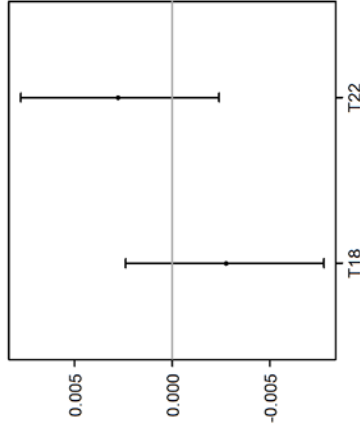
c) Population effects (b3)



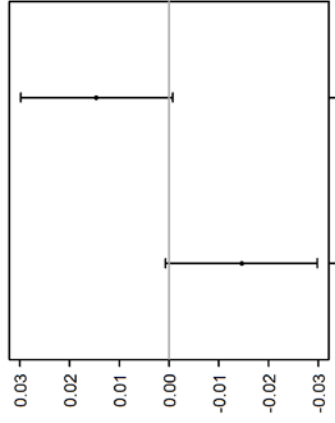
d) Variance ( $\sigma_{pop}^2$ ,  $\sigma_{tree}^2$  and  $\sigma_{Temp \times pop}^2$ )



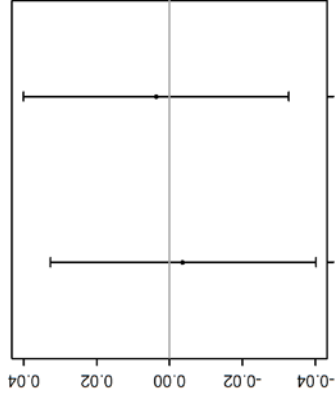
J Stem weight



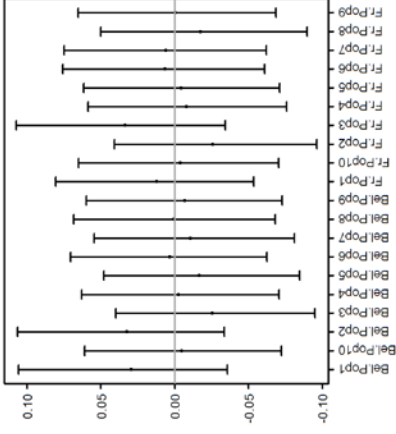
a) Temperature effects (b1)



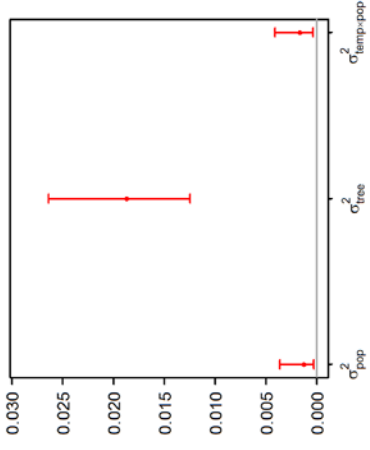
b) Range effects (b2)



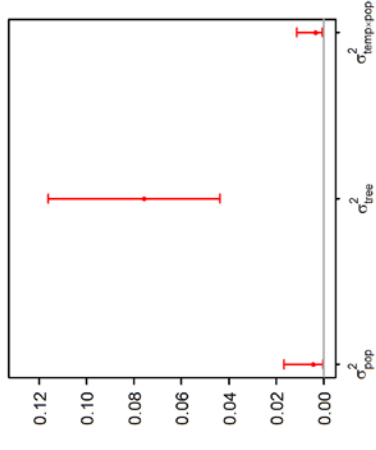
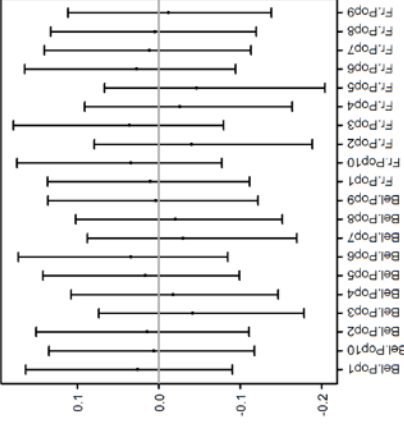
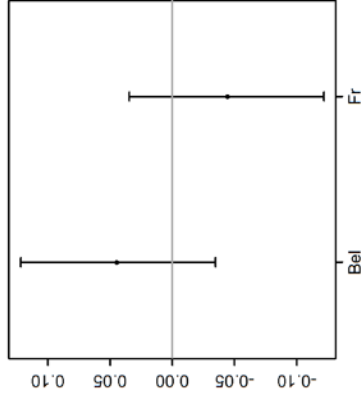
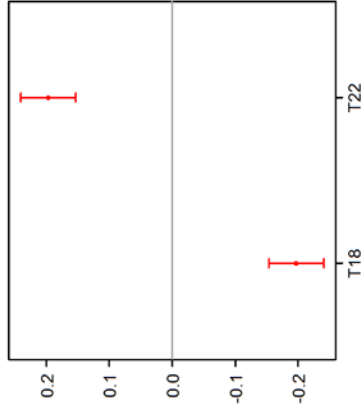
c) Population effects (b3)



d) Variance ( $\sigma_{pop}^2$ ,  $\sigma_{tree}^2$  and  $\sigma_{temp \times pop}^2$ )



L  
Diameter



## **Chapitre 4**

**The seeds of invasion: Enhanced germination among invasive European populations of black locust compared to native American populations.**

## TITLE

The seeds of invasion: Enhanced germination among invasive European populations of black locust compared to native American populations.

## STATUS

In prep (to be submitted to AoB PLANTS)

## AUTHORS

Bouteiller et al.

## ABSTRACT

Documenting population rapid evolution is a recent concern in invasion ecology, however these studies indicated that evolution occurred in the new range may have facilitated invasion. Additionally, phenotypic plasticity may facilitate invasiveness as it allows organisms to express different phenotypes depending on the environmental conditions and increase their fitness. However, to evaluate invasive tree adaptive capacities, quantifying the reproductive success is difficult; but germination and seedling recruitment are crucial life history stages for plants that directly impact their fitness and lead to local adaptation. This seems to be all the more true for newly introduced species suggesting the evolution of germination characteristics and juvenile traits could play a major role during the invasion process. Our objective was thus to evaluate the evolutionary mechanisms behind the invasiveness of *Robinia pseudoacacia* L. since its introduction from North America to Europe by analysing the genetic and environmental variability of germination and juvenile traits. We conducted a quantitative genetics experiment to compare both native and invasive black locust populations (3000 individuals from 20 populations) grown under 3 different temperature conditions (18°C, 22°C and 31°C). An enhanced germination rate was evidenced among European populations compared to that of the native American populations, whatever the environmental condition (88% vs 60%, in average). Other juvenile traits were highly plastic to temperature but responded similarly to the temperature increase: in general, the warmer the temperature, the higher or the faster the growth rate or the trait value whatever the population. Although no sign of departure from the neutral evolution was evidenced in our study based on  $Q_{ST}$   $F_{ST}$  comparisons, the genetic differences in germination rate between native and invasive populations could only partly reflect an ancestry effect as the population sources of the European black locust in the USA did not present such high germination rates. Thus, a possible scenario may be that Man would have selected the best seeds in the new range when propagating the species through Europe, thus favoring the evolution of germination rate that a key process that would likely have contributed to the success of black locust dynamics in Europe.

## KEYWORDS

Robinia pseudoacacia, black locust, biological invasion, quantitative genetics, germination, seedlings, juvenile traits, phenotypic plasticity,  $Q_{ST}$   $F_{ST}$  comparison, local adaptation, invasive species, native range, invasive range

## Acknowledgments

We would like to thank all people who helped us for sampling black locust populations: Dr Gideon Bradburd from the Michigan State University for his support in America (XPB field work), Dr Santiago Gonzalez-Martinez (BIOGECO - France) who helped us to organize sampling through Spain and to all scientists who sampled for us or with us wild populations of black locust all through Europe, i.e. Dennis Born (Technische Universität- Germany), Marianne Correard (INRA - Avignon -France), Christopher Galanos (Hellenic Open University - Greece), Dr Pablo Goikoetxea (Neiker - Spain), Dr Emilio Laguna (Generalitat Valenciana, Wildlife Service-CIEF, Valencia, Spain), Dr Roman Longaner and Antonin Martinik (Mendel University – Czech Republic), Dr Juan Pino (Universitat Autònoma de Barcelona - Spain), Alexandra Quenu and Sébastien Irola (SMABVC -France), Dr Volker Schneck (Thünen Institute of Forest Genetics -Germany), Dr Tomasz Wojda & Dr Martin Klisz (Forest Research Institute –Poland) and Dr Rafael Zas (Misión Biológica de Galicia - Spain) and Dr Kevin Walker and the Botanical Society of Britain and Ireland who kindly helped us to locate English populations of black locust. Special thanks to Régis Burtlett and Raphaël Segura for the help in sampling UK populations. We also thank Christophe Orazio (EFI – France), Nicolas Cheval and Dr Patrick Patuszka from (UE INRA – France) for helping us sampling European provenances within REINFFORCE project

This study was financially supported by the Forest and Nature Management Research Unit of Gembloux Agro-BioTech, The BioGeCo INRA-Univ Bordeaux Research Unit, the Special Research Fund of the University of Liège, the *Société Française d'Ecologie*, the French *Agence de l'Eau*, the ANR-10-EQPX-16 Xyloforest, the EU COST Action FP1403 "Non-Native Tree Species for European Forests - Experiences, Risks and Opportunities" and the Transnational Access to Research Infrastructures activity in the 7<sup>th</sup> Framework Program of the EC under the Trees4Future project (no.284181). Special thanks to Philippe Lejeune, for supporting the collaborative project.

## 1. Introduction

Several studies recently documented that recent evolution that had occurred in the new range may have facilitated invasion (Lee 2002; Blair and Wolfe 2004; Maron et al. 2004; Keller and Taylor 2008; Barrett et al. 2008; Colautti and Barrett 2013). For example, *Lythrum salicaria* populations were shown to evidence a latitudinal cline in flowering period in the invasive range (Barrett et al. 2008). It was further demonstrated using three common gardens along a latitudinal gradient, that differentiation among invasive populations for the fitness related trait of fruit production had a genetic basis ; this was attributed to a rapid evolution of the species in its new range in response to selective pressure leading to local adaptation along the latitudinal gradient (Barrett et al. 2008; Colautti and Barrett 2013). A recent meta-analysis evidenced that local adaptation in the new range is not rare for invasive plant species (Oduor et al. 2016), thus local adaptation could lead to divergence on phenotypic traits among introduced populations and could interact with phenotypic plasticity that is another mechanism that had been proposed to explain the invasiveness.

Phenotypic plasticity allows organisms to express different phenotypes depending in the environmental conditions (Pigliucci 2006; Ghalambor et al. 2007). In the context of invasion ecology, phenotypic plasticity could be defined here at the population level (Valladares et al. 2006), a level of genetic diversity being particularly relevant for understanding the ecological role of plasticity in natural populations (Richards et al. 2006; Monty et al. 2013). It was pointed out that plasticity could promote species invasiveness (Agrawal 2001; Richards et al. 2006; Davidson et al. 2011). Indeed it could facilitate establishment of newly introduced organisms allowing individuals to express beneficial phenotypes under a broad range of environmental conditions and thus maintaining a higher fitness (Richards et al. 2006; Davidson et al. 2011). Consequently, a conceptual framework has emphasized the different strategies of invasive plants related to plasticity that could allow to maintain a high fitness whatever the environment ('Jack-of-all-trades'), to reach a high fitness under favorable environments only ('Master-of some') or to join cases resulting in a high fitness enhanced under favorable environments ('Jack-and-Master', (Richards et al. 2006). However among invasive species, contrasting results on the amount of phenotypic plasticity have been found. For example, one study found that the invasive tree *Melaleuca quinquenervia* seems to be more plastic to variation in pH in the invasive range (Kaufman and Smouse 2001) and a meta-analysis had shown that invasive species are often more plastic than native especially when increasing resources availability, however this increase was only weakly associated with a gain in fitness

(Davidson et al. 2011). However another meta-analysis and a multiple species comparison demonstrated no difference in plasticity among invasive and native species (Godoy et al. 2011; Palacio-López and Gianoli 2011). Thus the role of phenotypic plasticity in invasion remains unclear and more empirical studies are needed. Within species, the comparison of native and invasive could allow to detect if plasticity evolved between ranges (Richards et al. 2006; Oduor et al. 2016).

Germination and seedling recruitment (Eriksson and Ehrlén 2008) are crucial life history stages for plants as they directly impact their fitness and lead to local adaptation (Donohue et al. 2010). Indeed it was demonstrated that germination timing regarding the season of seeds dispersal was under selection for the model plant *Arabidopsis thaliana* (Donohue et al. 2005). Furthermore, using the same species, a reciprocal transplant experiment evidenced that germination and juvenile stages experienced selection, which resulted in advantaging local genotypes and consequently promoting local adaptation (Postma and Ågren 2016). Thus, juvenile traits of plants during the germination process appear to be under strong selective pressures, indicating that they are key traits, critical to understanding local adaptation of populations (Postma and Ågren 2016).

This seems to be all the more true, for newly introduced species : germination rate, timing and speed have been recently documented comparing populations of both the native and invasive ranges, indicating the major role of germination and juvenile period during the invasion process (Blair and Wolfe 2004; Erfmeier and Bruehlheide 2005; Beckmann et al. 2011; Hirsch et al. 2012, 2017; Leiblein-Wild et al. 2014; Gioria and Pyšek 2016). Based on a literature review, Gioria and Pyšek (2017) underlined that germination is a crucial period for the success of invasive plants. An earlier germination in the season, a larger seed production, high germination rates and/or a tolerance to broad environmental conditions at the juvenile stage could contribute to plant invasiveness (Gioria and Pyšek 2016). Still few studies surveyed either life history traits (i.e. germination and seedling phenology) or functional traits (i.e. height, biomass production, effective quantum yield of photosystem II) of juveniles, comparing populations from both ranges. Furthermore, contrasting results are observed, depending on the species. On the one hand, for instance, invasive populations of the shrub *Rhododendron ponticum* exhibited faster seedlings phenology than native populations in response to temperature (Erfmeier and Bruehlheide 2005). Similarly, invasive populations of *Ulmus pumilla* always outperformed native ones for germination speed (Hirsch et al. 2012) comparing seeds development under various common gardens corresponding to different temperature conditions. However no difference in the plasticity of germination in response to temperature was

found among invasive and native populations (Hirsch et al. 2012), nor for seedling biomass production (Hirsch et al. 2016). However, the germination traits that were differing between native and invasive populations can vary with the species: higher germination rates were found for invasive populations of the shrub *Hypericum perforatum* and the herbaceous *Achillea millefolium* compared to that of their native counterparts while no difference was found for *Hieracium pilosella*. Yet, germination speeds were significantly higher for invasive populations of *H. pilosella* whereas no difference was found among populations of both ranges for *A. millefolium* and invasive populations of *H. perforatum* even presented a slower germination speed (Beckmann et al. 2011). Still, although sometimes contradictory, these indices suggest that enhanced germination within invasive populations would likely contribute to species invasiveness, thus questioning the potential role of evolution of these traits related to the introduction history.

To evaluate the role of local adaptation and plasticity in invasion success, temperature appear as an interesting abiotic variable to consider. Indeed, temperature is known to be one of the most important abiotic factors shaping tree species distribution on a wide scale (Woodward 1987; De Frenne et al. 2010) in particular by impacting their fitness (Prentice et al. 1992; Iverson and Prasad 1998). For invasive species, selective pressures acting during the stages of invasion process could be related to temperature in the new range and result in shifts the species climatic niche (Broennimann et al. 2007). Moreover, germination and seedlings recruitment strongly respond to temperature either as it interacts with dormancy initiation and break, leaf or fruit phenology and seedling development, survival and growth (Walck et al. 2011).

Observed among ranges phenotypic shifts could be due either to genetic differentiation, phenotypic plasticity or both factors interacting. In order to test if there is a genetic basis for the phenotypic variation or if it is environmentally induced, it is necessary to conduct a common garden experiment in which native and invasive populations are raised together allowing to separate genetic and environmental effects and even to evaluate their interaction if several common gardens are used (Keller and Taylor 2008; Colautti and Lau 2015). Moreover, the comparison of phenotypic  $Q_{ST}$  and genotypic  $F_{ST}$  differentiation indexes allows to propose some evolutionary inferences: under neutrality, the expectation is that  $Q_{ST}$  would be similar to  $F_{ST}$ , a  $Q_{ST}$  value significantly higher than a  $F_{ST}$  value would indicate a divergent selection acting among populations; on the contrary, a  $Q_{ST}$  value inferior to a  $F_{ST}$  value would signal a stabilizing selection (Keller and Taylor 2008; Leinonen et al. 2008, 2013; Whitlock and Guillaume 2009; Colautti and Lau 2015).



Black locust *Robinia pseudoacacia* L. (Fabaceae) is a tree native to North America and is considered invasive worldwide (eight regions among the 14 defined by Richardson & Rejmánek, 2011). In the United States, it naturally occurs in the eastern half of the country, over two disjoint geographic areas (Little, 1971): the largest native area corresponded to the Appalachian Mountains and partially encompassed several current States (i.e. Pennsylvania, Maryland, West Virginia, Virginia, North Carolina, South Carolina, Georgia, Alabama, Tennessee, Kentucky, and Ohio); the smallest area was located further west in the Ozark Mountains (i.e. Missouri, Arkansas and Oklahoma). In Europe, it was introduced at the beginning of the 17<sup>th</sup> century and is now recognized as one of the 100 worst invasive species (DAISIE 2006, 2009). A wide population genetics study demonstrated that only individuals from few populations of the Appalachian region were introduced to Europe. Thus a genetic bottleneck associated with a loss of genetic diversity in the introduced range was evidenced (Bouteiller et al. 2019). However, comparing populations within the introduced range, a recent study evidenced no sign of local adaptation in the response to temperature of germination and juvenile related traits (Bouteiller et al. 2018). In this study, we intended to test for local adaptation and phenotypic plasticity on germination related traits among populations from both native and invasive ranges raised in three different environments. Thus a large sampling of both native and invasive populations (20 populations, 3000 plants) was performed to evaluate germination progress in response to temperature and answer the following questions: i) Do we observe any signal of local adaptation of populations from the native and/or invasive range with respect to the juvenile traits surveyed? ii) Is there a genetic differentiation at the range or population level and iii) What is the extent of phenotypic plasticity to temperature? iv)

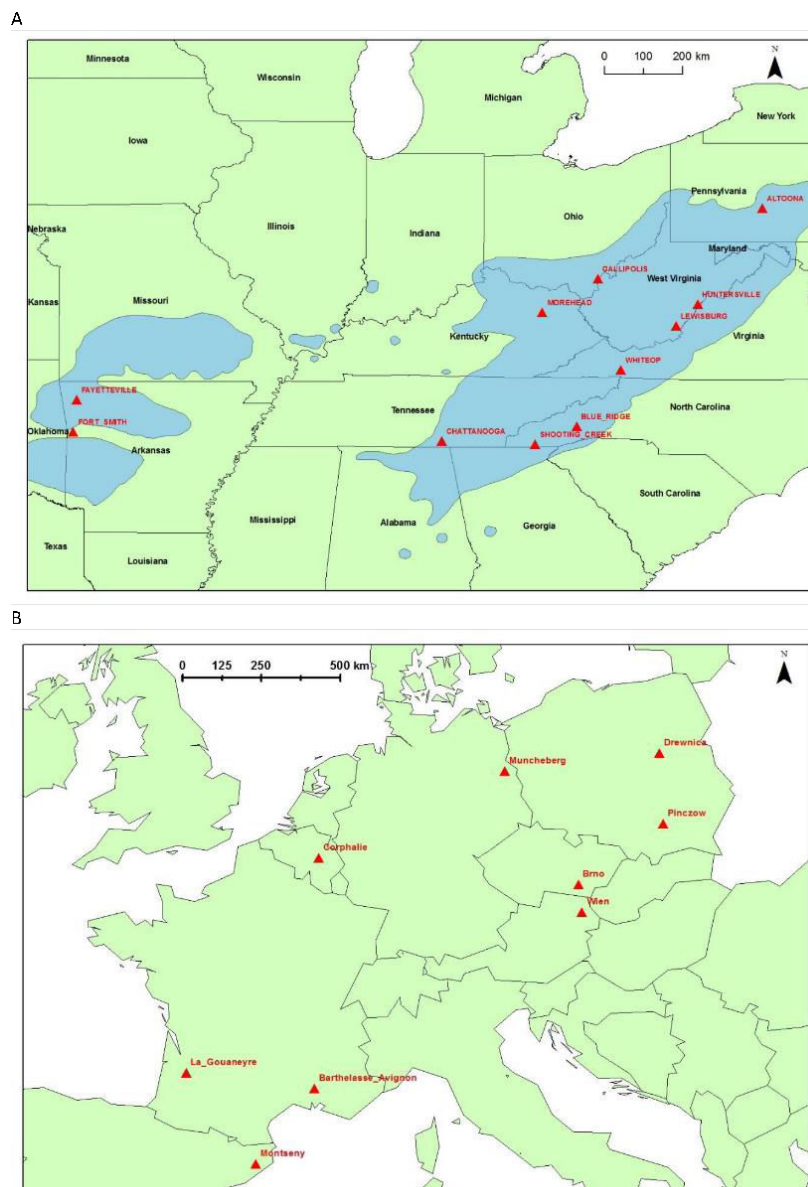
## **2. Material and Methods**

### **2.1. Sampling**

20 populations of *R. pseudoacacia* were sampled both in the native (11 populations) and invasive ranges (9 populations). Sampling was conducted between spring 2016 and fall 2016 by different collaborators using a similar protocol (Fig 1, Appendix 1). We sampled populations originated from natural regeneration. In North America, the different populations were spread all through the species native distribution from the Ozark Mountains to the Appalachian Mountains, and in Europe populations were dispatched all through the introduced range from Spain to Poland. In each population, 10 to 100 pods were collected on ten trees with a tree pruner.

Given that the species is known to spread by root-suckering (Cierjacks et al. 2013) a minimal distance of twenty meters was kept between two sampled trees in order to minimize the risk of collecting the same genotype. GPS coordinates either of the population or of each sampled tree were recorded using a portable GPS (GPSMAP62, Garmin, Olathe, KS, USA).

Seeds were manually extracted from pods and placed into a tea paper filter to be stored in a cold room at 0 - 5°C (Forest Research 2015; Kew Royal Botanic Garden 2015). Only well conserved seeds (i.e. seeds without mold stains or without damaged tegument) were kept and seeds with moist or empty appearance were systemically eliminated. Seeds were counted for each maternal tree and weighted with a 0.1mg accuracy (Practum 224-1S, Sartorius, Goettingen, Germany).



**Fig 1** Map of the sampled populations A) in the USA; the blue areas represent the native range of black locust (Little, 1971) B) in Europe.

## **2.2. Experimental design for phenotypic measurements**

A controlled experiment was set up using three climatic chambers (Micro-Clima Series Arabidopsis Chamber MCA1600, Snijders Labs, Tilburg, Netherland) with a total of 3000 individuals. We sowed 5 seeds per family (i.e. per maternal tree) in the same tray, using 10 families per populations with 9 European populations and 11 American populations. Within each climatic chamber, families were randomly dispatched.

Prior to seeding, seeds were mechanically scarified using an automated sand blasting technique (Bouteiller et al. 2017) to ensure a controlled dormancy breaking. Seeds were sowed into plastic trays (QuickPot 35RW, HerkuPlast Kubern GmbH, Ering, Germany); each pot was filled with 25 g of substrate (Substrate 307, Peltracom, Gent, Belgium) plus 6 g after sowing to cover the seeds (Bonner and Karrfalt 2008).

Environmental conditions within both chambers were set equal except for air temperature: a day/night 12/10 h photoperiod with progressive day/night transitions of one hour,  $60 \pm 5\%$  of air relative humidity,  $185 \pm 45 \mu\text{mol.m}^2.\text{s}^{-1}$  photosynthetic photon flux density (LiCOR Li190, Lincoln, USA).  $\text{CO}_2$  concentration was equal to ambient atmospheric concentration.

Temperature conditions were  $18^\circ\text{C}/16^\circ\text{C}$  day/night temperature in the first climatic;  $22^\circ\text{C}/20^\circ\text{C}$  day/night temperature in the second climatic chamber and  $31^\circ\text{C}/29^\circ\text{C}$  day/night temperature in the third climatic chamber to mimic a heat wave.

Watering (50 mL) was provided every two days in order to supply a non-limiting resource. After 31 days, seedlings were fertilized with a liquid fertilizer (NPK 7.5.6, Florendi Jardin SAS, Dinard, France), renewed twice every 10 days.

## **2.3. Phenotypic life history traits and functional traits**

### **2.3.1. Life history traits**

Germination and seedling phenology were monitored daily for each individual over a period of 372 Growing Degree Days (GDD). GDD were calculated by multiplying the climatic chamber day temperature condition (in  $^\circ\text{C}$ ) and the number of days spent in the climatic chamber, base temperature was considered equal to 0 since no reference value was found for black locust. Five phenological stages have been defined as follows: 1, emergence; 2, straight stem; 3, open cotyledons; 4, first leaf (see Bouteiller et al. 2017 for details and

photographs of the phenological stages). According to air temperature, the phenological survey was stopped after 12 days in the 31°C climatic chamber, after 17 days in the 22°C climatic chamber and after 21 days in the 18°C climatic chamber. For each phenological stage and each individual, the minimum number of days required to reach the stage was noted. When a stage was missing, the number of days was calculated as the average between the numbers of days required to reach the previous stage and the next one. When an individual died after germinating, it was further recorded as a missing value.

### **2.3.2.Functional traits**

At the end of the phenological survey (372 GDD), total height was measured from collar to the apical bud using a ruler (Height 1, cm ~1mm). Seedlings were thinned from 5 to 3 plants per pot in order to avoid competition, preferentially removing damaged individuals. After 1450 GDD corresponding to 49, 66 and 82 days in the 31°C, 22°C and 18°C climatic chamber, respectively, total height (Height 2, cm ~1mm), stem collar diameter (Diameter, mm ~ 0.01mm) and the effective quantum yield of photosystem II (PSII yield) were measured using a ruler, an electronic caliper and a portable chlorophyll fluorometer (PAM 2100, Heinz Walz GmbH, Effeltrich, Germany), respectively.

The PSII yield was measured on the terminal leaflet of the youngest fully-developed. This trait provides information about the proportion of photons captured by the PSII that is used in photochemistry (Genty et al. 1989; Santiso et al. 2015).

Then, seedlings were separated into leaves and stem, and oven dried at 65°C (Universal oven, Memmert, Swabach, Germany). After minimum one drying week to reach constant weight, leaves (Foliar weight) and stem (Stem weight) were weighted with a 0.1 mg accuracy (Entris 224-1S, Sartorius, Goettingen, Germany) and Total weight was calculated by summing both values. Growth Rates (GR) in height (Height GR 1 and 2), diameter (diameter GR) and leaf (Foliar weight GR), stem (Stem weight GR) and total (Total weight GR) biomass were calculated by dividing the trait value by the number of days spent in each climatic chamber. Data will be accessible on Open Science Framework repository after acceptance.

### **2.4.Genotyping data**

Genotyping of the seedlings was not performed in this study, but we extracted the genotypic information acquired in Bouteiller et al. (2019) using SNP molecular markers. 18 populations out of the 20 used in this study were genotyped previously. For two populations, no genotyping was performed previously so we used genotypic data from the closest geographic population (Ironton for Gallipolis and Slatyfork for Huntersville). Detailed protocol on DNA extraction and genotyping is available in Bouteiller et al. (2019) and more information on SNPs development in Verdu et al. (2016).

## 2.5. Data analysis

### Environment – traits correlations

To analyze the relationship between traits and environmental conditions, mean annual temperature averaged over the 1970 – 2000 period was extracted from WorldClim version 2.2 (Fick & Hijman 2017) using ARcGIS (ESRI, 2018) with the 30 seconds resolution raster (<http://worldclim.org/version2>) at the location of each sampled population according to the GPS coordinates of the sampled populations. A linear model (lm function and package tidyverse v1.2.1 and gapminder v0.3.0) was used to fit a regression between each trait averaged over each population against mean annual temperature of the population (R version 3.3.1, R Core Team, 2016). The significant slope estimates ( $p < 0.05$  Student t-test) were extracted and visualized in the form of a heatmap (package ggplot2 v2.2.1).

### Statistical model

To estimate differentiation between ranges and populations and between climatic chambers (i.e., the temperature effect) for each measured trait, a Bayesian mixed model of ANCOVA was fitted using the mothers' average seed weight as covariate as a proxy of maternal effect (Bouteiller et al. 2018). Differentiation between ranges and populations for the mothers' average seed weight was conducted using the same model but without the temperature effect (nor the interaction) as seeds were sampled in the field and without the tree effect as we only disposed of mothers' average seed weight.

The complete between-chambers model was defined as:

$$\underbrace{Y_{[h]}}_{\text{trait value}} = b_0 + \underbrace{b_{1,i[h]}}_{\text{temp}} + \underbrace{b_{2,j[h]}}_{\text{range}} + \underbrace{B_{3,k[h]}}_{\text{population}} + \underbrace{B_{4,l[h]}}_{\text{tree}} + \underbrace{B_{5,i[h]k[h]}}_{\text{temp} \times \text{pop}} + \underbrace{c(x_{l[h]} - \bar{x})}_{\text{maternal effect}} + \varepsilon_{m[h]} \quad (1)$$

With the residual as:  $\epsilon_{m[h]} \sim N(0, \sigma^2)$  (2)

Lowercases indicate the fixed effects and uppercases indicate the random effects, with  $i \in 1, 2, 3$  for the 18, 22 and 31°C temperature chamber conditions, respectively;  $j \in 1, 2$  for the two ranges: Europe, US;  $k \in [1, 9]$  for the 9 European populations and  $k \in [1, 11]$  for the 11 US populations;  $l \in [1, 10]$  for the 10 trees per population; and  $m \in [1, 5]$  for the 5 individuals (seeds) per tree.  $x_h$  represents the seed weight at the tree (i.e. family level) and  $\bar{x}$  the mean seed weight over all families. The code used for implementing this Bayesian mixed model is provided on the open source GitHub platform <https://github.com/xbouteiller/BayesMix>. More details can be found in Bouteiller et al. (2018) where it was tested and compared to frequentist methods. Additionally, for the germination rate corresponding to binary data (0 non germinated / 1 germinated), we used a Bernoulli distribution with a logit link function instead of a normal one for the model defined in equation 2. The code used for implementing this model is provided on the open source GitHub platform <https://github.com/xbouteiller/BayesMixBinary>.

In order to determine if the observed differentiation can be explained by the genetic proximity among populations, a second model similar to equation 2 was implemented with the additional incorporation of a matrix of pairwise genetic distance between populations (Hadfield and Nakagawa 2010). We used the complement of Cavalli-Sforza and Edwards Chord Distance (i.e. 1 – Distance) computed using Hierfstat v 0.04-28 R package (Goudet 2005) as the matrix of pairwise distances among populations as it is considered the best to retrieve the relationship among samples (Takezaki and Nei 1996). The code used for implementing this model is provided on the open source GitHub platform <https://github.com/xbouteiller/BayesMixGenetix>.

All Bayesian models were computed using JAGS 3.4.0 (Plummer 2005), the R2jags package of R (Su and Yajima 2012) and R v3.2.2 (R Development Core Team 2013). For each trait, 100,000 iterations were run with a burn-in of 75,000 iterations and a thinning interval of 50, using 4 chains. Autocorrelation and convergence were assessed using the “autocorr.plot” and “gelman.plot” native JAGS functions and Rhat convergence criterion was inferior to 1.01 as recommended (Kruschke 2015).

Codes provided by O’Hara et al. (2005) and Kruschke (2015) were adapted to our design using uninformative inverse gamma conjugate priors  $\Gamma(10^{-4}, 10^{-4})$  for random effects and normal priors  $N(0, 10^6)$  for fixed effects (O’Hara and Merilä 2005; Kruschke 2015).

### Calculation of genetic differentiation indexes

Genetic differentiation among populations can be defined from molecular markers analyses using  $F_{ST}$  indices calculated as follows (Wright, 1951):

$$F_{ST} = \frac{\sigma_B^2}{(\sigma_B^2 + \sigma_W^2)} \quad (3)$$

$F_{ST}$  within ranges were calculated using the Hierfstat v 0.04-28 R package (Goudet 2005) according to the Weir and Cockerham method (1984). 95% Confidence Intervals (CI) were estimated by performing 1000 bootstraps over all loci.

For each quantitative trait, the narrow sense heritability  $h^2$  and the  $Q_{ST}$  index of quantitative genetic differentiation among populations were calculated as follows (Spitze 1993):

$$h^2 = \frac{\sigma_W^2}{\sigma_P^2} \quad (4)$$

$$Q_{ST} = \frac{\sigma_B^2}{(\sigma_B^2 + 2\sigma_W^2)} \quad (5)$$

where  $\sigma_W^2$  is the within population genetic variance (i.e.  $V_A$  the additive variance),  $\sigma_B^2$  is the between populations genetic variance estimated using  $\hat{\sigma}_{pop}^2$  and  $\sigma_P^2$  the total phenotypic variance. Assuming that individuals obtained from seeds of the same tree were half-sibs, within population variance can be estimated as  $\hat{\sigma}_W^2 = 4\hat{\sigma}_{tree}^2$ .

$h^2$  values were estimated within range and between temperature conditions and  $Q_{ST}$  values were computed per temperature chamber condition using the same model than presented above (equation 1) but without the temperature and temperature x population interaction effects.

### Evaluation of plasticity

For each trait, plasticity among environments was evaluated using the Relative Distance Plasticity Index (RDPI, (Valladares et al. 2006; Godoy et al. 2012). RDPI is an index ranging between 0 and 1, defined as the absolute phenotypic distance between families across various environments, divided by the sum of the phenotypic values.

Thus, for a given family  $j$ , it was calculated following the mathematic formula:

$$RDPI = \sum_{\substack{i, i'=1 \\ i \neq i'}}^n \frac{|X_{ij} - X_{i'j}|}{(X_{ij} + X_{i'j})} \times \frac{1}{n} \quad (6)$$

where  $i$  and  $i'$  referred to two of the  $n$  levels of the environment treatment (here  $n$  is defined from 1 to 3 for each of the growing chamber temperature conditions), and  $j$  referred to a given family (i.e. a mother tree). Thus  $X_{ij}$  is

the phenotypic value for a given trait for the family  $j$  in the environment  $i$ .  $n$  is defined as the number of distances calculated (i.e. number of pairwise environments, 3 in our cases). Then one value of RDPI is obtained per mother tree. RDPI values were calculated using R, with a function available on the GitHub repository: <https://github.com/xbouteiller/Plasticity>.

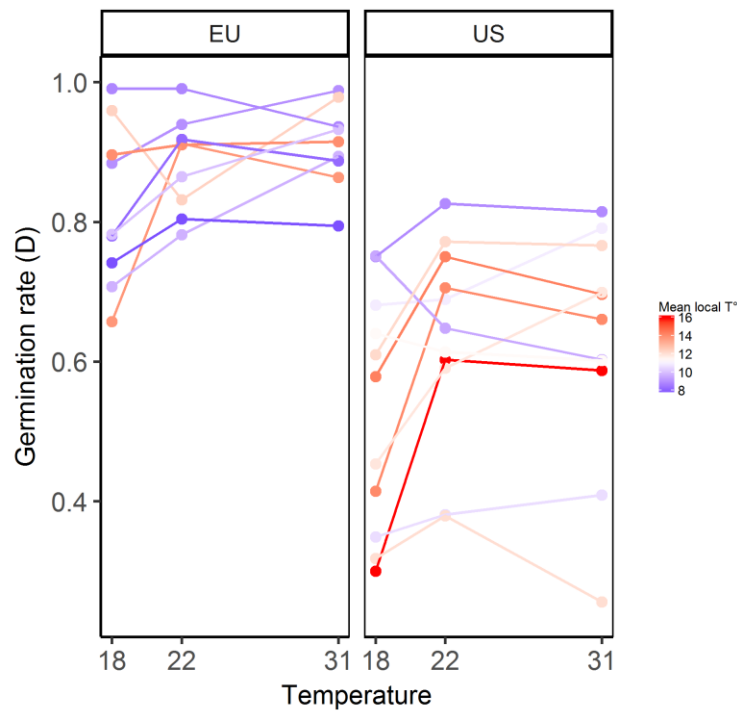
Among ranges and populations differentiation of plasticity was tested by using a similar Bayesian mixed model than defined above (equation 1) but without the temperature effect and the tree level effect.

### **3. Results**

#### **3.1. Phenotypic differences between ranges for germination rate and seed weight**

Overall, seed germination was higher in the European range compared to the native range whatever the growing chamber temperature (Fig 2, Table 1). The number of days for reaching stage 1 was also significantly shorter for European populations compared to the native ones (Table 1, 9.1/5.1/3.0 d versus 10.0/6.0/3.7 d in the 18/22/31°C temperature conditions) indicating a faster germination in the European range. The matrix of pairwise genetic distances between populations was incorporated into the model to account for the potential relatedness of the populations: when taking into account the information of the molecular similarity, differences between ranges for stage 1 was no more significant but it remained significant for germination rate (Appendix 6). Globally, over the 13 traits measured only two showed differences between the two ranges.





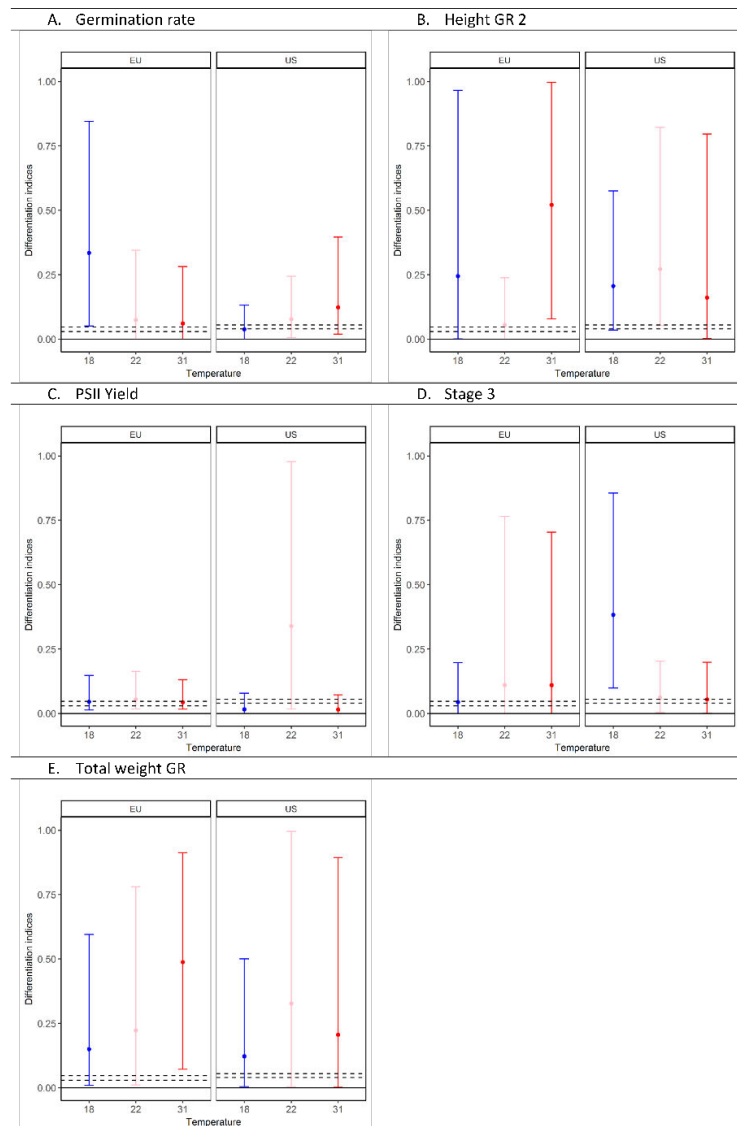
**Fig 2** Germination rate of black locust populations in response to growing temperature (18, 22 and 31°C). EU for European populations, US for American populations. Each line represents one population (averaged values over all individuals), the color indicating the mean annual temperature at the sampling location (from the coldest in blue, to the warmest in red).

Moreover, it is worth noticing that seed weight presented a significant difference between ranges (Appendix 2). Seed weight was significantly lower in the American families ( $17.13 \pm 4.30$  mg ranging from 4.32 mg to 25.51 mg) compared to the European ones ( $21.89 \pm 4.47$  mg ranging from 14.26 mg to 36.85 mg). There were also significant differences among populations: the American population presenting the heaviest seeds (Altoona: 19.89 mg) presented however values equivalent to the lightest seeds observed in European populations (Brno: 19.86 mg, Drewnica: 17.45 mg, Montseny: 19.73 mg).

### **3.2. Genetic differentiation among populations was evidenced more strongly in the native than in the invasive range**

At the molecular level, weak but significant genetic differentiation among populations was evidenced with similar levels of differentiation for both ranges. In the American native range,  $F_{ST}$  between populations was

equal to 0.0479 [0.0402 – 0.0558] and in the European range,  $F_{ST}$  between populations was equal to 0.0391 [0.0297 – 0.0476]. Significant differentiation among populations (Appendix 3c – A:M, ) was evidenced using a Bayesian mixed model for all traits except for the PSII yield. However, differentiation was mostly observed within the native American range, except regarding germination rate and early development stages (1-2). Populations that presented a significantly higher germination rate (90-100%) within Europe were Munchenberg and Brno whereas significantly lower germination rate were observed for Barthelasse and Corphalie under the 18°C temperature conditions (68% and 72% respectively). Differentiation was mostly evidenced under the 18°C conditions as suggested by the variability observed on the trait value (Fig 2) and by the  $Q_{ST}$  value (Fig 3). In the American range, two populations were differentiated with a particularly low germination rate, reaching 38-40% maximum (Lewisburg and Shooting Creek). For phenological stage 1, up to 3 days of delay was evidenced between the fastest (Munchenberg, 7.85 d under the 18°C condition) and the slowest populations of Barthelasse, Drewnica and Wien (10.7-10.9 d under the 18°C condition) in Europe. In the native range, the fastest populations to reach stage 1 were the ones from the northern Appalachian Plateaus (Altoona and Huntersville, 8.59 and 8.64 d under the 18°C condition) and the slowest populations were the ones from the Ozark Mountains with a 4 days delay (Fayetteville and Fort Smith, 12.2 and 12.3 d under the 18°C condition).



**Fig 3** Differentiation indices calculated using five phenotypic traits ( $Q_{ST}$ ) and using SNP markers ( $F_{ST}$ ).  $Q_{ST}$  values (mean and 95% confidence interval) were calculated either on the 9 European populations (left) and 11 US populations (right) per trait and growing chamber (blue circle: 18°C, pink circle: 22°C, red circle: 31°C). The dashed lines represent the  $F_{ST}$  95% confidence interval based on allelic variation in SNP's loci.  $Q_{ST}$  Estimations were calculated using Bayesian methods. GR: Growth Rate.

All traits were heritable (Table 1) except the number of leaves which had a very low heritability of 0.008 [0.000 – 0.051] in Europe and 0.027 [0.000 – 0.126] in America; however, no significant differences in heritability were found between ranges. Many  $Q_{ST}$  value were significantly higher than zero: indeed, among the 78 calculated  $Q_{ST}$  values (13 traits x 2 ranges x 3 environmental conditions), 33 were significantly higher than 0.005, evenly distributed between ranges (14 and 19 in the native and invasive range, respectively). Although  $Q_{ST}$  values could be high for some traits, the confidence interval over the mean was generally large and consequently,  $Q_{ST}$  values

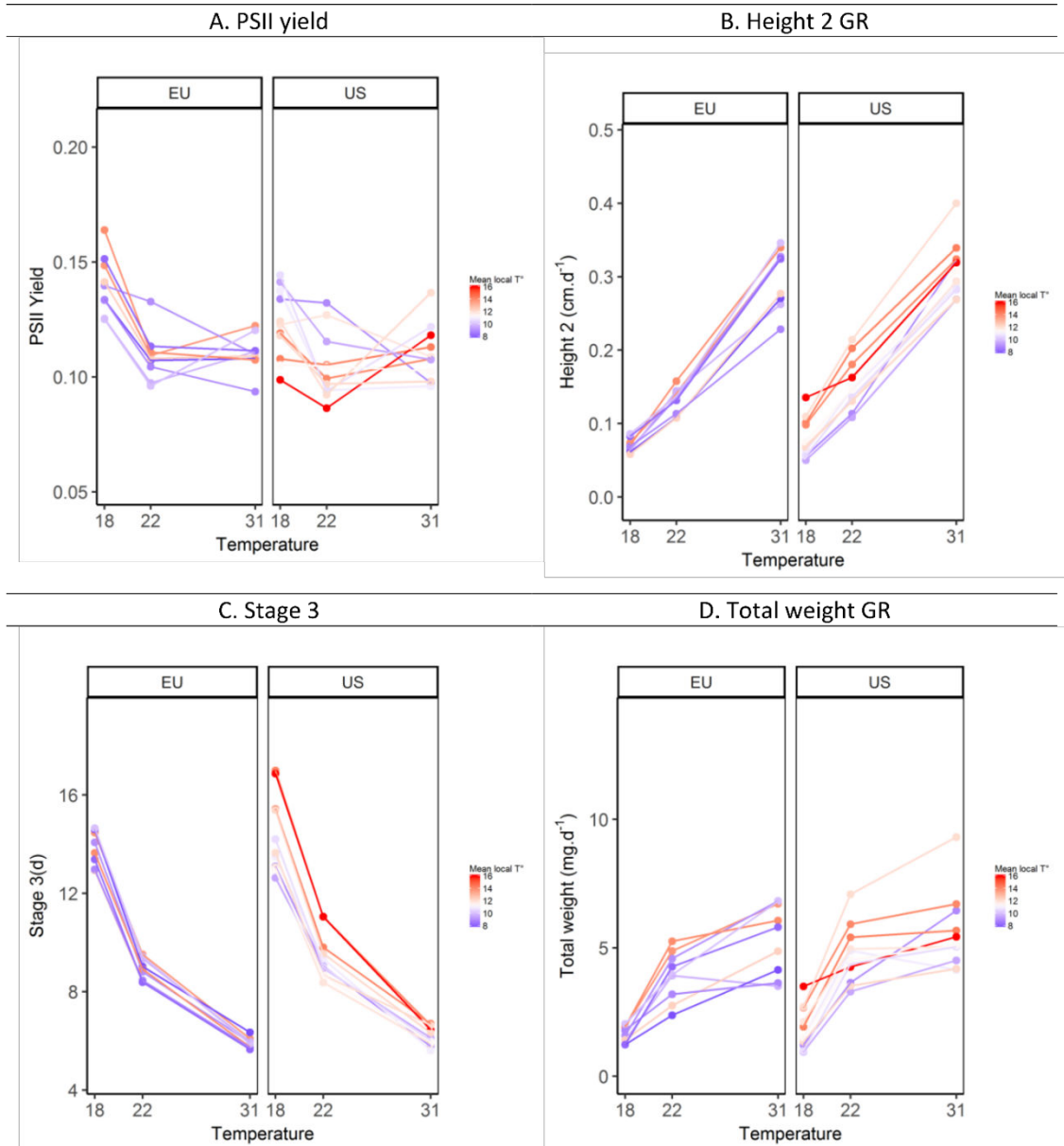
were not significantly different from  $F_{ST}$  values for most traits (Fig 3, Appendix 7). Still, in Europe,  $Q_{ST} > F_{ST}$  was found for Germination rate (0.236 [0.0500 – 0.634]) and early development (Stage 1, 0.236 [0.050 – 0.634]) both in the 18°C environment, for growth at 31°C (Height GR 2: 0.521 [0.0783 - 0.9976] and Total weight GR 2: 0.488 [0.0722 – 0.912]). In the native range only one  $Q_{ST}$  value was significantly higher than the  $F_{ST}$  value the Stage 3 at 18 °C (0.383 [0.0933 – 0.855]).

Table 1 Trait values (mean and standard deviation) within range and within temperature growing conditions (18, 22 and 31°C). Bold characters indicate the traits which present significant differences between ranges. Heritability and 95% credibility intervals of all traits. Heritability was estimated within range and between the three temperature conditions; & RDPI values (mean + sd)

		T18		T22		T31		h <sup>2</sup>			RDPI	
		Mean	sd	Mean	sd	Mean	sd	Mean	95% CI inf	95% CI sup	Mean	sd
Germination Rate	EU	<b>0.84</b>	<b>0.36</b>	<b>0.90</b>	<b>0.31</b>	<b>0.90</b>	<b>0.31</b>	0.47	0.08	1.04	<b>0.090</b>	<b>0.097</b>
	US	<b>0.55</b>	<b>0.50</b>	<b>0.64</b>	<b>0.48</b>	<b>0.60</b>	<b>0.49</b>	0.56	0.12	1.24	<b>0.218</b>	<b>0.229</b>
Stage 1	EU	<b>9.11</b>	<b>2.74</b>	<b>5.08</b>	<b>1.57</b>	<b>3.05</b>	<b>1.11</b>	0.11	0.05	0.20	0.346	0.069
	US	<b>10.01</b>	<b>3.27</b>	<b>6.01</b>	<b>2.23</b>	<b>3.68</b>	<b>1.40</b>	0.14	0.05	0.26	0.332	0.082
Stage 2	EU	11.12	2.46	6.76	1.60	4.28	1.02	0.18	0.11	0.29	0.303	0.043
	US	11.89	2.77	7.72	2.25	4.82	1.18	0.23	0.13	0.36	0.299	0.054
Stage 3	EU	13.84	2.55	8.90	1.88	5.89	1.03	0.09	0.04	0.17	0.274	0.044
	US	14.03	2.45	9.44	2.07	6.15	1.12	0.20	0.11	0.32	0.273	0.050
Stage 4	EU	20.47	0.82	13.56	1.54	7.82	0.93	0.05	0.03	0.08	0.315	0.022
	US	20.47	0.84	14.09	1.76	8.10	1.06	0.09	0.05	0.16	0.304	0.024
Height GR 1	EU	0.157	0.044	0.232	0.054	0.380	0.087	0.24	0.13	0.37	0.287	0.086
	US	0.152	0.045	0.225	0.063	0.362	0.098	0.22	0.11	0.38	0.294	0.088
Height GR 2	EU	0.070	0.039	0.131	0.050	0.301	0.103	0.08	0.02	0.16	0.439	0.072
	US	0.074	0.037	0.150	0.065	0.306	0.102	0.16	0.06	0.30	0.442	0.096
Diameter GR	EU	0.028	0.007	0.033	0.011	0.027	0.009	0.59	0.37	0.88	0.136	0.068
	US	0.025	0.006	0.033	0.012	0.027	0.009	0.77	0.48	1.19	0.147	0.064
N of Leaves	EU	4.65	1.73	7.56	2.08	11.21	2.16	0.01	0.00	0.05	0.292	0.092
	US	4.72	1.93	7.69	2.03	11.01	2.36	0.03	0.00	0.13	0.292	0.108
Foliar weight GR	EU	0.89	0.86	2.57	1.93	4.00	2.55	0.14	0.00	0.38	0.480	0.147
	US	1.09	1.16	3.11	2.35	4.02	2.67	0.39	0.04	0.84	0.466	0.163
Stem weight GR	EU	0.71	0.53	1.33	1.19	1.39	1.24	0.16	0.00	0.42	<b>0.318</b>	<b>0.141</b>
	US	0.58	0.38	1.51	1.31	1.35	1.20	0.26	0.00	0.62	<b>0.376</b>	<b>0.127</b>
Total weight GR	EU	1.60	1.24	3.91	3.02	5.39	3.73	0.14	0.00	0.37	0.410	0.140
	US	1.67	1.48	4.62	3.45	5.37	3.80	0.40	0.13	0.80	0.428	0.143
PSII yield	EU	0.140	0.044	0.109	0.039	0.110	0.039	0.39	0.17	0.68	0.173	0.083
	US	0.125	0.046	0.105	0.039	0.107	0.033	0.38	0.15	0.68	0.164	0.073

### 3.3. Functional traits strongly responded to temperature, without range distinction

Overall, a strong response to growing temperature was observed (Fig 4 A, B, D, Appendix 5). Globally, individuals performed better (i.e. faster growth or development) in one or two of the warmest temperature conditions (22°C and 31°C) compared to the coldest condition (18°C).



**Fig 4** Reaction norms to growing temperature conditions (18, 22, 31°C) for 4 of the measured traits. A: PSII yield; B: Height 2 GR; C: Stage 3; D: Total weight GR. Each line represents one population (averaged values over all individuals), the color indicating the mean annual temperature at the sampling location (from the coldest in blue, to the warmest in red).

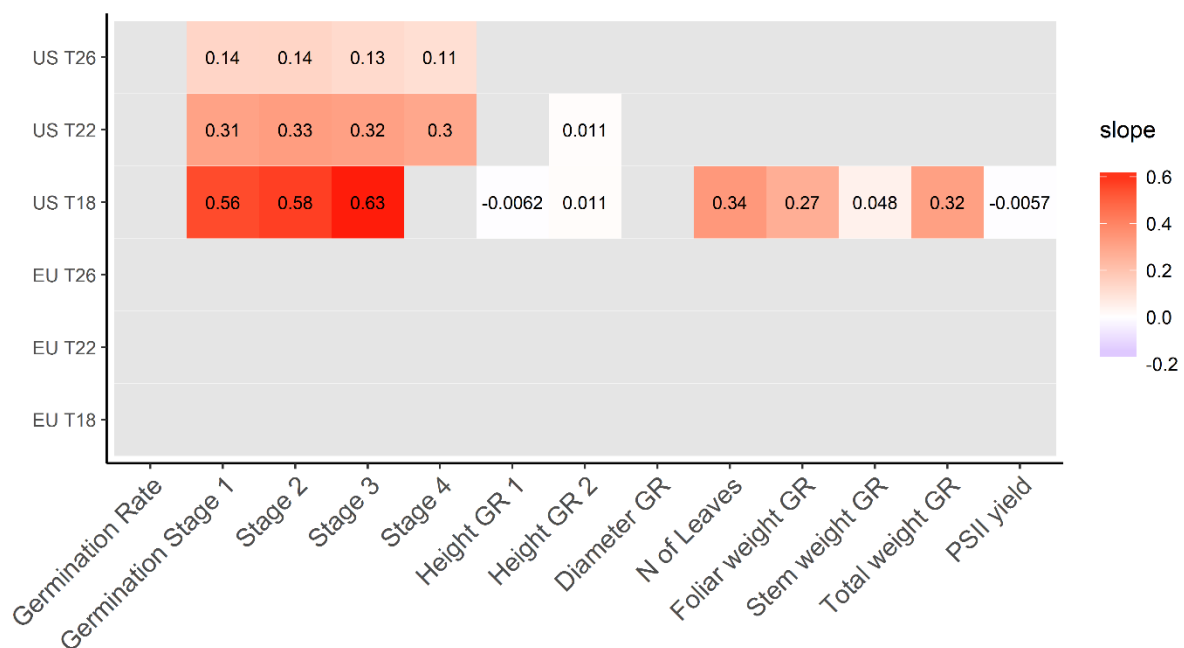
Germination rate responded strongly to temperature and was in average the highest in the 22 and 31°C temperature conditions (Fig. 2, EU: 0.84/0.90/0.90, US: 0.0.55/0.64/0.60 in the 18/22/31°C temperature conditions). For all traits we observed significant temperature effects (Fig 4, Appendix 3, Appendix 5) but with some slight variations in the optimal temperature. Significantly higher values of the PSII yield were observed in the coldest environment (0.133 / 0.107/ 0.109 at 18/22/31°C over both ranges). But in general, most traits continuously increased with increasing temperature (Appendix 3: Height GR 1 and 2, Number of leaves, Foliar and Total weight GR): for example, the height growth rate doubled with each temperature change, from 18 to 22°C and to 31°C (Height GR 0.072 cm.d<sup>-1</sup> at 18°C, 0.141 cm.d<sup>-1</sup> at 22°C and 0.303 cm.d<sup>-1</sup> at 31°C, over both ranges). For some traits, a maximum was reached at 22°C, with a further plateau at 31°C for Stem weight GR or even a slight decrease for Diameter GR. Fig 4 presents four of the traits as examples that illustrate that the response to temperature was similar between European and American populations: a positive response was observed for the Height growth rate (Height GR 2) (EU: 0.07/0.13/0.30 cm.d<sup>-1</sup>, US: 0.07/0.15/0.31 cm.d<sup>-1</sup> in the 18/22/31°C temperature conditions) and above ground biomass increment (Total weight GR, EU: 1.60/3.91/5.39 mg.d<sup>-1</sup>, US: 1.67/4.62/5.37 mg.d<sup>-1</sup> in the 18/22/31°C temperature conditions) whatever the range; a negative response was observed for populations of both ranges for leaf physiology assessed through PSII yield (EU: 0.14/0.11/0.11, US: 0.12/0.11/0.11 in the 18/22/31°C temperature conditions) or phenological development (time to reach stage 3 EU: 13.84/8.90/5.89 d, US: 14.03/9.44/6.15 in the 18/22/31°C temperature conditions, Fig 4C).

Based on the analysis of the plasticity index RDPI (Table 1, Appendix 4), we observed differences in plasticity between ranges, for germination rate and for stem weight: plasticity was significantly higher for these traits for populations of the native American range (Table 1, germination rate RDPI: EU = 0.090, US = 0.218; stem weight RDPI: EU = 0.318, US = 0.376). All the other measured traits exhibited similar amount of plasticity between ranges.

### **3.4. Only native black locust populations evidenced significant trait-environment relationships**

Whatever the trait, the relationship with temperature at the sampling location was not significant in the invasive range. On the contrary, significant relationships between traits and temperature conditions at the sampling site were evidenced within the native American range (Fig 5). This relationship was evidenced for almost all life

history and functional traits measured in the 18°C climatic chamber whereas in the 22 and 31°C chambers, only the rhythm of seedling emergence (stages 1-4) was significantly related to temperature at the sampling location. Thus life history traits (i.e. germination stages) were always positively associated with mean annual temperature at the sampling site irrespective of the growing conditions for the native populations, though the relationship was stronger in the coldest environment.



**Fig 5** Heatmap illustrating the slope of the regression between each trait phenotypic value averaged over all population within each range against the mean annual temperature at the sampled location of the population. Only slopes with significant p-value were plotted ( $p < 0.05$ ). The color scale is relative to the value of the slope, from the highest positive slope in red to the lowest negative slope in blue.

#### 4. Discussion

Biological invasions represent an opportunity for studying recent evolution (Lee 2002; Keller and Taylor 2008; Colautti and Lau 2015); some studies (e.g. Maron et al. 2004; Colautti and Barrett 2013) demonstrated that rapid evolution likely occurred in the new range for *Lythrum salicaria* and *Hypericum perforatum* and thus rapid recent evolution could explain the invasiveness of the studied species. However, demonstrating recently acquired

local adaptation requires an intensive common garden experiment with several environmental conditions and populations; it also necessitates to evaluate a fitness proxy such as seed production. This is especially difficult for long-lived organisms such as trees and consequently, these kind of studies have been less frequently conducted on invasive trees (Richardson and Rejmánek 2011). Nonetheless, juvenile stages have been recognized as critical stages in the invasion process and they interact strongly with fitness (Donohue 2002; Gioria and Pyšek 2016); thus, for invasive trees, survey of seedling development could allow to detect genetically-based shifts between ranges. In this study, using a climatic chamber experiment with three temperature conditions, we studied germination-related traits within twenty populations of black locust from both the native American range and the invasive European range. Our results evidenced that germination rate was significantly higher within the European range and that these differences were genetically based; moreover, germination rates were significantly more robust to environmental changes among European populations. Among the other studied traits, we found no differentiation between ranges and overall, traits responded similarly to a temperature increase. Finally, we observed a signal of local adaptation to temperature only in the native range.

#### **4.1. Enhanced germination rate in the European range could reflect anthropic selection**

We found an original pattern of improved germination rate within the invasive range. Indeed, increase in germination rate has already been observed for invasive populations of *H. perforatum* and *A. millefolium* (Beckmann et al. 2011). However, it seems that in the invasive range, it is generally not the rule in invasion and several studies failed to demonstrate this pattern; still, they evidenced a change in germination traits within the invasive range, but rather found a faster germination (Erfmeier and Bruelheide 2005; Hodgins and Rieseberg 2011; Hirsch et al. 2012). Consistently with these studies, Gioria & Pysek (2016) reviewed that germination rate was not a consistent predictor of invasiveness whereas they highlighted a trend among invasive species to germinate faster in the invasive range. We also found that germination speed (i.e. stage 1) was faster within the invasive range, but the range effect remained relatively weak, about ten times lower than the temperature effect. One hypothesis to explain the observed shift in germination rate in black locust could be the occurrence of artificial anthropic selection, in which human activities would have unintentionally selected individuals with better seed traits. Indeed, a previous study demonstrated that black locust was introduced to Europe from a limited number of individuals from the North Eastern Appalachian plateau (Bouteiller et al., 2019): the



American population that was genetically closest to the European ones was Altoona in Pennsylvania where historical records also indicated that many seeds were collected and sent to Europe (Cobbett, 1828; Bouteiller et al., 2019). Seeds originating from Altoona presented the highest and the more constant germination rate in response to temperature among the American populations used in this study; although this co-ancestry link could contribute to the increase in germination rate observed in Europe, it is however not sufficient compared to the magnitude of the observed increase. On the other hand, several historical records attested that black locust seeds were further used for disseminating the species across its new range in Europe (Saint-Jean de Crève Coeur, 1786; François de Neufchateau, 1807., Cobbett, 1828). This repeated process could have resulted in a mass selection towards a more successful germination rate. Consistently with this hypothesis, we found that (i) the  $Q_{ST}$  value was significantly higher than the  $F_{ST}$  value for germination rate in the invasive range within the coldest environment, indicating a signal of directional selection whereas nothing similar was observed in the native range (ii) when taking into account the information of the molecular similarity between populations, the range effect remained significant for the germination rate. Thus a possible scenario may be that human unwillingly selected and propagated the best seeds of black locust in the new range that would have favored an evolution of germination rate through European populations. In any case, the strong enhancement in germination among European populations would likely be a key process in the success and dynamics of black locust in Europe.

#### **4.2. Evidence for genetic differentiation among populations but no signal of local adaptation was detected in the invasive range**

Phenotypic differentiation was observed among populations for all traits in both ranges. As individuals were cultivated in the same environmental conditions, we can conclude that these differentiations had a genetic basis. Differentiation between populations of the native range was expected because local adaptation is common for plants in their natural range (Linhart and Grant 1996; Leimu and Fischer 2008) but it was not trivial between invasive populations. For example, based on common garden experiments, invasive populations of *Eschscholzia californica* did exhibit significant genetic differentiation on several quantitative traits. However more contrasting results were found among invasive populations of maple trees: populations of the tree *Acer negundo* exhibited genetic differentiation whereas no genetic differentiation was evidenced among invasive populations of *Acer platanoides* (Lamarque et al. 2014). Based on molecular markers, the herbaceous *Phylla canescens* exhibited genetic differentiation between populations of the native South American range and the

invasive Australian range but not within the invasive European range (Xu et al. 2015). Thus divergent results can be found depending on species and empirical studies are still needed to evaluate the role of evolution in biological invasions. When comparing Belgium and French populations of black locust, no significant differentiation was found between populations since all the variability was observed at the tree level (Bouteiller et al. 2018). The current study was able to draw evidence of differentiation by increasing the sampling among both native and invasive populations. Thus we can ask which evolutionary forces sharpened the differentiation between populations.

A signal of local adaptation to temperature was found among populations of the native range but not within the invasive range: significant correlations among phenotypic values averaged per population and the local environment (i.e. mean annual temperature at the original position of the population) were indeed observed only within the native range. Additionally, some native populations belonging to the warmest part of the native range (i.e. Fayetteville and Fort Smith) had a very poor performance regarding germination rate when sown within the coldest environment, whereas their germination rates were three or four times higher within the two warmest environments. Rigorously speaking, testing for local adaptation of germination would require a reciprocal sowing experiment using seeds collected on plants grown in the same environment to remove potential maternal effects (Donohue et al. 2010). However, this kind of experimental design would be difficult to undertake for a long lived species, for which the first flowering period is reached after several years, a minimum of 6 years for black locust (Cierjacks et al. 2013). In the framework of invasion ecology, Monty et al. (2009) tested for maternal effects on *Senecio inaequidens* but although seed weight presented some variability, no maternal effect was evidenced.

Using a  $Q_{ST} - F_{ST}$  comparison approach, no unidirectional support for a departure from neutral evolution was found in our study: mean  $Q_{ST}$  values were generally larger than  $F_{ST}$  values in both ranges but only a few were significantly different.  $Q_{ST}$  estimation often suffers of some limitations and it could be biased downward or upward. To account for this bias in  $Q_{ST}$  estimation, a Bayesian approach was employed (Bouteiller et al. 2018); indeed this method had been shown to minimize the bias in  $Q_{ST}$  estimation (O'Hara and Merilä 2005; Leinonen et al. 2013). Despite this precaution, for black locust, large confidence intervals around  $Q_{ST}$  mean were found for many traits. Precision on  $Q_{ST}$  estimate depends mainly on the number of populations used, with a minimum of 20 populations suggested (O'Hara and Merilä 2005; Goudet and Büchi 2006). Thus increasing  $Q_{ST}$  estimate precision would have required a much more intensive experimental design, by doubling the number of

populations; unfortunately, that would have been achieved at an unrealistic cost. Additionally,  $Q_{ST}$  estimation assumes that genetic effects are purely additive, still maternal effects, dominance and epistatic variance could bias downward  $Q_{ST}$  estimation making results more conservative (O'Hara and Merilä 2005; Whitlock 2008; Leinonen et al. 2013).

As a consequence, seed weight was used as covariate to account for maternal effects, and half-sibs were used to reduce dominance and epistatic effects. All put together, the fail in detecting  $Q_{ST}$  higher than  $F_{ST}$  could be due to the inherent difficulties of the estimation that would be hardly removed. Colautti et al. (2015) emphasized that  $Q_{ST}$   $F_{ST}$  comparisons could be hard to interpret for invasive species and that results should be interpreted cautiously at the light of a common garden experiment. These limits were illustrated through a typical example on *Lythrum salicaria*, where one  $Q_{ST}$   $F_{ST}$  comparison failed to demonstrate evidence of selection in the invasive range whereas local adaptation was demonstrated based on analysis of populations along natural clines and using a reciprocal transplant experiment. Interestingly, for black locust, many  $Q_{ST}$  were significantly higher than 0 attesting significant genetic differentiation on phenotypic traits among populations whereas Bouteiller et al. (2018) on the same traits found no  $Q_{ST}$  significantly different from zero. Once again enlarging the sampling both in the native and invasive ranges permitted to capture more finely the intra-specific genetic diversity. Thus, it is likely that selection on juvenile traits acted conjointly to neutral evolutionary forces for differentiating native populations leading to local adaptation.

Within the invasive range, no signal of local adaptation in response to temperature was evidenced for germination-related traits. A similar result was observed among twenty black locust populations from the introduced range (Bouteiller et al. 2018). Among invasive species, a recent meta-analysis highlighted that local adaptation was as common for invasive species in their new range that for native species (Oduor et al. 2016), but examples mostly came from herbaceous species (8 tree species out of 137). For local adaptation of perennial woody species, time since introduction and number of generations occurred in the new range are likely to be crucial. As black locust is a long lived species with a first generation timing occurring after 6 years and optimal seed production starting at 15 years of age (Huntley, 1990), it is possible that not enough time was elapsed since introduction. In comparison, *Lythrum salicaria* was introduced to America at the beginning of the XIX<sup>th</sup> century, resulting in approximately 200 generations since introduction, considering one generation per year. Considering a minimum generation time of 6 years for black locust, approximately only 65 generations would have occurred since the beginning of the XVII<sup>th</sup> century.

### **4.3. The strong positive response to temperature for all traits would favor the species facing climate warming**

For all phenotypic traits with the exception of germination rate for European populations, we observed a strong response to temperature: thus all those traits appeared to be plastic to temperature. However, except for germination rate, no difference in plasticity to temperature was detected between ranges. As evocated in the introduction, the role of plasticity in invasion success remains unclear. The analysis of germination response to a temperature increase (24/27°C) for invasive populations of two *Ludwigia* species resulted in opposite patterns; for one of the species, germination rate and rhythm increased with increasing temperature, for the other, germination rate was high and showed no plasticity to temperature (Gillard et al. 2017). Since no between ranges difference in plasticity was observed for most of the studied traits, we cannot argue that globally, plasticity evolved during or since introduction. Overall, the increase in temperature was beneficial to black locust development: the warmer the environment, the higher or the faster the growth rate or the germination trait whatever the population, even in the 31°C climatic chamber simulating an extreme heat wave. A similar pattern was evidenced by surveying germination timing among 20 black locust populations from the invasive range sowed at 18°C and 22°C (Bouteiller et al. 2018). Consequently, as long as water is available, germination of black locust is likely to be favored by the increase in temperature induced by climate change.

## **5. Conclusion**

We found a strong support for a genetic differentiation of germination rate between native and invasive populations of black locust. Germination rate was enhanced within the invasive range and was less plastic to temperature changes than in the native range where decrease in temperature resulted in decreased germination rate. All other studied traits of juvenile development were highly plastic and positively correlated to temperature whatever the population and range. Human activity was probably an important determinant of this shift in germination rate because seeds were used since the 18<sup>th</sup> century to disseminate black locust all through Europe. Assessment of seedling development in the field in both ranges would allow to determine if this shift in adaptive capacities would lead to effectively increased germination and success of black locust natural regeneration.

## 6. Bibliography

- Agrawal AA (2001) Phenotypic plasticity in the interactions and evolution of species. *Science* (80- ) 294:321–326 . doi: 10.1126/science.1060701
- Barrett SCH, Colautti RI, Eckert CG (2008) Plant reproductive systems and evolution during biological invasion. *Mol Ecol* 17:373–83 . doi: 10.1111/j.1365-294X.2007.03503.x
- Beckmann M, Bruelheide H, Erfmeier A (2011) Germination responses of three grassland species differ between native and invasive origins. *Ecol Res* 26:763–771 . doi: 10.1007/s11284-011-0834-3
- Blair AC, Wolfe LM (2004) The Evolution of an Invasive Plant : An Experimental Study with *Silene latifolia*. *Ecology* 85:3035–3042 . doi: 10.1890/04-0341
- Bonner FT, Karrfalt RP (2008) The Woody Plant Seed Manual, Agric. Han. U.S. Department of Agriculture, Forest Service, Washington, DC
- Bouteiller XP, Barraquand F, Garnier-géré P, et al (2018) No evidence for genetic differentiation in juvenile traits between Belgian and French populations of the invasive tree *Robinia pseudoacacia*. *Plant, Ecol Evol* 151:5–17 . doi: <https://doi.org/10.5091/plecevo.2018.1403>
- Bouteiller XP, Porté AJ, Mariette S, Monty A (2017) Using automated sanding to homogeneously break seed dormancy in black locust (*Robinia pseudoacacia* L., Fabaceae). *Seed Sci Res* 1–8 . doi: 10.1017/S0960258517000150
- Broennimann O, Treier UA, Müller-Schärer H, et al (2007) Evidence of climatic niche shift during biological invasion. *Ecol Lett* 10:701–709 . doi: 10.1111/j.1461-0248.2007.01060.x
- Cierjacks A, Kowarik I, Joshi J, et al (2013) Biological Flora of the British Isles: *Robinia pseudoacacia*. *J Ecol* 101:1623–1640 . doi: 10.1111/1365-2745.12162
- Colautti RI, Barrett SCH (2013) Rapid Adaptation to Climate Facilitates Range Expansion of an Invasive Plant. *Science* (80- ) 342:364–366 . doi: 10.1126/science.1242121
- Colautti RI, Lau JA (2015) Contemporary evolution during invasion : evidence for differentiation, natural selection, and local adaptation. *Mol Ecol* 24:1999–2017 . doi: 10.1111/mec.13162
- DAISIE (2006) Delivering Alien Invasive Species In Europe. <http://www.europe-aliens.org>. Accessed 26 Feb 2015
- DAISIE (2009) Handbook of alien species in Europe. Springer, Dordrecht, Netherlands
- Davidson AM, Jennions M, Nicotra AB (2011) Do invasive species show higher phenotypic plasticity than

- native species and, if so, is it adaptive? A meta-analysis. *Ecol Lett* 14:419–31 . doi: 10.1111/j.1461-0248.2011.01596.x
- De Frenne P, Graae BJ, Kolb A, et al (2010) Significant effects of temperature on the reproductive output of the forest herb *Anemone nemorosa* L. *For Ecol Manage* 259:809–817 . doi: 10.1016/j.foreco.2009.04.038
- Donohue K (2002) Germination timing influences natural selection on life-history characters in *Arabidopsis thaliana*. *Ecology* 83:1006–1016 . doi: 10.1890/0012-9658(2002)083[1006:GTINSO]2.0.CO;2
- Donohue K, Dorn L, Griffith C, et al (2005) The evolutionary ecology of seed germination of *Arabidopsis thaliana*: variable natural selection on germination timing. *Evolution* 59:758–770 . doi: 10.1554/04-419
- Donohue K, Rubio de Casas R, Burghardt L, et al (2010) Germination, Postgermination Adaptation, and Species Ecological Ranges. *Annu Rev Ecol Evol Syst* 41:293–319 . doi: 10.1146/annurev-ecolsys-102209-144715
- Erfmeier A, Bruelheide H (2005) Invasive and native *Rhododendron ponticum* populations: is there evidence for genotypic differences in germination and growth? *Ecography (Cop)* 28:417–428 . doi: 10.1111/j.0906-7590.2005.03967.x
- Eriksson O, Ehrlén J (2008) Seedling recruitment and population ecology. 239–254
- Forest Research (2015) Seed storage and pretreatment for *Robinia pseudoacacia*. <http://www.forestry.gov.uk>. Accessed 10 Feb 2015
- Genty B, Briantais J-M, Baker NR (1989) The relationship between the quantum yield of photosynthetic electron transport and quenching of chlorophyll fluorescence. *Biochim Biophys Acta - Gen Subj* 990:87–92 . doi: 10.1016/S0304-4165(89)80016-9
- Ghalambor CK, McKay JK, Carroll SP, Reznick DN (2007) Adaptive versus non-adaptive phenotypic plasticity and the potential for contemporary adaptation in new environments. *Funct Ecol* 21:394–407 . doi: 10.1111/j.1365-2435.2007.01283.x
- Gillard M, Grewell BJ, Deleu C, Thiébaud G (2017) Climate warming and water primroses: Germination responses of populations from two invaded ranges. *Aquat Bot* 136:155–163 . doi: 10.1016/J.AQUABOT.2016.10.001
- Gioria M, Pyšek P (2016) Early bird catches the worm: germination as a critical step in plant invasion. *Biol Invasions* 19:1055–1080 . doi: 10.1007/s10530-016-1349-1
- Godoy O, Valladares F, Castro-Díez P (2011) Multispecies comparison reveals that invasive and native plants differ in their traits but not in their plasticity. *Funct Ecol* 25:1248–1259 . doi: 10.1111/j.1365-2435.2011.01886.x

- Godoy O, Valladares F, Castro-Díez P (2012) The relative importance for plant invasiveness of trait means, and their plasticity and integration in a multivariate framework. *New Phytol* 195:912–922 . doi: 10.1111/j.1469-8137.2012.04205.x
- Goudet J (2005) HIERFSTAT, a package for R to compute and test hierarchical F-statistics. *Mol Ecol Notes* 5:184–186 . doi: 10.1111/j.1471-8278.2004.00828.x
- Goudet J, Büchi L (2006) The effects of dominance, regular inbreeding and sampling design on QST, an estimator of population differentiation for quantitative traits. *Genetics* 172:1337–1347 . doi: 10.1534/genetics.105.050583
- Hadfield JD, Nakagawa S (2010) General quantitative genetic methods for comparative biology: Phylogenies, taxonomies and multi-trait models for continuous and categorical characters. *J Evol Biol* 23:494–508 . doi: 10.1111/j.1420-9101.2009.01915.x
- Hirsch H, Hensen I, Wesche K, et al (2016) Non-native populations of an invasive tree outperform their native conspecifics. *AoB Plants* 8: . doi: 10.1093/aobpla/plw071
- Hirsch H, Richardson DM, Le Roux JJ (2017) Introduction to the special issue: Tree invasions: Towards a better understanding of their complex evolutionary dynamics. *AoB Plants* 9: . doi: 10.1093/aobpla/plx014
- Hirsch H, Wypior C, Wehrden H Von, et al (2012) Germination performance of native and non-native *Ulmus pumila* populations. *NeoBiota* 68:53–68 . doi: 10.3897/neobiota.15.4057
- Hodgins KA, Rieseberg L (2011) Genetic differentiation in life-history traits of introduced and native common ragweed (*Ambrosia artemisiifolia*) populations. *J Evol Biol* 24:2731–2749 . doi: 10.1111/j.1420-9101.2011.02404.x
- Huntley JC (1990). *Robinia pseudoacacia* L. black locust. In: Burns, Russell M.; Honkala, Barbara H., technical coordinators. *Silvics of North America. Volume 2. Hardwoods. Agric. Handb. 654.* Washington, DC: U.S. Department of Agriculture, Forest Service: 755-761.
- Iverson LR, Prasad AM (1998) Predicting Abundance of 80 Tree Species Following Climate Change in the Eastern United States. *Ecol Monogr* 68:465–485
- Kaufman SR, Smouse PE (2001) Comparing indigenous and introduced populations of *Melaleuca quinquenervia* (Cav.) Blake: response of seedlings to water and pH levels. *Oecologia* 127:487–494 . doi: 10.1007/s004420000621
- Keller SR, Taylor DR (2008) History, chance and adaptation during biological invasion: Separating stochastic phenotypic evolution from response to selection. *Ecol Lett* 11:852–866 . doi: 10.1111/j.1461-

0248.2008.01188.x

- Kew Royal Botanic Garden (2015) Seed Information Database. <http://data.kew.org>. Accessed 10 Feb 2015
- Kruschke JK (2015) Doing Bayesian data analysis: a tutorial with R, JAGS, and Stan, Edition 2. Academic Press, Boston
- Lamarque LJ, Lortie CJ, Porté AJ, Delzon S (2014) Genetic differentiation and phenotypic plasticity in life-history traits between native and introduced populations of invasive maple trees. *Biol Invasions*. doi: 10.1007/s10530-014-0781-3
- Lee CE (2002) Evolutionary genetics of invasive species. *Trends Ecol Evol* 17:386–391 . doi: 10.1016/S0169-5347(02)02554-5
- Leiblein-Wild MC, Kaviani R, Tackenberg O (2014) Germination and seedling frost tolerance differ between the native and invasive range in common ragweed. *Oecologia* 174:739–750 . doi: 10.1007/s00442-013-2813-6
- Leimu R, Fischer M (2008) A meta-analysis of local adaptation in plants. *PLoS One* 3:1–8 . doi: 10.1371/journal.pone.0004010
- Leinonen T, McCairns RJS, O’Hara RB, Merilä J (2013) Qst-Fst comparisons: evolutionary and ecological insights from genomic heterogeneity. *Nat Rev Genet* 14:179–90 . doi: 10.1038/nrg3395
- Leinonen T, O’Hara RB, Cano JM, Merilä J (2008) Comparative studies of quantitative trait and neutral marker divergence: A meta-analysis. *J Evol Biol* 21:1–17 . doi: 10.1111/j.1420-9101.2007.01445.x
- Linhart YB, Grant MC (1996) Evolutionary Significance of Local Genetic Differentiation in Plants. *Annu Rev Ecol Syst* 27:237–277 . doi: 10.1146/annurev.ecolsys.27.1.237
- Monty A, Lebeau J, Meerts P, Mahy G (2009). An explicit test for the contribution of environmental maternal effects to rapid clinal differentiation in an invasive plant. *J. Evol. Biol.* 22:917-926. <https://doi.org/10.1111/j.1420-9101.2009.01728.x>
- Maron JL, Vilà M, Bommarco R, et al (2004) Rapid evolution of an invasive plant. *Ecol Monogr* 74:261–280 . doi: 10.1890/03-4027
- Monty A, Bizoux J-P, Escarré J, Mahy G (2013) Rapid Plant Invasion in Distinct Climates Involves Different Sources of Phenotypic Variation. *PLoS One* 8:e55627 . doi: 10.1371/journal.pone.0055627
- O’Hara RB, Merilä J (2005) Bias and precision in QST estimates: Problems and some solutions. *Genetics* 171:1331–1339 . doi: 10.1534/genetics.105.044545
- Oduor AMO, Leimu R, van Kleunen M (2016) Invasive plant species are locally adapted just as frequently and at least as strongly as native plant species. *J Ecol* 104:957–968 . doi: 10.1111/1365-2745.12578



- Palacio-López K, Gianoli E (2011) Invasive plants do not display greater phenotypic plasticity than their native or non-invasive counterparts: A meta-analysis. *Oikos* 120:1393–1401 . doi: 10.1111/j.1600-0706.2010.19114.x
- Pigliucci M (2006) Phenotypic plasticity and evolution by genetic assimilation. *J Exp Biol* 209:2362–2367 . doi: 10.1242/jeb.02070
- Plummer M (2005) JAGS: just another Gibbs sampler. In: Proceedings of the 3rd International Workshop on Distributed Statistical Computing (DSC 2003)
- Postma FM, Ågren J (2016) Early life stages contribute strongly to local adaptation in *Arabidopsis thaliana*. *Proc Natl Acad Sci* 113:7590–7595 . doi: 10.1073/pnas.1606303113
- Prentice IC, Cramer W, Harrison S., et al (1992) A global biome model based on plant physiology and sominance, soil properties and climate. *J Biogeogr* 19:117–134 . doi: 10.2307/2845499
- R Development Core Team (2013) R Software. R A Lang. Environ. Stat. Comput.
- Richards CL, Bossdorf O, Muth NZ, et al (2006) Jack of all trades, master of some? On the role of phenotypic plasticity in plant invasions. *Ecol Lett* 9:981–993 . doi: 10.1111/j.1461-0248.2006.00950.x
- Richardson DM, Rejmánek M (2011) Trees and shrubs as invasive alien species - a global review. *Divers Distrib* 17:788–809 . doi: 10.1111/j.1472-4642.2011.00782.x
- Santiso X, López L, Gilbert KJ, et al (2015) Patterns of genetic variation within and among populations in *Arbutus unedo* and its relation with selection and evolvability. *Perspect Plant Ecol Evol Syst* 1–8 . doi: 10.1016/j.ppees.2015.02.006
- Spitze K (1993) Population structure in *Daphnia obtusa*: Quantitative genetic and allozymic variation. *Genetics* 135:367–374 . doi: 10.1007/s10709-011-9564-2
- Su Y, Yajima M (2012) R2jags: A Package for Running jags from R. <http://CRAN.R-project.org/package=R2jags>
- Takezaki N, Nei M (1996) Genetic distances and the setting of conservation priorities. *Biol Conserv* 75:311 . doi: 10.1016/0006-3207(96)85988-X
- Valladares F, Sanchez-Gomez D, Zavala M a. (2006) Quantitative estimation of phenotypic plasticity: Bridging the gap between the evolutionary concept and its ecological applications. *J Ecol* 94:1103–1116 . doi: 10.1111/j.1365-2745.2006.01176.x
- Walck JL, Hidayati SN, Dixon KW, et al (2011) Climate change and plant regeneration from seed. *Glob Chang Biol* 17:2145–2161 . doi: 10.1111/j.1365-2486.2010.02368.x

Whitlock MC (2008) Evolutionary inference from QST. *Mol Ecol* 17:1885–1896 . doi: 10.1111/j.1365-294X.2008.03712.x

Whitlock MC, Guillaume F (2009) Testing for spatially divergent selection: Comparing QST to FST. *Genetics* 183:1055–1063 . doi: 10.1534/genetics.108.099812

Woodward FI (1987) *Climate and plant distribution*. Cambridge University Press, Cambridge [Cambridgeshire] ; New York

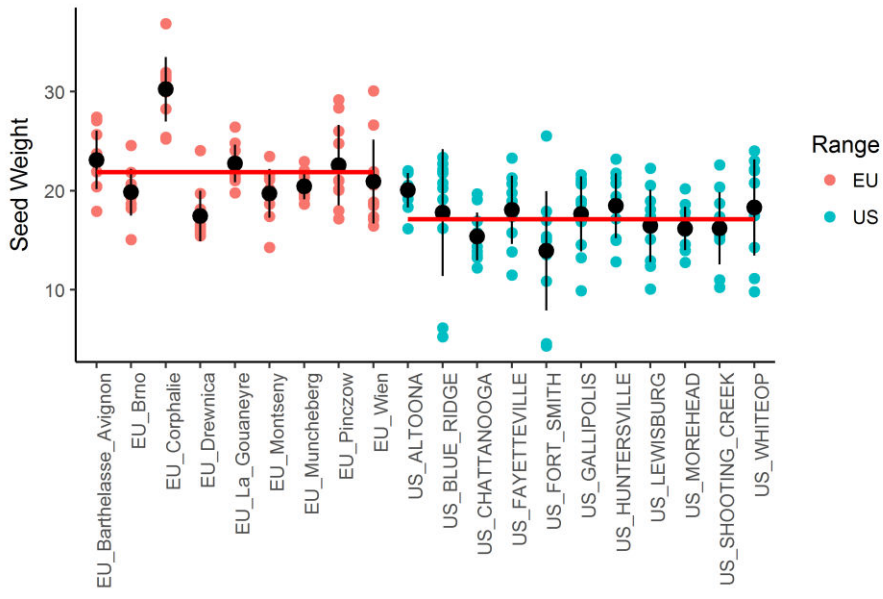
Xu C, Tang S, Fatemi M, et al (2015) Population structure and genetic diversity of invasive *Phyla canescens*: implications for the evolutionary potential. *Ecosphere* 6:Art162 . doi: 10.1890/ES14-00374.1

Appendix 1 Information about the sampled populations. X and Y (longitude and latitude respectively) are for GPS coordinates in WGS84 system. Altitude and annual mean temperature are extracted from worldclim database.

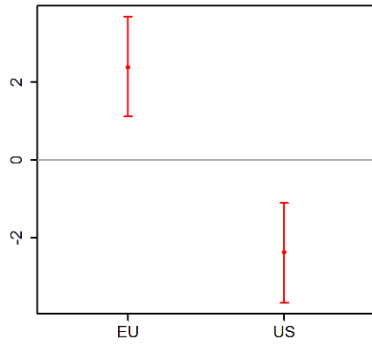
Range	Pop	Name	X (WGS84)	Y (WGS84)	Altitude (m)	Annual mean T (°C)	Seed weight mean (mg)
EU	Pop1	Corphalie	5.2588	50.5389	100	9.7	29.34669613
EU	Pop2	Barthelasse_Avigno	4.81781	43.9646	19	13.8	23.10009782
EU	Pop3	Montseny	2.51236	41.8306	403	14	19.73035994
EU	Pop4	Munchenberg	14.04553	52.5595	85	8.8	20.41925351
EU	Pop5	Brno	16.5178	49.0423	223	9	19.85603244
EU	Pop6	Wien	16.47286	48.2515	158	10	20.91338314
EU	Pop7	Pinczow	20.7022	50.2647	203	8.2	22.57100531
EU	Pop8	Drewnica	21.2508	52.2528	105	7.7	17.44892855
EU	Pop9	La_Gouaneyre	-0.27251	44.376	95	12.4	22.73042193
US	Pop1	ALTOONA	-78.3832	40.4885	465	8.9	19.88850155
US	Pop2	HUNTERSVILLE	-79.8799	38.2785	785	9.5	18.4811567
US	Pop3	LEWISBURG	-80.3812	37.7825	608	10.6	16.4445876
US	Pop4	WHITEOP	-81.6562	36.7694	679	11.4	18.30203788
US	Pop5	GALLIPOLIS	-82.1834	38.8615	230	11.9	17.65908571
US	Pop6	MOREHEAD	-83.4663	38.0909	237	12.2	15.55373141
US	Pop7	BLUE_RIDGE	-82.6722	35.4567	1033	10.9	18.40286273
US	Pop8	SHOOTING_CREEK	-83.6284	35.0546	761	12.1	16.20277805
US	Pop9	CHATTANOOGA	-85.783	35.1198	309	14.3	15.3674099
US	Pop10	FORT_SMITH	-94.2895	35.3431	121	16	13.92534984
US	Pop11	FAYETTEVILLE	-94.205	36.0711	387	14.1	18.06440674

Appendix 2 A: Seed weight differentiation among ranges and populations, black circles represent mean population seed weight averaged over families associated with standard deviations (vertical black lines), each mean family seed weight is plotted with a colored circle (red: European families and blue US families), horizontal red lines indicate the mean seed weight per range. B and C: Results of the seed weight analysis using the Bayesian model testing for B: range effects, C: population effects. Each dot represents the mean value of the estimate, and bars represent the Bayesian 95% credible interval. Effects significantly different from 0 are drawn in red. Both range and population effects are significant.

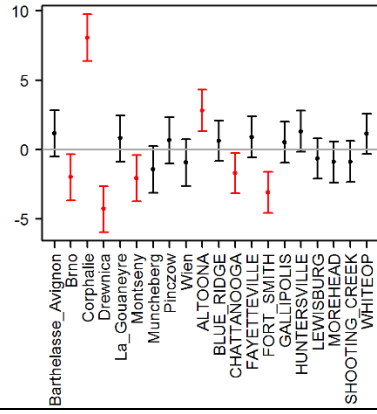
A



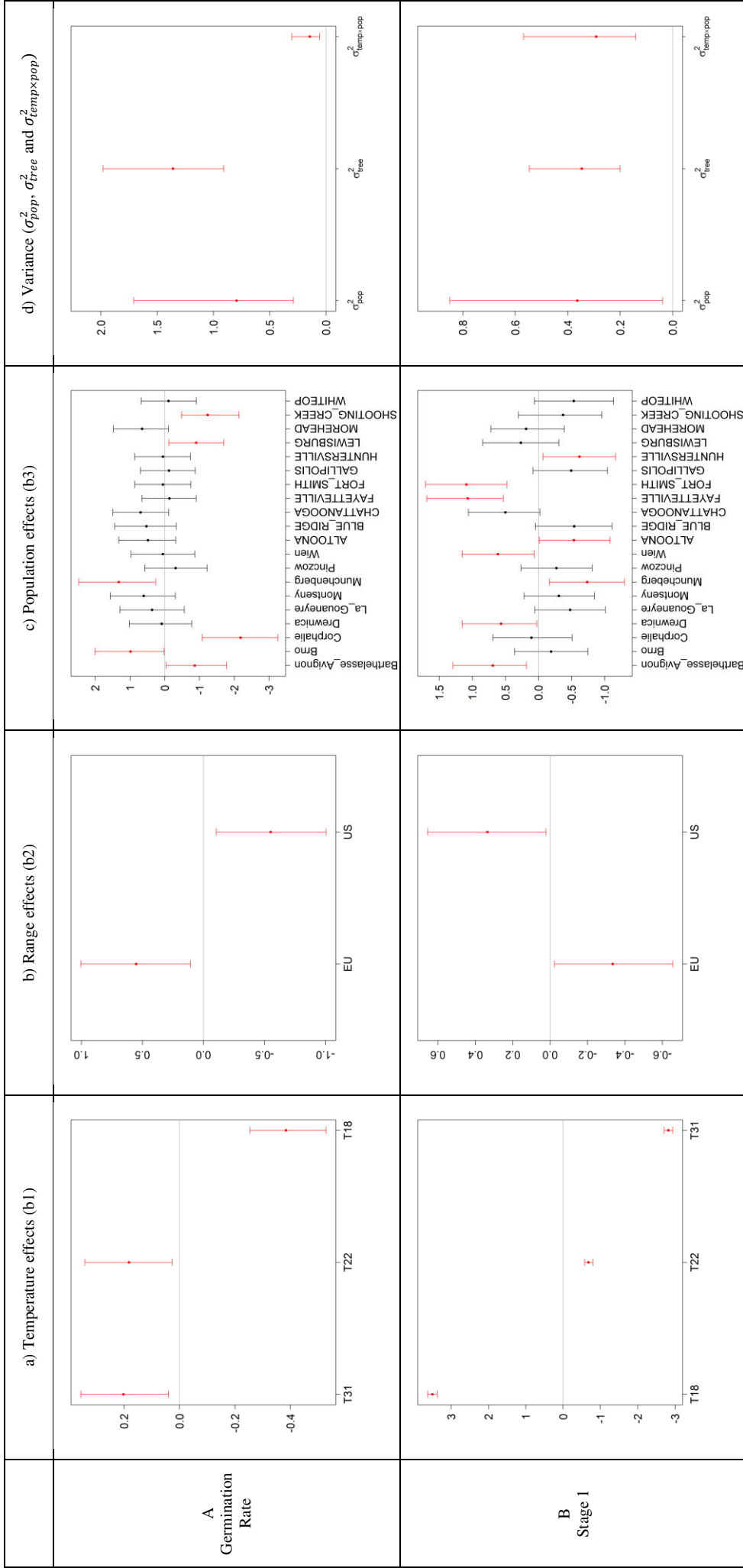
B



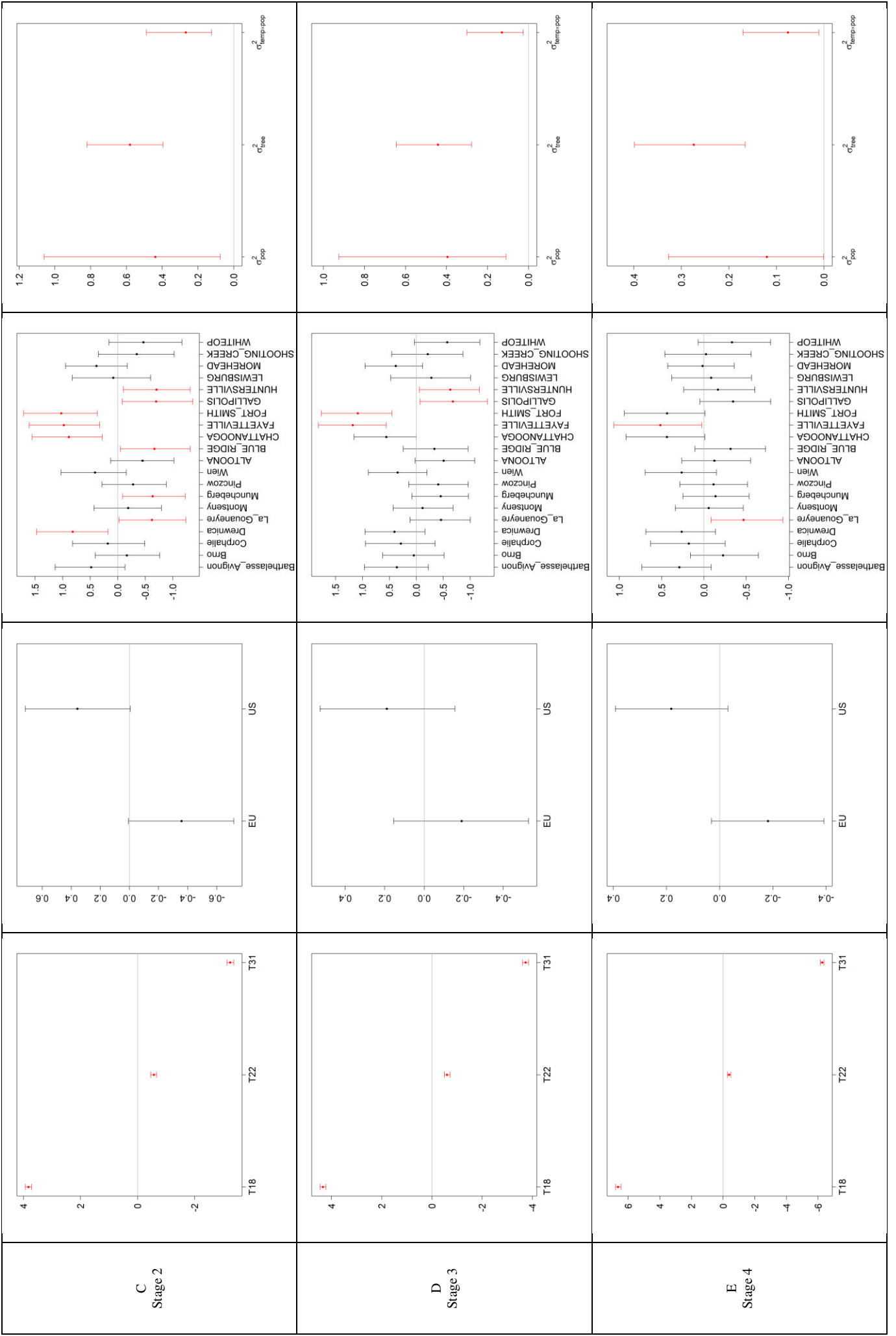
C

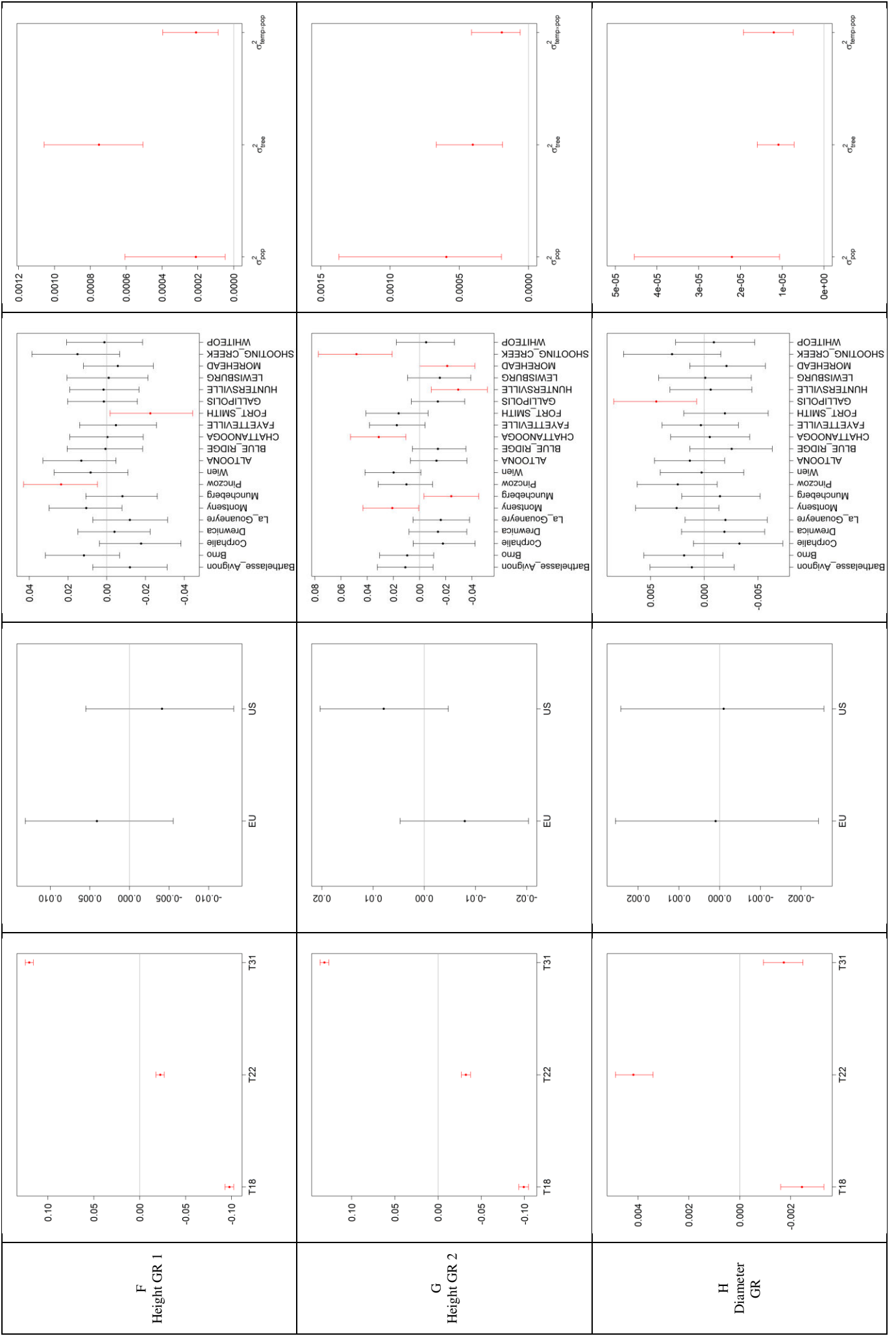


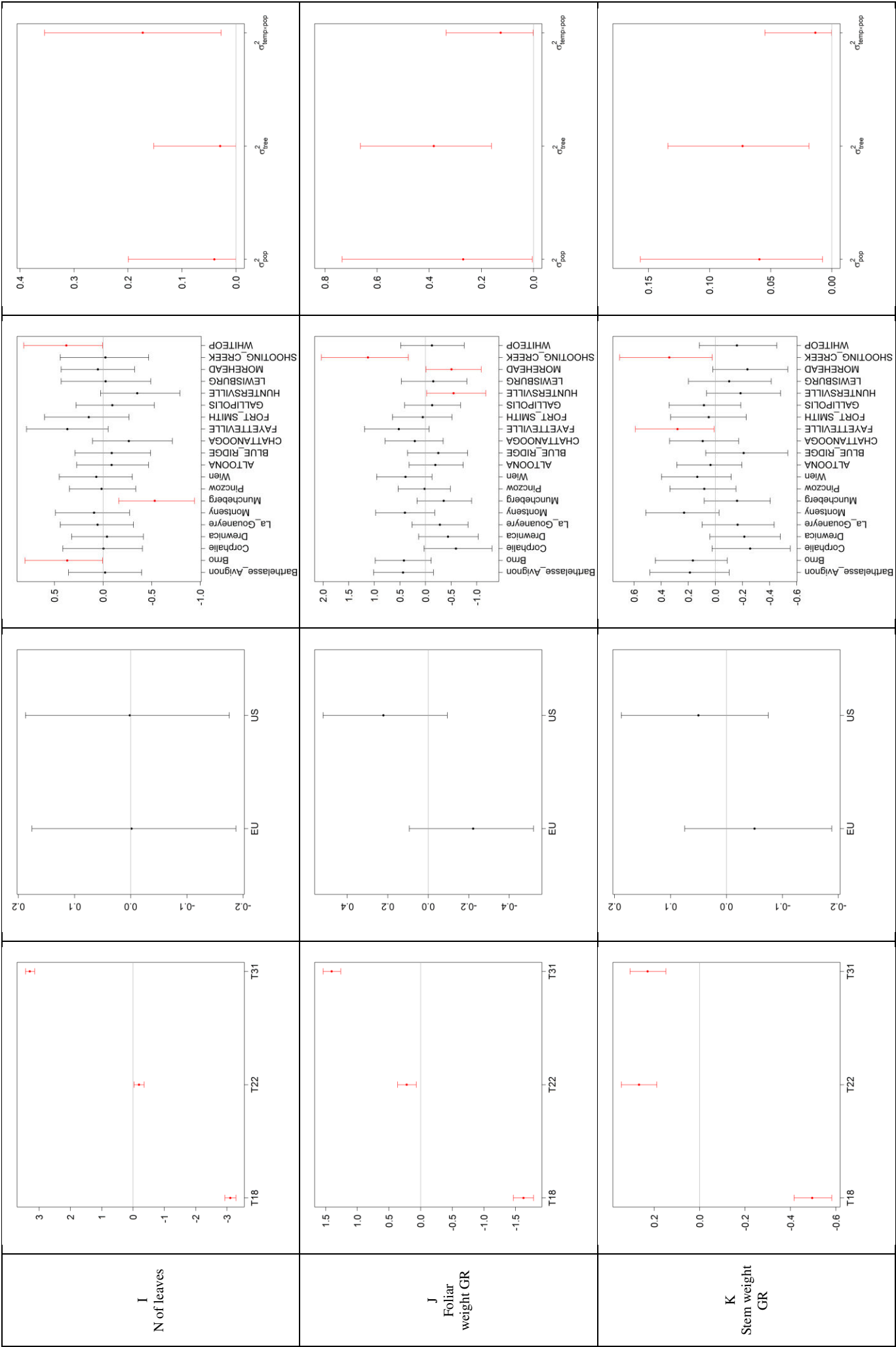
Appendix 3 Results of the Bayesian analyses for all phenotypic traits using the full between chamber model presented in the main document. A, Germination rate; B–E, phenological stages 1 to 4; F–G, height Growth Rate (GR) 1 and 2; H, diameter GR; I, Number of leaves; J–L, leaf, stem and total weight GR; M, PSII yield; with a: temperature effects (b1), b: range effect (b2), c: population effects (b3), d: estimated population, tree and temperature × population interaction variance. Each dot represents the mean value of the estimated effect, and bars represent the 95% credible interval over the mean value. Effects significantly different from 0 are drawn in red.

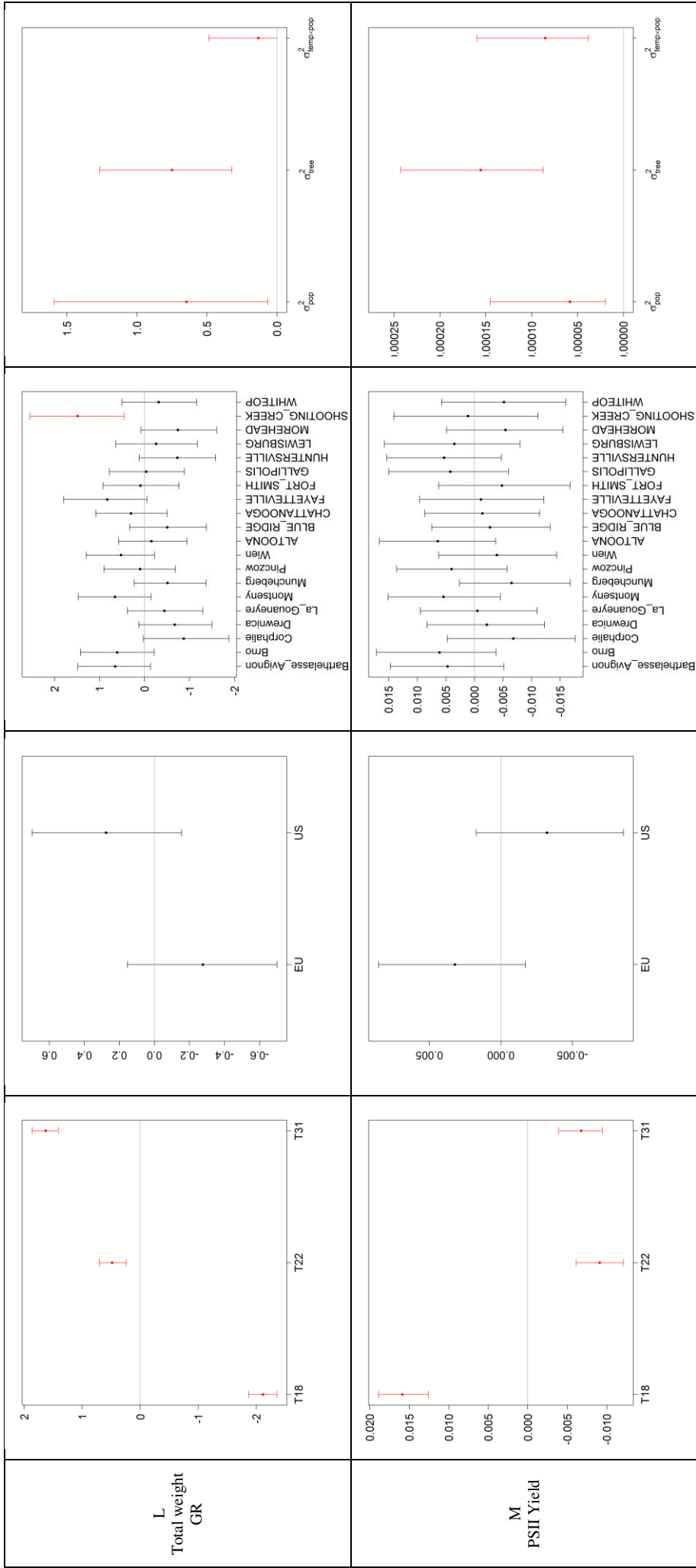




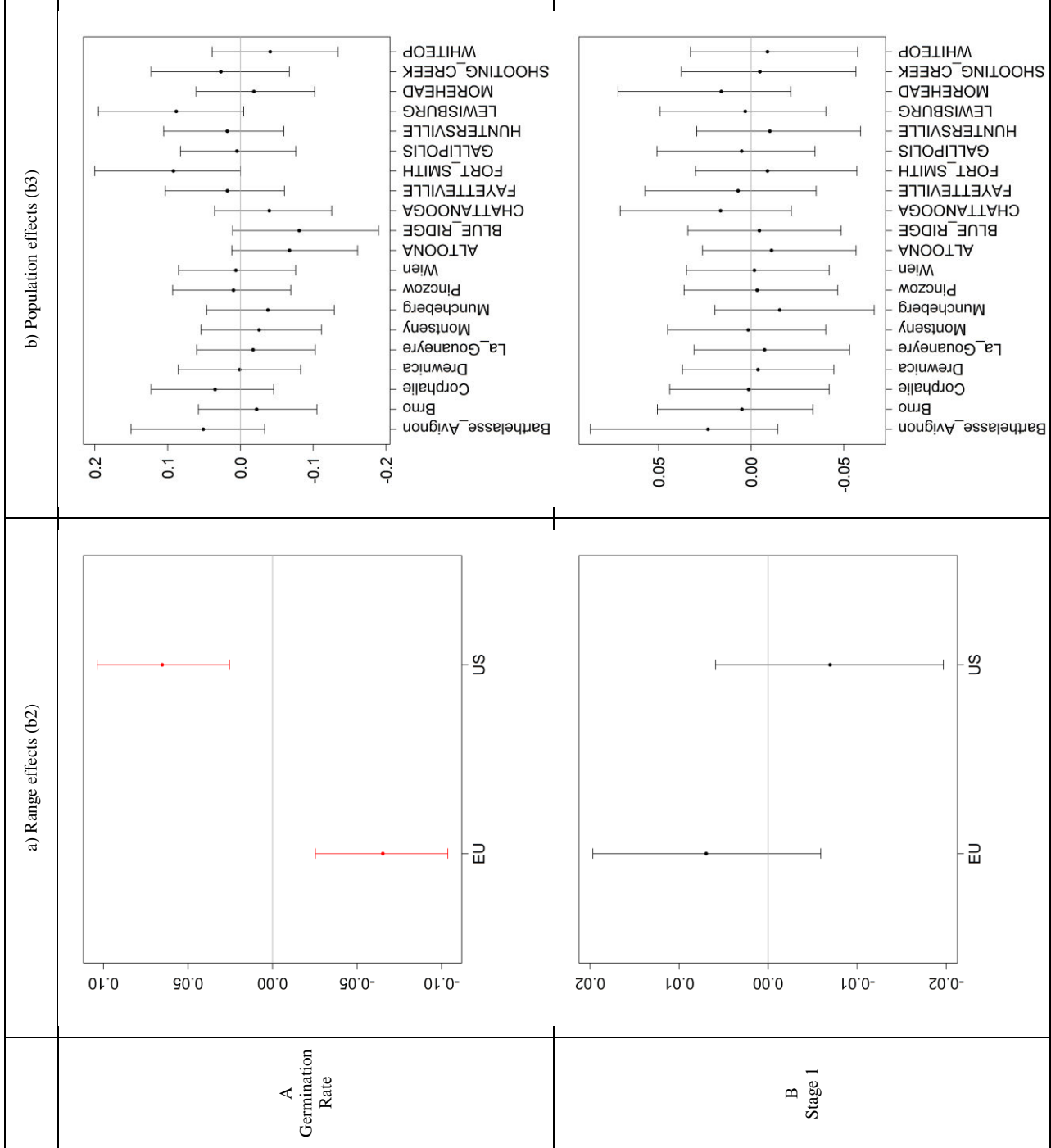


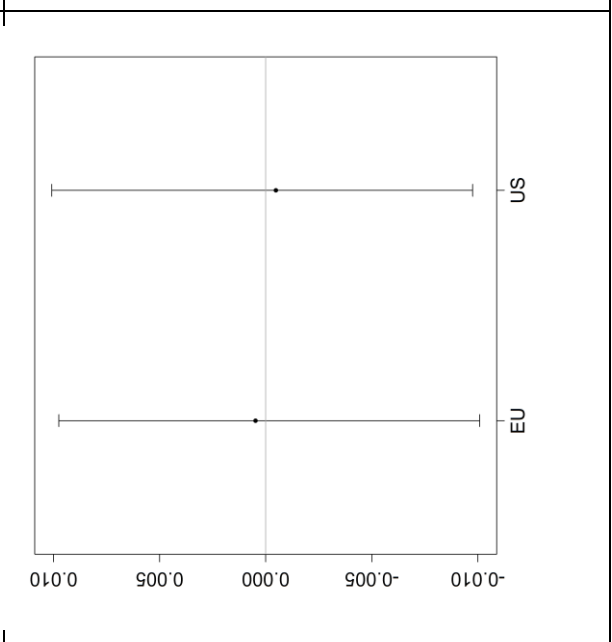
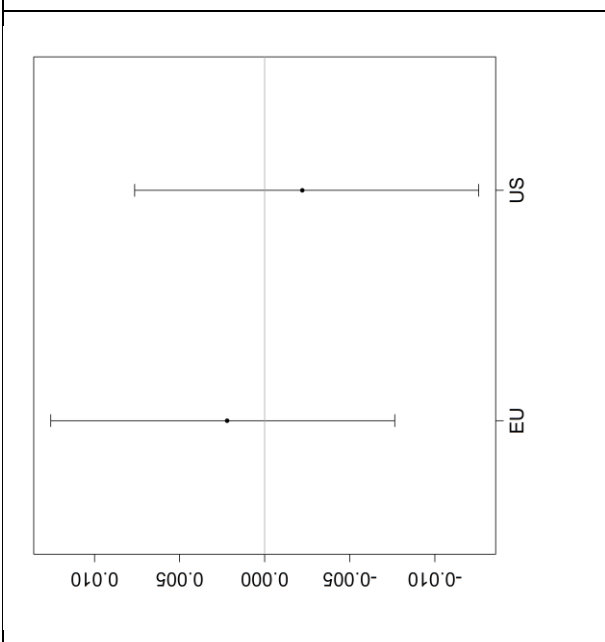
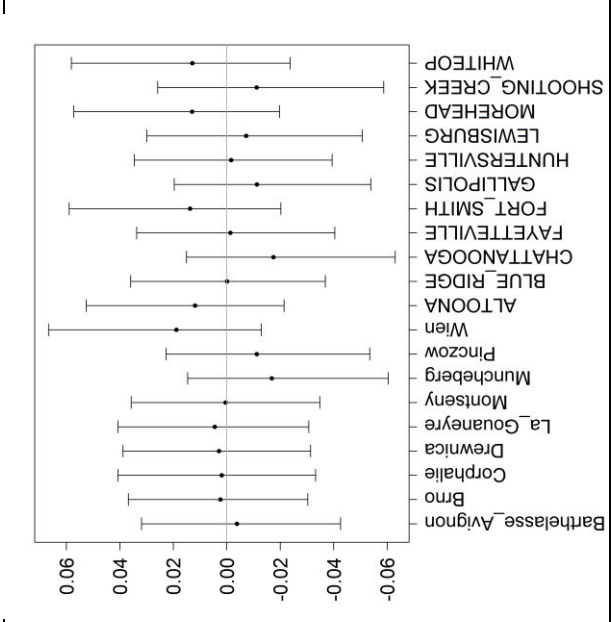
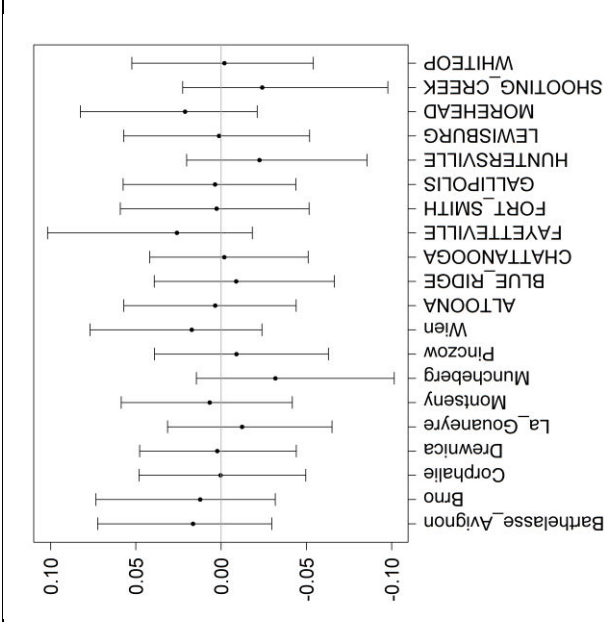






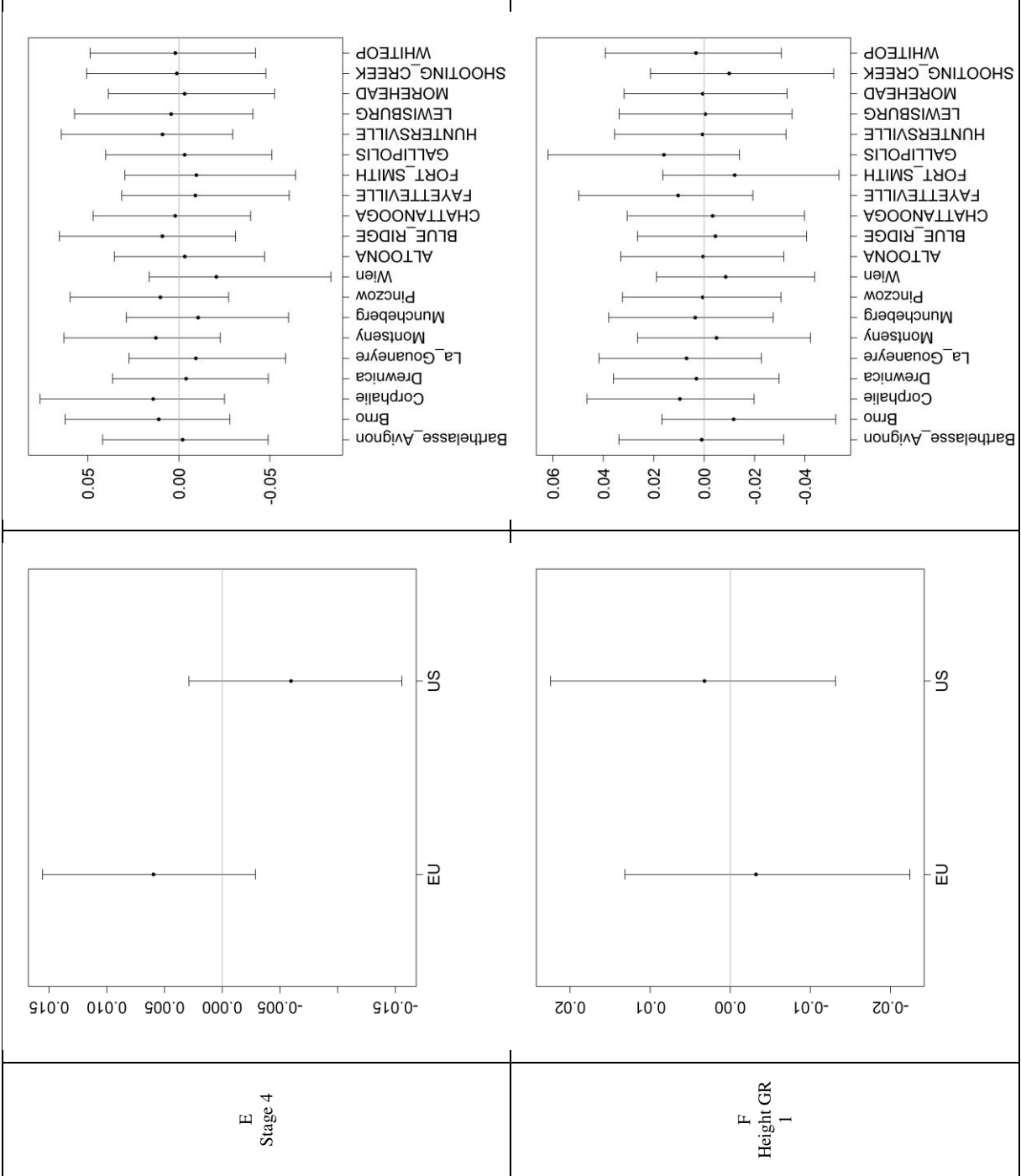
Appendix 4 Results of the Bayesian analyses for the RDPI. A, Germination rate; B–E, phenological stages 1 to 4; F–G, height Growth Rate (GR) 1 and 2; H, diameter GR; I, Number of leaves; J–L, leaf, stem and total weight GR; M, PSII yield; with a: temperature effects (b1), b: range effect (b2), c: population effects (b3), d: estimated population, tree and temperature  $\times$  population interaction variance. Each dot represents the mean value of the estimated effect, and bars represent the 95% credible interval over the mean value. Effects significantly different from 0 are drawn in red.



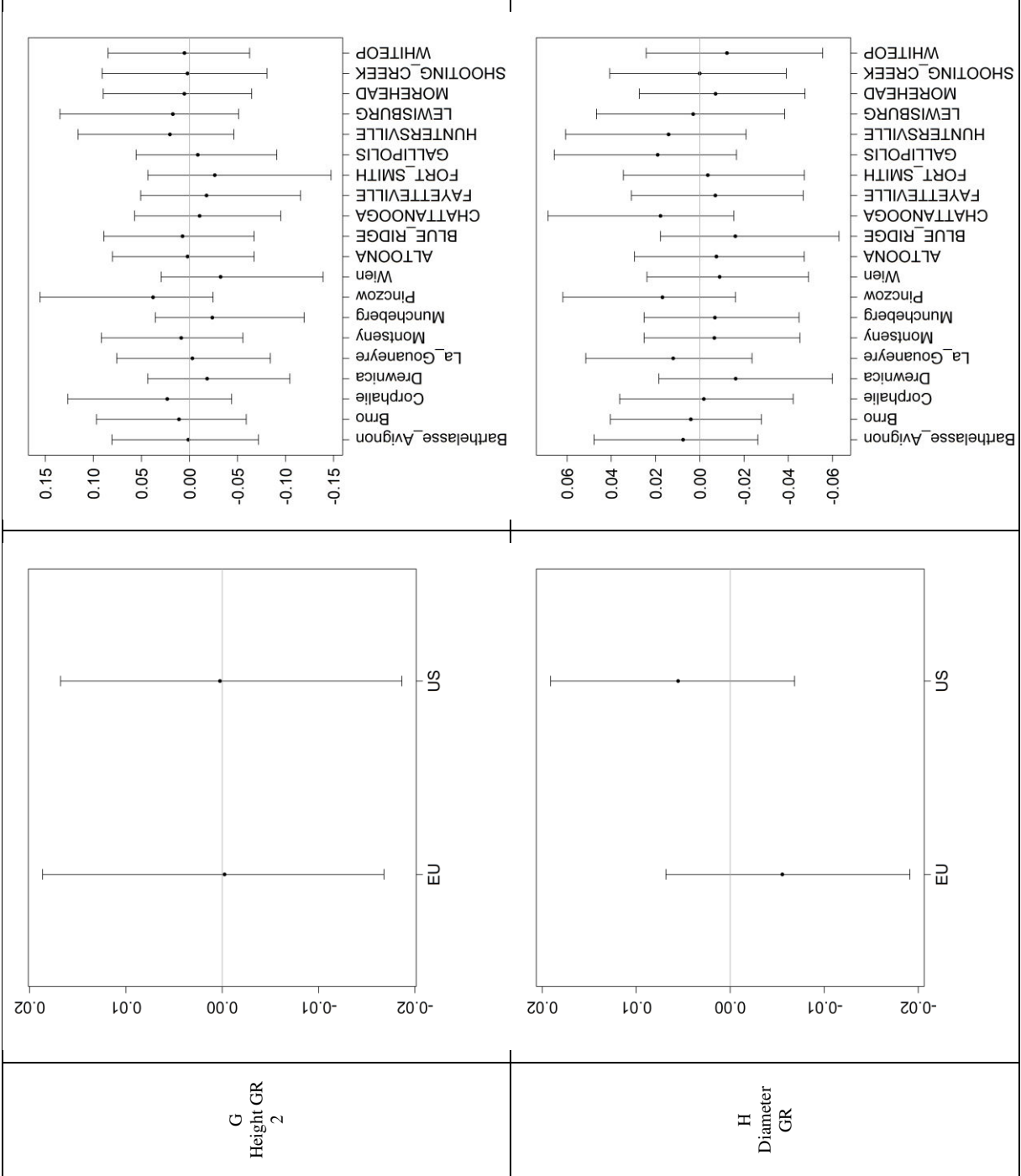


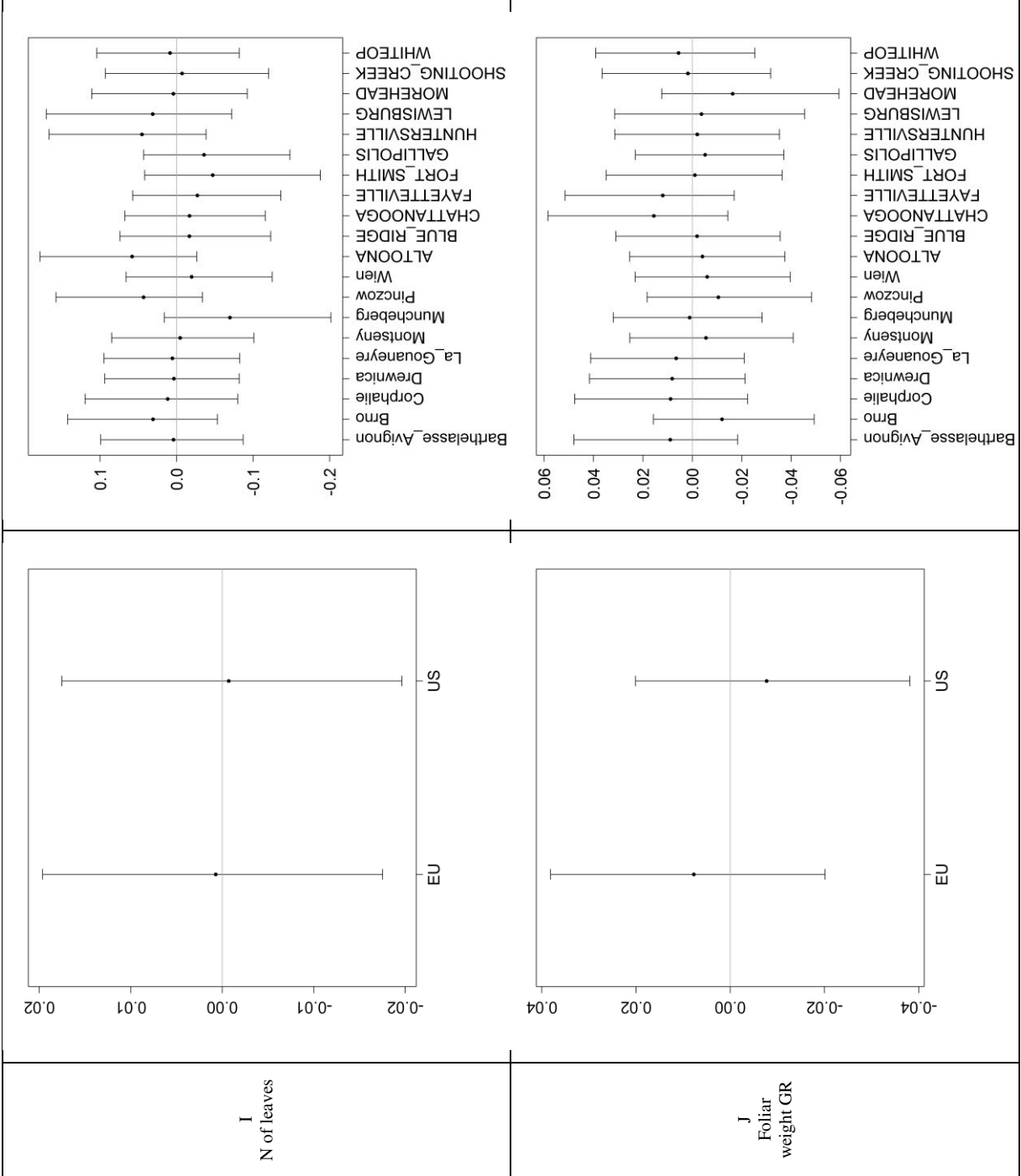
C Stage 2

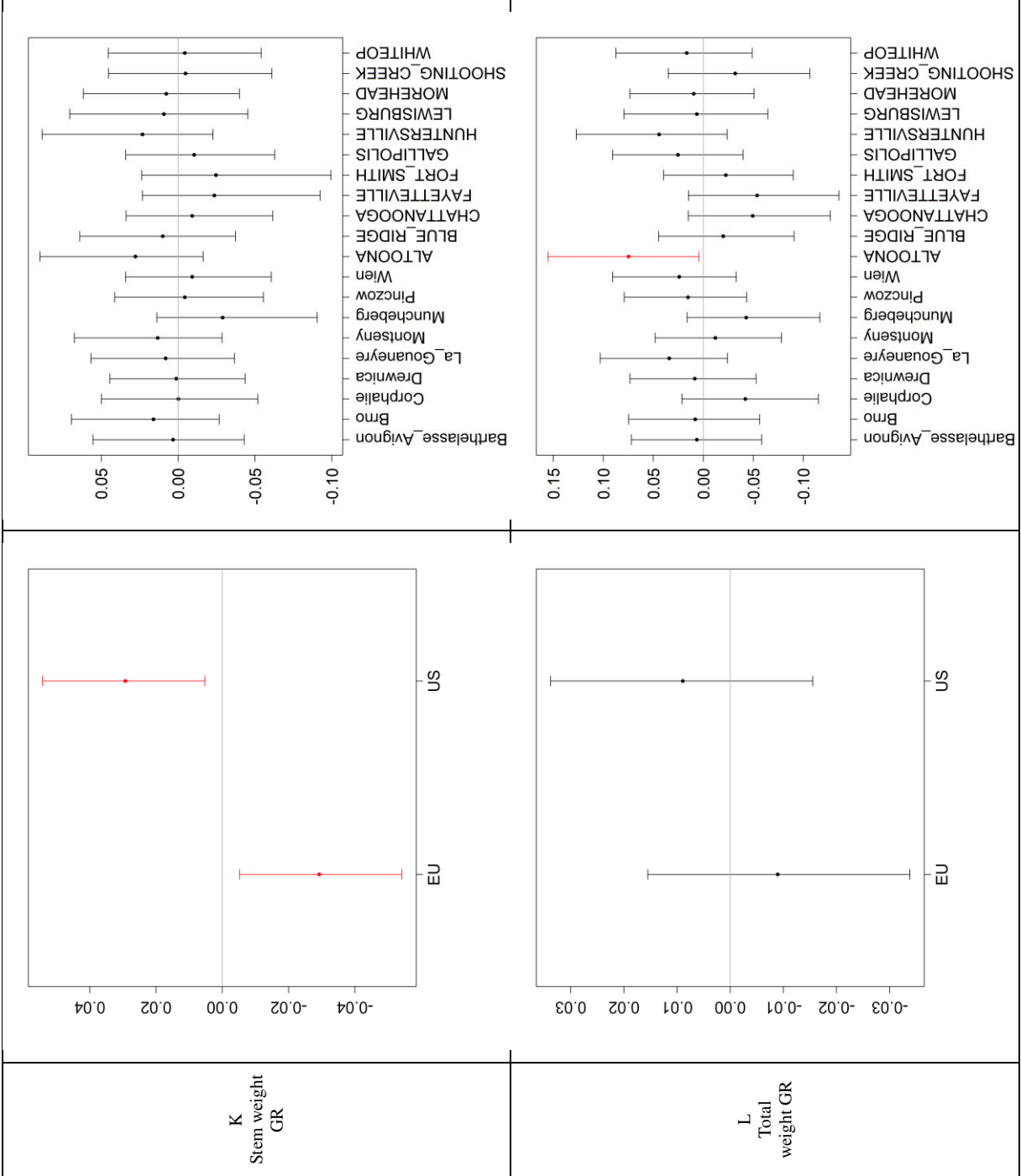
D Stage 3



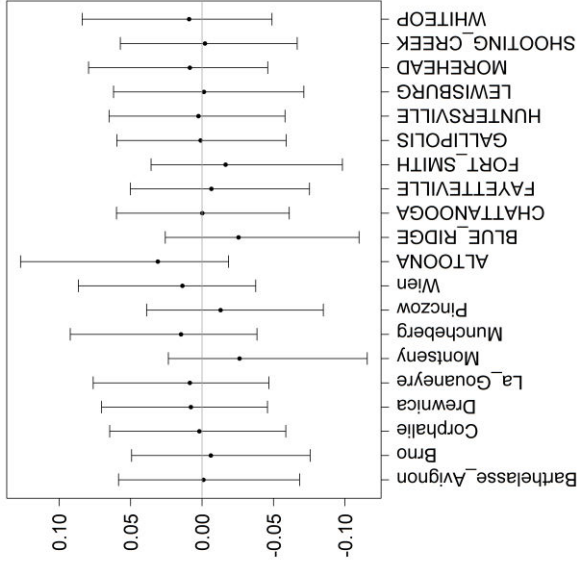
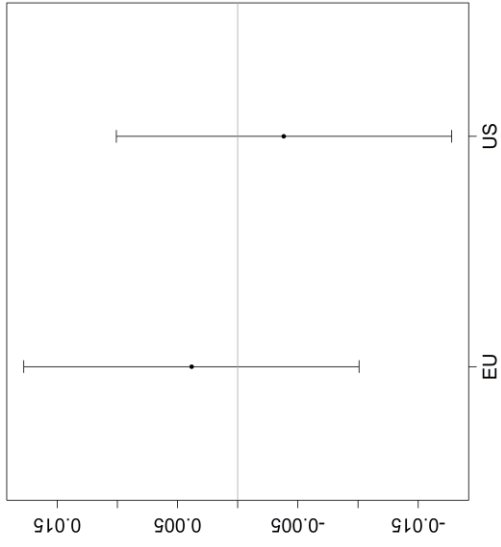






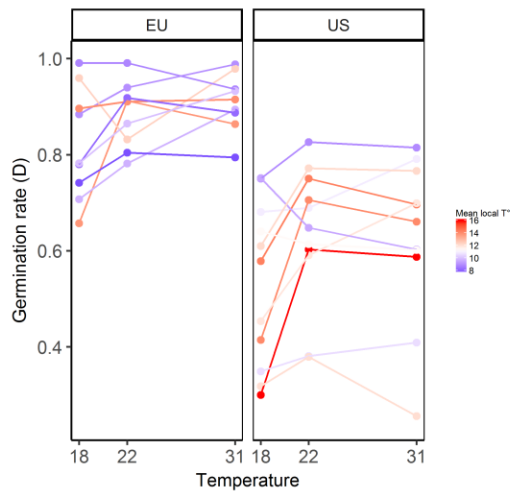


M  
PSII Yield

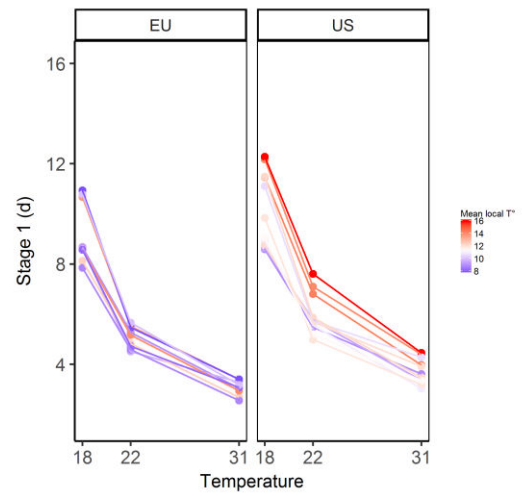


Appendix 5 Reaction norms for all studied traits. A, Germination rate; B–E, phenological stages 1 to 4; F–G, height Growth Rate (GR) 1 and 2; H, diameter GR; I, Number of leaves; J–L, leaf, stem and total weight GR; M, PSII yield. Each line represents one population (averaged values over all individuals), the color indicating the mean annual temperature at the sampling location (from the coldest in blue, to the warmest in red).

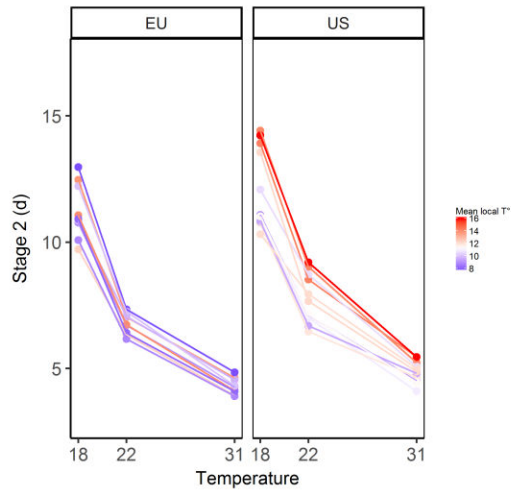
A. Germination Rate



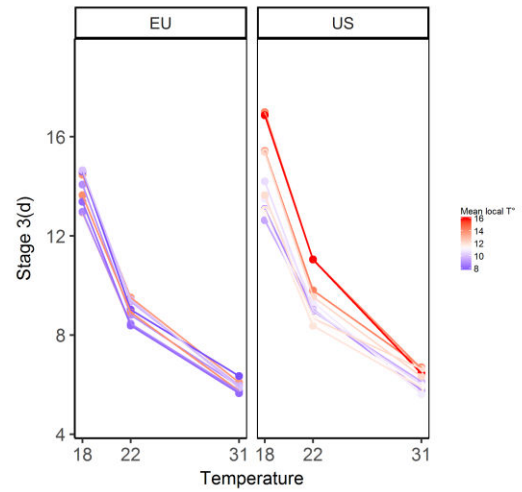
B. Stage 1



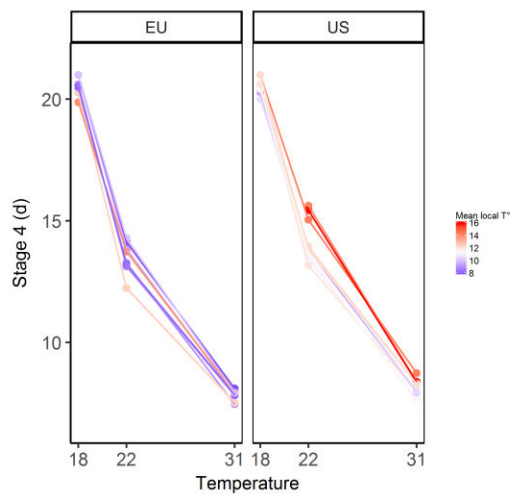
C. Stage 2



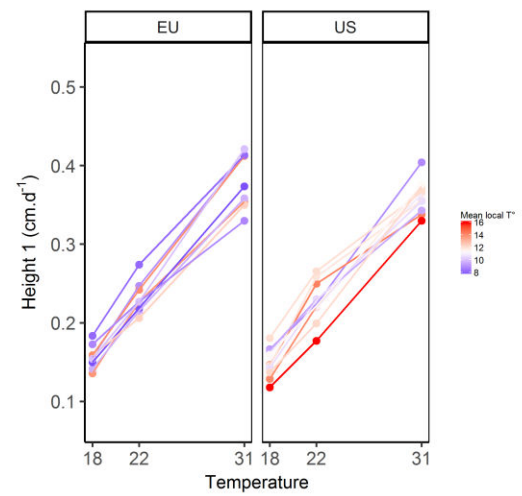
D. Stage 3



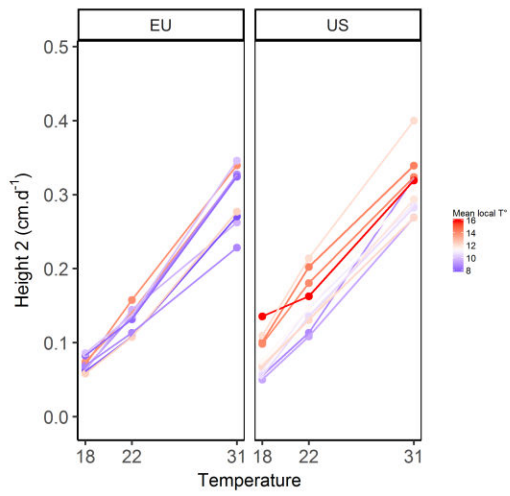
E. Stage 4



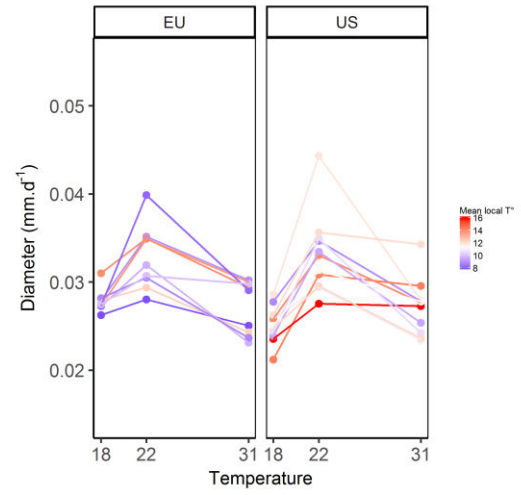
F. Height 1 GR



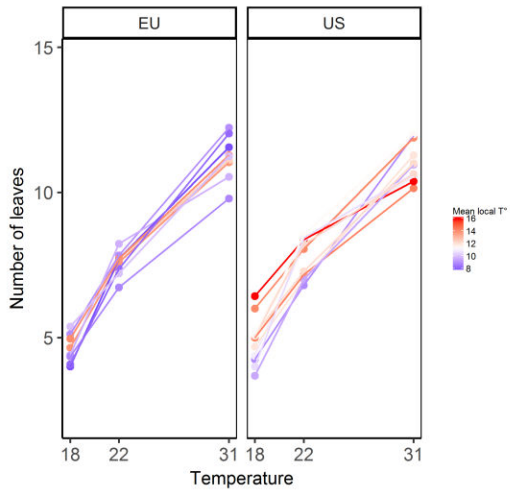
G. Height 2 GR



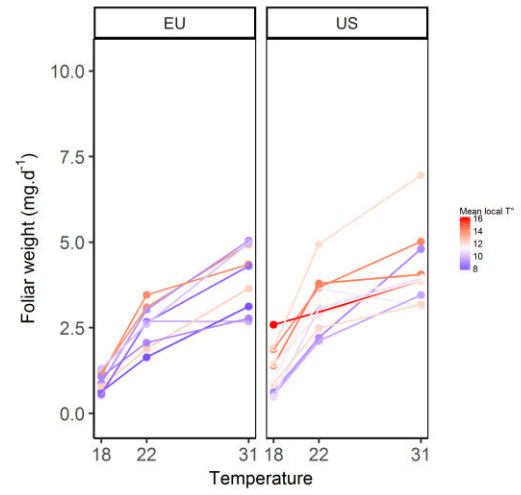
H. Diameter GR



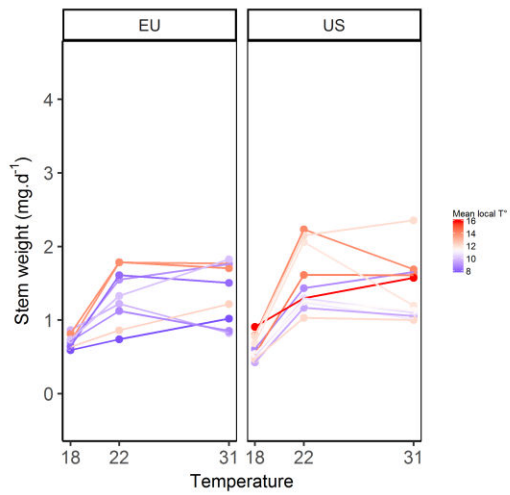
I. Number of leaves



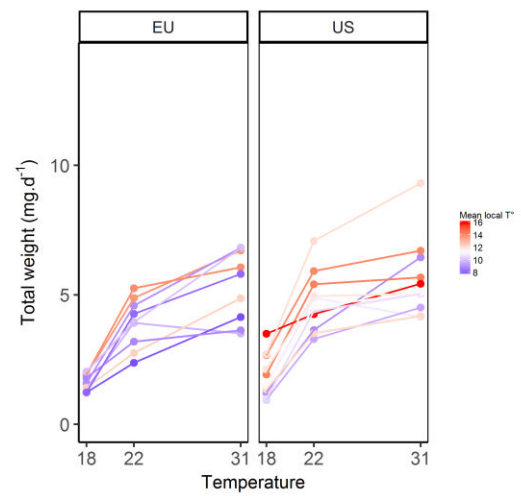
J. Foliar weight GR



K. Stem weight GR



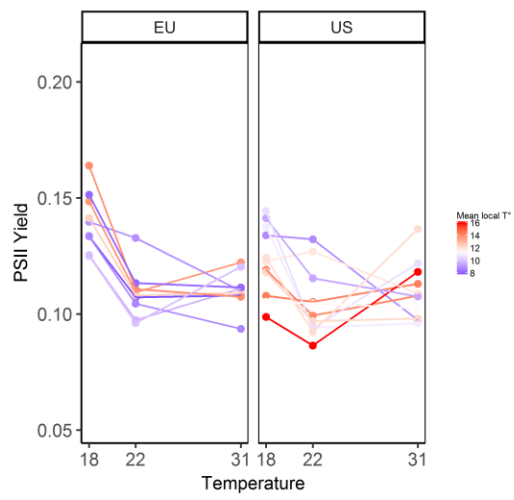
L. Total weight GR



---

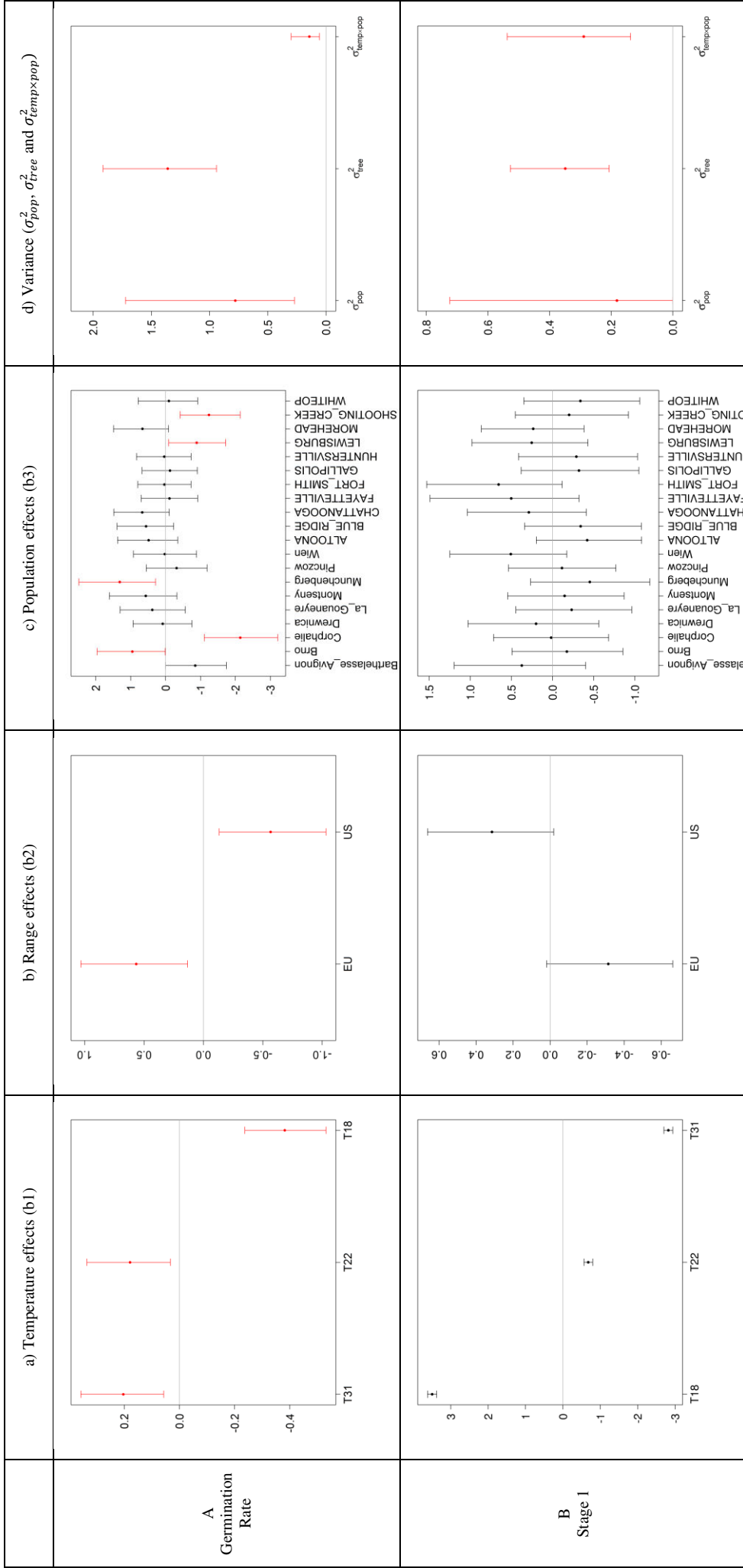
### M. PSII Yield

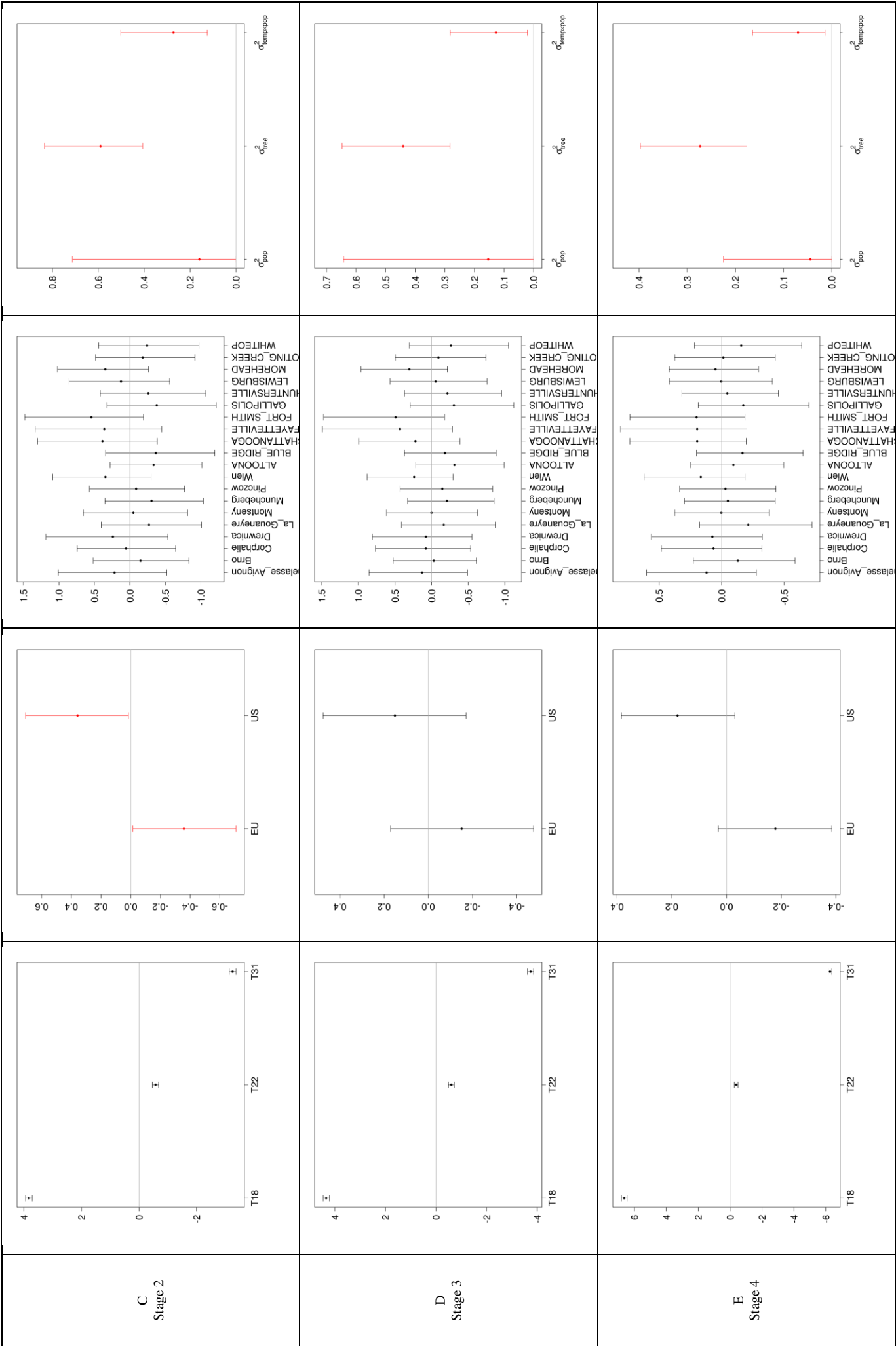
---



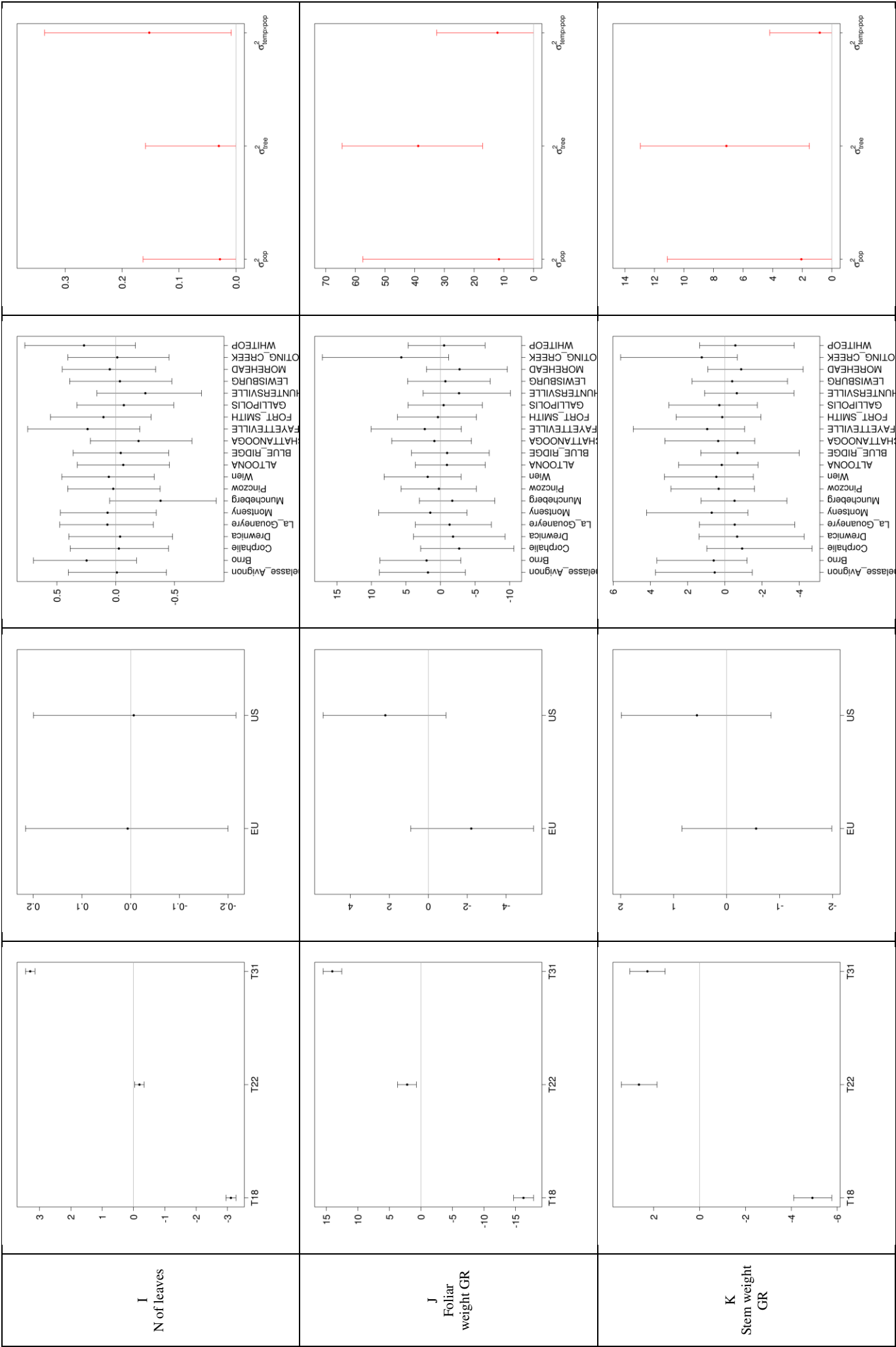


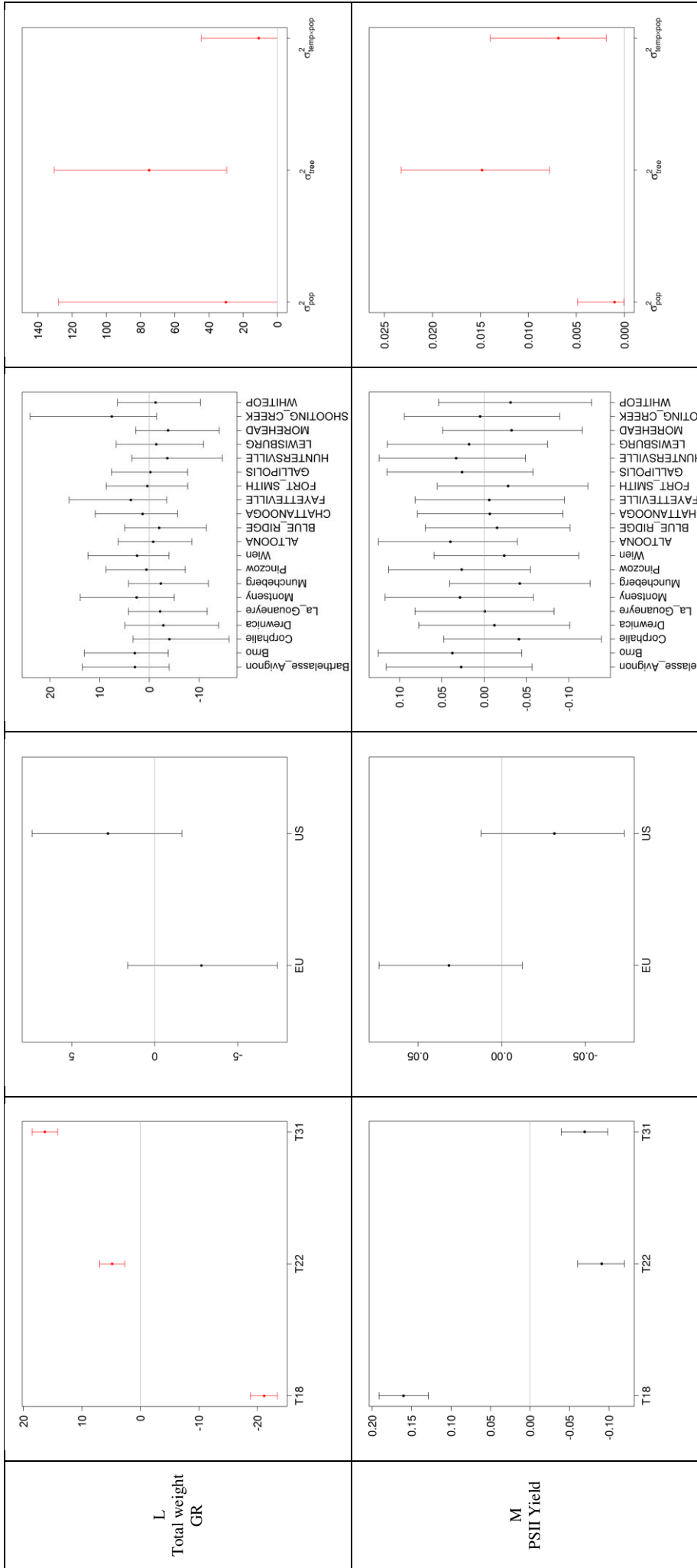
Appendix 6 Results of the Bayesian analyses for all phenotypic traits using the full between chamber model incorporating a matrix of pairwise genetic distance between populations. A, Germination rate; B–E, phenological stages 1 to 4; F–G, height Growth Rate (GR) 1 and 2; H, diameter GR; I, Number of leaves; J–L, leaf, stem and total weight GR; M, PSII yield; with a: temperature effects (b1), b: range effect (b2), c: population effects (b3), d: estimated population, tree and temperature  $\times$  population interaction variance. Each dot represents the mean value of the estimated effect, and bars represent the 95% credible interval over the mean value. Effects significantly different from 0 are drawn in red.









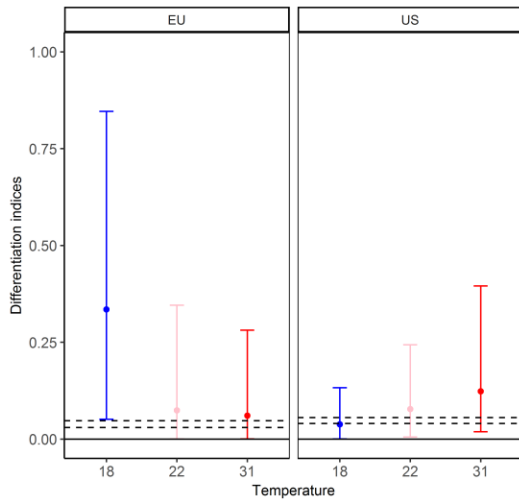


Appendix 7 Differentiation indices calculated for all studied phenotypic traits ( $Q_{ST}$ ) and using SNP markers ( $F_{ST}$ ).  $Q_{ST}$  values (mean and 95% confidence interval) were calculated either on the 9 European populations (left) and 11 US populations (right) per trait and growing chamber (blue circle: 18°C, pink circle: 22°C, red circle: 31°C). The dashed lines represent the  $F_{ST}$  95% confidence interval based on allelic variation in SNP's loci.  $Q_{ST}$  estimations were calculated using Bayesian methods. A, Germination rate; B–E, phenological stages 1 to 4; F–G, height Growth Rate (GR) 1 and 2; H, diameter GR; I, Number of leaves; J–L, leaf, stem and total weight GR; M, PSII yield. GR: Growth Rate.

---

**A. Germination Rate**

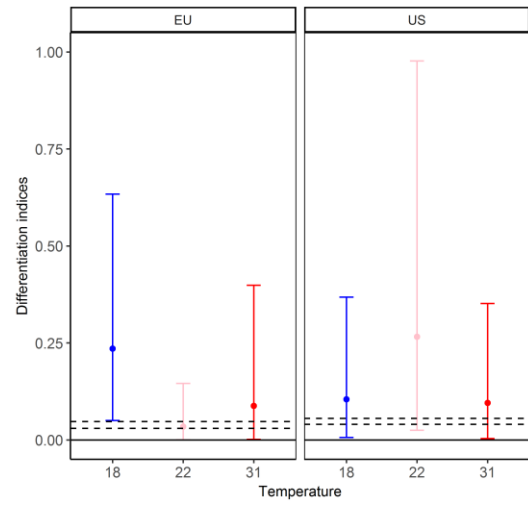
---



---

**B. Stage 1**

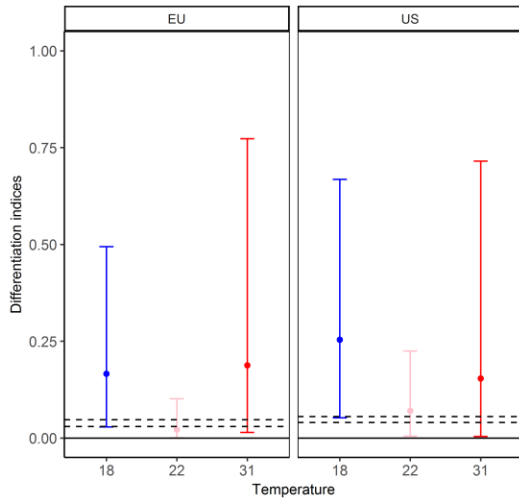
---



---

**C. Stage 2**

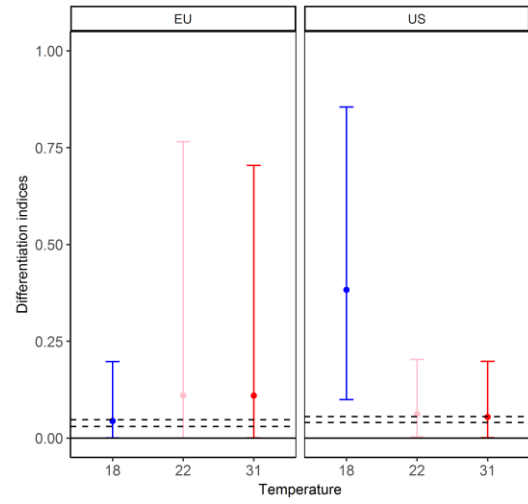
---



---

**D. Stage 3**

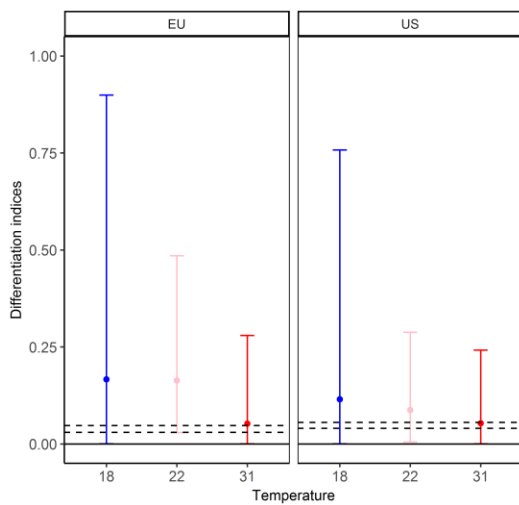
---



---

**E. Stage 4**

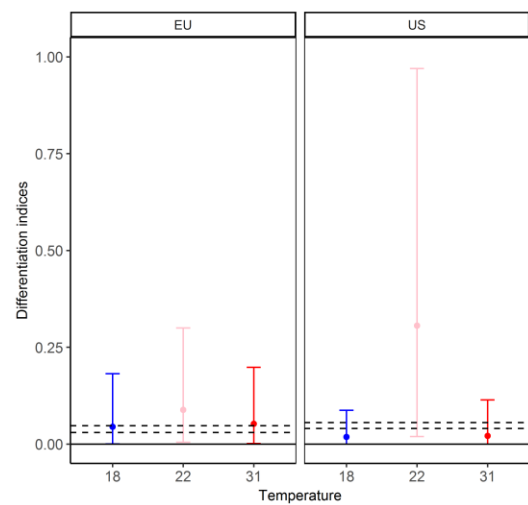
---



---

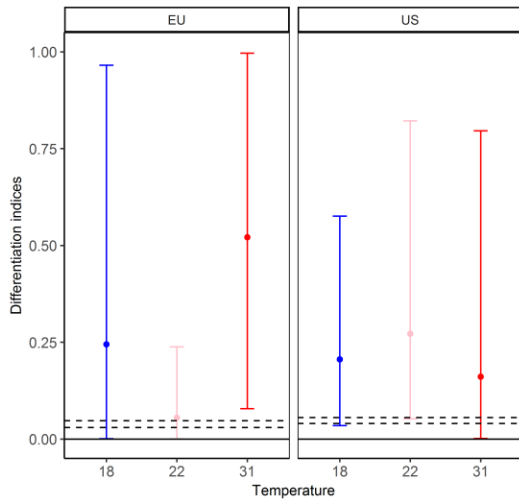
**F. Height 1 GR**

---

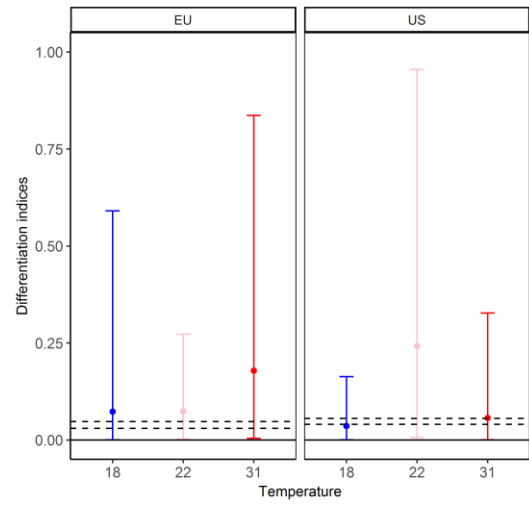




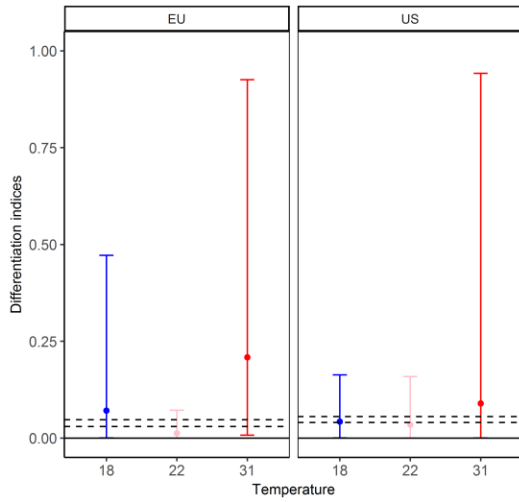
G. Height 2 GR



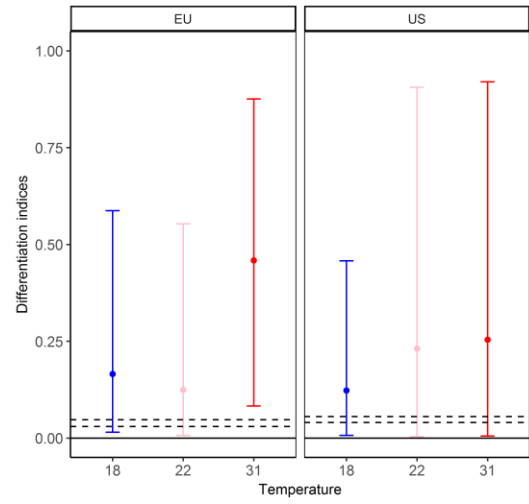
H. Diameter GR



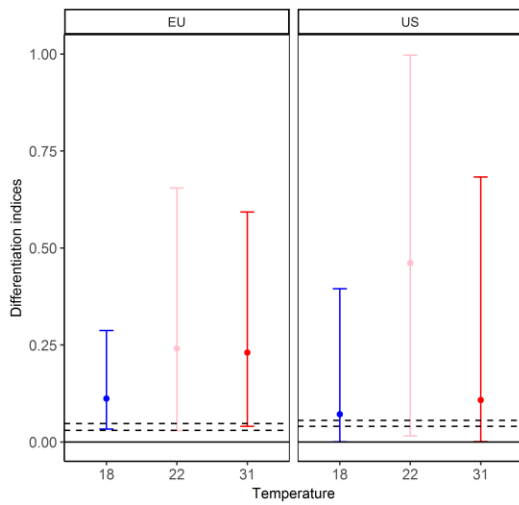
I. Number of leaves



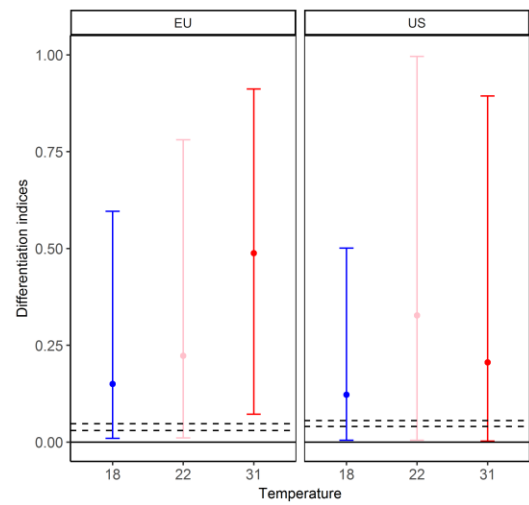
J. Foliar weight GR



K. Stem weight GR



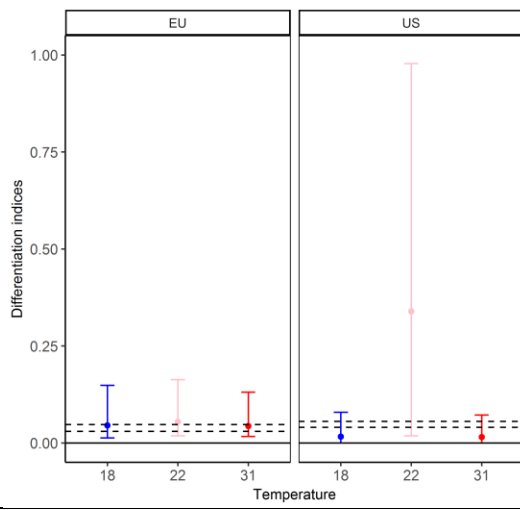
L. Total weight GR



---

## M. PSII Yield

---



## **Discussion générale**

# Discussion générale

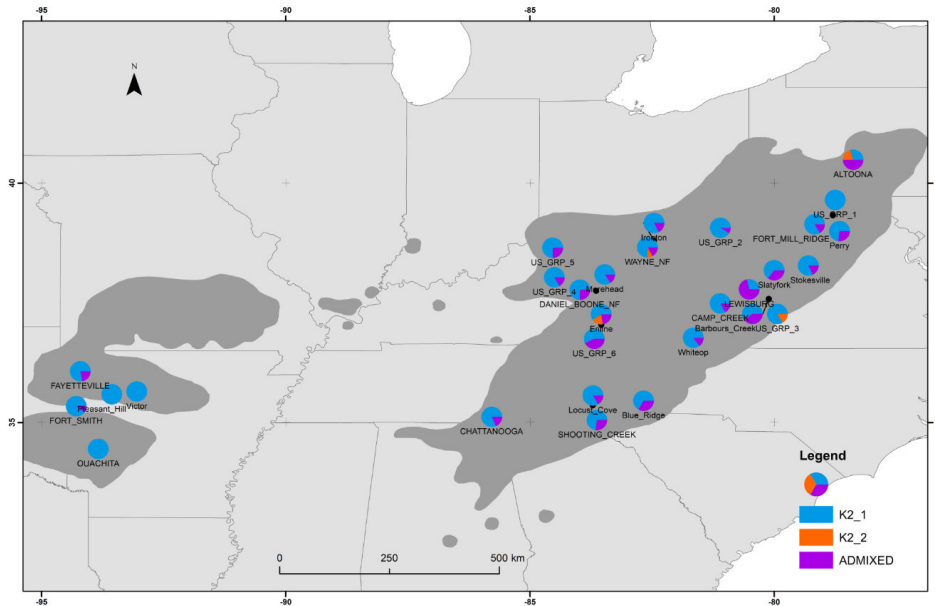
## 3. Discussion

### 3.1. D'où vient le robinier européen ou comment retracer le scénario de l'introduction ?

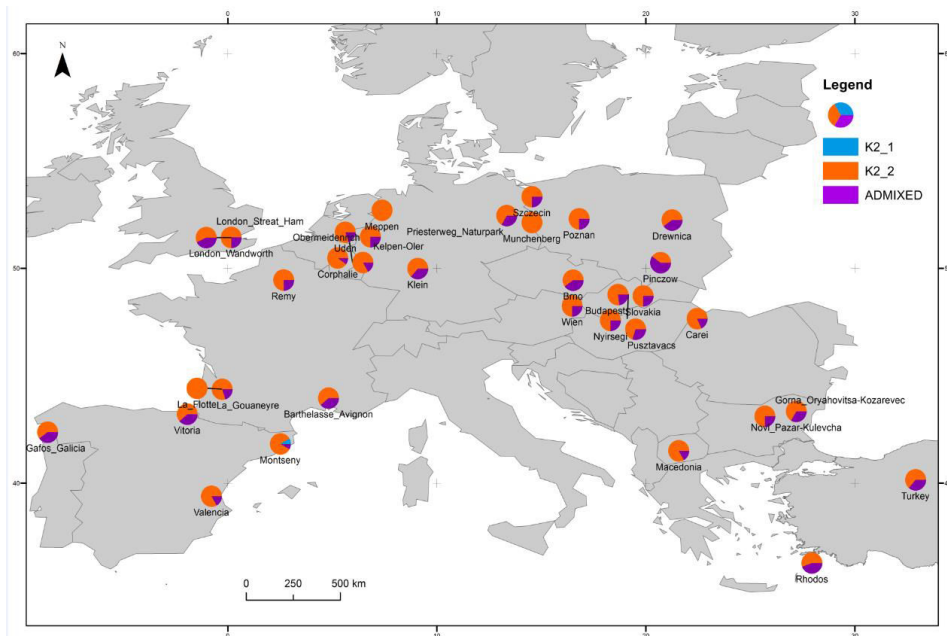
#### 3.1.1. Une introduction restreinte depuis les États-Unis suivie d'une multiplication en Europe

Nous avons pu mettre en évidence que le robinier a été introduit en Europe depuis un nombre limité de populations américaines situées dans les Appalaches (Fig 4). En effet, quand nous avons utilisé la méthode implémentée dans le logiciel Structure, les individus européens appartenaient au même groupe génétique que des individus provenant d'un nombre limité de populations américaines : 4 populations seulement sur 63 échantillonnées (groupe « orange » ci-dessous). Elles appartiennent à la partie nord du plateau des Appalaches (Kentucky, Ohio, Pennsylvanie et Virginie).

A



B



**Fig 4** Diagrammes circulaires représentant la proportion d'individus assignés dans chaque population à chacun des groupes génétiques identifiés (K2\_1 en bleu ou K2\_2 en orange) ou encore *admixed* en violet (c'est à dire non assigné significativement à un groupe) avec A) dans l'aire native et B) dans l'aire d'introduction européenne. Notez la présence de quatre populations dans le nord du plateau des Appalaches qui contiennent des individus assignés au cluster K2\_2.

Ainsi les données moléculaires concordent avec la revue historiographique que nous avons effectuée : John Tradescant, un pionnier anglais de la fin du XVI<sup>ème</sup> siècle, aurait probablement rapporté les tous premiers robiniers introduits depuis la Virginie. La Pennsylvanie aurait été ensuite une source secondaire importante

d'approvisionnement du robinier. Michaux notait déjà l'usage important du bois de robinier qui y était fait en citant la ville de Harrisburg comme un haut lieu de l'exploitation de cet arbre dans ces termes « Ainsi, dans les environs de Lancaster, et à Harrisburgh, petite ville située sur la Susquehannah, où il se fait un commerce considérable des bois qui descendent par cette rivière » (Michaux, 1813). William Cobbett reconnaitra ensuite la région de Harrisburg comme la source des graines qu'il importa « par millions » en Europe (Cobbett, 1825). Un scénario vraisemblable est donc qu'une fois le robinier introduit en Europe, il y ait été acclimaté puis dispersé à partir de graines ou boutures obtenues dans cette nouvelle aire. Nous retrouvons plusieurs indices historiographiques qui sont en faveur de ce scénario. Le tome 38 de la Biographie Universelle Ancienne et Moderne (1824) exprime « qu'il est certain que c'est aux soins des Robin qu'on doit la multiplication de ce bel arbre, ainsi que de plusieurs autres. ». Un peu plus tard, le Catalogue des plantes, arbres, arbrisseaux et arbustes dont on trouve des graines, des bulbes et du plant chez les sieurs Andrieux et Vilmorin par Philippe Victoire Leveque de Vilmorin (1783) mentionne la disponibilité de semences de robinier, sous le nom d'Acacia ou *Robinia pseudoacacia*, les auteurs font également état d'une variété sans épines *Robinia pseudoacacia inermis*. La dynastie Vilmorin, car l'entreprise fut reprise de père en fils jusqu'en 1962, possédèrent plusieurs vergers à graines en France à partir desquelles ils obtinrent les semences qu'ils commercialisaient. Le Catalogue des végétaux ligneux indigènes et exotiques existant sur le domaine forestier des Barres-Vilmorin (Administration des Forêts, 1878) renseigne sur les arbres que l'on retrouve dans le domaine forestier des Barres situé dans la commune de Nogent-sur-Vernisson dans le Loiret. Plusieurs variétés de robinier y sont plantées, et recensées dans le catalogue : « 68. *Robinia pseudo-acacia*. Lin. [Robinier faux acacia]. Amérique du Nord. *Ancienne Pépinière* (cinq arbres de 22 mètres de hauteurs) ; 69. *Robinia pseudo-acacia spectabilis*. [Acacia remarquable]. Cailloutière (huit lignes, de 12 mètres de hauteur et de 60 centimètres de circonférence, mélangées de Pins Laricios de Calabre et de Pins Sylvestres)[...] ». William Cobbett lui aussi lors de son exil aux Etats-Unis établit une pépinière sur l'île de Long Island à New York à partir des graines qu'il obtint de Pennsylvanie, puis à son retour en Angleterre, il y forma un arboretum où il continua à cultiver le robinier (Cobbett, 1828; Michener, 1988). Ajoutons à cela que l'on trouve dans les nombreux ouvrages sur le robinier rédigés au XIX<sup>ème</sup> siècle, que ce soient des bulletins de société d'horticulture ou des éloges, des témoignages de cultivateurs à propos de leurs travaux pour multiplier le robinier à partir de boutures ou de graines (voir par exemple Mémoire sur les *robinia pseudo-acacia, viscosa et hispida* par Charles Gombault (1837) dans Mémoires de la société royale des sciences, belles-lettres et arts d'Orléans ou encore Lettre sur le robinier, connu sous le nom impropre de faux acacia par François de Neufchâteau (1807). Le scénario d'introduction du robinier depuis un nombre limité de populations de l'aire native situées dans le nord du plateau des Appalaches puis sa

multiplication, par graines ou boutures, en Europe, nous apparaît donc comme le plus vraisemblable et cohérent à la fois aux données moléculaires et aux données historiques.

### **3.1.2.Considérations méthodologiques quant à l'inférence de l'histoire de l'invasion**

D'un point de vue méthodologique, nous avons travaillé à partir des coefficients d'appartenance aux K groupes déterminés pour chaque individu par le logiciel STRUCTURE à partir des données de génotypage. Nous avons identifié le nombre de groupes K le plus probable selon la méthode dite du delta K (Evanno et al., 2005) qui est recommandée. L'inconvénient de cette méthode est qu'elle ne permet pas directement de tester si le meilleur nombre de groupe est  $K = 1$ , autrement dit l'absence de structure. Nous avons donc suivi les recommandations des auteurs du logiciel, émises dans la documentation (Pritchard, 2010) et vérifié la convergence des paramètres. De plus, nous avons subdivisé notre jeu de données et relancé les analyses en retirant certains individus ou certaines populations afin de vérifier que la structure identifiée était réelle et n'était pas un artefact dû à une population ou à quelques individus qui seraient sortis du lot. Enfin, nous avons travaillé à partir des intervalles de confiance obtenus pour les coefficients d'appartenance de chaque individu aux K groupes identifiés. En effet par défaut, STRUCTURE fournit la valeur moyenne du coefficient d'appartenance aux K groupes, mais il est possible de l'assortir d'un intervalle de confiance. Nous avons ainsi démontré qu'il y avait toujours une majorité d'individus attribuée significativement à l'un des K groupes, ce qui représente un argument fort pour dire que la structure est bien réelle. Nous encourageons vivement cette approche notamment dans le cas où le nombre de groupe détecté le plus probable est 2. Certes, l'interprétation est un peu alourdie mais cela permet de s'assurer de la robustesse de la structure observée.

Dans notre étude, lors de l'analyse de STRUCTURE portant sur les 720 individus génotypés provenant des deux aires, nous avons montré qu'ils se répartissaient en deux groupes génétiques significatifs. En travaillant à partir des intervalles de confiance obtenus pour les coefficients d'appartenance de chaque individu aux deux groupes identifiés, nous avons pu mettre en évidence que s'il y avait toujours des individus assignés significativement à l'un des deux groupes, certains individus n'étaient attribués à aucun groupe, nous les avons regroupés sous le vocable d'individus « *admixed* ». En analysant les individus significativement assignés, nous avons pu observer que les individus européens et américains formaient des groupes génétiques différents bien que pour quatre populations américaines, nous retrouvions des individus attribués aux mêmes groupes que les individus

européens. Une explication parcimonieuse est que le robinier a été introduit en Europe depuis un nombre limité de populations américaines. L'interprétation des individus *admixed* est plus délicate et peut correspondre à des situations différentes. Par exemple, en Europe les individus *admixed* pourraient correspondre à des individus issus de la reproduction entre individus provenant de populations américaines isolées dans l'aire native ; aux Etats-Unis les *admixed* pourraient être dus à des individus déplacés par l'homme, mais comme nous avons échantillonné des populations en conditions naturelles cela apparait peu probable. Le fait que le structure génétique soit globalement faible avec un  $F_{ST}$  de 5.23 % (95% CI: 4.77 % - 5.70 %) peut également induire le fait que certains individus ne soient pas attribués significativement à un groupe.

D'autres méthodes existent pour reconstruire le scénario des invasions (Cristescu, 2015). Une méthode communément employée fait appel aux *Approximate Bayesian Computation* ABC (Beaumont, 2010; Beaumont et al., 2002). De façon simplifiée, les méthodes bayésiennes consistent à intégrer nos connaissances *a priori* sur un paramètre, qui peuvent être faibles, et l'information issue des données de l'expérience (ex.  $F_{ST}$ , nombres d'allèles, distances génétiques entre paires de populations...), qu'on appelle la vraisemblance, afin d'estimer une connaissance *a posteriori* de ce paramètre. Cependant, dans le cas de modèles complexes, la vraisemblance peut être difficile à calculer. Les méthodes ABC proposent d'estimer une valeur approchée de la vraisemblance en effectuant un grand nombre de simulations (Cornuet et al., 2014). Les distributions *a posteriori* sont ensuite estimées à partir des valeurs des paramètres, qui fournissent les données simulées les plus proches des données réelles sur la base d'un certain nombre d'indicateurs statistiques synthétisant le jeu de données (Cornuet et al., 2014).

Les méthodes ABC sont souvent employées pour reconstruire les scénarios d'invasion car elles permettent d'estimer différents paramètres démographiques tels la taille des populations efficaces, le nombre de générations depuis la séparation des populations, de détecter des goulots d'étranglement ou des introductions multiples; il est également possible de confronter différents scénarios d'introduction afin d'estimer le plus crédible (Cornuet et al., 2014). Le logiciel DIYABC a ainsi été mis au point afin de faciliter l'implémentation de ces analyses (Cornuet et al., 2014). Cependant, ces méthodes ont été initialement développées pour analyser des données de génotypage de microsatellites qui sont des marqueurs très polymorphes. Dans le cas de SNPs, les marqueurs sont moins polymorphes, c'est-à-dire que l'on s'attend à beaucoup de marqueurs avec une faible fréquence de l'allèle rare (*Minor Allele Frequency* ou MAF). Cependant lorsque l'on sélectionne les SNPs pour faire du génotypage, on choisit des SNPs avec une fréquence allélique de l'allèle rare élevée, car en effet on conserve les SNPs avec au

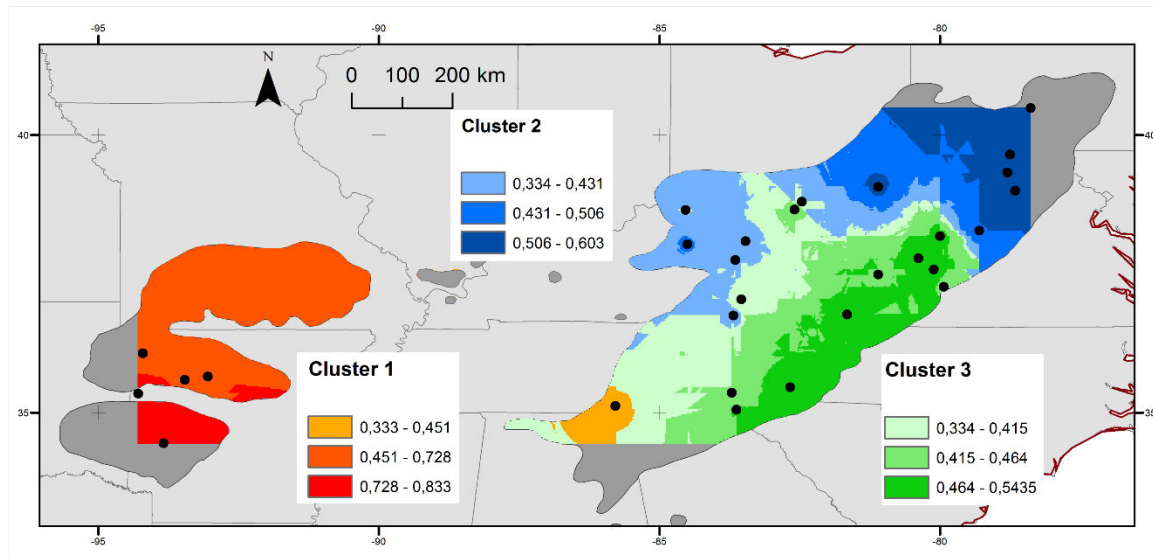


moins deux clusters hétérozygotes et homozygotes clairs, ce qui biaise la distribution de la MAF dans les populations génotypées (Helyar et al., 2011). Le problème est que des jeux de données de SNPs avec une faible MAF sont simulés par DIYABC, conformément à l'attendu théorique, et les données observées ne convergent pas avec les données simulées. Pour tenter de corriger ce problème, il est possible de filtrer les SNPs d'après leur MAF, ce qui induit une diminution du nombre de marqueurs utilisés dans l'analyse et par conséquent la puissance de l'analyse. Lors de nos essais, en filtrant les SNPs avec une MAF inférieure à 0.1, nous passons de 113 marqueurs à 73 et ceux avec une MAF inférieure à 0.2 à 42 marqueurs. Dans ce dernier cas uniquement, les données observées et simulées avaient convergé ; mais les possibilités d'inférences étaient plus limitées, par le moindre nombre de marqueurs moléculaires et par conséquent une puissance d'analyse plus faible. Ces analyses ont mis en évidence un scénario d'introduction du robinier depuis les Appalaches, mais ce scénario était déjà mis clairement en évidence par les analyses de structure. L'utilisation des méthodes ABC n'apportait pas alors d'informations complémentaires sur ce point. De plus, les paramètres démographiques d'intérêts (taille des populations efficace) étaient assortis d'intervalles de confiance extrêmement larges. Ainsi, les analyses ABC semblent plus efficaces dans le cas de données microsatellites (Besnard et al., 2014; Keller et al., 2014) ou avec un grand jeu de SNPs, de l'ordre du millier (Yang et al., 2017). Pour ces raisons, suite à nos analyses exploratoires, nous avons décidé de ne pas employer de méthodes ABC pour tester les scénarios d'introduction du robinier en Europe.

### **3.2. Structure génétique des populations des Etats-Unis à l'Europe**

#### **3.2.1. Des processus naturels anciens ont façonné la diversité aux États-Unis tandis que son introduction en Europe est marquée par l'empreinte de l'Homme**

Aux Etats-Unis, le robinier est structuré en trois groupes génétiques (Fig 5). Le premier groupe génétique correspond aux monts Ozark tandis que les deux autres groupent divisent les Appalaches en deux selon l'axe nord sud. En considérant les valeurs des coefficients d'appartenance, nous pouvons observer que le groupe des Ozark se distingue nettement des deux groupes génétiques des Appalaches entre lesquels la différenciation est plus ténue.

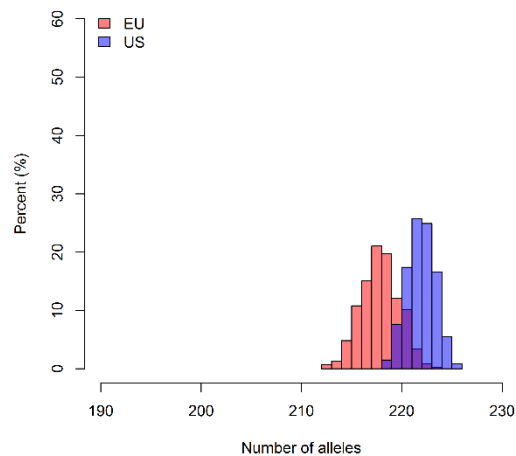


**Fig 5** Interpolation permettant de visualiser les groupes génétiques identifiés aux Etats-Unis d'après le génotypage de 386 individus américains échantillonnés dans 29 populations (cercles noirs). L'interpolation a consisté en une pondération inverse de la distance sur les coefficients d'appartenance aux groupes génétiques obtenus d'après l'analyse de STRUCTURE. La zone grisée correspond à l'aire native du robinier d'après Little (1971).

Ces trois groupes résultent très probablement des processus de glaciation déglaciation qui ont lieu il y a quelques dizaines de milliers d'années (Hewitt, 2000). En effet, des refuges glaciaires ont déjà été identifiés de part et d'autre du fleuve Mississippi pour d'autres espèces d'arbres, comme par exemple chez *Pinus taeda* (Lu et al., 2016). Lors des mouvements de migration suivant le dernier maximum glaciaire il y a 21 000 ans, le bassin alluvial du Mississippi a agi comme une barrière physique et a limité la migration, contribuant à conserver la différenciation génétique entre les populations des Ozarks et celles Appalaches. De même au sein des Appalaches, cette structure génétique le long d'un axe nord-sud a été relevée chez plusieurs espèces ligneuses comme *Scirpus ancistrochaetus* (Cipollini et al., 2017), *Tsuga caroliniana* (Potter et al., 2017) et *Pinus strobus* (Nadeau et al., 2015). Des refuges glaciaires ont également été mis en évidence au nord et au sud des Appalaches chez *P. strobus* (Nadeau et al., 2015) et la structure génétique des populations observée correspond à un mouvement de recolonisation à la suite du maximum glaciaire à partir de refuges nord et sud des Appalaches.

Si des processus naturels ont forgé la structure génétique des populations de robinier dans l'aire native, l'introduction en Europe est marquée par l'action humaine, tout d'abord par une introduction partielle de la diversité existante au sein de l'aire native puis par l'influence probable de la sélection d'origine anthropique sur

l'évolution du robinier dans sa nouvelle aire. Seuls des individus appartenant aux deux groupes génétiques identifiés dans les Appalaches ont été introduits en Europe. Le robinier européen a subi un goulot d'étranglement génétique suite à cet effet de fondation attesté par la perte de diversité génétique mise en évidence. Ainsi, nous observons un nombre d'allèles plus faible en Europe qu'aux États-Unis (Fig 6). Ce résultat n'a été révélé qu'en élargissant le nombre de marqueurs sur un nombre plus restreint d'individus, en prenant soin d'échantillonner plus largement les marqueurs à faible MAF.

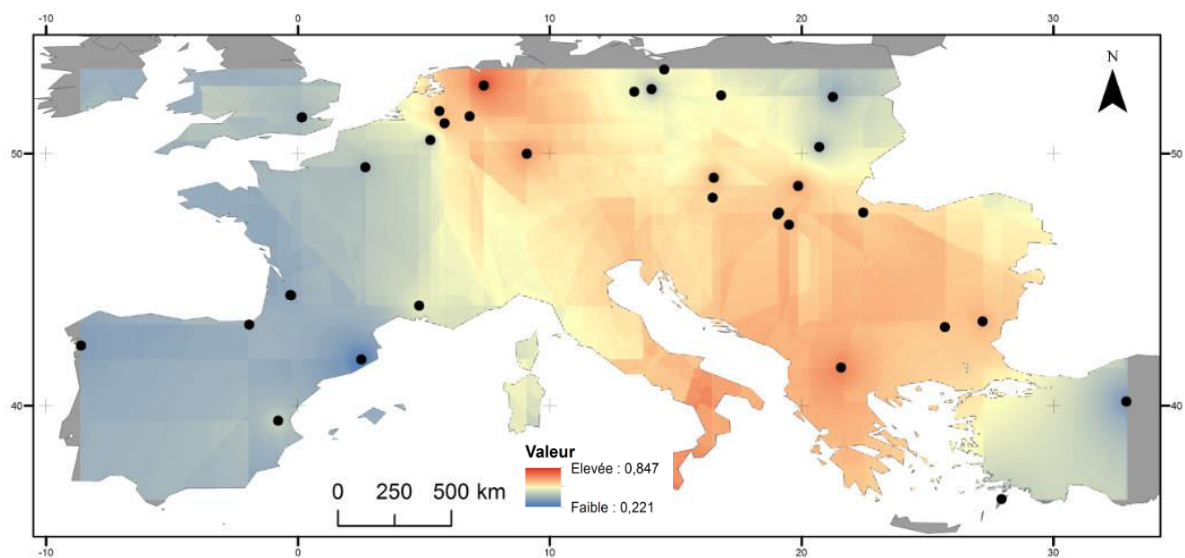


**Fig 6** Bootstrap sur le nombre d'allèles identifiés en Europe et aux Etats-Unis à partir du génotypage de 163 individus avec 251 SNP. Le nombre d'allèles est significativement inférieur en Europe (Wilcoxon,  $p < 0.0001$ ).

Au sein de l'aire européenne, deux groupes génétiques ont été identifiés (Fig 7). Cette structure génétique observée en Europe avec deux groupes génétiques identifiés aurait pu correspondre aux deux groupes génétiques qui ont été identifiés aux Etats-Unis dans les Appalaches qui ont été introduits et qui seraient alors restés ségrégués en Europe. Afin de tester cette hypothèse, nous avons subdivisé le jeu de données en ne conservant que les populations américaines identifiées comme sources du robinier ainsi que les populations européennes avec les coefficients d'appartenance les plus élevés en Europe. Ces nouvelles analyses n'ont mis en évidence aucun patron particulier permettant de conforter cette hypothèse.

Une hypothèse alternative serait alors que cette organisation génétique en Europe ait été modélisée par l'action humaine. En effet, depuis la seconde moitié du XVIII<sup>ème</sup> siècle, des vastes programmes de plantation d'arbres forestiers ont été entrepris en Europe centrale afin de favoriser l'usage du robinier en foresterie (Cierjacks et al., 2013; Vítková et al., 2017). De plus, la Hongrie a conduit un large programme d'amélioration génétique du robinier

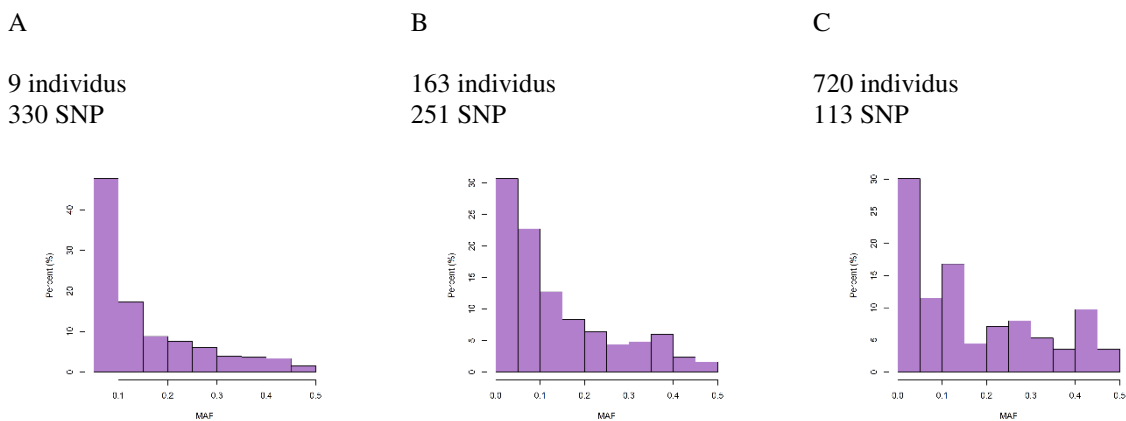
à partir du début du XX<sup>ème</sup> siècle (Keresztesi, 1983). Les variétés de robinier hongroises sont maintenant renommées, par exemple les variétés Puztavacs et Nyirsegi (Keresztesi, 1983), et elles sont présentes dans la plupart des catalogues des pépinières en Europe (ex. France : Naudet, 2018 ; ONF, 2018 ; Belgique : Wallonie Environnement, 2018 ; Allemagne : Energie Pflanzen, 2018 ; Hongrie : Silvanus Forestry, 2018). La structure observée en Europe pourrait alors être une trace d'évolution causée par la sélection anthropique réalisée par l'homme depuis le XVIII<sup>ème</sup> siècle en Europe centrale.



**Fig 7** Interpolation permettant de visualiser les groupes génétiques identifiés en Europe d'après le génotypage de 334 individus européens échantillonnés dans 34 populations (cercles noirs). Comme le nombre de groupes génétiques identifiés en Europe était deux, les coefficients d'appartenance ont été représentés sur une échelle continue comprise entre 0 et 1, le premier groupe génétique étant compris entre 0 et 0.5 et le second entre 0.5 et 1. Plus le coefficient d'appartenance est proche de 0 ou de 1, plus la différenciation génétique est marquée. L'interpolation a consisté en une pondération inverse de la distance sur les coefficients d'appartenance aux groupes génétiques obtenus d'après l'analyse de STRUCTURE. L'échelle des coefficients d'appartenance est comprise entre 0.221 et 0.847 et non entre 0 et 1 ce qui permet de mieux visualiser la différenciation entre les deux groupes génétiques identifiés mais peut faire paraître cette différenciation plus importante qu'elle ne l'est en réalité (voir par comparaison la figure dans le chapitre 1 « *Few North-Appalachian populations are the sources of European black locust* »)

### 3.2.2.Considérations méthodologiques sur l'estimation de la diversité à l'aide de marqueurs SNPs

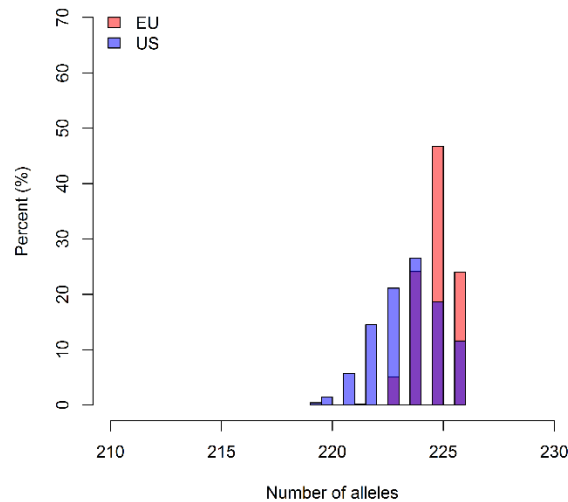
Comme je l'ai évoqué précédemment, la distribution attendue théoriquement pour les SNPs est dite en « L » avec beaucoup de SNPs dont la fréquence de l'allèle minoritaire (MAF) est faible. Cependant, lorsque nous avons sélectionné les SNPs pour les analyses, nous avons préférentiellement choisi au départ les SNPs présentant trois groupes de génotypes clairs. Ce faisant, la distribution globale de MAF est biaisée vers les valeurs intermédiaires. La figure 8 illustre les différentes distributions observées de la MAF dans trois cas. Dans la première situation (Fig 8A), 9 individus ont été génotypés avec 330 marqueurs SNP (Verdu et al., 2016), nous pouvons voir ici que la distribution de la MAF observée est proche de la forme théorique avec beaucoup d'allèles rares. En revanche, comme nous pouvons l'observer lorsque 251 et 113 SNPs sont utilisés (Fig 8B et C), plus le nombre de marqueurs utilisés diminue, plus l'allure de la distribution s'écarte de l'attendu théorique. Le jeu de données SNPs (113 SNPs) que nous avons utilisé pour étudier la structure des populations présente un excès de SNPs avec une MAF élevée (Fig 8C), nous avons en partie corrigé l'échantillonnage pour mieux estimer la perte de diversité en utilisant 251 SNPs (Fig 8B).



**Fig 8** Distributions observées de la MAF dans trois situations différentes. A) 9 individus et 330 SNP, B) 163 individus et 251 SNP et C) 720 individus et 113 SNP. Nous pouvons observer que plus on augmente le nombre de marqueurs, plus la distribution observée s'approche de la distribution attendue théoriquement.

La distribution est donc biaisée par le fait que nous avons sélectionné les SNPs les plus polymorphes pour notre analyse. Cela affecte moins l'analyse de la structure génétique des populations, mais un biais peut survenir lorsque l'on compare la diversité génétique des deux aires. La figure 9 représente le nombre d'allèles identifiés en

Europe et aux Etats-Unis calculé d'après le génotypage des 720 individus avec 113 marqueurs (Fig 9). Si nous avons utilisé ce jeu de données sans prendre en compte le biais de la distribution de la MAF, les conclusions auraient été tout autres et nous aurions improprement conclu à l'absence de goulot d'étranglement et à une diversité génétique supérieure en Europe.



**Fig 9** Bootstrap sur le nombre d'allèles identifiés en Europe et aux Etats-Unis en se basant sur le génotypage des 720 individus avec 113 marqueurs SNPs. Le nombre d'allèles est significativement supérieur en Europe (Wilcoxon,  $p < 0.0001$ )

Ainsi les analyses de la diversité génétique réalisées avec des SNPs peuvent mener à des conclusions trompeuses si l'on ne prend pas garde au biais de recrutement<sup>1</sup> des SNPs (Morin et al., 2004; Helyar et al., 2011). Dans une étude utilisant 5 SNPs pour comparer la diversité génétique entre populations natives et invasives du poisson *Gambusia holbrooki*, les auteurs attribuent l'absence de différence de diversité observée entre les deux aires au biais de recrutement des SNPs (Vidal et al., 2012). En utilisant 62 SNPs, aucune différence de richesse allélique n'a été mise en évidence entre l'aire native et l'aire invasive de la plante *Mimulus guttatus*, toutefois les auteurs n'ont pas pris en compte le biais de recrutement, et vu le nombre modéré de SNPs utilisés, ce résultat devrait être considéré avec prudence (Pantoja et al., 2017). Des études plus robustes utilisant des SNPs ont été faites sur des poissons invasifs et n'ont pas révélé de perte de diversité génétique pour le poisson lion (*Pterois*

---

<sup>1</sup> *Ascertainment bias* en anglais

spp.) avec plus de 1000 SNPs employés (Pérez-Portela et al., 2018) et pour différents salmonidés avec plus de 10000 SNPs génotypés (Narum et al., 2017). Des résultats similaires ont été obtenus avec des microsatellites, la plante *Heracleum mantegazzianum* génotypée avec 8 microsatellites présentent une réduction importante de la diversité génétique dans l'aire d'introduction (Henry et al., 2009), mais *Ambrosia artemisiifolia* génotypée avec 5 microsatellites ne présente pas de réduction de la diversité génétique entre les deux aires (Genton et al., 2005) similairement à *Alliaria petiolata* génotypée avec huit microsatellites (Durka et al., 2005). Ainsi la perte de diversité n'est pas systématique dans l'aire d'introduction, et une méta-analyse a montré que celle-ci était souvent modérée, de l'ordre de 15 – 20 % (Dlugosch et al., 2015; Dlugosch and Parker, 2008). Certaines études empiriques ont mis en évidence que les SNPs étaient de meilleurs marqueurs que les microsatellites pour étudier la structure des populations et que réciproquement les microsatellites étaient de meilleurs marqueurs que les SNPs concernant l'estimation de la diversité (Singh et al., 2013). Cependant d'autres études ont abouti aux mêmes conclusions quels que soient les types de marqueurs utilisés (van Inghelandt et al., 2010; Filippi et al., 2015).

### **3.3. Le caractère invasif du robinier peut-il être expliqué par l'évolution de certains traits phénotypiques ?**

#### **3.3.1. Un robinier qui se reproduit davantage de façon asexuée dans l'aire d'introduction**

Le robinier pratique à la fois la reproduction sexuée et asexuée par drageonnage (Cierjacks et al., 2013), et nous avons mis en évidence que le robinier se reproduit davantage de façon asexuée dans l'aire d'introduction. Lors de notre étude visant à reconstruire l'histoire de l'introduction en Europe, nous voulions éviter d'échantillonner des clones car leur présence dans les populations génotypées pourrait biaiser l'estimation de la diversité génétique. Un protocole d'échantillonnage a donc été conçu afin de minimiser le risque de prélever des clones en laissant un minimum de 25 mètres entre chaque individu sélectionné. Des génotypes identiques avec notre jeu de SNPs, et que nous avons interprétés comme des clones, ont pourtant été prélevés que ce soit dans l'aire native ou dans l'aire d'introduction. Parmi les individus prélevés dans des conditions naturelles, c'est-à-dire ne provenant pas d'un jardin commun, nous avons génotypé 280 individus européens et 356 américains. Parmi eux, 98 clones ont été détectés, 30 dans l'aire américaine et 68 dans l'aire européenne. La différence est statistiquement significative ( $\chi^2=29.04$ ,  $df = 1$ ,  $p = 7.1 \cdot 10^{-8}$ ), indiquant davantage de clonalité en Europe.

L'évolution des systèmes de reproduction vers davantage de reproduction asexuée est une hypothèse fréquemment émise pour expliquer le potentiel invasif d'une espèce. En effet, la reproduction asexuée offre une assurance reproductive dans l'aire d'introduction (Barrett et al., 2008). D'un point de vue génétique, il a été démontré que les populations clonales ne subissaient pas nécessairement de forte réduction de la diversité génétique et elles peuvent même conserver une forte diversité génétique (Barrett, 2010). Cependant, une modélisation a mis en évidence que le potentiel invasif d'une espèce n'était pas linéairement corrélé à son taux de reproduction asexuée, mais plutôt qu'il existait un optimum théorique pour une espèce clonale pratiquant une faible part de reproduction sexuée, de l'ordre de 5% (Bazin et al., 2013). Des modifications du système de reproduction ont été observées chez plusieurs espèces invasives (Barrett et al., 2008; Petanidou et al., 2012; Rambuda and Johnson, 2004). Par exemple, la modification du système de reproduction sexuée vers une reproduction clonale due suite à un effet de fondation a été mis en évidence chez *Eichhornia crassipes* (Barrett et al., 2008) et *Fallopia japonica* (Hollingsworth and Bailey, 2000). Cette tendance vers davantage de reproduction asexuée n'est pourtant pas universelle et des espèces pratiquant uniquement la reproduction sexuée ont été rapportées comme hautement invasives telle *Ambrosia artemisiifolia* (Friedman and Barrett, 2008). Au Japon, il a été mis en évidence que l'invasion par le robinier était favorisé à la fois par la production sexuée et asexuée à une échelle locale (Kurokochi and Hogetsu, 2014). Il est possible qu'un changement vers davantage de reproduction asexuée soit survenue dans l'aire d'introduction chez le robinier suite à la sélection artificielle : en effet, une façon traditionnelle de cultiver le robinier a été de stimuler le drageonnage en endommageant les racines des arbres plantés (Saint Jean de Crève Coeur, 1786, Bergeron, 1802, de Neufchateau, 1807), ce qui a pu représenter une pression de sélection. D'autre part, le drageonnage est favorisé par la perturbation d'origine naturelle et anthropique ; notamment, les pratiques de gestion sylvicoles telles que le taillis favorisent sa dynamique de drageonnement (Cierjacks et al., 2013; Radtke et al., 2013) et donc des différences de pratiques entre les deux continents pourraient contribuer à ces différences d'expression de la clonalité en milieu naturel. Nos résultats ne permettent pas de trancher entre ces deux hypothèses. Cependant, résoudre définitivement cette question nécessiterait d'autres études complémentaires. Par exemple, à la manière de ce qu'ont réalisé Kurokochi & Hogetsu (2014) sur l'intégralité des robiniers situés dans trois placettes (soit environ 200 individus par placette, avec huit microsatellites), on pourrait analyser la clonalité en réalisant le génotypage exhaustif d'individus situés dans plusieurs placettes réparties dans l'aire d'introduction et dans l'aire native, en sélectionnant des placettes où les perturbations anthropiques ou naturelles sont connues et minimisées ou stratifiées. Une approche expérimentale

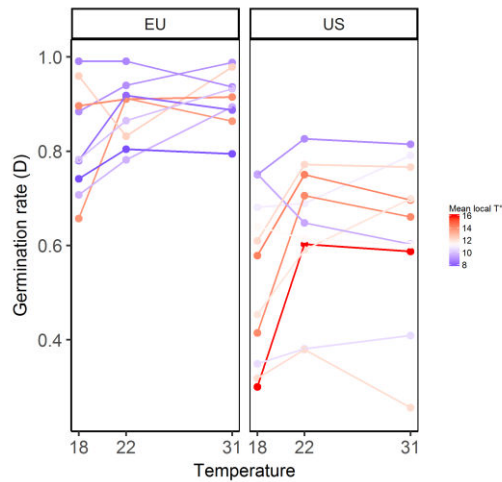


est en cours au laboratoire pour tester si on détecte une différence de rejet de tige entre populations et entre les deux aires.

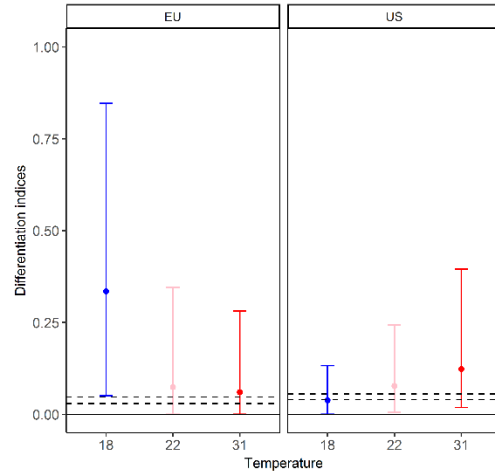
### **3.3.2. Une germination potentielle améliorée dans l'aire d'introduction**

Nous avons réalisé une expérience de génétique quantitative pour évaluer la variabilité des traits phénotypiques juvéniles de populations de robinier (3000 individus provenant de 20 populations). Les semis ont été cultivés dans 3 enceintes climatiques définies à des températures différentes (18°, 22°C et 31°C). Les résultats ont révélé que tous les traits étudiés répondaient à la température et qu'il existait une différenciation génétique entre populations. A l'exception du taux de germination, aucun trait n'était cependant différencié entre l'aire native et l'aire invasive. Le taux de germination est significativement supérieur dans l'aire invasive par rapport à l'aire native (88% versus 60%). La figure 10A représente les normes de réaction pour le taux de germination de chacune des populations étudiées des aires (EU : Europe, US : Amérique) en fonction de l'environnement de croissance mais la couleur du trait fait référence à la température moyenne annuelle observée au lieu de prélèvement de la population. La population américaine qui a le meilleur taux de germination (76%/84%/80% respectivement à 18°C/22°C/31°C) ainsi que le plus constant au changement de température est Altoona située en Pennsylvanie, état américain depuis lequel de nombreuses graines ont été importées en Europe (chapitre 1).

A



B



**Fig 10** A) Normes de réaction pour le taux de germination de chacune des populations étudiées des aires (EU : Europe, US : Amérique) en fonction de l'environnement de croissance (18°C, 22°C, 31°C). La couleur du trait fait référence à la température moyenne annuelle observée au lieu de prélèvement de la population, comprise entre 8°C (en bleu) et 16°C (en rouge). B) Comparaison  $Q_{ST}$   $F_{ST}$  pour le taux de germination en fonction de l'environnement de croissance (18°C, en bleu ; 22°C, en rose ; 31°C, en rouge). Les limites supérieures et inférieures de l'intervalle de confiance à 95% du  $F_{ST}$  sont représentés en pointillées, les  $Q_{ST}$  moyens sont indiqués par un point et sont assortis de leur intervalle de confiance à 95%.

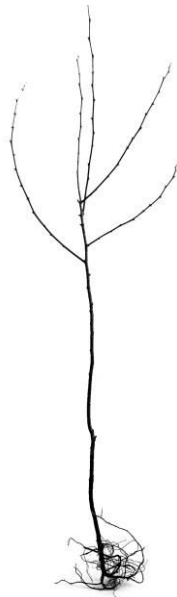
Une augmentation du taux de germination entre l'aire native et l'aire invasive telle que l'on a pu l'observer (Fig 10A) a été mise en évidence chez *Hypericum perforatum* et *Achillea millefolium* (Beckmann et al., 2011). Il semble toutefois que ce ne soit pas la règle chez les espèces invasives. Bien que la germination et les traits juvéniles soient cruciaux pour le succès des espèces invasives (Gioria and Pyšek, 2016), un taux de germination amélioré dans l'aire d'introduction n'est pas un bon prédicteur d'un bon potentiel invasif. C'est plutôt une tendance à l'augmentation de la vitesse de germination qui a été observée (Gioria and Pyšek, 2016) comme chez *Rhododendron ponticum*, *Ambrosia artemisiifolia* ou encore *Ulmus pumila* (Erfmeier and Bruehlheide, 2005; Hirsch et al., 2012; Hodgins and Rieseberg, 2011). D'autre part, la Pennsylvanie a été une zone active du commerce inter continental de robiniers et comme rappelé précédemment, nos résultats montrent que les graines issues de la population de Pennsylvanie ont le meilleur potentiel germinatif parmi les populations natives étudiées. La comparaison  $Q_{ST}$   $F_{ST}$  montre cependant que pour le taux de germination, le  $Q_{ST}$  est significativement supérieur au  $F_{ST}$  pour les populations européennes (Fig 10B). Cela pourrait indiquer un signal de sélection sur ce trait. Une

explication pour cette augmentation de la germination pourrait être la sélection artificielle d'origine anthropique appliquée sur le pouvoir germinatif des graines, de façon plus ou moins consciente notamment lors de la propagation de l'espèce dès le 18<sup>ème</sup> siècle. Aux vues de ces considérations, il semble donc vraisemblable que cette sélection soit en partie due à l'action anthropique.

Un point important à garder à l'esprit est que nous avons mesuré la germination potentielle, c'est-à-dire la faculté de germer qu'a la graine une fois sa dormance levée, car nous avons levé la dormance des graines avant de les semer et mis en culture en conditions trophiques (lumière, eau, nutriments) non limitants. En effet, les graines de robinier ont une dormance physique en raison de leurs téguments durs et imperméables, et sans traitement préalable le taux de germination est très faible, inférieur à 10% (Cierjacks et al., 2013). Dans la nature, les conditions environnementales pour lever la dormance diffèrent souvent de celles requises pour la germination. Aussi la dormance et le déroulement de la germination doivent-ils être étudiés séparément (Gioria and Osborne, 2014). La dormance induit souvent la création d'une banque de graines dans le sol ce qui semble être particulièrement vrai pour le robinier qui produit un grand nombre de graines, une masse produite de 6 – 12 kg par arbre et par année a été rapportée pour un poids moyen d'une graine d'environ 20 mg (Cierjacks et al., 2013).

Lors de nos échanges auprès des différents acteurs autour du robinier, exploitants forestiers et gestionnaires de milieux naturels, nous avons fréquemment entendu que le robinier ne se multiplie pas par graines dans son milieu naturel en Europe mais uniquement par reproduction asexuée. D'une part, nos travaux en génétique des populations suggèrent que la reproduction sexuée participe à la régénération des populations de robinier. De plus, notre expérience en conditions contrôlées démontre que le potentiel germinatif des graines est supérieur en Europe par rapport aux Etats-Unis. Considérer que les graines ne germent jamais dans la nature revient à supposer que la levée de dormance ne se produit pas en conditions naturelles. Cela serait surprenant car d'une part le robinier produit un nombre élevé de graines et d'autre part l'abrasion des téguments est un événement plausible dans la nature, où les graines transportées par le vent ou l'eau sont potentiellement abrasées par frottements lors du transport. Pour explorer cette question autour de la germination du robinier en conditions naturelles, une étude préliminaire a été menée : nous avons extrait du sol de jeunes robiniers (i.e. diamètre de 1cm maximum à 1.3 m de hauteur) afin d'analyser leur système racinaire et déterminer s'ils provenaient de drageons ou de graines (Fig 11). Dans chaque placette, entre 15 et 30 robiniers étaient sélectionnés aléatoirement en respectant une distance minimale de 5m entre eux. A l'heure actuelle, huit sites ont été échantillonnés dans le sud-ouest de la France.

A

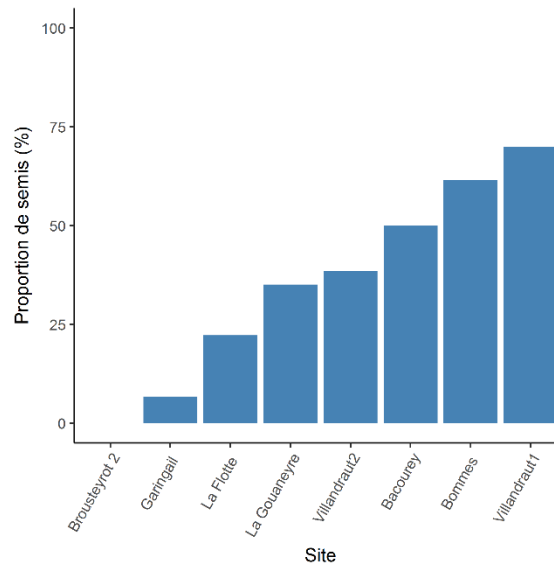


B



**Fig 11** Jeunes robiniers issus A) d'une graine et B) d'un drageon. Notez la différence du système racinaire ou la racine traçante horizontale s'observe distinctement dans le cas du drageon (Photo XP Bouteiller)

Les résultats ont montré une grande disparité entre sites, avec des proportions de semis variant entre 0% et 70% (Fig 12). Des analyses similaires avaient été menées en Espagne par extraction de juvéniles, avec un taux de reproduction sexuée du robinier variable entre sites, mais beaucoup plus faible en moyenne (6%) (Cabra-Rivas and Castro-Diez, 2016).



**Fig 12** Proportion de semis observés dans les sites de l'étude préliminaire

Ces résultats confirment que les reproductions sexuée et asexuée jouent toutes les deux un rôle efficace dans la persistance et l'expansion du robinier dans son aire d'introduction. Toutefois la portée de cette étude était limitée et il conviendrait d'en augmenter l'échelle, et d'évaluer les différences à travers l'aire native et l'aire d'introduction.

#### 4. Conclusion

Cette thèse a permis de résoudre un certain nombre de questions concernant l'histoire écologique et évolutive du robinier faux-acacia (*Robinia pseudoacacia* L.) depuis son introduction en Europe. J'ai tout d'abord démontré que le robinier avait été introduit en Europe depuis un nombre limité de populations américaines localisées dans le plateau des Appalaches. Un goulot d'étranglement génétique associé à une perte de diversité génétique a été établi dans l'aire d'introduction. De plus la structure des populations reflète des processus évolutifs et démographiques différents dans les deux aires. Aux Etats-Unis, elle résulte de processus historiques au long cours suite aux mouvements de migration post glaciaire tandis qu'en Europe, la structure est beaucoup plus ténue, et semble influencée par les actions anthropiques, notamment la sélection réalisée durant les phases de propagation active de l'espèce en Europe, et plus récemment les programmes d'amélioration conduits en Europe de l'Est.

Un autre résultat majeur concerne le système de reproduction du robinier qui semble lui aussi empreint de l'activité humaine. Davantage de reproduction clonale a été observée parmi les populations européennes et le potentiel germinatif est également supérieur au sein de l'aire d'introduction. Là encore le rôle de l'Homme semble prépondérant levant la question d'une espèce en voie de domestication.

Plusieurs champs d'études restent en suspens. L'analyse approfondie de l'évolution du système de reproduction du robinier semble particulièrement intéressante et d'autres études sont nécessaires pour élucider cet aspect et son rôle dans la promotion de l'invasion. Conjointement, la comparaison du développement juvénile avec d'autres espèces natives ou invasives offre une voie de recherche intéressante. Enfin, la mise en place de jardins communs pour suivre sur une longue durée le développement permettrait d'avoir accès à d'autres traits reliés à la *fitness* tels la floraison ou la fructification.

# **Bibliographie**

# Bibliographie

- Administration des Forêts. 1878. Catalogue des végétaux ligneux indigènes et exotiques existant sur le domaine forestier des Barres-Vilmorin. France.
- Agrawal, A.A., 2001. Phenotypic plasticity in the interactions and evolution of species. *Science* (80- ). 294, 321–326.
- Alberto, F.J., Aitken, S.N., Alía, R., González-Martínez, S.C., Hänninen, H., Kremer, A., Lefèvre, F., Lenormand, T., Yeaman, S., Whetten, R., Savolainen, O., 2013. Potential for evolutionary responses to climate change - evidence from tree populations. *Glob. Chang. Biol.* 19, 1645–1661.
- Ayres, D.R., Zaremba, K., Strong, D.R., 2004. Extinction of a common native species by hybridization with an invasive congener. *Weed Technol.* 18, 1288–1291.
- Bais, H.P., 2003. Allelopathy and Exotic Plant Invasion: From Molecules and Genes to Species Interactions. *Science* (80- ). 301, 1377–1380.
- Baker, H.G., 1967. Support for Baker's Law as a rule. *Evolution* (N. Y). 21, 853–856.
- Balogh, L., Dancza, I., Király, G., 2007. Preliminary report on the grid-based mapping of invasive plants in Hungary. *Neobiota* 7, 105–114.
- Barrett, S.C.H., 2010. Why Reproductive Systems Matter for the Invasion Biology of Plants, Fifty Years of Invasion Ecology: The Legacy of Charles Elton.
- Barrett, S.C.H., Colautti, R.I., Eckert, C.G., 2008. Plant reproductive systems and evolution during biological invasion. *Mol. Ecol.* 17, 373–83.
- Bazin, É., Mathé-Hubert, H., Facon, B., Carlier, J., Ravigné, V., 2013. The effect of mating system on invasiveness: some genetic load may be advantageous when invading new environments. *Biol. Invasions* 16, 875–886.
- Beaumont, M.A., 2010. Approximate Bayesian Computation in Evolution and Ecology. *Annu. Rev. Ecol. Evol. Syst.* 41, 379–406.
- Beaumont, M.A., Zhang, W., Balding, D.J., 2002. Approximate Bayesian computation in population genetics. *Genetics* 162, 2025–2035.
- Beckmann, M., Bruelheide, H., Erfmeier, A., 2011. Germination responses of three grassland species differ between native and invasive origins. *Ecol. Res.* 26, 763–771.
- Beilharz, R.G., Luxford, B.G., Wilkinson, J.L., 1993. Quantitative genetics and evolution: Is our understanding of



- genetics sufficient to explain evolution? *J. Anim. Breed. Genet.* 110, 161–170.
- Besnard, G., Dupuy, J., Larter, M., Cuneo, P., Cooke, D., Chikhi, L., 2014. History of the invasive African olive tree in Australia and Hawaii: Evidence for sequential bottlenecks and hybridization with the Mediterranean olive. *Evol. Appl.* 7, 195–211.
- Blackburn, T.M., Pyšek, P., Bacher, S., Carlton, J.T., Duncan, R.P., Jarošík, V., Wilson, J.R.U., Richardson, D.M., 2011. A proposed unified framework for biological invasions. *Trends Ecol. Evol.* 26, 333–9.
- Blair, A.C., Wolfe, L.M., 2004. The Evolution of an Invasive Plant : An Experimental Study with *Silene latifolia*. *Ecology* 85, 3035–3042.
- Blondel, J., Hoffmann, B., Courchamp, F., 2013. The end of Invasion Biology: intellectual debate does not equate to nonsensical science. *Biol. Invasions* 16, 977–979.
- Bock, D.G., Caseys, C., Cousens, R.D., Hahn, M.A., Heredia, S.M., Hübner, S., Turner, K.G., Whitney, K.D., Rieseberg, L.H., 2014. What we still don't know about invasion genetics. *Mol. Ecol.* 24, 1–22.
- Boero, F., 2011. New species are welcome, but. . . what about the old ones? *Ital. J. Zool.* 78, 1–2.
- Botta-Dukát, Z., 2008. Invasion of alien species to Hungarian (semi-)natural habitats. *Acta Bot. Hung.* 50, 219–227.
- Brown, J.H., Sax, D.F., 2004. An essay on some topics concerning invasive species. *Austral Ecol.* 29, 530–536.
- Cabra-Rivas, I., Castro-Diez, P., 2016. Comparing the sexual reproductive success of two exotic trees invading Spanish riparian forests vs. A native reference. *PLoS One* 11, 1–16.
- Callaway, R.M., Ridenour, W.M., 2004. Novel weapons: Invasive success and the evolution of increased competitive ability. *Front. Ecol. Environ.* 2, 436–443.
- Chen, C., Durand, E., Forbes, F., François, O., 2007. Bayesian clustering algorithms ascertaining spatial population structure: a new computer program and a comparison study. *Mol. Ecol. Notes* 7, 747–756.
- Cheptou, P.-O., 2012. Clarifying Baker's Law. *Ann. Bot.* 109, 633–41.
- Chun, Y.J., Nason, J.D., Moloney, K.A., 2009. Comparison of quantitative and molecular genetic variation of native vs. invasive populations of purple loosestrife (*Lythrum salicaria* L., Lythraceae). *Mol. Ecol.* 18, 3020–3035.
- Cierjacks, A., Kowarik, I., Joshi, J., Hempel, S., Ristow, M., Von Der Lippe, M., Weber, E., 2013. Biological Flora of the British Isles: *Robinia pseudoacacia*. *J. Ecol.* 101, 1623–1640.
- Cipollini, K., Lavretsky, P., Cipollini, D., Peters, J.L., 2017. Genetic Population Structure across the Range of Endangered Northeastern Bulrush, *Scirpus ancistrochaetus*. *Int. J. Plant Sci.* 178, 67–78.

- Clavero, M., García-Berthou, E., 2005. Invasive species are a leading cause of animal extinction. *Trends Ecol. Evol.* 20, 110.
- Cobbett, W. 1828. *A Year's Residence in the United States of America* (3<sup>rd</sup> ed.). London, England.
- Colautti, R.I., Barrett, S.C.H., 2013. Rapid Adaptation to Climate Facilitates Range Expansion of an Invasive Plant. *Science* (80-. ). 342, 364–366.
- Colautti, R.I., Grigorovich, I.A., MacIsaac, H.J., 2006. Propagule pressure: A null model for biological invasions. *Biol. Invasions* 8, 1023–1037.
- Colautti, R.I., Lau, J.A., 2015. Contemporary evolution during invasion: evidence for differentiation, natural selection, and local adaptation. *Mol. Ecol.* 24, 1999–2017.
- Conner, J.K., Hartl, D.L., 2004. *A Primer of Ecological Genetics*. Bioscience.
- Corander, J., Marttinen, P., 2006. Bayesian identification of admixture events using multilocus molecular markers. *Mol. Ecol.* 15, 2833–2843.
- Corander, J., Marttinen, P., Sirén, J., Tang, J., 2008. Enhanced Bayesian modelling in BAPS software for learning genetic structures of populations. *BMC Bioinformatics* 9, 539.
- Cornuet, J.M., Pudlo, P., Veyssier, J., Dehne-Garcia, A., Gautier, M., Leblois, R., Marin, J.M., Estoup, A., 2014. DIYABC v2.0: A software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism, DNA sequence and microsatellite data. *Bioinformatics* 30, 1187–1189.
- Cristescu, M.E., 2015. Genetic reconstructions of invasion history. *Mol. Ecol.* 24, 2212–2225.
- Crosti, R., Agrillo, E., Ciccacese, L., Guarino, R., Paris, P., Testi, A., 2016. Assessing escapes from short rotation plantations of the invasive tree species *Robinia pseudoacacia* L. in Mediterranean ecosystems: a study in central Italy. <http://www.sisef.it/iforest> 9, 822.
- Daehler, C.C., 2003. Performance Comparisons of Co-Occurring Native and Alien Invasive Plants: Implications for Conservation and Restoration. *Annu. Rev. Ecol. Evol. Syst.* 34, 183–211.
- DAISIE, 2009. *Handbook of alien species in Europe, Invading nature*: Springer series in invasion ecology. Springer, Dordrecht, Netherlands.
- Davidson, A.M., Jennions, M., Nicotra, A.B., 2011. Do invasive species show higher phenotypic plasticity than native species and, if so, is it adaptive? A meta-analysis. *Ecol. Lett.* 14, 419–31.
- Davis, M.A., Chew, M.K., Hobbs, R.J., Lugo, A.E., Ewel, J.J., Vermeij, G.J., Brown, J.H., Rosenzweig, M.L., Gardener, M.R., Carroll, S.P., Thompson, K., Pickett, S.T.A., Stromberg, J.C., Tredici, P. Del, Suding, K.N.,

- Ehrenfeld, J.G., Philip Grime, J., Mascaro, J., Briggs, J.C., 2011. Don't judge species on their origins. *Nature* 474, 153–154.
- Davis, M.A., Grime, J.P., Thompson, K., 2000. Fluctuating resources in plant communities: a general theory of invasibility. *J. Ecol.* 88, 528–534.
- Denslow, J.S., 2003. Weeds in paradise : thoughts on the invasibility of tropical islands. *Ann. Missouri Bot. Gard.* 90, 119–127.
- Dlugosch, K.M., Anderson, S.R., Braasch, J., Cang, F.A., Gillette, H.D., 2015. The devil is in the details: genetic variation in introduced populations and its contributions to invasion. *Mol. Ecol.* 24, 2095–2111.
- Dlugosch, K.M., Parker, I.M., 2008. Founding events in species invasions: Genetic variation, adaptive evolution, and the role of multiple introductions. *Mol. Ecol.* 17, 431–449.
- Donohue, K., Rubio de Casas, R., Burghardt, L., Kovach, K., Willis, C.G., 2010. Germination, Postgermination Adaptation, and Species Ecological Ranges. *Annu. Rev. Ecol. Evol. Syst.* 41, 293–319.
- Drake, J.M., Bagginstos, P., Lodge, D.M., 2005. Propagule pressure and persistence in experimental populations. *Biol. Lett.* 1, 480–483.
- Durka, W., Bossdorf, O., Prati, D., Auge, H., 2005. Molecular evidence for multiple introductions of garlic mustard (*Alliaria petiolata*, Brassicaceae) to North America. *Mol. Ecol.* 14, 1697–1706.
- Edelaar, P., Burraco, P., Gomez-Mestre, I., 2011. Comparisons between Q(ST) and F(ST) --how wrong have we been? *Mol Ecol* 20, 4830–9.
- Elton, C.S., 1958. *The ecology of invasions by animals and plants*, University. ed. University of Chicago Press, Chicago.
- Energie pflanzen. <https://www.energiepflanzen.com/en/product/robinia-seedling-acacia/> (accès le 25 juin 2018)
- Erfmeier, A., Bruelheide, H., 2005. Invasive and native *Rhododendron ponticum* populations: is there evidence for genotypic differences in germination and growth? *Ecography (Cop.)*. 28, 417–428.
- Essl, F., Dullinger, S., Rabitsch, W., Hulme, P.E., Hülber, K., Jarošík, V., Kleinbauer, I., Krausmann, F., Kühn, I., Nentwig, W., Vilà, M., Genovesi, P., Gherardi, F., Desprez-Loustau, M.-L., Roques, A., Pyšek, P., 2011. Socioeconomic legacy yields an invasion debt. *Proc. Natl. Acad. Sci. U. S. A.* 108, 203–7.
- Estoup, A., Guillemaud, T., 2010. Reconstructing routes of invasion using genetic data: why, how and so what? *Mol. Ecol.* 19, 4113–4130.
- Evanno, G., Regnaut, S., Goudet, J., 2005. Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Mol. Ecol.* 14, 2611–2620.

- Falconer, D., Mackay, T., 1996. Introduction to Quantitative Genetics. Longmans Green, Harlow, Essex, UK Ed 4., 464.
- Falque, M., 2011. Evolution des methodes de génotypage. Journée A.S.F 1–16.
- Felker-Quinn, E., Schweitzer, J.A., Bailey, J.K., 2013. Meta-analysis reveals evolution in invasive plant species but little support for Evolution of Increased Competitive Ability (EICA). *Ecol. Evol.* 3, 739–751.
- Fitzpatrick, B.M., Fordyce, J.A., Niemiller, M.L., Reynolds, R.G., 2012. What can DNA tell us about biological invasions? *Biol. Invasions* 14, 245–253.
- Foxcroft, L.C., Rouget, M., Richardson, D.M., Mac Fadyen, S., 2004. Reconstructing 50&nbsp;years of *Opuntia stricta* invasion in the Kruger National Park, South Africa: environmental determinants and propagule pressure. *Divers. Distrib.* 10, 427–437.
- François, O., Ancelet, S., Guillot, G., 2006. Bayesian clustering using hidden Markov random fields in spatial population genetics. *Genetics* 174, 805–16.
- François de Neufchateau, N. 1807. Lettre sur le robinier, Galland, Paris.
- Friedman, J., Barrett, S.C.H., 2008. High outcrossing in the annual colonizing species *Ambrosia artemisiifolia* (Asteraceae). *Ann. Bot.* 101, 1303–1309.
- Genton, B.J., Shykoff, J.A., Giraud, T., 2005. High genetic diversity in French invasive populations of common ragweed, *Ambrosia artemisiifolia*, as a result of multiple sources of introduction. *Mol. Ecol.* 14, 4275–4285.
- Ghalambor, C.K., McKay, J.K., Carroll, S.P., Reznick, D.N., 2007. Adaptive versus non-adaptive phenotypic plasticity and the potential for contemporary adaptation in new environments. *Funct. Ecol.* 21, 394–407.
- Gilbert, B., Levine, J.M., 2013. Plant invasions and extinction debts. *Proc. Natl. Acad. Sci.* 110, 1744–1749.
- Gilbert, K.J., Whitlock, M.C., 2014. Qst - Fst comparisons with unbalanced half-sib designs. *Mol. Ecol. Resour.* 15, 262–267.
- Gioria, M., Osborne, B.A., 2014. Resource competition in plant invasions: emerging patterns and research needs. *Front. Plant Sci.* 5, 1–21.
- Gioria, M., Pyšek, P., 2016. Early bird catches the worm: germination as a critical step in plant invasion. *Biol. Invasions* 19, 1055–1080.
- Godoy, O., Valladares, F., Castro-Díez, P., 2011. Multispecies comparison reveals that invasive and native plants differ in their traits but not in their plasticity. *Funct. Ecol.* 25, 1248–1259.
- Godoy, O., Valladares, F., Castro-Díez, P., 2012. The relative importance for plant invasiveness of trait means, and their plasticity and integration in a multivariate framework. *New Phytol.* 195, 912–922.

- Gombault, C. 1837. « Mémoire sur les robinia pseudo-acacia, viscosa et hispida » in Mémoires de la société royale des sciences, belles-lettres et arts d'Orléans, Orléans, France.
- Grime, J.P., 2001. Plant strategies, vegetation processes, and ecosystem properties, 2nd ed. ed. Wiley, Chichester, West sussex ; New York, NY.
- Guillot, G., Estoup, A., Mortier, F., Cosson, J.F., 2005. A spatial statistical model for landscape genetics. *Genetics* 170, 1261–80.
- Gunther, R.T. 1922. Early British botanists and their gardens, Oxford, England.
- Henry, P., Le Lay, G., Goudet, J., Guisan, A., Jahodová, S., Besnard, G., 2009. Reduced genetic diversity, increased isolation and multiple introductions of invasive giant hogweed in the western Swiss Alps. *Mol. Ecol.* 18, 2819–2831.
- Hewitt, G., 2000. The genetic legacy of the quaternary ice ages. *Nature* 405, 907–913.
- Hierro, J.L., Maron, J.L., Callaway, R.M., 2005. A biogeographical approach to plant invasions : the importance of studying exotics in their introduced and native range. *J. Ecol.* 93, 5–15.
- Hirsch, H., Richardson, D.M., Le Roux, J.J., 2017. Introduction to the special issue: Tree invasions: Towards a better understanding of their complex evolutionary dynamics. *AoB Plants* 9.
- Hirsch, H., Wypior, C., Wehrden, H. Von, Wesche, K., Renison, D., Hensen, I., 2012. Germination performance of native and non-native *Ulmus pumila* populations. *NeoBiota* 68, 53–68.
- Hodgins, K.A., Rieseberg, L., 2011. Genetic differentiation in life-history traits of introduced and native common ragweed (*Ambrosia artemisiifolia*) populations. *J. Evol. Biol.* 24, 2731–2749.
- Hollingsworth, M.L., Bailey, J.P., 2000. Evidence for massive clonal growth in the invasive weed *Fallopia japonica* (Japanese Knotweed). *Bot. J. Linn. Soc.* 133, 463–472.
- Jablonski, N.G., Chaplin, G., 2000. The evolution of human skin coloration. *J. Hum. Evol.* 39, 57–106.
- Jäger, H., Kowarik, I., Tye, A., 2009. Destruction without extinction: Long-term impacts of an invasive tree species on Galápagos highland vegetation. *J. Ecol.* 97, 1252–1263.
- Jombart, T., 2008. adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* 24, 1403–5.
- Jombart, T., Devillard, S., Balloux, F., 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genet.* 11, 94.
- Kaufman, S.R., Smouse, P.E., 2001. Comparing indigenous and introduced populations of *Melaleuca quinquenervia* (Cav.) Blake: response of seedlings to water and pH levels. *Oecologia* 127, 487–494.

- Keane, R.M., Crawley, M.J., 2002. Exotic plant invasions and the enemy release hypothesis. *Trends Ecol. Evol.* 17, 164–170.
- Keller, S.R., Fields, P.D., Berardi, A.E., Taylor, D.R., 2014. Recent admixture generates heterozygosity-fitness correlations during the range expansion of an invading species. *J. Evol. Biol.* 27, 616–627.
- Keller, S.R., Taylor, D.R., 2008. History, chance and adaptation during biological invasion: Separating stochastic phenotypic evolution from response to selection. *Ecol. Lett.* 11, 852–866.
- Keller, S.R., Taylor, D.R., 2010. Genomic admixture increases fitness during a biological invasion. *J. Evol. Biol.* 23, 1720–1731.
- Kennedy, T.A., Naeem, S., Howe, K.M., 2002. Biodiversity as a barrier to ecological invasion. *Nature* 417, 636–638.
- Keresztesi, B., 1983. Breeding and cultivation of black locust, *Robinia pseudoacacia*, in Hungary. *For. Ecol. Manage.* 6, 217–244.
- Kowarik, I., Säumel, I., 2007. Biological flora of Central Europe: *Ailanthus altissima* (Mill.) Swingle. *Perspect. Plant Ecol. Evol. Syst.* 8, 207–237.
- Kurokochi, H., Hogetsu, T., 2014. Fine-scale initiation of non-native *Robinia pseudoacacia* riparian forests along the Chikumagawa River in central Japan. *J. Ecol. Environ.* 37, 21–29.
- Lamarque, L.J., Delzon, S., Lortie, C.J., 2011. Tree invasions: a comparative test of the dominant hypotheses and functional traits. *Biol. Invasions* 13, 1969–1989.
- Lamarque, L.J., Porté, A.J., Eymeric, C., Lasnier, J.-B., Lortie, C.J., Delzon, S., 2013. A Test for Pre-Adapted Phenotypic Plasticity in the Invasive Tree *Acer negundo* L. *PLoS One* 8, e74239.
- Larkin, D.J., 2011. Lengths and correlates of lag phases in upper-Midwest plant invasions. *Biol. Invasions* 14, 827–838.
- Lavergne, C., Rameau, J., Figier, J., 1999. The invasive woody weed *Ligustrum robustum* subsp. *walkeri* threatens native forests on La Réunion. *Biol. Invasions* 1, 377–392.
- Lavergne, S., Molofsky, J., 2007. Increased genetic variation and evolutionary potential drive the success of an invasive grass. *Proc. Natl. Acad. Sci. U. S. A.* 104, 3883–8.
- Lee, C.E., 2002. Evolutionary genetics of invasive species. *Trends Ecol. Evol.* 17, 386–391.
- Leiblein-Wild, M.C., Kaviani, R., Tackenberg, O., 2014. Germination and seedling frost tolerance differ between the native and invasive range in common ragweed. *Oecologia* 174, 739–750.
- Leimu, R., Fischer, M., 2008. A meta-analysis of local adaptation in plants. *PLoS One* 3, 1–8.

- Leinonen, P.H., Sandring, S., Quilot, B., Clauss, M.J., Mitchell-Olds, T., Agren, J., Savolainen, O., 2009. Local adaptation in European populations of *Arabidopsis lyrata* (Brassicaceae). *Am. J. Bot.* 96, 1129–1137.
- Leinonen, T., McCairns, R.J.S., O'Hara, R.B., Merilä, J., 2013. Qst-Fst comparisons: evolutionary and ecological insights from genomic heterogeneity. *Nat. Rev. Genet.* 14, 179–90.
- Leveque de Vilmorin, P.V. 1783. Catalogue des plantes, arbres, arbrisseaux et arbustes dont on trouve des graines, des bulbes et du plant chez les sieurs Andrieux et Vilmorin par Philippe Victoire Leveque de Vilmorin (1783).
- Levine, J.M., 2000. Species diversity and biological invasions: relating local process to community pattern. *Science* (80-. ). 288, 852–854.
- Liesebach, H., Yang, M., Schneck, V., 2004. Genetic diversity and differentiation in a black locust (*Robinia pseudoacacia* L.) progeny test. *For. Genet.* 11, 151–161.
- Little, E.L. 1971. Atlas of United States trees, vol. 1. Conifers and important hardwoods. U.S. Department of Agriculture, Miscellaneous Publication 1146. Washington, DC. 9 p., 313 maps.
- Lonsdale, W.M., 1999. Global Patterns of Plant Invasions and the Concept of Invasibility. *Ecology* 80, 1522–1536.
- Lu, M., Krutovsky, K. V., Nelson, C.D., Koralewski, T.E., Byram, T.D., Loopstra, C.A., 2016. Exome genotyping, linkage disequilibrium and population structure in loblolly pine (*Pinus taeda* L.). *BMC Genomics* 17, 1–11.
- Maron, J.L., Vilà, M., Bommarco, R., Elmendorf, S., Beardsley, P., 2004. Rapid evolution of an invasive plant. *Ecol. Monogr.* 74, 261–280.
- McKay, J.K., Latta, R.G., 2002. Adaptive population divergence: Markers, QTL and traits. *Trends Ecol. Evol.* 17, 285–291.
- Menges, E.S., 2000. Population viability analyses in plants: challenges and opportunities. *Trends Ecol. Evol.* 15, 51–56.
- Merceron, N.R., Leroy, T., Chancerel, E., Romero-Severson, J., Borkowski, D., Ducouso, A., Monty, A., Porté, A.J., Kremer, A., 2017. Back to America: tracking the origin of European introduced populations of *Quercus rubra* L. *Genome* 790, 1–54.
- Merilä, J., Crnokrak, P., 2001. Comparison of genetic differentiation at marker loci and quantitative traits. *J. Evol. Biol.* 14, 892–903.
- Metzker, M.L., 2010. Sequencing technologies the next generation. *Nat. Rev. Genet.* 11, 31–46.
- Michaux, A. 1810. Histoire des arbres forestiers de l'Amérique septentrionale. Paris, France.

- Michener, D.C., 1988. The Introduction of Black Locust (*Robinia pseudoacacia* L.) to Massachusetts. *Arnoldia* 48, 52–57.
- Miller, P. 1756-1759. *The Gardeners Dictionary* (7<sup>th</sup> ed). London, England.
- Monty, A., Lebeau, J., Meerts, P., Mahy, G., 2009. An explicit test for the contribution of environmental maternal effects to rapid clinal differentiation in an invasive plant. *J. Evol. Biol.* 22, 917–926.
- Morimoto, J., Kominami, R., Koike, T., 2010. Distribution and characteristics of the soil seed bank of the black locust (*Robinia pseudoacacia*) in a headwater basin in northern Japan. *Landsc. Ecol. Eng.* 6, 193–199.
- Müller, G.B., 2017. Why an extended evolutionary synthesis is necessary. *Interface Focus* 7, 20170015.
- Nadeau, S., Godbout, J., Lamothe, M., Gros-Louis, M.-C., Isabel, N., Ritland, K., 2015. Contrasting patterns of genetic diversity across the ranges of *Pinus monticola* and *P. strobus*: A comparison between eastern and western North American postglacial colonization histories. *Am. J. Bot.* 102, 1342–1355.
- Naudet. 2017. Pépinières, catalogue 2017-2018. [http://www.pepinieres-naudet.com/fic\\_bdd/administre\\_fichier\\_annexe\\_fr\\_fichier/Catalogue-Pepinieres-Naudet-17-18\\_15124075431.pdf](http://www.pepinieres-naudet.com/fic_bdd/administre_fichier_annexe_fr_fichier/Catalogue-Pepinieres-Naudet-17-18_15124075431.pdf) (accès le 25 juin 2018).
- Narum, S.R., Gallardo, P., Correa, C., Matala, A., Hasselman, D., Sutherland, B.J.G., Bernatchez, L., 2017. Genomic patterns of diversity and divergence of two introduced salmonid species in Patagonia, South America. *Evol. Appl.* 10, 402–416.
- O’Hara, R.B., Merilä, J., 2005. Bias and precision in QST estimates: Problems and some solutions. *Genetics* 171, 1331–1339.
- Oduor, A.M.O., Leimu, R., van Kleunen, M., 2016. Invasive plant species are locally adapted just as frequently and at least as strongly as native plant species. *J. Ecol.* 104, 957–968.
- ONF. 2017. Le catalogue 2017/2018 des semences d’arbres et d’arbustes forestiers. [http://www.onf.fr/produits\\_prestations/++oid++11ec/@@display\\_media.html](http://www.onf.fr/produits_prestations/++oid++11ec/@@display_media.html) (accès le 25 juin 2018).
- Pairon, M., Petitpierre, B., Campbell, M., Guisan, A., Broennimann, O., Baret, P. V., Jacquemart, A.L., Besnard, G., 2010. Multiple introductions boosted genetic diversity in the invasive range of black cherry (*Prunus serotina*; Rosaceae). *Ann. Bot.* 105, 881–890.
- Palacio-López, K., Gianoli, E., 2011. Invasive plants do not display greater phenotypic plasticity than their native or non-invasive counterparts: A meta-analysis. *Oikos* 120, 1393–1401.
- Panetta, F.D., Randall, R.P., 1994. An assessment of the colonizing ability of *Emex australis*. *Aust. J. Ecol.* 19, 76–82.



- Pantoja, P.O., Simón-Porcar, V.I., Puzey, J.R., Vallejo-Marín, M., 2017. Genetic variation and clonal diversity in introduced populations of *Mimulus guttatus* assessed by genotyping at 62 single nucleotide polymorphism loci. *Plant Ecol. Divers.* 10, 5–15.
- Paolucci, E.M., Macisaac, H.J., Ricciardi, A., 2013. Origin matters: Alien consumers inflict greater damage on prey populations than do native consumers. *Divers. Distrib.* 19, 988–995.
- Pappert, R.A., Hamrick, J.L., Donovan, L.A., 2000. Genetic variation in *Pueraria lobata* (Fabaceae), an introduced, clonal, invasive plant of the southeastern United States. *Am. J. Bot.* 87, 1240–1245.
- Pauchard, A., Kueffer, C., Dietz, H., Daehler, C.C., Alexander, J., Edwards, P.J., Arévalo, J.R., Cavieres, L.A., Guisan, A., Haider, S., Jakobs, G., McDougall, K., Millar, C.I., Naylor, B.J., Parks, C.G., Rew, L.J., Seipel, T., 2009. Ain't no mountain high enough: plant invasions reaching new elevations. *Front. Ecol. Environ.* 7, 479–486.
- Pérez-Portela, R., Bumford, A., Coffman, B., Wedelich, S., Davenport, M., Fogg, A., Swenarton, M.K., Coleman, F., Johnston, M.A., Crawford, D.L., Oleksiak, M.F., 2018. Genetic homogeneity of the invasive lionfish across the Northwestern Atlantic and the Gulf of Mexico based on Single Nucleotide Polymorphisms. *Sci. Rep.* 8, 5062.
- Petanidou, T., Godfree, R.C., Song, D.S., Kantsa, A., Dupont, Y.L., Waser, N.M., 2012. Self-compatibility and plant invasiveness: Comparing species in native and invasive ranges. *Perspect. Plant Ecol. Evol. Syst.* 14, 3–12.
- Pigliucci, M., 2006. Phenotypic plasticity and evolution by genetic assimilation. *J. Exp. Biol.* 209, 2362–2367.
- Pimentel, D., McNair, S., Janecka, J., Wightman, J., Simmonds, C., O'Connell, C., Wong, E., Russel, L., Zern, J., Aquino, T., Tsomondo, T., 2000. Economic and environmental threats of alien plant, animal, and microbe invasions 84, 1–20.
- Porté, A.J., Lamarque, L.J., Lortie, C.J., Michalet, R., Delzon, S., 2011. Invasive *Acer negundo* outperforms native species in non-limiting resource environments due to its higher phenotypic plasticity. *BMC Ecol.* 11, 28.
- Postma, F.M., Ågren, J., 2016. Early life stages contribute strongly to local adaptation in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci.* 113, 7590–7595.
- Potter, K.M., Campbell, A.R., Josserand, S.A., Nelson, C.D., Jetton, R.M., 2017. Population isolation results in unexpectedly high differentiation in Carolina hemlock (*Tsuga caroliniana*), an imperiled southern Appalachian endemic conifer. *Tree Genet. Genomes* 13.
- Prentis, P.J., Wilson, J.R.U., Dormontt, E.E., Richardson, D.M., Lowe, A.J., 2008. Adaptive evolution in invasive

- species. *Trends Plant Sci.* 13, 288–294.
- Price, T.D., 2006. Phenotypic plasticity, sexual selection and the evolution of colour patterns. *J. Exp. Biol.* 209, 2368–2376.
- Pritchard, J.K., 2010. Documentation for structure software : Version 2 . 3 6, 321–326.
- Pritchard, J.K., Stephens, M., Donnelly, P.J., 2000. Inference of population structure using multilocus genotype data. *Genetics* 155, 945–949.
- Radtke, A., Ambraß, S., Zerbe, S., Tonon, G., Fontana, V., Ammer, C., 2013. Traditional coppice forest management drives the invasion of *Ailanthus altissima* and *Robinia pseudoacacia* into deciduous forests. *For. Ecol. Manage.* 291, 308–317.
- Rambuda, T.D., Johnson, S.D., 2004. Breeding systems of invasive alien plants in South Africa: does Baker’s rule apply? *Divers. Distrib.* 10, 409–416.
- Randall, J.M., 1997. Defining Weeds of Natural Areas. In: Luken, J.O., Thieret, J.W. (Eds.), . Springer New York, New York, NY, pp. 18–25.
- Rejmánek, M., 2000. Invasive plants : approaches and predictions. *Austral Ecol.* 25, 497–506.
- Ricciardi, A., 2007. Are modern biological invasions an unprecedented form of global change? *Conserv. Biol.* 21, 329–336.
- Ricciardi, A., Gurevitch, J., Padilla, D.K., 2004. Assessing species invasions as a cause of extinction [1] (multiple letters). *Trends Ecol. Evol.* 19, 619–620.
- Richards, C.L., Bossdorf, O., Muth, N.Z., Gurevitch, J., Pigliucci, M., 2006. Jack of all trades, master of some? On the role of phenotypic plasticity in plant invasions. *Ecol. Lett.* 9, 981–993.
- Richardson, D.M., Allsopp, N., D’Antonio, C.M., Milton, S.J., Rejmánek, M., 1999. Plant invasions – the role of mutualisms. *Biol. Rev.* 75, 65–93.
- Richardson, D.M., Pyšek, P., 2006. Plant invasions: merging the concepts of species invasiveness and community invasibility. *Prog. Phys. Geogr.* 30, 409–431.
- Richardson, D.M., Pyšek, P., 2008. Fifty years of invasion ecology – the legacy of Charles Elton. *Divers. Distrib.* 14, 161–168.
- Richardson, D.M., Pyšek, P., Rejmánek, M., Barbour, M.G., Panetta, F.D., West, C.J., 2000. Naturalization and invasion of alien plants : concepts and definitions. *Divers. Distrib.* 6, 93–107.
- Richardson, D.M., Rejmánek, M., 2004. Conifers as invasive aliens: a global survey and predictive framework. *Divers. Distrib.* 10, 321–331.

- Richardson, D.M., Rejmánek, M., 2011. Trees and shrubs as invasive alien species - a global review. *Divers. Distrib.* 17, 788–809.
- Richardson, D.M., Ricciardi, A., 2013. Misleading criticisms of invasion science: A field guide. *Divers. Distrib.* 19, 1461–1467.
- Saiki, R.K., Scharf, S., Faloona, F., Mullis, K.B., Horn, G.T., Erlich, H.A., Arnheim, N., 1985. Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science* 230, 1350–1354.
- Saint-Jean de Crève Coeur, J.H. 1786. Mémoire sur la culture et les usages du faux Acacia dans les États-Unis de l'Amérique septentrionale. Paris, France.
- Santiso, X., López, L., Gilbert, K.J., Barreiro, R., Whitlock, M.C., Retuerto, R., 2015. Patterns of genetic variation within and among populations in *Arbutus unedo* and its relation with selection and evolvability. *Perspect. Plant Ecol. Evol. Syst.* 1–8.
- Säumel, I., Kowarik, I., 2013. Propagule morphology and river characteristics shape secondary water dispersal in tree species. *Plant Ecol.* 214, 1257–1272.
- Silvanus Forestry. <http://silvanusforestry.com/turboakac-en.html> (accès le 25 juin 2018).
- Simberloff, D., 1996. Hybridization between native and introduced wildlife species: importance for conservation. *Wildlife Biol.* 2, 143–150.
- Simberloff, D., 2011. How common are invasion-induced ecosystem impacts? *Biol. Invasions* 13, 1255–1268.
- Simberloff, D., Souza, L., Nunez, M.A., Barrios-Garcia, M.N., Bunn, W., 2012. The natives are restless, but not often and mostly when disturbed. *Ecology* 93, 598–607.
- Sitzia, T., Cierjacks, A., Rigo, D. De, Caudullo, G., 2016. *Robinia pseudoacacia* in Europe: distribution, habitat, usage and threats, European Atlas of Forest Tree Species.
- Sol, D., 2000. Are islands more susceptible to be invaded than continents? Birds say no. *Ecography (Cop.)*. 23, 687–692.
- Spitze, K., 1993. Population structure in *Daphnia obtusa*: Quantitative genetic and allozymic variation. *Genetics* 135, 367–374.
- Straker, K.C., Quinn, L.D., Voigt, T.B., Lee, D.K., Kling, G.J., 2015. Black Locust as a Bioenergy Feedstock: a Review. *Bioenergy Res.* 8, 1117–1135.
- Sultan, S.E., Horgan-Kobelski, T., Nichols, L.M., Riggs, C.E., Waples, R.K., 2013. A resurrection study reveals rapid adaptive evolution within populations of an invasive plant. *Evol. Appl.* 6, 266–278.

- Thomas, C.D., 2013. The Anthropocene could raise biological diversity. *Nature* 502, 7.
- Thompson, G.D., Bellstedt, D.U., Richardson, D.M., Wilson, J.R.U., Le Roux, J.J., 2015. A tree well travelled: Global genetic structure of the invasive tree *Acacia saligna*. *J. Biogeogr.* 42, 305–314.
- Uller, T., Leimu, R., 2011. Founder events predict changes in genetic diversity during human-mediated range expansions. *Glob. Chang. Biol.* 17, 3478–3485.
- Valéry, L., Fritz, H., Lefeuvre, J.C., 2013. Another call for the end of invasion biology. *Oikos* 122, 1143–1146.
- Valladares, F., Sanchez-Gomez, D., Zavala, M. a., 2006. Quantitative estimation of phenotypic plasticity: Bridging the gap between the evolutionary concept and its ecological applications. *J. Ecol.* 94, 1103–1116.
- van Kleunen, M., Dawson, W., Dostal, P., 2011a. Research on invasive-plant traits tells us a lot. *Trends Ecol. Evol.* 26, 317.
- van Kleunen, M., Schlaepfer, D.R., Glaetli, M., Fischer, M., 2011b. Preadapted for invasiveness: do species traits or their plastic response to shading differ between invasive and non-invasive plant species in their native range? *J. Biogeogr.* 38, 1294–1304.
- Verdu, C.F., Guichoux, E., Quevauvillers, S., De Thier, O., Laizet, Y., Delcamp, A., Gévaudant, F., Monty, A., Porté, A.J., Lejeune, P., Lassois, L., Mariette, S., 2016. Dealing with paralogy in RADseq data: in silico detection and single nucleotide polymorphism validation in *Robinia pseudoacacia* L. *Ecol. Evol.* 6, 7323–7333.
- Vidal, O., Sanz, N., Araguas, R.-M., Fernández-Cebrian, R., Diez-del-Molino, D., García-Marín, J.-L., 2012. SNP diversity in introduced populations of the invasive *Gambusia holbrooki*. *Ecol. Freshw. Fish* 21, 100–108.
- Vilà, M., Pino, J., Montero, A., Font, X., 2010. Are island plant communities more invaded than their mainland counterparts? *J. Veg. Sci.* 21, 438–446.
- Visscher, P.M., Hill, W.G., Wray, N.R., 2008. Heritability in the genomics era — concepts and misconceptions. *Nat. Rev. Genet.* 9, 255–266.
- Vítková, M., Müllerová, J., Sádlo, J., Pergl, J., Pyšek, P., 2017. Black locust (*Robinia pseudoacacia*) beloved and despised: A story of an invasive tree in Central Europe. *For. Ecol. Manage.* 384, 287–302.
- Vitousek, P.M., D'Antonio, C.M., Loope, L.L., Westbrooks, R., 1996. Biological Invasions as Global Environmental changes. *Am. Sci.* 84, 468–478.
- von der Lippe, M., Kowarik, I., 2006. Long-Distance Dispersal of Plants by Vehicles as a Driver of Plant Invasions. *Conserv. Biol.* 21, 986–996.
- Wallonie Environnement. 2017. Catalogue des graines forestières.

- [http://environnement.wallonie.be/orvert/docs/Catalogue\\_comptoir\\_complet.pdf](http://environnement.wallonie.be/orvert/docs/Catalogue_comptoir_complet.pdf) (accès le 25 juin 2018).
- Weir, B.S., Cockerham, C.C., 1984. Estimating F-statistics for the Analysis of Population Structure. *Evolution* (N. Y.), 38, 1358–1370.
- Whitlock, M.C., 2008. Evolutionary inference from QST. *Mol. Ecol.* 17, 1885–1896.
- Whitlock, M.C., Guillaume, F., 2009. Testing for spatially divergent selection: Comparing QST to FST. *Genetics* 183, 1055–1063.
- Whitney, K.D., Gabler, C.A., 2008. Rapid evolution in introduced species, “invasive traits” and recipient communities: Challenges for predicting invasive potential. *Divers. Distrib.* 14, 569–580.
- Williams, D.G., Mack, R.N., Black, R.A., 1995. Ecophysiology of Introduced *Pennisetum setaceum* on Hawaii: The Role of Phenotypic Plasticity. *Ecology* 76, 1569–1580.
- Williams, J.L., Auge, H., Maron, J.L., 2008. Different gardens, different results: native and introduced populations exhibit contrasting phenotypes across common gardens. *Oecologia* 157, 239–248.
- Williams, P.A., 2003. Weed management for developing countries. *FAO* 1–15.
- Williamson, M., Fitter, A., 1996. The varying success of invaders. *Ecology* 77, 1661–1666.
- Wolfe, L.M., 2002. Why Alien Invaders Succeed: Support for the Escape-from-Enemy Hypothesis. *Am. Nat.* 160, 705–711.
- Wright, S., 1951. The genetical structure of natural populations. *Ann. Eugen.* 15, 323–354.
- Wu, S.-H., Hsieh, C.-F., Chaw, S.-M., Rejmánek, M., 2004. Plant invasions in Taiwan: Insights from the flora of casual and naturalized alien species. *Divers. Distrib.* 10, 349–362.
- Yang, M., He, Z., Huang, Y., Lu, L., Yan, Y., Hong, L., Shen, H., Liu, Y., Guo, Q., Jiang, L., Zhang, Y., Greenberg, A.J., Zhou, R., Ge, X., Wu, C.I., Shi, S., 2017. The emergence of the hyperinvasive vine, *Mikania micrantha* (Asteraceae), via admixture and founder events inferred from population transcriptomics. *Mol. Ecol.* 26, 3405–3423.
- Zenni, R.D., Dickie, I.A., Wingfield, M.J., Hirsch, H., Crous, C.J., Meyerson, L.A., Burgess, T.I., Zimmermann, T.G., Klock, M.M., Siemann, E., Erfmeier, A., Aragon, R., Montti, L., Le Roux, J.J., 2016. Evolutionary dynamics of tree invasions: complementing the unified framework for biological invasions. *AoB Plants* 9, plw085.
- Zenni, R.D., Lamy, J.B., Lamarque, L.J., Porté, A.J., 2014. Adaptive evolution and phenotypic plasticity during naturalization and spread of invasive species: Implications for tree invasion biology. *Biol. Invasions* 16, 635–644.

*This is the end  
Beautiful friend  
This is the end  
My only friend  
The end  
The end of our elaborate plans  
The end of everything that stands  
The end  
The Doors, The End, 1967*



## Résumé

*Robinia pseudoacacia* L. est un arbre nord-américain qui est largement répandu à travers l'Europe. Afin d'évaluer les mécanismes évolutifs pouvant expliquer son potentiel invasif, il est indispensable d'identifier les populations sources de l'introduction et de comprendre quels traits ont contribué à son succès dans l'aire européenne.

Une étude de génétique des populations a été réalisée, un large échantillonnage a été conduit au sein des deux aires et 818 individus provenant de 63 populations ont été génotypés grâce à 113 marqueurs SNPs. En premier lieu, nous avons démontré que le robinier européen provenait d'un nombre restreint de populations situées au nord du plateau des Appalaches, ce qui est en accord avec les données historiques. Au sein des États-Unis la structure reflète des processus évolutifs au long cours tandis qu'en Europe, la structure est plus ténue et pourrait être due à l'activité humaine, notamment aux programmes de sélection entrepris en Europe centrale depuis le 18<sup>ème</sup> siècle. En deuxième lieu, un goulot d'étranglement génétique associé à une perte de diversité a été mis en évidence dans l'aire d'introduction. Enfin un plus fort taux de clonalité a été observé au sein des populations d'Europe.

Conjointement à cette étude de génétique des populations. Une expérimentation de génétique quantitative a été réalisée afin d'évaluer la différenciation de traits juvéniles entre populations des deux aires (3000 individus de 20 populations) cultivées dans 3 conditions de température différentes (18°C ; 22°C et 31°C). Les résultats ont révélé une augmentation du taux de germination parmi les populations européennes par rapport aux américaines (88% vs 60%) quelles que soient les conditions environnementales. Un scénario possible est que l'Homme aurait sélectionné et propagé les meilleures graines en Europe favorisant alors l'évolution du taux de germination.

De plus, les traits phénotypiques juvéniles sont extrêmement plastiques à la température avec une tendance générale à l'augmentation de la valeur du trait avec la température. Cela suggère qu'un réchauffement climatique favoriserait le développement de l'espèce, au moins en conditions d'alimentation en eau non limitante. Cependant, seules les populations américaines montrent un signal une adaptation locale à la température d'origine.

Dans tous les cas, ils semblent que l'action de l'Homme sur le potentiel reproductif, sexué ou asexué, a probablement influencé le potentiel invasif du robinier en Europe.

## Abstract

*Robinia pseudoacacia* L. is a North American tree which has now broadly spread in Europe. In order to evaluate the evolutionary mechanisms behind its invasiveness, it is crucial to identify the population sources of the introduction and to understand which traits contributed to its success in the European range.

To undertake a population genetics study, we performed a large sampling both in the invasive and native ranges; 818 individuals from 63 populations were genotyped using 113 SNPs. First, we demonstrated that European black locust was introduced from only a limited number of populations located in the Northern plateau of the Appalachians Mountains; this is in agreement with historical records. Within America, population structure reflected long time evolutionary processes whereas in Europe, it was largely impacted by human activities. In the European range, the genetic clustering may be a signal of evolution caused by artificial selection due to human oriented mass selection or tree breeding initiated in Central Europe since the 18<sup>th</sup> century. Second, we evidenced a genetic bottleneck among ranges with a decrease in allelic richness and in the total number of alleles in Europe. Lastly, we found more clonality within the European populations.

Conjointly to the population genetic analysis, we conducted a quantitative genetics experiment to evaluate juvenile traits of both native and invasive black locust populations (3000 individuals from 20 populations) grown under 3 different temperature conditions. Results revealed an enhanced germination rate among European populations compared to that of the native American populations whatever the environmental condition (88% vs 60%). Thus a possible scenario may be that Man would have selected and propagated the best seeds in the new range that would have favored an evolution of germination rate through European populations.

Phenotypic traits of juvenile development are extremely plastic in response to temperature (18/22/31°C), with a general tendency to increasing trait values with increasing temperature. This suggests that global warming would favor the development of the species, at least under non-limiting water supply conditions. However, only American populations demonstrated a signal of local adaptation to the temperature at the sampling location.

In any case, the human role on the reproductive potential, both sexual and asexual, within the introduced range would likely be a key process in the success of black locust dynamics in Europe.