Understanding the origin and predicting adaptive genetic variation at large scale in the genomic era : a case study in maritime pine

Juliette Archambeau

THÈSE PRÉSENTÉE
POUR OBTENIR LE GRADE DE

DOCTEUR
DE L'UNIVERSITÉ DE BORDEAUX

ECOLE DOCTORALE SCIENCES ET ENVIRONNEMENTS
ECOLOGIE ÉVOLUTIVE, FONCTIONNELLE ET DES COMMUNAUTÉS

Par **Juliette ARCHAMBEAU**

Comprendre l'origine et prédire la variation génétique adaptative à large échelle à l'ère de la génomique : une étude de cas chez le pin maritime

Sous la direction de Santiago C. GONZÁLEZ-MARTÍNEZ
Co-directrice Marta BENITO GARZÓN

Soutenue le 17 janvier 2022

Membres du jury :

Dr. Delphine Grivet	Chargée de recherche	INIA (Madrid)	Rapportrice
Dr. Sylvie Muratorio	Directrice de recherche	INRAE (St-Pée-sur-Nivelle)	Rapportrice
Dr. Christian Rellstab	Chargé de recherche	WSL (Birmensdorf)	Examinateur
Pr. Matt Fitzpatrick	Professeur	Université du Maryland (Frostburg)	Examinateur
Dr. Antoine Kremer	Directeur de recherche	INRAE (Bordeaux)	Président du jury
Dr. Santiago C. González-Martínez	Directeur de recherche	INRAE (Bordeaux)	Directeur de thèse
Dr. Marta Benito Garzón	Directrice de recherche	INRAE (Bordeaux)	Co-directrice de thèse

Comprendre l'origine et prédire la variation génétique adaptative à large échelle à l'ère de la génomique : une étude de cas chez le pin maritime

Résumé : Le changement climatique impacte déjà les populations d'arbres forestiers, comme en témoignent les évènements de mortalité de plus en plus fréquents et les migrations vers le nord et en altitude. Cependant, les populations pourraient ne pas migrer assez rapidement face au rythme sans précédent du changement climatique. Par conséquent, à des fins de conservation et de gestion, évaluer le potentiel des populations d'arbres forestiers à persister face au changement climatique est nécessaire. Chez les arbres forestiers, une longue histoire de jardins communs a fourni un cadre unique afin d'associer la variation des traits quantitatifs à de larges gradients environnementaux, permettant ainsi de mieux comprendre l'origine de la variation des traits quantitatifs et d'identifier les populations qui pourraient grandir et survivre mieux, ou moins bien, sous les climats futurs. Les quantités massives de données génomiques provenant des outils de séquençage de nouvelle génération révolutionnent actuellement notre compréhension de la composante génétique des traits quantitatifs et stimulent le développement de nouvelles méthodes statistiques visant à anticiper les réponses des populations aux conditions changeantes. Dans les approches basées sur les traits, la combinaison des données phénotypiques et climatiques des jardins communs avec les données génomiques semble être une approche particulièrement pertinente afin de séparer les composantes plastiques et génétiques de la variation des traits, ainsi que les processus neutres et adaptatifs derrière la composante génétique, ce qui est prometteur vis-à-vis de l'amélioration des prédictions de la variation des traits à grande échelle. En génomique du paysage, les données génomiques et environnementales peuvent être combinées afin d'identifier les relations gènes-environnement actuelles, qui servent ensuite à estimer le changement génétique nécessaire au maintien des relations gènes-environnement dans les climats futurs, une métrique appelée 'décalage génomique'. Dans cette thèse, le pin maritime (*Pinus pinaster* Ait), un conifère à longue durée de vie originaire de la partie occidentale du bassin méditerranéen, est utilisé comme étude de cas afin d'évaluer comment les données génomiques pourraient contribuer à anticiper les réponses des populations au changement climatique. Le premier chapitre vise à comprendre comment la variation génétique quantitative est maintenue au sein des populations en testant trois hypothèses concurrentes, mais non mutuellement exclusives, sur plusieurs traits : (i) les populations admixtes présentent une variation génétique quantitative plus élevée en raison de l'introgression en provenance d'autres pools génétiques, (ii) la variation génétique quantitative est plus faible dans les populations provenant d'environnements plus difficiles (c'est-à-dire subissant une sélection plus forte), et (iii) la variation génétique quantitative est plus élevée dans les populations provenant d'environnements spatialement hétérogènes. Le deuxième chapitre vise à déterminer si des modèles combinant des données climatiques et génomiques pourraient capturer les facteurs sous-jacents de la variation de la croissance en hauteur, et ainsi améliorer les prédictions à grande échelle, en particulier par rapport aux prédictions des fonctions de réponse des populations basées sur le climat qui sont actuellement couramment utilisées chez les arbres forestiers. Le troisième chapitre a pour but d'identifier les populations dont les relations gène-environnement seront les plus perturbées par le changement climatique (c'est-à-dire les populations à risque de maladaptation climatique à court terme) en utilisant l'approche du décalage génomique, et à valider les prédictions qui en résultent (c'est-à-dire que les populations avec un décalage génomique élevé devraient avoir une valeur adaptative plus faible) à la fois dans les populations naturelles et dans des conditions de jardins communs.

Mots-clés : Génétique quantitative et des populations, Arbres forestiers, Modélisation à grande échelle, Génomique du paysage, Variation génétique adaptative, Plasticité phénotypique

Understanding the origin and predicting adaptive genetic variation at large scale in the genomic era: a case study in maritime pine

Abstract: Climate change is already affecting forest tree populations, as evidenced by increased forest die-off events, background mortality and the northward and upward migration of tree populations. However, forest tree populations may not be able to migrate fast enough to track the unprecedented rate of climate change. Therefore, for conservation and breeding purposes, we have to assess the potential of forest tree populations to persist under climate change. In forest trees, a long history of common gardens has provided a unique framework to associate population-specific quantitative-trait variation with large environmental gradients, resulting in a better understanding of the origin of quantitative-trait variation and the identification of populations that may grow and survive better, or worse, under future climates. The huge amount of genomic data from the next-generation sequencing tools is currently revolutionizing our understanding of the genetic component of quantitative traits and is subsequently driving the development of new statistical methods to anticipate the population responses to changing conditions. In trait-based approaches, combining phenotypic and climatic data from common gardens with genomic data appears to be a particularly relevant approach to separate the plastic and genetic components of trait variation, as well as the neutral and adaptive processes behind the genetic component, which is promising towards improving the predictions of trait variation across the species ranges. In landscape genomics, genomic and environmental data can be combined to identify current gene-environment relationships across the landscape, which are then used to estimate the genetic change required to maintain the current gene-environment relationships under future climates, a metric often referred to as genomic offset. In this PhD, maritime pine (*Pinus pinaster* Ait), a long-lived conifer native to the western part of the Mediterranean Basin, is used as a case study to investigate how genomic data could contribute to anticipating population responses to climate change. The first chapter aims to understand how quantitative genetic variation is maintained within populations by testing three competing, but not mutually exclusive, hypotheses for several traits: (i) admixed populations have higher quantitative genetic variation due to introgression from other gene pools, (ii) quantitative genetic variation is lower in populations from harsher environments (i.e. experiencing stronger selection), and (iii) quantitative genetic variation is higher in populations from spatially heterogeneous environments. The second chapter investigates whether models combining climate and genomic data could capture the underlying drivers of height-growth variation, and thus improve predictions at large geographic scales, especially compared to the predictions from climate-based population response functions that are currently commonly used in forest trees. The third chapter aims to identify the populations whose gene-environment relationships will be the most disrupted under climate change (i.e. populations at risk of short-term climate maladaptation) using the genomic offset approach, and to validate the resulting predictions (i.e. populations with high genomic offset are expected to show a decrease in fitness) both in natural populations and in common garden conditions. Finally, the present PhD work investigates different ways to integrate genomic information into current modeling approaches, therefore contributing to the development of a much-needed robust framework to make reliable predictions and to determine when and to what extent genomics can help in making decisions in conservation strategies or in anticipating population responses to climate change.

Keywords: Population and quantitative genetics, Forest trees, Large-scale modeling, Landscape genomics, Adaptive genetic variation, Phenotypic plastiticty

Remerciements

Avant d'aborder les questions scientifiques qui m'ont captivée ces trois dernières années, j'aimerais remercier un certain nombre de personnes m'ayant, de près ou de loin, épaulée au cours de cette aventure. Grâce à vous, j'ai passé trois années mémorables et extrêmement enrichissantes tant du point de vue scientifique, intellectuel, qu'humain.

J'ai mis beaucoup de cœur à rendre la lecture de cette thèse agréable, mais je dois admettre qu'il y a probablement mieux comme livre de chevet durant les fêtes de fin d'année. J'aimerais donc remercier chaleureusement les deux rapportrices, **Sylvie Muratorio** et **Delphine Grivet**. Je remercie également les examinateurs de mon jury de thèse qui m'ont fait l'honneur d'évaluer mon travail, **Antoine Kremer**, **Matt Fitzpatrick** et **Christian Rellstab**.

Un merci qui ne sera jamais assez grand pour mes deux merveilleux et remarquablement complémentaires encadrants de thèse, **Marta Benito-Garzón** et **Santiago C. González-Martínez**.

Marta, il y a maintenant cinq années que j'ai frappé à ta porte à la recherche d'un stage de master 1 en modélisation. Ah, je m'en souviendrai de mon été 2016 à tenter de comprendre la signification des warnings de R et des raisons pour lesquelles ces satanés modèles de Hurdle ne convergeaient pas ! Grâce à toi Marta, j'ai pu partir en 2017 à Alcalá de Henares dans le cadre de mon stage de master 2, un séjour extraordinaire et formateur, notamment marqué par la découverte des statistiques bayésiennes. J'y ai rencontré **Sophia Ratcliffe** et surtout **Paloma Ruiz-Benito**, que je tiens également à remercier ; merci pour votre bienveillance, votre passion contagieuse de la recherche, votre rigueur et toutes les connaissances et compétences que vous m'avez apportées en statistiques. Marta, sans toi cette thèse n'aurait jamais eu lieu puisque c'est ensemble que nous avons pris la décision de se lancer dans l'aventure et de tenter le concours de l'École doctorale au mérite. J'ai vite perdu le compte de toutes les répétitions que tu avais organisées pour préparer le concours ! Enfin, tu m'as permis de travailler dans des conditions matérielles excellentes et tu as toujours été extrêmement réactive et à l'écoute, jusqu'à même t'excuser quand tu ne pouvais pas relire dans la journée ce que je venais juste de t'envoyer… quel scandale !

Santi, nous avons mis un peu de temps à trouver un bon mode de fonctionnement, mais une fois trouvé, c'est devenu un réel plaisir et une source de motivation de travailler avec toi. Tu m'as énormément appris et encouragée à me surpasser, et je trouve qu'on forme une paire de choc pour la conception des papiers et la rédaction (même si, oh détresse, il faut pour cela que j'utilise Google Docs). J'ai adoré ton optimisme (je dirais même ton *quokka*-optimisme), ton humour, ton honnêteté intellectuelle, nos réunions alliant discussions scientifiques et courses à pied, le terrain au milieu des tiques et des ajoncs, et j'en passe. Santi, merci pour tout, je suis vraiment heureuse de t'avoir eu comme encadrant et j'espère que nous aurons l'occasion de travailler encore longtemps ensemble !

J'aimerais également adresser un merci particulier à **Benoit Pujol**, qui était dans mon comité de thèse mais qui me suit en réalité depuis maintenant neuf ans. En effet, c'est avec toi que j'ai fait mon premier stage dans la recherche académique, à suivre les populations de mufliers (et expérimenter les brûlures de la rue officinale !). En nous parlant avec passion de tes recherches sur les rivages ventés de Bages, tu as éveillé en moi un intérêt vif pour la

biologie de l'évolution, et plus généralement, pour le monde de la recherche. Intérêt qui n'a pas décru puisque j'ai ensuite passé quatre étés consécutifs dans ton équipe, naviguant entre mise en place et maintien de dispositifs expérimentaux contrôlés, biologie moléculaire, suivi de populations naturelles et familiarisation avec les bases de la génétique quantitative. Et surtout, Benoit, tu m'as rattrapé *in extremis* alors que j'étais sur le point d'arrêter mes études à un passage à vide de ma vie, me redonnant l'élan et la motivation qui ne m'ont jamais quittée depuis. Je t'en serais toujours profondément reconnaissante et tu resteras toujours pour moi un mentor, une source d'inspiration.

Je ne peux parler de mes stages avec Benoit sans évoquer **Sara Marin**, une membre de l'équipe qui est devenue une amie chère. Que de bons souvenirs avec toi sur le terrain, dans le jardin commun-camping à Mérens-les-Vals, le jardin commun-limace à Garin, dans la voiture avec la musique à fond. Merci pour ta rigueur et ton sérieux lors des manips, ta bonne humeur et ton soutien à de nombreuses étapes de ma vie.

Je tiens également à remercier les deux autres membres de mon comité de thèse, **Laura Leites** et **Stephen Cavers** pour le temps qu'ils m'ont consacré et leurs précieux conseils et encouragements.

En comptabilisant les deux stages de master, je serais restée plus de cinq ans au sein du laboratoire BIOGECO, dont quasiment un an et demi de télétravail dû à la pandémie de Covid-19. C'est dans ce contexte que je tiens à brosser un tableau (forcément incomplet) des collègues, pour certains devenus des amis proches, qui auront marqué positivement mon passage à BIOGECO.

Je ne peux que commencer par **Frédéric Barraquand**. Fred, je suis venue te voir très vite au début de ma thèse car tu étais 'le matheux du premier étage qui fait des stats bayésiennes'. Je suis entrée dans ton bureau en ayant une question technique sur une histoire d'effets fixes et aléatoires dans mes modèles, j'en suis ressortie nageant dans un flou complet quant aux objectifs desdits modèles... Cette première interaction déroutante avec un matheux bayésien ne m'a finalement pas découragée puisque, trois ans plus tard, tu as contribué à la partie modélisation de tous mes papiers de thèse et est devenu un ami proche. Tu m'auras énormément appris en statistiques, mais plus largement en termes de méthode scientifique, de rédaction et d'organisation. Tu as été très présent lorsque j'avais des baisses de moral, notamment pendant les périodes de confinement, durant lesquelles nos petits cafés hebdomadaires étaient plus que ressourçants. Sans toi, cette thèse aurait pris une tournure très différente c'est certain, merci pour tout Fred.

Le lien est vite fait entre le matheux bayésien et le petit groupe de motivés qui s'est construit progressivement afin de s'entraider pour progresser dans la programmation de modèles bayésiens avec le langage *Stan*. J'ai pris énormément de plaisir à participer et à animer nos workshops, que je trouvais toujours très intéressants et stimulants, et qui ont été particulièrement bénéfiques pour mon moral en périodes d'isolement dues au Covid-19. Merci donc au petit noyau dur des *Stan*-users de BIOGECO, **Sylvain Schmitt**, **Simon Labarthe**, **Coralie Picoche**, **Frédéric Barraquand** et **Guillaume Ravel**. Et un merci particulier à Sylvain Schmitt et Coralie Picoche pour m'avoir aider à apprivoiser Github.

Autres événements ayant eu un effet très positif sur ma motivation et l'avancée de mes recherches, les réunions de l'équipe E4E. **Benjamin Brachi**, un grand merci à toi d'avoir mis

en place et poussé au maintien de ces réunions stimulantes scientifiquement et très chouettes d'un point de vue humain.

Virgil Fievet, merci pour les - très très très - longues discussions poussant à la réflexion, les notions de biologie évolutive que tu m'auras apprise (même longtemps après le master !), le thé népalais qui donne mal à la tête, les conseils en escalade et dans la vie en général, qui, bien qu'ils me fassent d'abord râler, s'avèrent toujours avisés.

Myriam Heuertz, un immense merci pour m'avoir relancée dans la course à pied. Je garderai des souvenirs géniaux de nos escapades sportives matinales, de ton banana bread et de tes falafels, de ton courtisan le dindon, des randonnées avec Helena et Flor, des soirées à papoter et tellement d'autres moments chouettes.

Sophie Gerber, ou la philosophe-actrice-amie des plantes et des vélos-palme d'or pour la préparation de la vinaigrette, merci pour l'ouverture vers la philosophie en m'entraînant dans des conférences passionnantes (alors, les végétaux sont-ils sentients finalement ?), les mails amusants ('Finalement, soyons dégenrées, dérengées'), les pauses partagées.

Charlie Pauvert et **Tania Fort**, mes colocataires de bureau dans la vie avant-covid, les journées étaient bien plus tristes une fois que vous êtes partis. Heureusement qu'il y a encore sur les murs du bureau tes dessins et devises Shadok, Tania, et ta courbe de productivité, Charlie, c'est comme si vous étiez encore un peu là. Merci pour les séances de sport, en visio avec Tania pendant les confinements, les footings à la pause midi avec Charlie. Charlie, un merci particulier pour m'avoir motivée à apprendre le LaTeX(et pour le template du manuscrit de thèse, rien que ça !), pour tes conseils en programmation (et pour devenir un as des pains au levain), ton aide pour débugger mon ordinateur (quand j'étais en panique) ou installer la nouvelle version de Debian (le dimanche, quand j'étais de nouveau en panique).

Alexandre Changenet, merci pour les séances de cardio-training-footing inoubliables, tes houmous délicieux (sauf celui à la feta ! !) et surtout pour les échanges autour de la pensée critique qui m'ont amenée à complètement changer ma manière de mettre à jour mes connaissances, raisonner, débattre, échanger. Tu n'es jamais à cours de proposition de contenu à écouter, lire, regarder, c'est fantastique de discuter avec toi. Tu as été une réelle source d'inspiration et d'admiration, surtout, ne change rien héhé.

Deborah Corso, merci de ne pas m'avoir laissée seule finir ma thèse au rez-de-chaussée du B2. Merci pour les séances de sport en visio avec Tania lors des confinements, quand se refait-on le jeu de cartes ? **Thomas Caignard**, merci pour tes encouragements réguliers, les parties de Codenames et la relecture du manuscrit de thèse. **Marine Cambon**, merci pour ta compagnie brève mais fun, et pour les conseils *Tikz*. **Mathieu Crétet**, merci pour les nombreux joggings (mon premier marathon !) le soir à la frontale (avec les salamandres, les crapauds, les chevreuils et les sangliers) ou le dimanche dans le froid matinal, les discussions toujours chouettes et drôles ; tu m'as manqué quand tu es parti !

Enfin, un merci à tous les autres collègues de l'Université pour les discussions à la pause café ou autres circonstances : **Xavier Bouteiller** (notamment pour les conseils en statistique et programmation au début de la thèse), **Annabel Porté** (sacrée coach pour la préparation du concours de l'École doctorale), **Thibaut Fréjaville**, **Natalia Vizcaíno-Palomar**, **Alice Sauve**, **Laurent Lamarque**, **Anne-Isabelle Gravel**, **Sylvain Delzon**, **Gaëlle Capdeville**, **Régis**

Burlett, **Corinne Vacher**, etc. Ainsi que les collègues de Pierroton, en particulier **Marina de Miguel** (un immense merci notamment pour le formatage des données CLONAPIN) et **Katharina Budde**.

Qu'on se le dise, on n'avancerait sûrement pas aussi loin dans la vie sans les amis. Merci **Marie Savignac** (*Guiit-Guiit*) pour la relecture du manuscrit de thèse, les appels réguliers depuis l'Afrique et les séjours ressourçants à la Coulinière. Merci **Aila José** pour égayer de ta bonne humeur notre appartement, les séances de brossage de dents, de sport, de ménage et les discussions pendant les balades de Falco. Merci **Tanguy Aubé** pour les weekends maraîchages (particulièrement les après-midis à étaler du fumier). Merci **Cédric Lemaire** (et les parents Lemaire, et Mireille !) pour les séjours à Saint-Bazerque, et **Pierre Lemaire** pour les séances d'escalade. Merci **Thibaut Poiret** pour avoir été le premier à me traîner à la salle d'escalade et pour les discussions sur les équilibres de Nash, l'altruisme efficace et l'illusionnisme. Merci à **Valentin Hivert** pour les explications sur la décomposition de la variance génétique et les skypes depuis Brisbane. Merci à **Julien Dlubala** pour les séances de sport en visio. Merci à **Maxime de Ronne** pour les skypes depuis Montréal et les ballades au bois de Pau. Merci aux amis sportifs du trail (en particulier **Nicole Bach**) et de l'escalade.

Un merci spécial à **François Pero** (et la famille Pero) qui m'a épaulée pendant cinq années, fait réciter nombre de présentations, aidée à relativiser, tout en me faisant partager sa passion de la musique et du cinéma. Je te dois énormément, merci pour ta patience, ton soutien et tous les beaux moments partagés ensemble.

Merci à ma famille et à mes parents, pour votre soutien sans faille, vos encouragements, vos petits messages réguliers, votre patience, les livraisons de purée d'amande, les tartes aux poireaux de Papa et les taboulés de Maman.

Enfin, merci à **Benjamin Sozeau** pour l'élan de motivation que tu m'as insufflé sur la fin de la thèse, tes précieux encouragements, les moments de détente et sportifs, les sorties en montagne 'incroyables' (vivement les prochaines !). Merci d'avoir été là, tout simplement, et d'être prêt et optimiste pour les prochaines aventures.

CONTENTS

I

Résumé substantiel

1 Introduction

Les arbres forestiers sont des espèces clés de voûte essentielles au fonctionnement et au maintien des écosystèmes, de la biodiversité et de multiples services écosystémiques. Prédire comment les populations d'arbres forestiers s'adapteront *in situ* aux conditions environnementales futures, notamment celles engendrées par le changement climatique, est aujourd'hui un enjeu critique et urgent, nécessitant une compréhension approfondie des processus évolutifs en jeu.

En outre, certaines populations ne seront pas en mesure de s'adapter assez rapidement face au rythme du changement climatique et sont donc susceptibles de connaître des déclins démographiques, voire des extinctions, dans un avenir proche. Identifier de telles populations en amont et les classer par ordre de priorité est nécessaire afin de mettre en œuvre des stratégies de conservation et de gestion pertinentes. De plus, implémenter des stratégies tenant compte des processus adaptatifs, comme le déplacement d'individus vers des environnements au sein desquels ils sont supposés être mieux adaptés (stratégie de flux génétique assisté) ou vers des populations menacées en manque de variation génétique (stratégie de sauvetage évolutif), nécessite d'anticiper la réponse des individus transplantés aux nouveaux environnements.

Chez les arbres forestiers, une longue histoire de test de provenances (désormais communément appelés jardins communs) a fourni un cadre unique pour associer la variation des traits quantitatifs entre populations à de larges gradients environnementaux. L'estimation de fonctions de réponse des populations ('*population response functions*') a permis d'évaluer l'origine de la variation des traits quantitatifs (notamment en séparant la part plastique et génétique de la variation des traits) et d'identifier les populations présentant des risques de lags phénotypiques face au changement climatique, et donc potentiellement à risque de maladaptation (par exemple Fréjaville et al. 2020, Pedlar et McKenney 2017, Rehfeldt et al. 2002, Savolainen et al. 2007). Cependant, les jardins communs sont coûteux et chronophages à entretenir, et limités à quelques espèces ou populations. Parallèlement, la disponibilité croissante des données génomiques issues des nouvelles technologies de séquençage à des coûts abordables (et en constante diminution) pour les espèces non modèles offre de nouvelles perspectives pour comprendre les processus adaptatifs, identifier les populations à risque de maladaptation ou améliorer les prédictions des traits quantitatifs à l'échelle individuelle. En conséquence, en biologie de l'évolution et génétique quantitative, les approches de modélisation statistique incorporant des informations génomiques se développent rapidement (par exemple Gienapp

et al. 2017, Meuwissen et al. 2001), mais la plupart manquent encore de validations robustes et indépendantes, et peuvent être considérablement améliorées.

Les principaux objectifs de la présente thèse visent à contribuer à deux préoccupations majeures du domaine de la biologie évolutive : (i) la compréhension des mécanismes sous-jacents à l'adaptation des populations à leur environnement local, et (ii) l'amélioration des prédictions des réponses des populations aux changements environnementaux, tels que le changement climatique. Le pin maritime (*Pinus pinaster* Ait), un conifère à longue durée de vie originaire de la partie occidentale du bassin méditerranéen, est utilisé comme étude de cas. Le chapitre 1 vise à comprendre comment la variation génétique quantitative est maintenue au sein des populations de pin maritime en testant trois hypothèses concurrentes, mais non mutuellement exclusives, sur plusieurs traits phénotypiques : (i) les populations à forts niveaux d'introgression entre différents pools génétiques présentent une variation génétique quantitative plus élevée, (ii) la variation génétique quantitative est plus faible dans les populations issues d'environnements plus rudes (car subissant potentiellement une sélection plus forte ; Fisher 1930), et (iii) la variation génétique quantitative est plus élevée dans les populations issues d'environnements spatialement hétérogènes (McDonald et Yeaman 2018, Yeaman et Jarvis 2006). Le chapitre 1 renseigne donc sur les populations qui pourraient être en mesure de s'adapter plus rapidement face au changement climatique puisque le potentiel adaptatif des populations est proportionnel à leur variation génétique ('*breeder's equation*' ; Falconer et Mackay 1996, Lush 1937). Le chapitre 2 a pour but de déterminer si des modèles combinant des données génomiques, climatiques et phénotypiques peuvent capturer les facteurs sous-jacents de la variation de croissance en hauteur, et ainsi améliorer les prédictions phénotypiques à large échelle, en particulier en comparaison avec les prédictions des fonctions de réponse des populations basées sur le climat d'origine des populations et qui sont actuellement généralement utilisées pour les arbres forestiers (par exemple Leites et al. 2012a, Rehfeldt et al. 1999). Enfin, le premier objectif du chapitre 3 est d'identifier les populations de pin maritime dont les relations gène-environnement seront les plus altérées par le changement climatique (c'est-à-dire les populations à risque de maladaptation climatique à court terme) en appliquant l'approche du décalage génomique ('*genomic offset*' ; Fitzpatrick et Keller 2015). Le second objectif est de tester une hypothèse clé de l'approche du décalage génomique, à savoir que les populations pour lesquelles les prédictions de décalage génomique sont les plus élevées présentent une diminution de leur valeur adaptative absolue moyenne ou des tendances démographiques en déclin (Capblancq et al. 2020a). Les chapitres 2 et 3 fournissent donc des informations précieuses à la mise en œuvre d'une gestion des populations de pin maritime tenant compte des processus adaptatifs. De plus, en étudiant comment combiner les données phénotypiques, génomiques et environnementales dans deux cadres de modélisation très différents (respectivement basés sur les traits et la génomique du paysage), ces deux chapitres contribuent à l'objectif ambitieux de prédire comment les populations d'arbres forestiers répondront au changement climatique et quelles stratégies de gestion et de conservation seront les plus efficaces pour sauver les populations en déclin.

2 Matériels & méthodes

L'espèce modèle utilisée dans cette thèse est le pin maritime (*Pinus pinaster* Ait., Pinaceae), une espèce d'arbre forestier écologiquement et économiquement importante, largement exploitée pour son bois (Viñas et al. 2016), stabilisant les dunes côtières atlantiques et, en tant qu'espèce clé de voûte, soutenant la biodiversité forestière. Originaire de la partie occidentale du bassin méditerranéen, des montagnes de l'Atlas au Maroc et de la côte atlantique sud-ouest

de l'Europe, sa répartition naturelle s'étend des montagnes du Haut Atlas au sud (Maroc) à la Bretagne française au nord, et de la côte du Portugal à l'ouest à l'Italie occidentale à l'est. Il a également été introduit à des fins commerciales en Australie où il est désormais considéré comme une espèce hautement invasive (Viñas et al. 2016).

Le pin maritime est une espèce d'arbre pollinisé par le vent, allogame et à longue durée de vie. Il pousse sur une grande variété de substrats, des sols sableux et acides aux sols plus calcaires. Il peut également tolérer des climats variés : le climat sec des côtes nord du bassin méditerranéen (du Portugal à l'Italie occidentale), le climat montagneux du sud-est de l'Espagne et du Maroc, le climat plus humide de la région atlantique (de la région ibérique espagnole à l'ouest de la France) et le climat continental du centre de l'Espagne. Comme de nombreuses espèces d'arbres méditerranéennes, le pin maritime présente une forte structure de population et des populations très fragmentées (Alberto et al. 2013). Ses populations peuvent être regroupées en six pools génétiques (Jaramillo-Correa et al. 2015), c'est-à-dire des groupes génétiques qui ne peuvent être différenciés sur la base de marqueurs génétiques neutres et qui dérivent probablement d'un refuge glaciaire commun (Bucci et al. 2007, Santos-del-Blanco et al. 2012).

Les données phénotypiques et génomiques utilisées dans la présente thèse proviennent du réseau de jardins communs clonaux CLONAPIN, composé de cinq sites situés dans des environnements différents. Trois sites (Bordeaux, Asturias et Portugal) se trouvent dans la région atlantique, caractérisée par des hivers doux, des précipitations annuelles élevées et des étés relativement humides. Les deux autres sites (Cáceres et Madrid) sont localisés dans la région méditerranéenne continentale, caractérisée par des étés chauds et très secs et des hivers froids. En 2010 ou 2011 selon le site, des réplicats clonaux de 34 populations ont été plantés selon un plan expérimental en blocs aléatoires complets. Entre 2 et 28 clones (génotypes), en moyenne 15, représentaient chaque population . Pour obtenir des clones non apparentés, des arbres distants d'au moins 50 m ont été échantillonnés dans des peuplements naturels, et une graine par arbre a été plantée dans une pépinière et propagée végétativement par bouturage (voir Rodríguez-Quilón et al. 2016 pour plus de détails).

La mortalité et la hauteur des arbres ont été mesurées dans tous les jardins communs et à différents âges des arbres : 10, 21 et 37 mois à Asturias, 25, 37 et 49 mois à Bordeaux (ainsi que 13 mois pour la mortalité et 85 mois pour la hauteur), 8 mois à Cáceres, 13 mois à Madrid et 11, 15, 20 et 27 mois au Portugal. La hauteur des arbres n'a été mesurée que sur les arbres vivants, déséquilibrant considérablement les données de hauteur à Cáceres et à Madrid, où 92% et 75% des arbres sont morts, respectivement (en partie à cause des sols argileux et d'une forte sécheresse estivale). Deux traits liés à la phénologie, la date moyenne de débourrement et la durée moyenne du débourrement, ont été mesurés à Bordeaux lorsque les arbres avaient 2, 3, 4 et 6 ans. Le débourrement correspond à la date d'émergence des brachyblastes en degrés-jours cumulés (avec une température de base de 0°C) à partir du premier jour de l'année, ce qui permet de tenir compte de la variabilité interannuelle des températures. La durée du débourrement correspond au nombre de degrés-jours entre le début de l'élongation des bourgeons et l'élongation totale des aiguilles (voir Hurel et al. 2019). Enfin, deux traits fonctionnels, δ^{13}C et la surface foliaire spécifique (SLA), ont été mesurés au Portugal.

Les 34 populations plantées dans les cinq jardins communs représentent un échantillon de populations naturelles couvrant l'ensemble des pools génétiques connus du pin maritime. 523 clones collectés dans le jardin commun se trouvant à Asturias ont été génotypés avec le test Infinium d'Illumina, ce qui a permis d'obtenir 5 165 SNPs polymorphes de haute qualité. Il n'y avait en moyenne que 3,3 valeurs manquantes par génotype (entre 0 et 142). Des détails

sur l'extraction de l'ADN et le génotypage peuvent être trouvés dans Plomion et al. (2016b). Ce premier ensemble de données génomiques a été utilisé dans les chapitres 1 et 2. Dans le chapitre 3, nous avons combiné ce premier ensemble de données génomiques avec un autre obtenu avec le test de génotypage d'Affymetrix et développé dans le cadre du projet H2020 EU B4EST (4Tree ; https://b4est.eu). Les SNPs ayant une fréquence des allèles mineurs inférieure à 1% ou plus de 20% de données manquantes ont été filtrés, ce qui a permis d'obtenir 454 clones et 9 817 SNPs polymorphes de haute qualité. Parmi ces derniers, 2 855 étaient génotypés par les deux tests de génotypage (Infinium d'Illumina et Affymetrix), garantissant ainsi l'identité de l'échantillon et permettant d'estimer les erreurs de génotypage. Le pourcentage de données manquantes par clone était inférieur à 12% pour tous les clones, avec une moyenne de 2,5%.

3 Chapitre 1

La plupart des traits complexes présentent une variation héritable substantielle dans les populations naturelles. Comment l'interaction des forces évolutives maintient une telle variation reste un dilemme de longue date en biologie évolutive et en génétique quantitative, qui a fait l'objet d'un vaste corpus de travaux théoriques mais manque de preuves empiriques à ce jour (Johnson et Barton 2005). Alors que la mutation et la migration augmentent la variation génétique au sein des populations, la sélection naturelle et la dérive génétique sont supposées l'appauvrir (Walsh et Lynch 2018). Seulement, la variation génétique élevée des populations naturelles est difficile à expliquer sans tenir compte d'autres processus, tels que la sélection balancée au sein d'environnements hétérogènes (Mitchell-Olds et al. 2007). Chez les arbres forestiers, à ma connaissance, seules deux études ont examiné empiriquement comment les forces évolutives façonnent la variation génétique au sein des populations, et suggèrent un effet positif de l'hétérogénéité environnementale (Yeaman et Jarvis 2006) et un effet négatif de la sélection induite par le climat (Ramírez-Valiente et al. 2019).

Dans ce chapitre, nous avons testé des hypothèses concurrentes sur l'origine et le maintien de la variation génétique quantitative au sein des populations de pin maritime. Nous avons utilisé des mesures phénotypiques de traits de croissance (hauteur), phénologiques (date et durée du débourrement) et fonctionnels (δ^{13}C et surface foliaire spécifique, SLA) issues des trois jardins communs du réseau CLONAPIN situés dans la région Atlantique (Bordeaux, Asturias et Portugal). Les mesures phénotypiques ont été réalisées sur 522 clones (génotypes) provenant de 33 populations, couvrant tous les pools génétiques connus de l'espèce (Jaramillo-Correa et al. 2015) et génotypés pour 5 165 SNPs (Plomion et al. 2016b). Pour chaque trait, nous avons comparé les estimations de modèles hiérarchiques bayésiens de l'association entre variance génétique au sein des populations et facteurs sous-jacents potentiels, à savoir la rigueur du climat sur le lieu d'origine des populations (intensité de la sécheresse et épisodes de froid intense), l'hétérogénéité environnementale dans les régions environnantes des populations, et le degré et origine du mélange génétique dans les populations (estimés avec les marqueurs SNPs). Les hypothèses concurrentes, mais non mutuellement exclusives, testées étaient les suivantes : (i) les populations présentant les plus forts degrés d'introgression en provenance d'autres pools génétiques ont une variation génétique quantitative plus élevée, et cette relation est proportionnelle à la divergence entre les pools génétiques sources et puits ; (ii) la variation génétique quantitative est plus faible dans les populations qui ont évolué dans des environnements plus rudes, en raison des pressions de sélection plus élevées dans ces régions ; et (iii) la variation génétique quantitative est plus élevée dans les populations qui ont évolué dans des environnements spatialement hétérogènes.

Le résultat le plus intéressant de ce chapitre était que les populations de pin maritime soumises à des événements de froid intense présentaient une variation génétique plus faible pour la hauteur dans les trois jardins communs. Ce résultat appuie l'hypothèse selon laquelle la variation génétique quantitative des traits liés à la valeur adaptative (fitness) est plus faible dans les populations soumises à une forte sélection (Fisher 1930), ici une sélection induite par le climat. Ce résultat a été validé sur des données de hauteur indépendantes provenant d'un ensemble supplémentaire de dispositifs expérimentaux (aimablement fournies par des collègues). La robustesse de nos résultats a également été confirmée par l'absence d'association entre le degré d'introgression des populations et leur variation génétique quantitative, suggérant l'absence d'influence du flux génétique entre des pools de gènes distincts sur la variation génétique des traits considérés. En revanche, nous n'avons trouvé pour aucun des traits étudiés de variation génétique plus élevée dans les populations situées dans des environnements hétérogènes, ce qui va à l'encontre des prédictions de certains modèles théoriques (McDonald et Yeaman 2018, Walsh et Lynch 2018) et d'une étude empirique chez le pin tordu (Yeaman et Jarvis 2006).

En conclusion, ce chapitre contribue au débat sur le maintien de la variation génétique au sein des populations en apportant un appui empirique au rôle de la sélection naturelle dans la réduction de la variation génétique au sein des populations d'un arbre forestier à longue durée de vie. Plus largement, la variation génétique étant une brique essentielle de la réponse adaptative des populations à des changements de conditions environnementales, ce chapitre renseigne sur le potentiel adaptatif à court terme des populations, ce qui est d'une grande utilité pour prédire quelles populations sont en mesure de s'adapter rapidement face au changement climatique.

4 Chapitre 2

Anticiper la croissance des individus et populations dans de nouveaux environnements est essentiel pour guider les stratégies de conservation des arbres forestiers, notamment les translocations d'individus visant à compenser le changement climatique rapide (Aitken et Whitlock 2013). À ce jour, les fonctions de réponse des populations basées sur le climat d'origine des populations restent la méthode la plus couramment utilisée afin d'anticiper les valeurs des traits des populations transplantées dans de nouveaux environnements (Fréjaville et al. 2020, O'Neill et al. 2008, Pedlar et McKenney 2017, Rehfeldt et al. 2003, 1999, Wang et al. 2010). L'intégration des informations génomiques dans les modèles prédictifs des traits d'intérêts apparaît attrayante car elle permettrait de distinguer les contributions relatives de la variation génétique adaptative ou neutre dans les prédictions, et de prendre en compte la variabilité intraspécifique à une échelle plus fine que les modèles actuels, gagnant ainsi en précision de prédiction (Holliday et al. 2017).

L'objectif de ce chapitre était d'identifier les facteurs sous-jacents potentiels des composantes plastiques et génétiques de la croissance en hauteur des populations de pins maritimes et d'étudier comment les données phénotypiques des jardins communs peuvent être combinées avec des données génomiques afin d'améliorer les prédictions de la variation de la croissance en hauteur à l'échelle de l'aire de répartition de l'espèce. Nous avons comparé des modèles hiérarchiques bayésiens inférant les variations de croissance en hauteur du pin maritime en fonction de variables climatiques et génomiques, en utilisant les mesures de hauteur issues de 34 populations (523 génotypes et 12 841 arbres) plantées dans les cinq jardins communs du réseau CLONAPIN. Nous avons d'abord évalué l'importance relative des facteurs sous-

jacents potentiels des variations de croissance en hauteur. Nous nous attendions à ce que : (i) la composante plastique (environnementale) explique la plus grande partie de la variation des traits et soit associée au climat des jardins communs, (ii) la composante génétique soit déterminée à la fois par des processus adaptatifs, tels que l'adaptation au climat, et des processus neutres, tels que l'histoire démographique des populations. Deuxièmement, nous avons comparé la capacité de prédiction hors échantillon (sur des observations ou des populations non incluses lors du fit des modèles) de modèles basés exclusivement sur le design expérimental des jardins communs avec celle de modèles incluant (séparément ou conjointement) des prédicteurs potentiels de la composante génétique de la croissance en hauteur. Ces prédicteurs potentiels incluaient le climat d'origine des populations (un indicateur de l'adaptation au climat), l'assignement de chaque génotype aux différents pools génétiques (un indicateur de l'histoire démographique des populations et de la dérive génétique, reflétant probablement aussi les différentes histoires sélectives des pools génétiques) et des comptages d'allèles spécifiques à chaque génotype et ayant un effet positif sur la hauteur ('*positive-effet alleles*', PEAs ; identifiés via des études d'association pangénomique, 'genome-wide association studies', GWAS).

La composante plastique expliquait la proportion majeure des écarts à la trajectoire moyenne de croissance en hauteur (47%), ayant probablement pour origine de multiples facteurs environnementaux (confondus), dont le climat. La composante génétique expliquait 11% des déviations de la trajectoire moyenne de croissance en hauteur et était principalement associée au climat d'origine des populations dont l'effet est partiellement confondu avec l'assignement à des pools génétiques distincts. Les modèles combinant informations génomiques et climatiques capturaient bien la composante génétique de la croissance en hauteur. De façon importante, ils prédisaient mieux la croissance en hauteur de nouvelles populations (non incluses lors du fit des modèles) que les modèles basés exclusivement sur le design expérimental des jardins communs (c'est à dire uniquement sur les données phénotypiques) ou les modèles incluant séparément informations climatiques et génomiques (comme les fonctions de réponse des populations basées uniquement sur le climat d'origine des populations). Il est également intéressant de relever que les PEAs qui avaient été identifiés à une échelle régionale dans les GWAS avaient une plus grande capacité de prédiction que les PEAs identifiés globalement à l'échelle de l'aire de répartition de l'espèce.

Ce chapitre est un pas de plus vers l'intégration des connaissances récentes apportées par les avancées de la génomique à la modélisation de la variation des traits quantitatifs chez les arbres forestiers. La combinaison des jardins communs avec les outils génomiques est très prometteuse afin d'accélérer et améliorer les prédictions de traits à grande échelle et pour un large éventail d'espèces et de populations. Cependant, un cadre solide est nécessaire afin de générer des prédictions fiables et de déterminer quand et dans quelle mesure la génomique peut aider à prendre des décisions dans les stratégies de conservation ou à anticiper les réponses des populations au changement climatique.

5 Chapitre 3

Un objectif majeur de la biologie de l'évolution est de comprendre comment les populations s'adaptent à leur environnement et de prédire comment elles répondront aux conditions futures, en particulier celles découlant du changement climatique. L'approche du décalage génomique ('*genomic offset*') est de plus en plus populaire et vise à identifier les populations pour lesquelles les relations gène-environnement seront les plus perturbées face aux nouvelles conditions

climatiques, c'est-à-dire les populations à risque de maladaptation climatique à court terme (Fitzpatrick et Keller 2015, Rellstab et al. 2015). Elle apparaît comme une méthode prometteuse pour guider les stratégies de conservation et de gestion, en particulier pour les espèces sessiles et à longue durée de vie comme les arbres forestiers, pour lesquelles l'adaptation *in situ* ou la migration des allèles adaptatifs peuvent ne pas être assez rapides face au rythme du changement climatique (Fitzpatrick et Keller 2015). Cependant, cette approche repose sur un certain nombre d'hypothèses clés nécessitant une validation empirique solide (discutées dans Capblancq et al. 2020a, Rellstab et al. 2021).

L'objectif principal de ce chapitre était double : (1) identifier les populations de pins maritimes à risque de maladaptation climatique à court terme, (2) vérifier l'hypothèse selon laquelle les populations présentant les prédictions de décalage génomique les plus élevées présentent une diminution de leur valeur adaptative absolue ou des tendances démographiques en déclin (Capblancq et al. 2020a). Pour cela, une première étape de validation a consisté à détecter des associations négatives entre performance des populations (hauteur et taux de mortalité) et prédictions de décalage génomique dans des jardins communs (plutôt que dans les climats futurs), et à les comparer aux associations entre performance des populations et cinq distances de transfert climatique (différence absolue entre le climat des populations et celui des jardins communs). Une deuxième étape de validation a consisté à estimer les associations entre les taux de mortalité récents dans les populations naturelles (sur la base des données des inventaires forestiers nationaux français et espagnols) et les prédictions de décalage génomique sous les climats futurs, en supposant que les populations dont les prédictions suggèrent une maladaptation climatique dans un futur proche connaissent déjà des taux de mortalité plus élevés que la moyenne. Les prédictions de décalage génomique dans les jardins communs et les populations naturelles ont été dérivées pour toutes les combinaisons possibles de quatre ensembles de SNPs (un ensemble de SNPs de référence et trois ensembles de SNPs candidats plus ou moins strictement sélectionnés à l'aide de deux analyses d'association génétique-environnement, GEAs), deux approches de modélisation en génomique du paysage ('*Gradient Forest*', GF, et '*Generalised Dissimilarity Modelling*', GDM) et deux scénarios climatiques futurs plus ou moins alarmants (uniquement pour les prédictions dans les populations naturelles).

En ce qui concerne l'objectif (1), les prédictions de décalage génomique basées sur les SNPs candidats communs aux deux méthodes GEAs (c'est-à-dire ceux qui ont été sélectionnés avec le plus de confiance) indiquent un risque de maladaptation plus élevé pour les populations qui connaissent actuellement des conditions hivernales douces (la plupart des populations atlantiques et les populations méditerranéennes du sud-est de la France et du nord-ouest de l'Italie), mais pour lesquelles la transition vers des températures légèrement plus élevées impliquerait une étape évolutive importante. Le risque de maladaptation climatique dans un futur proche de ces populations pourrait s'expliquer par leur adaptation passée aux températures hivernales froides, contraignant une croissance et une survie optimales en cas de réchauffement des températures. Il est également important de noter que ces populations, qui se trouvent à l'extrémité chaude du gradient des températures froides hivernales, ne pourront pas bénéficier de la migration d'allèles adaptatifs provenant d'autres populations. Il semble donc crucial de suivre leur dynamique démographique et leurs trajectoires adaptatives dans les années à venir, sachant notamment que les populations du sud-ouest de la France et du nord-est de l'Ibérie sont celles qui ont la plus grande valeur commerciale. Ainsi, une maladaptation climatique affectant leurs traits phénotypiques d'intérêt pourrait avoir un impact substantiel sur l'économie locale.

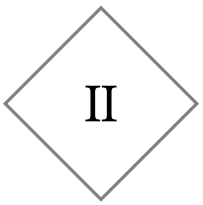
En ce qui concerne l'objectif (2), les prédictions de décalage génomique étaient généralement négativement associées à la performance des populations dans les jardins communs et les

populations naturelles, suggérant ainsi que les prédictions de décalage génomique peuvent être indicatives de déclins (futurs) de la valeur adaptative des populations, et donc validant nos résultats chez le pin maritime. Néanmoins, les prédictions de décalage génomique étaient très sensibles aux ensembles des SNPs considérés (c'est-à-dire à la rigueur de leur sélection en tant que SNP candidat) et à l'approche de modélisation utilisée (GDM vs GF), alors qu'elles étaient très similaires entre les deux scénarios de climats futurs. En particulier, aucune des modalités testées pour prédire le décalage génomique n'avait une meilleure capacité prédictive à travers toutes les étapes de validation.

Nos résultats confirment donc que l'approche du décalage génomique est prometteuse, mais suggèrent également qu'une validation plus poussée de ses prédictions, notamment basée sur des données expérimentales et d'observations indépendantes, est nécessaire. En particulier, déterminer quelles méthodes de modélisation et quels critères de sélection de la composante génétique adaptative conduisent aux prédictions les plus robustes et fiables possibles est indispensable avant que de telles prédictions soient utilisées pour guider les stratégies de conservation et de gestion forestière.

6 Discussion

Utilisant le pin maritime comme étude de cas, la présente thèse combine de façon innovante des approches de modélisation basées sur les traits phénotypiques (chapitres 1 et 2) avec des approches de génomique du paysage (chapitre 3). S'appuyant sur les données phénotypiques, environnementales et génomiques provenant d'un vaste réseau de cinq jardins communs et 34 populations, l'ensemble des résultats obtenus fournissent (i) une image globale et détaillée des patrons d'adaptation du pin maritime à large échelle et des processus évolutifs sous-jacents, et (ii) une évaluation du risque de maladaptation à court terme des populations de pin maritime face au changement climatique. Les conclusions des différents chapitres convergent vers le rôle clé des températures froides dans l'histoire adaptative du pin maritime, ayant un impact à la fois sur les états adaptatifs actuels des populations, mais aussi potentiellement sur la variance de certains traits quantitatifs au sein même des populations (par exemple la hauteur des arbres). Les données génomiques apparaissent comme particulièrement prometteuses pour améliorer les prédictions des réponses à court terme des populations à des changements environnementaux, notamment climatiques. En effet, le chapitre 2 montre que les prédictions phénotypiques pour des individus transférés dans de nouveaux environnements sont améliorées en incorporant des informations génomiques dans les modèles, ce qui est d'un grand intérêt pour les stratégies de conservation ou de gestion (par exemple, flux de gène assisté ou sauvetage évolutif). De plus, le chapitre 3 met en évidence que l'approche du décalage génomique peut s'avérer être un outil très pertinent dans l'identification rapide des populations à risque de maladaptation climatique dans un futur proche, nécessitant cependant une validation empirique plus poussée avant sa généralisation. Plus généralement, la présente thèse contribue à une meilleure compréhension des processus adaptatifs, au développement de méthodes statistiques robustes nécessaires à la mise en œuvre de stratégies de gestion basées sur l'évolution, et à la progression vers l'objectif ambitieux mais urgent de prédire la réponse des populations au changement climatique.



Introduction

How do species and populations adapt to their environment? This question has held the attention of countless scientists for centuries and is at the heart of this PhD work. Four centuries BC, Aristotle, who believed in 'finalism' and 'fixism', had already observed that species were particularly well adapted to their habitats (Gelber 2015) and described adaptation as a *state*. Mechanisms of populations' adaptation to their environment only began to be elucidated centuries later, after the work of Charles Darwin and Alfred Russel Wallace, who believed in 'transformism' and saw adaptation as a *process* (Darwin 1859, Darwin and Wallace 1858). Depending on the context, the term **adaptation** is thus defined as the average phenotypic change that enhances fitness and has a genetic basis (i.e. adaptation as a *process*) or any condition/trait that enhances fitness in a given environment relative to other possible conditions/traits in that environment (i.e. adaptation or **adaptive trait** as a *state*) (Hendry 2017). Today, new sequencing technologies provide access to the genetic basis of adaptive traits, allowing us to investigate past adaptations but also to follow evolution in real time. We are therefore living in an extremely exciting time scientifically as these new genomic data open the way to tremendous progress in our understanding of the evolutionary processes. Yet, reading the genome is not as simple as reading a book and we remain far from elucidating the mechanisms underlying the genotype-to-phenotype relationship. We are also living in worrying times, as global change induced by humans is already causing the extinction of many species and populations around the world (Butchart et al. 2010, Steffen et al. 2011). Thus, the question of how species adapt to their environment is now of primary importance and understanding the adaptive processes is necessary to anticipate how populations will adapt in the future. This is the context of this PhD. In the introduction, I will start by presenting the concepts and mechanisms of population and quantitative genetics necessary to understand the work presented here. I will then discuss in the second part the underlying mechanisms of the population responses to environmental changes, with a particular focus on changing climatic conditions. In the third part, I will discuss some current statistical methods integrating genomic data in predictive modelling of short-term population response to new environments. Finally, the fourth part will be dedicated to the specificities of forest trees, which are excellent models to study adaptation to the environment but also present important challenges.

1 Concepts and mechanisms in population and quantitative genetics

'There is no thing more practical than a good theory' (Lewin 1943, McCain 2015).

1.1 From Fisher's infinitesimal model to the omnigenic model

In the 1900's, a bitter debate was ongoing between the **Mendelians**, who were interested in monogenic (discrete) phenotypes and thought that evolution occurred via major new mutations, and the **biometricians**, who were interested in the inheritance and variation of continuous traits and believed that evolution consisted in very small steps. Fisher's *infinitesimal model* brilliantly settled the debate by showing that if many genes contribute to phenotypic variation, then, according to the central limit theorem, random sampling of alleles at each gene produces continuous and normally distributed phenotypes in the population (Fisher 1918). This model states that quantitative traits are determined by an infinitely large number of genes, each with infinitely small and additive contributions to the phenotype, and by environmental factors (Fisher 1918). It has been highly effective in describing inheritance patterns, especially in plant and animal breeding, and formed the basis for quantitative genetics in the future (Visscher and Goddard 2019). In particular, the theory built in Fisher's 1918 paper allows the partitioning of phenotypic variance into genetic and environmental components, a point developed in detail in the next section.

In the genomic era, Fisher's *infinitesimal model* has proved highly successful in the face of accumulating empirical observations from new genomic tools, which now provide cheap genotyping of hundreds of thousands of common allelic markers or even whole genome sequencing for more and more species (see Section 3.1). **Genome-wide association studies** (GWAS) have been widely used to statistically associate genetic variants (usually SNPs, single nucleotide polymorphisms) with quantitative traits. They have resulted in the discovery of a huge number of adaptive variants in humans (Sella and Barton 2019), but also in other species (e.g. in plant model species; Brachi et al. 2011), thus confirming that quantitative traits are under the control of a large number of genes (i.e. polygenic), each having a small effect on the phenotype (Tam et al. 2019, Visscher et al. 2017). Genetic variants with large effect sizes may also influence quantitative traits but they are extremely rare and are often associated with diseases that have a strong impact on fitness (Gibson 2012), such as autism and schizophrenia (De Rubeis et al. 2014, Purcell et al. 2014).

GWAS results have also highlighted that genetic variants are generally associated with multiple phenotypes, thus suggesting pervasive **pleiotropy** in quantitative traits (Gratten and Visscher 2016). This is supported by the widespread genetic correlations among traits observed in pedigree studies, implying that sets of genetic variants affect two or more traits in a consistent direction (Visscher et al. 2017). Other noteworthy observations come from GWAS outputs: (i) most GWAS hits are noncoding variants probably influencing gene regulation (Li et al. 2016); (ii) genetic variants are spread broadly across the genome (e.g. variants significantly associated with height can be found almost every 100 kb on the genome; Boyle et al. 2017), which is supported by the correlation between the length of a chromosome and its heritability (Shi et al. 2016); (iii) genes with putatively relevant functions have only marginally higher genetic contributions to phenotypes (Boyle et al. 2017); and (iv) cell type-specific regulatory elements

and generically active regions contribute almost equally to genetic variance (Boyle et al. 2017). All these observations have paved the way for the recent development of a new conceptual framework for understanding the genetic architecture of complex traits: the *omnigenic model* (Boyle et al. 2017, Liu et al. 2019; criticized in Wray et al. 2018). Under this model, genetic contributions are partitioned into direct effects from core genes and indirect effects from peripheral genes, which are far more numerous and drive the expression of core genes via weak *trans* effects, thus explaining most of the genetic variance (Liu et al. 2019). Extending the omnigenic model to the *omni-environmental model* has been recently proposed to consider that some environmental factors have direct effects on phenotypes, which are likely to be constant across populations, while others have more peripheral effects, which are likely to vary unpredictably (Mathieson 2021).

The genetic architecture of complex traits will not be explored further in this manuscript. However, readers may find it useful to bear in mind that behind the genetic component of quantitative trait variation lie extremely interconnected gene regulatory networks, whose functioning and organization are still very poorly understood. We will now explore in more details how variation in quantitative traits can be partitioned.

1.2 Components of quantitative trait variation

As previously mentioned, the phenotypic variance (V_P) can be partitioned into genetic (V_G) and environmental (V_E) components, plus their potential interactions (V_{GE}) and covariances (2cov_{GE}), such as:

$$V_P = V_G + V_E + V_{GE} + 2\text{cov}_{GE}$$

In practice, covariances can often be neglected in randomized experiments, which break down potential correlations between environmental deviations and genotypic values (Falconer and Mackay 1996). Note that V_{GE} should be included in the environmental variance as, although the environmental variance is specific to each genotype, the source of the variation is environmental, and not genetic (Falconer and Mackay 1996).

The genetic variance (V_G) can itself be partitioned into the **additive** (V_A), **dominance** (V_D) and **epistatic** (V_I) variances (Falconer and Mackay 1996, Lynch and Walsh 1998), such as:

$$V_G = V_A + V_D + V_I$$

The additive variance is the variance of the **breeding values** (i.e. the additive genetic value of an individual based on the mean additive genetic value of its progeny), which accounts for the influence of the additive effects of the alleles on the phenotype. This has to be differentiated from V_G, the variance of the **genetic values**, which includes non-additive effects, such as dominance and epistasis. The dominance deviations stem from within-locus interactions while the epistatic deviations come from among-locus interactions. Importantly, all these quantities depend on the gene frequencies and therefore are properties of a given population (Falconer and Mackay 1996).

The relative contributions of dominance and epistasis to genetic variance have been debated in the face of the seemingly conflict between on the one hand, estimates of genetic variance suggesting the overall predominance of additive variance (Falconer and Mackay 1996, Lynch and Walsh 1998) and on the other hand, the deepening understanding of gene networks and

interactions, suggesting extensive epistasis (Carlborg and Haley 2004). However, this debate may stem mainly from two sources of confusion: (i) statistical epistasis (i.e. interaction variance due to deviation from additive effects, as coined by Fisher 1918) does not imply functional epistasis (i.e. a biological phenomenon in which the effect of a particular locus depends on the genotype at another locus, as described by Bateson 1909), and vice versa (Cordell 2002); (ii) models estimating additive effects do not assume the absence of gene interactions, since by definition the average effect of an allele is a function of both dominance and epistasis effects (Hivert et al. 2021b). Recent empirical and theoretical work suggests that interactions at the gene level are unlikely to generate much interaction variance, and, therefore, that the bulk of genetic variance is additive (Hill et al. 2008, Hivert et al. 2021a). An exception can be noted in the case of inbred lines with high levels of heterozygosity, which maximize the variance from non-additive effects (Hivert et al. 2021b).

Based on the partitioning of the phenotypic variance, two key parameters can be estimated for a given population: H^2, the **broad-sense heritability** or **degree of genetic determination** (the ratio of genetic variance to total variance) and h^2, the **narrow-sense heritability** (the ratio of additive variance to total variance). H^2 reflects to what extent the phenotypes of individuals are determined by their genotypes while h^2 reflects to what extent the phenotypes are determined by the genes passed on by their parents. h^2 is therefore the major determinant of the resemblance among relatives and a key component of the short-term population response to selection (Falconer and Mackay 1996). A general trend across species is that h^2 is higher for morphological traits and lower for life-history traits, which are more closely related to fitness (Charmantier and Garant 2005, Merilä and Sheldon 2000).

The environmental variance (V_E) can also itself be partitioned into the **special environmental variance** (V_{Es}), the **general environmental variance** (V_{Eg}) and the **genotype-by-environment interaction** (V_{GE}), such as:

$$V_E = V_{Es} + V_{Eg} + V_{GE}$$

The special environmental variance refers to the within-individual variance, which originates from two main sources: **developmental noise** and **temporal fluctuations**. Developmental noise, also called stochastic developmental variation, refers to phenotypic variation that is not explained by genetic or environmental factors. It results from stochastic cellular and developmental process, which occur during cleavage, cell differentiation, patterning or morphogenesis but also tissue regeneration and life history attributes in adulthood (Vogt 2015). The temporal fluctuations refer to phenotypic variation of successive measurements on the same individual, e.g. variation in phenology-related traits among years in perennial plants, variation in milk yield or number of offspring among litters in cattle (Falconer and Mackay 1996). In plant and animal breeding, the special environmental variance is considered as random noise in phenotype expression, which cannot be eliminated by experimental design and therefore interferes with artificial trait selection by weakening the association between the genotype and the phenotype. In contrast, this variation is of particular interest in evolutionary biology because it could itself be the object of selection. For instance, proteins responding to environmental changes show higher expression noise than those involved in protein synthesis, suggesting that protein expression noise levels may be under selection (Newman et al. 2006). Similarly, higher within-individual variation in labile traits can be selected in fluctuating environments, thus allowing individuals to respond more flexibly to changing environmental conditions across their lifespan (Westneat et al. 2015).

The general environmental variance refers to the between-individual variance caused by external variations in the environment, e.g. nutritional or climatic factors (Falconer and Mackay 1996). It typically constitutes most of the variance in natural populations and can be partially controlled under experimental conditions. A potentially large part of the general environmental variance comes from maternal effects, i.e. prenatal or postnatal causal influences of the mother on the phenotypes of the offspring (Wolf and Wade 2009).

Genotype-by-environment interaction refers to the variance in the response of individual genotypes to general environmental variance. It may arise from changes in variance or in the ranking of genotype performance across different environments, i.e. one genotype may perform better than other genotypes in one environment but worse in other environments (de Jong 1990, Falconer and Mackay 1996, Lynch and Walsh 1998). Classic examples include the genotype-specific changes in bristle number across changing temperatures in *Drosophila* (Gupta and Lewontin 1982), the genotype-specific larval development of a leaf-mining insect on two different host plant species (Via 1984) and the genotype-specific growth of tobacco plants in environments with varying sowing dates and plant densities (Falconer and Mackay 1996).

Individual responses of genotypes to the environment are modeled with **reaction norms**, in which trait values are a function of an environmental variable. Fig. II.1 shows three possible configurations. In the first panel, the two genotypes do not show phenotypic responses to environmental changes and their differences in trait expression are purely genetic. In the second panel, the two genotypes respond similarly to environmental changes, i.e. they do not show genetic differences in trait expression along the environmental gradient. In such a case, the phenotypic variance comes entirely from the general environmental variance and there is no genotype-by-environment interaction variance. In the third panel, the two genotypes respond differently to environmental changes, i.e. there is genetic variation in trait expression along the environmental gradient (genotype-by-environment interaction). In this simplified example, reaction norms are linear but more realistic and informative reaction norms can be fitted with nonlinear functions (Arnold et al. 2019).

General environmental variance, special environmental variance associated to temporal fluctuations and the genotype-by-environment interaction are all generated by a mechanism called **phenotypic plasticity**. Phenotypic plasticity is defined as the ability of a genotype to express different phenotypes across environments, i.e. one genotype may code for different environment-dependent phenotypes (DeWitt and Scheiner 2004). Fig. II.1 can thus be interpreted as follows: genotypes show no phenotypic plasticity in the first panel, genotypes show the same plastic response to the environment in the second panel, and genotypes show a different plastic response to the environment in the third panel (i.e. genetic variation in phenotypic plasticity). In quantitative genetics, phenotypic plasticity was first considered as a source of noise affecting the precision of genetic studies and leading to unpredictable performance of genotypes in untested environments (Bradshaw 2006, Pigliucci 2005). In contrast, it is now considered as a rapid-response process of major importance for individuals to cope with changing environmental conditions, which may even constrain or boost adaptive processes (Fox et al. 2019, Nicotra et al. 2010; see Section 2.3).

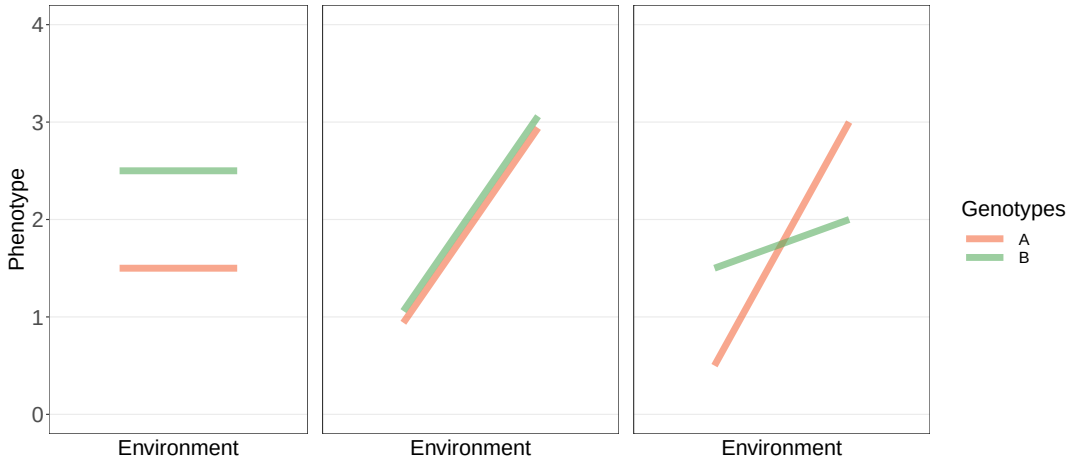


Figure II.1. Simplified example of linear reaction norms for two genotypes. Phenotypic variation comes from: genetic differences among genotypes in the first panel, environmental differences in the second panel (i.e. plastic component of trait variation) and interacting genetic and environmental differences in the third panel (i.e. genetic variation in the plastic response to the environment).

In the present PhD, the focus was mainly on understanding the genetic rather than the plastic component of quantitative trait variation. That's why I will now concentrate mainly on the various forces and mechanisms driving the genetic component of quantitative trait variation.

1.3 Evolutionary forces underlying quantitative genetic variation

A population or a species is evolving when its genetic composition (often measured by allele frequencies) is changing over time. Four main evolutionary processes affect the genetic composition of populations, and thereby their quantitative genetic variation: mutation, genetic drift, gene flow and natural selection.

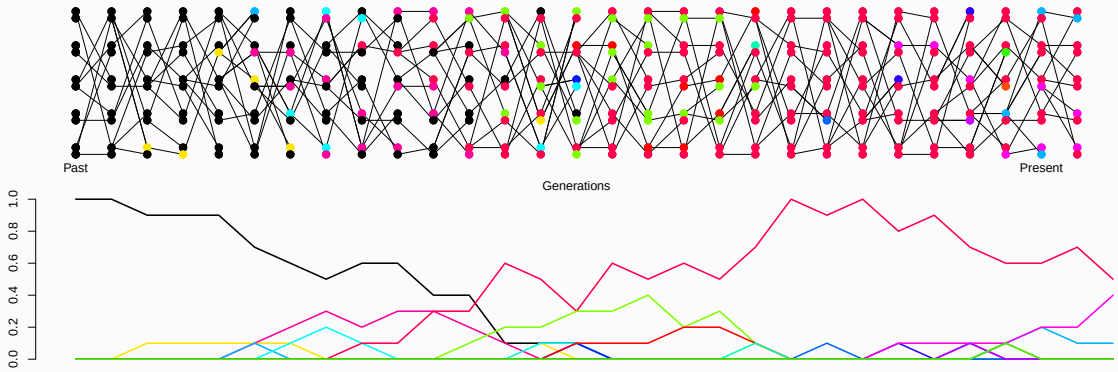
Mutations originate from errors in the replication process of DNA sequences and are the only evolutionary force generating new alleles (i.e. new genetic variants). Mutations are rare, with mutation rates about 10^{-5} and 10^{-6} per generation for most loci in most organisms (Falconer and Mackay 1996). A mutation appearing in an individual has a high probability of being lost (with a zero probability of survival in an infinite population). Mutations therefore generate tiny changes in allele frequency (and thereby tiny increases in genetic variation), which might be important on an evolutionary scale but is difficult to detect on a ecological timescale (except in microorganisms). Importantly, for a mutation to be heritable, it has to be passed to a gamete (i.e. a reproductive cell, that is a haploid cell carrying a single copy of each chromosome).

Genetic drift is the random changes in allele frequency in finite populations due to sampling error between generations. For instance, the gametes of a sexually reproducing diploid organism contain only one copy of each gene, so that only one of the two copies is transmitted to the offspring. This reshuffling of the combinations of genes between parents and offspring is achieved through the process of **recombination**. In a broad sense, recombination is any genetic mechanism (e.g. independent assortment and crossing overs during meiosis, gene conversion) that can create new combinations of alleles or haplotypes (Templeton 2006). Importantly, recombination (and therefore genetic drift) is neutral, as it does not change allele frequencies in any specific direction over time, and results in a progressive erosion of genetic variation.

The role of genetic drift in the allele dynamics, and therefore population and species evolution, has been a matter of intense debate (Ohta and Gillespie 1996), which persists to date (Jensen et al. 2019, Kern and Hahn 2018). The *neutral theory of molecular evolution* states that the vast majority of new mutations are either neutral (e.g. mutations arising in non-coding regions of the DNA) or deleterious (e.g. mutations disrupting important protein functions), and have therefore low probability of becoming fixed in the population (Kimura 1968). Whether the neutral theory is right or wrong will not be discussed further here, what is relevant is that it can serve as a null model from which various hypotheses can be tested (e.g. Box 1).

Box 1. Mutation-drift balance

Here is an example of a simple null model derived from the neutral theory: the **mutation-drift balance**. While mutation is generating new genetic variation, genetic drift slowly erodes neutral (and weakly selected) genetic variation, as alleles drift to high or low frequencies until they get lost or fixed over time. In a panmictic diploid population of size N and mutation rate μ, drift dominates whenever $4\mu N < 1$ (i.e. resulting in the fixation of most alleles) while mutation dominates whenever $4\mu N > 1$ (i.e. maximizing genetic variation). The figure below shows an example of the dynamic of the two evolutionary forces in a very small population with high mutation rate, in which mutation-drift balance consists of a constantly evolving set of alleles maintaining an equilibrium level of polymorphism.



An initial population of five diploid individuals, each having the same black allele, evolves under 29 generations. Each transmitted allele can mutate between generations (thereby changing colour). In this example, the mutation rate is implausibly high to counter-balance the strong effect of genetic drift (as the population is very small). The bottom plot shows the allelic frequencies over time. From the 'Population and quantitative genetics' course of Graham Coop (University of California, Davis). Code available here.

Gene flow (e.g. through migrating individuals or dispersal of reproductive material such as gametes or seeds) can change the genetic composition of a population and increase its genetic variation through alleles carried by immigrants from the surrounding populations, or decrease its genetic variation by losing alleles carried by emigrants (Slatkin 1985). At the metapopulation level, gene flow homogenizes allele frequencies among populations.

The combination of the neutral (i.e. not affected by natural selection) evolutionary processes presented above (i.e. mutation, genetic drift and gene flow) and the demographic history of the populations generates **population structure** across the species ranges. This is particularly true in species with fragmented populations (reduced gene flow and higher effect of genetic drift in small populations) and has to be accounted for in studies aiming at detecting adaptation patterns (e.g. the case study of human height differences across Europe; Barton et al. 2019).

Natural selection can trigger evolution when (i) there is phenotypic variation, (ii) **fitness** (i.e. defined here as the amount of successful DNA replication of an individual; Templeton 2006) is non-random with respect to this phenotypic variation and (iii) this phenotypic variation is heritable. Importantly, changes in allele frequency induced by natural selection tend to increase the average fitness of the population, resulting in **adaptation** (Templeton 2006).

Unlike other evolutionary forces, the influence of natural selection on quantitative genetic variation within populations is far more complex and still under debate (Pélabon et al. 2010, Walsh and Lynch 2018). Selection is expected to deplete genetic variation when the fitness function is concave, e.g. under stabilizing selection, but increase it when the fitness function is convex, e.g. under disruptive selection (Layzer 1980). Linear directional selection is unlikely to influence genetic variation of complex traits with near-Gaussian distributions (Barton and Turelli 1987), except in the presence of directional epistasis (i.e. when epistasis consistently affects the effect of an allele in a given direction; Hansen et al. 2006). Finally, how selection can maintain genetic variation (i.e. balancing selection) remains unclear and has been the subject of extensive theoretical work, albeit lacking empirical validation (Delph and Kelly 2014, Johnson and Barton 2005). For example, high levels of genetic variation may be maintained under selection pressures that fluctuate in time or space (Felsenstein 1976, McDonald and Yeaman 2018), although some mechanisms may mitigate the expression of genetic variation under such conditions (e.g. genetic canalization; Kawecki 2000).

2 How populations respond to environmental change

2.1 (Mal)adaptation within the adaptive landscape framework

How populations respond to their environment can be conceptualized within the framework of the **phenotype adaptive landscape**, a n-dimensional surface (for n traits) describing the relationship between mean population fitness and mean population phenotypes, assuming a constant variance in phenotypic traits (Arnold et al. 2001, Hendry and Gonzalez 2008, Lande 1976, 1979, Schluter 2000b, Simpson 1944, Svensson and Calsbeek 2012). The direction and steepness of the surface at a given location on the landscape reflect the change over time in the mean phenotype and mean fitness of a population (Hendry 2017), which is shown in Fig. II.2 with blue vectors tangent to an hypothetical adaptive landscape. Far from the optimum (phenotype value Z_A in Fig. II.2), the selection is stronger and therefore the evolution towards the optimum faster, especially for traits with higher additive genetic variance relative to their phenotypic variance (i.e. higher heritability). Importantly, selection decreases as the mean population phenotype gets closer to the optimum (phenotype value Z_B in Fig. II.2), so that evolution towards the optimum is slower and selection becomes almost undetectable once the population have adapted to a fitness peak (selection 'erases its traces'; Estes and Arnold 2007, Haller and Hendry 2014).

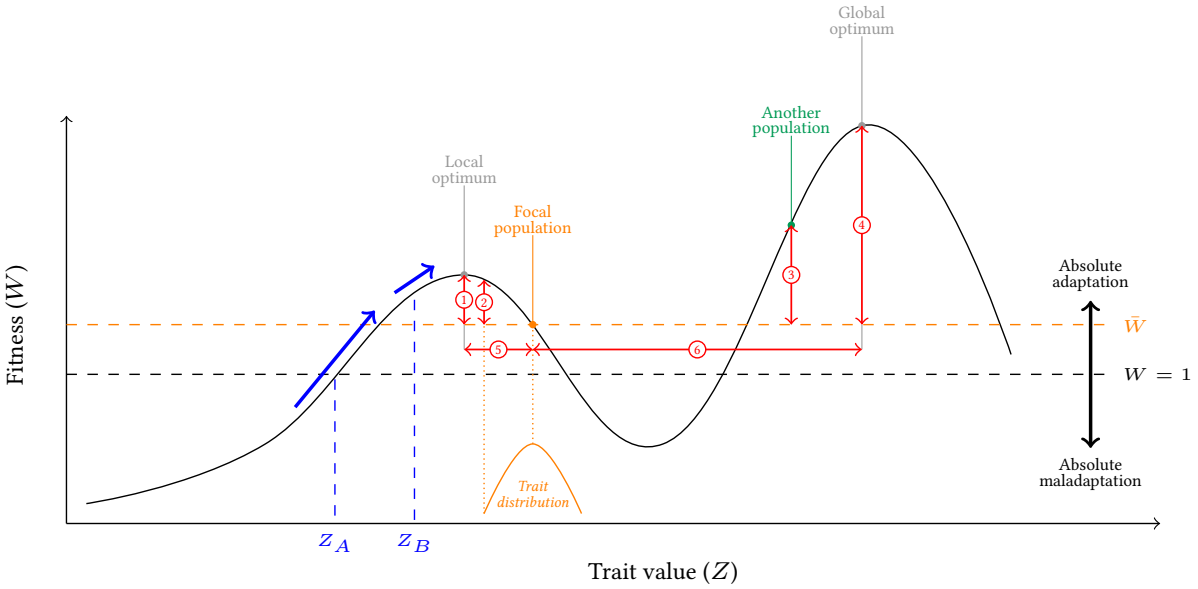


Figure II.2. Hypothetical adaptive landscape (black curve) with two fitness peaks (a local and a global optimum). The blue vectors tangent to the adaptive landscape represent the strength of selection for populations with mean phenotypes Z_A and Z_B. The strength of selection is expected to decrease near the fitness peak, as shown by different vector lengths. The average absolute fitness of a focal resident population is represented by an orange dot, and its trait distribution is illustrated below by an orange curve. The absolute fitness (mal)adaptation of the focal population is calculated as the distance between the population mean fitness and a threshold mean fitness value of 1 that corresponds to the replacement rate: absolute adaptation when $W \geq 1$ and absolute maladaptation when $W < 1$. The magnitude of the relative fitness (mal)adaptation of the focal population can be estimated by comparing along the Y-axis (vertical red arrows) the population mean fitness with (1) the local optimum, (2) the fitness of the most fit individual from the focal population, (3) another population (in green), or (4) the global optimum. Trait-based measures of (mal)adaptation can be obtained by comparing along the X-axis (horizontal red arrows) the mean trait values of the focal population with (5) the trait value at the local optimum, or (6) the trait value at the global optimum. Adapted from Brady et al. (2019a).

The phenotypic adaptive landscape framework can help to understand the different ways of defining the *state* of adaptation or maladaptation, and both concepts are referred under the term **(mal)adaptation** (Capblancq et al. 2020a). A key parameter is the **mean absolute fitness** $\bar{W}$ of a population, which corresponds to the mean expected lifetime reproductive success in the population (see details in Brady et al. 2019b on how this quantity can be calculated). The **absolute (mal)adaptation** of a population can be calculated by comparing its mean absolute fitness $\bar{W}$ with a threshold $W = 1$, that corresponds to the absolute fitness of a population at demographic equilibrium (i.e. each individual of the population gives on average one individual that survives and reproduces; Brady et al. 2019b). In this case, a population is adapted to its environment if $\bar{W} \geq 1$, i.e. the population is at the equilibrium or in expansion and a population is maladapted to its environment if $\bar{W} < 1$, i.e. the population is decreasing (Fig. II.2). Estimates of absolute (mal)adaptation are mainly used in the context of ecology and conservation biology and suggest pervasive maladaptation in nature (Brady et al. 2019b, Hendry and Gonzalez 2008). This conclusion follows, for example, from the observation that population declines and range contractions, eventually leading to population or species extinction, occur continuously over macroevolutionary time (Muscente et al. 2018, Novacek and Wheeler 1992), with increasing rates due to human activities (Ceballos et al. 2017, Dirzo et al. 2014).

Relative (mal)adaptation can be characterized via the **mean relative fitness** $\bar{w}$ of a population, which is the mean absolute fitness of the population $\bar{W}$ divided by the absolute fitness of another entity w_e, such as other populations (e.g. in reciprocal transplant experiments), the individual with the highest fitness in the population (with the fittest individual having a relative fitness of one; Crow and Kimura 1970) or a local or global optimum (vertical red arrows in Fig. II.2; Brady et al. 2019b). A population is then considered maladapted when $\bar{w} < w_e$. Relative (mal)adaptation can also be characterized with trait-based approaches, in which the

population mean phenotype is compared to the optimal phenotype (horizontal red arrows in Fig. II.2). However, inferring fitness from phenotypes rather than directly may be biased by trade-offs among traits (Shoval et al. 2012), e.g. between reproduction and other life-history traits (Obeso 2002).

The concept of relative (mal)adaptation has been mostly used in evolutionary biology and the general view is that both adaptation and maladaptation are widespread states (Brady et al. 2019b), though the latter has been largely less studied (Brady et al. 2019a). On the one hand, empirical evidence of the prevalence of adaptation mainly comes from: (i) reciprocal transplant studies showing that local individuals have generally higher fitness than foreign ones (Hereford 2009, Leimu and Fischer 2008, Nosil et al. 2005), a pattern known as **local adaptation** or **home-site advantage** (Kawecki and Ebert 2004), but whose prevalence may be overestimated (Schluter 2000a), (ii) difficulties in detecting selection in the field, which suggest that populations are generally close to fitness peaks (Estes and Arnold 2007, Haller and Hendry 2014, Hendry 2017), (iii) levels of additive genetic variance measured in wild populations for most traits (Hansen et al. 2006, Houle 1998, Mousseau and Roff 1987) and fitness (Burt 1995, Hendry 2017), which are high enough to respond rapidly to selection, (iv) numerous examples of rapid adaptive evolution due to selection in natural populations (Gingerich 2009, Hendry and Kinnison 1999, Kinnison and Hendry 2001, Reznick and Ghalambor 2001, Thompson 1998), and (v) invasive species that can successfully colonize new environments through rapid adaptation (Colautti and Barrett 2013, Phillips et al. 2006). On the other hand, the ubiquity of maladaptation is supported by: (i) the obvious observation that, for selection to act, not all individuals within a population can be at the fitness peak (Barton and Partridge 2000, Brady et al. 2019b), (ii) a considerable proportion of reciprocal transplant studies that do not find patterns of local adaptation (Brady et al. 2019b, Kooyers et al. 2019, Rogalski 2017, Samis et al. 2016), (iii) theoretical work suggesting that population mean phenotypes may constantly be tracking an optimum moving within stable limits (Estes and Arnold 2007).

2.2 Global change as the main driver of contemporary maladaptation?

Global change encompasses any anthropogenic environmental change that alters ecosystems (Vitousek 1992), thus threatening their ability to sustainably provide goods and services, especially for future generations (Millennium Ecosystem Assessment 2005). Global change has five major components: climate change, land-use change (i.e. habitat loss and fragmentation), overexploitation, pollution, and invasive species (Matesanz et al. 2010, Soulé 1991).

The components of global change can alter the phenotypic adaptive landscapes in multiple ways, detailed in Svensson and Calsbeek (2012) from which the following examples are taken. Invasive species can either induce the emergence (e.g. introduction of a new host plant representing a new resource) or loss of a fitness peak (e.g. competition for resources leading to resource depletion), smooth the valley between two fitness peaks (e.g. introduction of non-native plants increasing the relative abundance of intermediate-sized seeds; Hendry et al. 2006), increase the dimensionality of the adaptive landscape by causing selection to act on a new trait (e.g. introduction of new predators on isolated islands), and alter the phenotypes and thus change their position on the adaptive landscape (e.g. through hybridization). Hunting, through the removal of larger individuals, can shift the position of the fitness peak (toward smaller body size) and sharpen the fitness peak (by reducing trait variance). Change in habitat quality can impact the elevation of a fitness peak, thus impacting the absolute fitness of the

population (scenario of the degraded target in Brady et al. 2019a) and may also change the dimensionality of the landscape (e.g. new pollutants requiring new adaptations to persist in the new habitat). Habitat loss and fragmentation can reduce the diversity in habitats and resources, and thus the number of fitness peaks, or change the fitness peak position (e.g. selection for lower seed dispersal in urban environments; Cheptou et al. 2008).

The components of global change may have cumulative effects, thus bringing species to their adaptive capacity threshold faster than if only one of the components were involved (e.g. Drouineau et al. 2018), or alternatively, their effects may counterbalance each other (e.g. Morelli et al. 2012). In addition, components of global change may interact, e.g. the long-lasting effects of land-use change and human-altered fire regimes on vegetation dynamics and biodiversity may interact with the climate change impacts (Franklin et al. 2016, Hansen et al. 2001), which adds considerable uncertainty to their long-term effects on ecosystems (Sala et al. 2000). The Mediterranean area, a biodiversity hotspot, may experience a major change in biodiversity due to its high sensitivity to all components of global change, especially land-use and climate change (Bellard et al. 2014, IPCC 2018, Sala et al. 2000). Indeed, land-use change might be the most important factor affecting terrestrial ecosystems in this biogeographical region, followed by climate change, nitrogen deposition, biotic exchange and elevated carbon dioxide (Sala et al. 2000).

2.3 Maladaptation induced by climate change: tracking a moving optimum

Climate change is the component of global change for which the most extensive data are available and whose effects on ecosystems are best documented (Foden et al. 2019, Gattuso et al. 2015, Parmesan 2006, Urban 2015, Visser 2008, Wiens 2016). To date, the Earth's climate has warmed by 1.5 °C above pre-industrial levels, which results from increased greenhouse gas emissions by human activities (IPCC 2015, 2018, 2021). Projections for mean temperatures and heat extremes show an increase in almost all locations, both on land and oceans, while projections for precipitation are more uncertain, predicting a likely increase in droughts in the Mediterranean region (IPCC 2018, 2021).

Climate change primarily alters adaptive landscapes by moving the optimal phenotype (i.e. the fitness peak) away from the population mean phenotype (scenario of the moving target in Brady et al. 2019a), thus decreasing the mean fitness of the population (both absolute and relative). In face of climate change, species and populations within species may migrate to other geographical locations in which they have a higher fitness (e.g. through habitat choice by individuals or seed dispersal; Edelaar and Bolnick 2019), thereby shifting their distribution range to track their climatic niche (Hughes 2000, Parmesan 2006, Peñuelas et al. 2013). Shifts in species distribution and abundance have already been observed (Chen et al. 2011, Dobrowski et al. 2011, Hughes 2000), which notably create changes in species interactions and community dynamics (Ockendon et al. 2014, Prober et al. 2012, Walther et al. 2002); but some species might not be able to migrate to more favorable environments (Burrows et al. 2014, Liang et al. 2018, Rehm and Feeley 2015). Alternatively, populations may persist within their current geographical location by responding to new climates through genetic change and phenotypic plasticity (Hendry et al. 2017, Merilä and Hendry 2014, Parmesan 2006), two processes I will now focus on.

Phenotypic plasticity is an immediate phenotypic response to changing conditions that operates at the scale of an individual's lifetime. If adaptive, it may allow at least some individuals to track a moving fitness peak by adjusting their phenotype to the altered environment (West-Eberhard 2003), thereby increasing the population mean fitness. This process, called the **Baldwin effect** or **plastic rescue** (Baldwin 1896), may be particularly beneficial in resisting abrupt environmental shifts (e.g. extreme drought events during which most existing phenotypes may not be able to cope with the new environment) and has therefore a major influence on the ability of populations to persist and colonize new environments (Crispo 2007, Ghalambor et al. 2007, Hendry 2017). In face of a negative demographic trend induced by climate change, a population may also recover and avoid extirpation through adaptive evolutionary change, a process called **evolutionary rescue** (Carlson et al. 2014, Gomulkiewicz and Holt 1995, Gonzalez et al. 2013, Kinnison and Hairston 2007). Accumulating evidence from empirical studies in natural populations shows that evolutionary change can be fast in natural populations in response to strong selection pressures, e.g. during native range expansion (Lustenhouwer et al. 2018), colonization of new environments (Colautti and Barrett 2013, Losos et al. 1997, Reznick et al. 1997), soil contamination by heavy metals (Antonovics 2006, Antonovics and Bradshaw 1970), changes in food supply (Grant and Grant 1995) and multiyear drought (Franks et al. 2007). Evolutionary rescue is facilitated by high standing genetic variation and mutation rate (Barrett and Schluter 2008, Bell 2013, Gomulkiewicz and Holt 1995, Orr and Unckless 2008), and also depends on the genetic architecture and the genetic correlations among fitness-related traits (Chevin 2013, Gomulkiewicz et al. 2010). Populations that are large - and therefore less prone to demographic stochasticity - and do not have excessive initial maladaptation are more likely to escape extinction through adaptation (Carlson et al. 2014, Gomulkiewicz and Holt 1995). Finally, the faster the environmental change relative to the population generation time, the shorter the time for adaptive genetic change to spread through the population and restore positive growth (Bell 2013, Carlson et al. 2014).

Importantly, phenotypic plasticity and adaptive evolution might have interactive effects on population persistence in face of climate change. First, if there is genetic variation in plasticity (third panel of Fig. II.1) and this variation induces differences in fitness between individuals, then plasticity can be under selection and, if heritable, can evolve (Pigliucci 2005, Scheiner 1993, Tufto 2000, Via and Lande 1985). Second, phenotypic plasticity can influence the adaptive trajectories of populations (Ghalambor et al. 2007, West-Eberhard 2003, Wund 2012). Indeed, it may constrain or slow the rate of adaptation by shielding genotypes from selection (e.g. Huey et al. 2003). Conversely, by allowing some individuals to cope with novel environments, plasticity can give time for selection to act (i.e. for adaptive mutations to appear and spread through the population), thus promoting evolutionary change (Ghalambor et al. 2007, Pennisi 2018). In a first step, such evolutionary change may result in favoring the most plastic genotypes if the latter bring phenotypes closer to the new optimum, a process know as **genetic accommodation** (Crispo 2007, Kelly 2019). For instance, populations in heterogeneous environments with reliable environmental clues are expected to evolve towards higher plasticity (Bonamour et al. 2019, Ghalambor et al. 2007, Kleunen and Fischer 2005, Schmitt et al. 2003). In a second step, if the environment is stable, a plastic trait (i.e. whose variation is environmentally induced) may be converted to a genetically determined and canalized trait (either fixed or expressed constitutively in the population), a process called **genetic assimilation** (Crispo 2007, Ehrenreich and Pfennig 2016, Waddington 1952, 1953, West-Eberhard 2003). Finally, phenotypic plasticity may be maladaptive (e.g. in rare and extreme environments; Chevin and Hoffmann 2017, Schlichting 2008), which may boost adaptive evolution by increasing the strength of directional selection and may result in the loss of plasticity as an adaptation to counterbalance maladaptive phenotypic change (i.e. **genetic compensation**; Ghalambor

et al. 2007, Ghalambor et al. 2015, Grether 2005). However, empirical evidence of selection on plasticity remains scarce (but see for instance Nussey et al. 2005), and a meta-analysis found no evidence for selection on thermal phenotypic plasticity (Arnold et al. 2019). Indeed, evolution of plasticity may be constrained by costs, i.e. reduction in fitness when a phenotype is expressed through plastic rather than fixed development, and limits, i.e. the inability to express the optimal phenotype (Aubret and Shine 2010, DeWitt et al. 1998, Kleunen and Fischer 2005, Murren et al. 2014). That's why, in the face of climate change, there may be a threshold after which plasticity is not enough for population survival and genetic change will be required. This is expected, in particular, when the environmental change is large.

Finally, the unprecedented rate and magnitude of climate change may push populations to the limits of their persistence ability, even with genetic adaptation. For all the reasons mentioned above, disentangling the contribution of phenotypic plasticity and genetic changes to observed phenotypic changes and determining the maximum rates of climate change that populations can cope with remains very difficult but, nevertheless, necessary to predict population responses to future climates and risks of extinction and extirpation (Bradshaw and Holzapfel 2006, Chevin et al. 2010, Hendry et al. 2008, Merilä and Hendry 2014).

3 Predicting short-term population responses to climate change in the genomic era

3.1 Next-generation sequencing approaches

The development and generalization of the so-called 'next-generation sequencing' (NGS) approaches provide, at decreasing costs, access to thousands of markers for a number of individuals, more or less densely scattered across the genome depending on its size (Stapley et al. 2010). The gold standard high-throughput genotyping method is whole-genome sequencing, which consists in determining nearly the entire DNA sequence and therefore offers the largest number of markers and the denser genotyping. However, performing whole-genome sequencing is currently cost-prohibitive, requires high DNA quality and quantity, and involves high computational power and data storage capacity (de Villemereuil et al. 2016, Tam et al. 2019). Consequently, to date, this method is not appropriate for non-model species with large or highly repetitive genomes, or in studies in which genomic sequence data for all individuals is unnecessary (e.g. many studies in ecological and conservation genomics; Narum et al. 2013). In such cases, reduced representation library (RRL) sequencing approaches have been extensively used and the most popular techniques are restriction site associated DNA sequencing (RADseq; Baird et al. 2008, Miller et al. 2007) and genotype-by-sequencing (GBS; Elshire et al. 2011). These approaches do not involve the costly process of genome assembly and decrease the sequencing effort by sequencing only restriction fragments resulting from restriction enzyme digestion (Davey et al. 2011). In addition, the genomic regions to be sequenced can be either randomly selected (as in RADseq or GBS), or targeted based on previous analyses or *a priori* knowledge of gene function (i.e. candidate gene or targeted resequencing approaches; Stapley et al. 2010, Tabor et al. 2002). This different kind of sequencing thus provides cheaper and less computationally heavy genotyping for huge number of markers, even in species with limited or no previous genomic information, and, in some cases, a greater depth of coverage per locus (and thus improved genotyping reliability) than whole-genome sequencing (Andrews et al.

2016, Tam et al. 2019). However, as sequencing costs decline, computational and bioinformatics methods develop, and more and more species have their whole genome sequenced, whole-genome sequencing may become the predominant method in the future, even for non-model species.

3.2 Predicting phenotypes based on genomic markers

Trait-based approaches that aim at predicting short-term responses of natural populations to changing conditions (e.g. climate change) have started to integrate genome-wide molecular markers provided by high-throughput genotyping (Gienapp et al. 2017). Genomic markers have been used in GWAS to investigate the genetic architecture of quantitative traits (see Section 1.1). This is particularly relevant for predicting the response of populations to selection because, for a given h^2, polygenic traits are likely to have higher evolutionary potential than traits with large effect alleles (Chevalet 1994, Kardos and Luikart 2021, Walsh and Lynch 2018). Genomic markers are also more and more used to predict phenotypes based on genotypes, which is appealing to predict trait values of populations across the species ranges and under future climates, but remains highly challenging, as I will elaborate in this section.

In humans, GWAS outputs are increasingly used to predict the phenotype of individuals through the use of **polygenic scores** (PRS). PRS are calculated by adding up the effects of the alleles associated with the trait of interest, under the assumption that alleles have additive effects (Lynch and Walsh 1998). They are highly promising towards identifying individuals that are more likely to be at risk of some diseases such as breast cancer (Khera et al. 2018, Mavaddat et al. 2019). However, GWAS have been criticized as the loci involved in significant associations often explain only a small fraction of the heritability of quantitative traits estimated from the resemblance between relatives in classical quantitative genetics analyses, such as twin or family studies (i.e. the missing heritability problem; Manolio et al. 2009). For instance, for height and body mass, a GWAS meta-analysis in humans found significant associations for 3,290 and 941 near-independent SNPs, respectively explaining ~ 24.6% and ~ 5% of the phenotypic variance, respectively (Yengo et al. 2018). By contrast, classical quantitative genetics analyses yield heritabilities of 80% for height and 40-60% for body mass. In reality, this apparent paradox is in accordance with the infinitesimal model (Walsh and Lynch 2018), or translations of it in the genomic era, such as the omnigenic model (see Section 1.1; Boyle et al. 2017, Liu et al. 2019). Indeed, as the vast majority of alleles have tiny effects on quantitative traits, they do not reach the stringent significant threshold common in GWAS and are therefore excluded from the models (Walsh and Lynch 2018). This is supported by the increased h^2 explained by GWAS as sample sizes become larger, e.g. the landmark study of Yang et al. 2015 which showed that 45% of the additive variance in human height can be explained by including all SNPs in h^2 estimation. The remaining missing heritability may be explained by the incomplete linkage disequilibrium between markers and causative alleles (Walsh and Lynch 2018), inflated heritability estimates in classical quantitative genetic studies (Mayhew and Meyre 2017, Zuk et al. 2012) or gene–gene and gene–environment interactions (Aschard et al. 2012, Frazer et al. 2009).

In plant and animal breeding, genome-wide markers have first been used to estimate the relatedness between individuals via **genomic relationship matrices** (GRM), which contain the realized proportion of genome shared among all pairs of individuals in a population (VanRaden 2008). Incorporating a GRM in an animal model (Box 2) provides more accurate estimates of key genetic parameters (i.e. breeding values, and additive and non-additive genetic

variance) than previous pedigree-based approaches (Bouvet et al. 2016, El-Dien et al. 2016, Jannink et al. 2010, Muñoz et al. 2014). Further, since Meuwissen's landmark paper (2001), plant and animal breeders have turned to genomic selection to predict the additive genetic value of individuals (Box 2), which is based on estimating **genomic estimated breeding values** (GEBVs; conceptually equivalent to PRS; Wray et al. 2019). This has led to tremendous progress in livestock improvement (García-Ruiz et al. 2016). Importantly, these methods do not require to breed individuals in costly experiments and can be applied directly in natural populations and in virtually any species (Beaulieu et al. 2014, Bérénos et al. 2014, Gienapp et al. 2019, 2017, Robinson et al. 2013).

Box 2. From the animal model to genomic selection

The animal model (Henderson 1975) is most often used to estimate genetic parameters such as the additive genetic variance and the narrow-sense heritability and can be written as follows:

$$y_i = \mu + a_i + e_i$$

where y_i is the phenotype of the individual i, μ is the average population phenotype, a_i is the breeding value of the individual i and e_i is the residual variation. a_i are unknown and can be estimated from the covariance among relatives in additive genetic effects, such as:

$$a_i \sim \mathcal{N}(0, \mathbf{A}V_A)$$

where V_A is the additive genetic variance and $\mathbf{A}$ is the relatedness matrix which can either be estimated from a pedigree or from genomic markers. In this latter case, the relatedness matrix is called a genomic relationship matrix (GRM) and is often referred as $\mathbf{G}$.

This model has been the subject of many interesting extensions, and I will mention two here that I used during my PhD work. Since the presence of different genetic groups (e.g. genetically differentiated breeds) can bias the estimates of the additive genetic variance, Wolak and Reid (2017) proposed a model that allows for differences in mean breeding values among individuals from different genetic groups, which can be written as follows (for r genetic groups):

$$y_i = \mu + u_i + e_i$$
$$u_i = \sum_{j=1}^{r} q_{ij} g_j + a_i$$

where u_i replaces the breeding value a_i of the basic animal model without genetic groups. More precisely, u_i corresponds to the total additive genetic effect of individual i, that is separated between a weighted sum of the group-specific means g_j and a breeding value a_i that accounts for deviations from the weighted sum.

Muff et al. (2019) further extended this model by allowing the genetic groups to have different additive genetic variances, such as:

$$y_i = \mu + \sum_{j=1}^{r} q_{ij} g_j + \sum_{j=1}^{r} a_{ij} + e_i$$

where $a_{ij} \sim \mathcal{N}(0, V_{A_j}\mathbf{A_j})$. V_{A_j} is the additive genetic variance in the genetic group j and $\mathbf{A_j}$ is a relatedness matrix specific to the genetic group j (see details in Muff et al. 2019 on how the $\mathbf{A_j}$ matrices are calculated). a_{ij} can be considered as a partial breeding value, as it accounts for the contribution to the breeding value a_i of individual i that is inherited from the genetic group j.

In breeding, predicting individuals with the best performing progeny is mainly done within the framework of genomic selection (Meuwissen et al. 2001). Genomic estimated breeding values (GEBVs) are estimated in a training population with the following model:

$$y_i = \mu + \sum(SNP_{ij}\gamma_j) + e_i$$

where SNP_{ij} is the genotype of the individual i at the locus j and γ_j is the effect of the SNP on the phenotype. GEBVs are then used to estimate the additive genetic values of individuals in a candidate population of related individuals. GEBVs are generally more accurate than the breeding values obtained from the animal model.

A major limitation of phenotypic predictions with PRS and GEBVs is that their accuracy greatly decreases when predicting the phenotype of individuals from genetic groups only weakly or not related to the one used to adjust the models. As such, the low transferability of PRS has been shown when predicting the phenotype of individuals with different genetic ancestry (Martin et al. 2017, 2019), or individuals within the same ancestry but with different characteristics such as age, sex, and socio-economic status (Mostafavi et al. 2020), and their accuracy improves when incorporating relatives in the discovery sample (Lee et al. 2017).

Similarly, in breeding, GEBVs usually predict phenotypes well within-population (or within-breed), as a result of the low effective population size (Wray et al. 2019), but very poorly across populations or breeds (Hayes et al. 2009, Hidalgo et al. 2016, Moghaddar et al. 2014, Resende et al. 2012) or across environments (Resende Jr et al. 2012). Potential explanations include population structure (Barton et al. 2019), variable patterns of linkage disequilibrium, inconsistent allelic effects across genetic groups, genetic interactions (Dai et al. 2020) or genotype-by-environment interaction (Resende Jr et al. 2012). Interestingly, Mathieson (2021) recently suggested that the low transferability of PRS provides further support for the omnigenic model. Indeed, under the assumption that the omnigenic model is true, the phenotypic effects of peripheral alleles are expected to vary across populations as a result of variation among populations in the structure and complexity of the gene networks and in the interactions with environmental factors (Mathieson 2021). This limitation on the use of PRS and GEBVs for phenotypic prediction across populations may therefore be hard, if not impossible, to overcome for some traits (and we have limited clues as to which traits will be predictable or not; Mathieson 2021), which would constrain their usefulness to be population specific (Resende et al. 2012).

Finally, combining multi-site common gardens (i.e. randomized controlled experiments where populations from different geographical locations are planted; also known as provenance tests) with genomic information may prove particularly valuable for phenotypic prediction and detection of adaptive variants/traits (de Villemereuil et al. 2016, Josephs et al. 2019). First, common gardens are the gold standard to separate the plastic and genetic components of trait variation and thus directly allow to control for the confounding effect of plasticity, which is otherwise very hard to do *in situ*. Second, genomic data can be used to infer the underlying population structure resulting from the population demographic and evolutionary history (Luikart et al. 2003, Nicholson et al. 2002). This is particularly valuable for methods aiming at detecting alleles under selection (e.g. genome scans, GWAS, genotype–environment association methods discussed in the next section) as they have to account for population structure to avoid false positives, which can be done for instance via latent factors (Frichot et al. 2013) or a covariance matrix based on the population allele frequencies (Gautier 2015). However, as already mentioned before, accounting for population structure is hard (Barton et al. 2019) and available methods still show high rates of false positives (de Villemereuil et al. 2014, Lotterhos and Whitlock 2014).

To conclude, incorporating genomic information in trait-based approaches is already largely improving the predictions of individual phenotypes within-population (i.e. individual genetic values) and on the way to expand to natural populations. A first potential application consists in the identification of *pre-adapted* individuals, i.e. individuals with adaptive variants or trait values allowing them to persist under future climates, which could be potentially used in conservation strategies based on adaptive *states* such as assisted migration (Derry et al. 2019). Another application could be to identify a set of individuals maximising the diversity in trait-associated variants and trait values, which could be potentially used to boost genetic variation -and thus maybe evolution- in populations at risk under climate change, i.e. in conservation strategies based on adaptive *processes* such as evolutionary rescue (Derry et al. 2019).

3.3 Landscape genomics

Landscape genomics is the study of the processes shaping the geographical patterns of adaptive genetic variation across the landscape (Storfer et al. 2018), and stems from landscape genetics, which focuses on patterns of neutral genetic variation (Manel et al. 2003, Storfer

et al. 2007). In the same vein as trait-based approaches using GWAS, landscape genomics approaches are widely used to detect adaptive variants through **genotype-environment association analyses** (GEAs; Rellstab et al. 2015), which test for associations between genetic variants and environmental variables measured at the sample location. GEAs assume that genetic variants whose frequency varies along an environmental gradient (once the underlying population structure is accounted for) are linked to fitness, and focus on the choice of the spatial scale and the particular environmental variables that potentially create a selection pressure (e.g. extreme cold or heat events). GEAs have the benefit of not relying on phenotypes, thus avoiding potential biases arising from the choice of which phenotypes to measure (and in which ontogenic stage or environment) and the uncertain link between fitness and the phenotypes under study (Capblancq et al. 2020a).

A space-for-time substitution approach, known as **genomic offset** (Fitzpatrick and Keller 2015) or **risk of nonadaptedness** (Rellstab et al. 2016), has recently become popular for identifying populations at risk of maladaptation under climate change (Capblancq et al. 2020a). This approach aims at predicting the change in genomic composition required to maintain the current relationships between a set of putatively adaptive alleles and the environment, and typically consists in four main steps: (1) identifying candidate adaptive SNPs via either genome scans, GEAs or GWAS, or a combination of methods; (2) modelling the turnover in adaptive allele (or genotype) frequencies along current environmental gradients; (3) projecting the current and future genomic composition across the species range, including areas where climate data are available but no individuals or populations were genotyped (i.e. extrapolation); (4) estimating the magnitude of genetic change required to maintain the current gene-environment relationships (i.e. the genomic offset metric). Importantly, this approach relies on four key assumptions (Capblancq et al. 2020a, Rellstab et al. 2021): (1) adaptive alleles have been correctly identified, which requires validation of the GWAS and GEAs results (Ioannidis et al. 2009, Oetting et al. 2017), for instance with gene knock-out experiments (Curtin et al. 2017, Monroe et al. 2018, Rohde et al. 2018); (2) populations with a higher genomic offset are expected to experience a decrease in fitness, which may be tested by determining whether they show decreasing demographic trends (Bay et al. 2018, Ruegg et al. 2018), or whether they have lower fitness in common gardens (Fitzpatrick et al. 2021); (3) the populations are currently at their fitness optimum, which can be evaluated in reciprocal transplant experiments (see Section 2.1; Browne et al. 2019, Leimu and Fischer 2008); (4) the current gene-environment relationships (spatial patterns) remain unchanged over space (spatial extrapolations), over time (space-for-time approach) and in a changing climate. This latter assumption strongly depends on the variation in genetic background across the landscape and processes such as migration, demographic trends and admixture (Capblancq et al. 2020a, Hoffmann et al. 2021, Rellstab et al. 2021). In this line, recent studies have incorporated information related to evolutionary processes that may mitigate future adaptive mismatch of populations, thus rendering the genomic offset approach more realistic and robust, e.g. by incorporating the migration potential of adaptive alleles in the predictions (Aguirre-Liguori et al. 2019, Capblancq et al. 2020b, Gougherty et al. 2020a). Last, climate forecasts are highly uncertain, and although some studies have used the average of different general circulation models (e.g. Gougherty et al. 2020b), they remain a serious source of uncertainty that propagates into the predictions (Hallingbäck et al. 2021). To conclude, the genomic offset approach is still under development and in need of further experimental validation (Rellstab et al. 2021). Like other phenotype-free approaches, the links with fitness-related traits and processes such as phenotypic plasticity are missing, which limits the assessment of the ability of populations to persist under climate change (Capblancq et al. 2020b, Rellstab et al. 2021). Therefore, the use of genomic offset approaches in conservation strategies and restoration projects remains subject to caution

(Hoffmann et al. 2021), but also holds great promise for identifying potential recipient and donor populations or suitable future habitats (e.g. Borrell et al. 2020).

4 Investigating past, current and future adaptation in forest trees

4.1 Forest trees are ecologically key species threatened by climate change

Forests cover 31% of the land area worldwide (i.e. 4.06 billion ha; FAO 2020). Forest trees, as foundation and keystone species, provide habitats for many species and thus contribute to maintaining biodiversity (Brockerhoff et al. 2017, Gibson et al. 2011), which in turn promotes numerous ecosystem functions (e.g. primary production, decomposition, nutrient cycling, trophic interactions) and services (e.g. water retention and purification, pollination, pest regulation, prevention of soil erosion) (Balvanera et al. 2006, Cardinale et al. 2012, Hooper et al. 2005, Mori et al. 2017). Forests play a key role in the global carbon cycle, storing ~ 45% of terrestrial carbon and absorbing annually ~ 33% of anthropogenic carbon emission from fossil fuel and land-use change (Bonan 2008). Forests have also biophysical effects, such as evapotranspiration, which tends to produce cooling, and albedo, which has a warming effect (Anderson et al. 2011, Gibbard et al. 2005, Li et al. 2015, Marland et al. 2003). Forest trees are extensively exploited for timber and fiber production, and 7% of the forested area worldwide is planted (i.e. 290 million ha; FAO 2020). Forests also provide other economic services, such as the provision of food and medicinal products, along with social and aesthetic services (e.g. through recreational uses).

Global change, including climate change, is already impacting forests worldwide, and thereby the ecosystem services they provide (Bonan 2008), e.g. the southeastern Amazonian forest now acts as a net carbon source (due to both climate change and deforestation; Gatti et al. 2021). Increasingly frequent and hotter droughts, often associated with pathogen and pest outbreaks (Weed et al. 2013) and more frequent fires (Seidl et al. 2017), cause an increase in tree mortality, through both die-backs and background mortality, across all terrestrial biomes (Allen et al. 2015, 2010, Anderegg et al. 2015, Mantgem et al. 2009). In face of climate change, tree migration northward and in altitude is already underway (Boisvert-Marsh Laura et al. 2014, Davis and Shaw 2001, Woodall et al. 2009). However, trees may not be able to migrate fast enough to keep pace with the unprecedented rate of climate change (Aitken et al. 2008, Dauphin et al. 2021, Johnstone and Chapin 2003, McLachlan et al. 2005, Sittaro et al. 2017, Zhu et al. 2012), which may be explained by their population dynamics (i.e. slow biomass increase and long-generation time) and interspecific competition, and only marginally by dispersal limitation (Scherrer et al. 2020). Determining whether forest trees will be able to persist in their current locations is therefore particularly important and urgent, and represents an area of research with a long history (Alberto et al. 2013), in which the contribution of new genomic data is very promising.

4.2 A long history of common gardens

In forest trees, investigating adaptation clines has mainly relied on multi-site common gardens, which were initially settled to identify populations with the highest values of some commercial traits (e.g. height, diameter, straightness) in different environments (Langlet 1971, Morgenstern 1996, Savolainen et al. 2007). These large networks of common gardens have provided a unique framework for quantitative geneticists as they are the gold standard to separate the plastic (i.e. environment) and genetic components of quantitative trait variation. Further, phenotypic data from common gardens have been combined with climatic data to derive population reaction norms along climatic gradients or climatic-transfer distances (i.e. climatic distances between the provenance and the common garden locations), often called **population response** (or transfer) **functions** (see details in Box 3; Rehfeldt et al. 2002, 1999, Wang et al. 2010). The development of the **universal response function** was a step forward to jointly evaluate, in a single step, the relative contribution of climate-driven plasticity (associated with the climate in the common gardens) and genetic differentiation (associated with the climate-of-origin of the populations) in explaining quantitative trait variation (O'Neill et al. 2008, Wang et al. 2010). This approach have been extensively used to investigate potential climatic drivers of the genetic and plastic components of trait variation for a large variety of traits (Benito Garzón et al. 2019, Leites et al. 2012a,b) and determine the climatic optimum of the populations (e.g. Fréjaville et al. 2020).

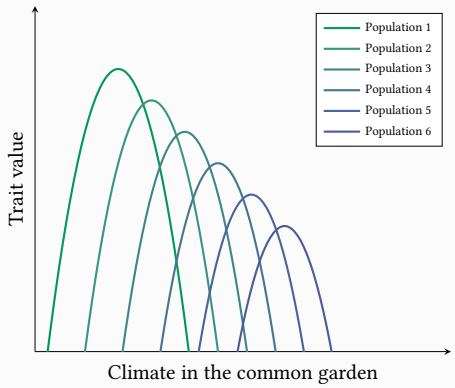
Box 3. Population reaction norms and site-specific functions

The relative influence of climate-associated plasticity and genetic differentiation on quantitative trait variation can be modeled with population reactions norms (also referred as population response functions; Rehfeldt et al. 2002, 1999, Wang et al. 2006) and site-specific functions, which can be combined into an universal response function (O'Neill et al. 2008, Wang et al. 2010). These functions rely on phenotypic data (e.g. height, survival) from multi-site common gardens in which populations (i.e. provenances) from different parts of the species range are included, and can be applied within the mixed model framework (Leites et al. 2012a,b). For instance, a given trait y can be expressed as a function of a climatic variable of interest (e.g. minimum or mean temperature) such as:

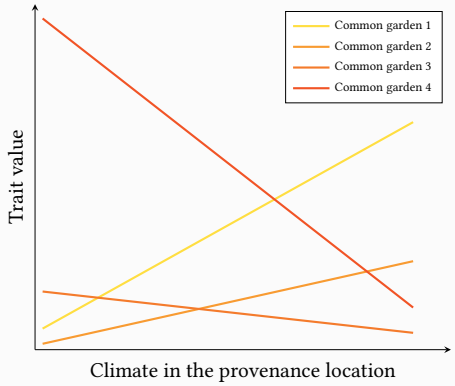
$$y_{ipsb} = \beta_0 + P_p + S_s + B_{s(b)} + \beta_{1s} PC_p + \beta_{2p} SC_s + \beta_{3p} SC_s^2 + \beta_4 PC_p SC_s + \epsilon_{ipsb}$$

- β_0 is the global intercept.
- P_p are the varying intercepts of the populations, associated with the genetic component of the trait and theoretically capturing the effects not accounting for by the climatic variable at the population location (PC_p).
- S_s and $B_{s(b)}$ are the varying intercepts of the common gardens (i.e. sites) and blocks nested within common gardens, associated with the plastic component of the trait and theoretically capturing the effects not accounting for by the climatic variable in the common gardens.
- PC_p is the value of the climatic variable at the location of the population p, and β_{1s} are the associated regression coefficients which are specific to each common garden.
- SC_s is the value of the climatic variable at the location of the common garden s, and β_{2p} and β_{3p} are the associated regression coefficients which are specific to each population.
- β_4 is the regression coefficient associated with the interaction between the values of the climatic variables at the location of the populations and the common gardens.
- ϵ_{ipsb} is the residual variation.

From this model, we can derive quadratic response functions specific to each population along the gradient of the climate in the common gardens, such as shown in the figure below, where each curve corresponds to a population with a different climatic niche. According to Rehfeldt et al. (2018), the climatic optimum of each population corresponds to its *physiological* optimum (i.e. the climate in which the population shows the highest trait values but is most often competitively excluded), while the climate in which the population is found corresponds to its *ecological* optimum (i.e. the climate in which the population is competitively exclusive).



This model can also be used to derive linear functions specific to each common garden along the gradient of the climate in the provenance location, as shown in the figure below. This is useful for identifying the best performing populations in a given common garden.



Noticeably, incorporating environmental sources of variation (either at the location of the populations or the common gardens) in quantitative genetic models can also be done with matrices of environmental similarity, either based on Euclidean distances (e.g. Thomson et al. 2018) or on covariances (e.g. Jarquín et al. 2014). This method has the benefit of not assuming linearity in the relationship between environment and phenotypes, but it does rely on the assumption that the environmental components included in the matrix have equal weight. Therefore, the choice of including climate variables directly in the model (as in the model above) or via similarity matrices should be decided based on the goals of the study (e.g. whether the researchers are interested in a particular environmental variable or not).

Quantitative traits in forest trees (and plants more generally) show large plastic variation and less (but still considerable) genetic variation, with notable differences across traits (Anderson et al. 2012, Benito Garzón et al. 2019, Cornelius 1994, Franks et al. 2014, Morgenstern 1996). Reproduction and phenology traits generally show higher heritabilities than growth traits (Caignard et al. 2018, Howe et al. 2003, Lind et al. 2018, Scotti-Saintagne et al. 2004). Population reaction norms revealed that populations of the same species have different climatic optima and niches, the latter being narrower than the whole species climatic niche (Rehfeldt et al. 2018, 1999; but see Gömöry et al. 2012). Growth traits generally show cogradient variation, i.e. the genetic and plastic effects influence the phenotypes in the same direction along geographical or environmental gradients, and therefore plasticity is considered adaptive and the change in trait mean value along the gradient is accentuated (Conover et al. 2009; e.g. Caignard et al. 2021, Ensing and Eckert 2019, Kremer et al. 2014). A general pattern derived from population reaction norms in temperate and boreal forests is that populations at the cold limit of the species range would grow and survive more under warmer temperatures, while populations at the warm limit would benefit from colder temperatures (Fréjaville et al. 2020, Pedlar and McKenney 2017, Rehfeldt et al. 2003, 2002, 1999, Reich and Oleksyn 2008; but see Savolainen et al. 2007 for a different response to warming temperatures for height and survival in Scots pine). This suggests that populations at the climatic edges of the species range are currently maladapted, which may

be explained by gene flow from populations at the center of the distribution (García-Ramos and Kirkpatrick 1997, Kirkpatrick and Barton 1997, Kremer et al. 2012) or, for populations at the cold margins, by adaptation lags along post-glacial colonization routes (García-Valdés et al. 2013, Johnstone and Chapin 2003). In contrast to latitudinal gradients, adaptation along altitudinal gradients for growth and survival traits appears to be more variable among species, as some species show adaptation lags (e.g. Ponderosa pine in Martínez-Berdeja et al. 2019) while others show patterns of home-site advantage (e.g. Sakhalin fir in Ishizuka and Goto 2012; Jeffrey pine in Martínez-Berdeja et al. 2019), suggesting that they will not benefit from increased temperatures.

Phenology traits show either cogradient or countergradient variation depending on the species (e.g. Gauzere et al. 2020, Vitasse et al. 2009; and see the meta-analysis of Radersma et al. 2020 on plant reciprocal transplant experiments). Countergradient patterns arise when the genetic and plastic effects influence the phenotype clines in opposite directions (Conover and Schultz 1995), which may be an adaptation to counteract maladaptive plasticity (Crispo 2008, Grether 2005) or can also emerge under spatially and temporally fluctuating environments (King and Hadfield 2019, Scheiner 2013). To my knowledge, only one study derived population reaction norms for reproduction traits in forest trees. Caignard et al. (2021) found a countergradient in reproductive effort and a cogradient in growth along an altitudinal gradient in the white oak *Q. petraea*, i.e. trees from higher elevations grew less and produced more and larger fruits in a low-elevation common garden, whereas in field conditions, they still grew less but produced smaller and fewer fruits than trees from lower elevations. This pattern suggests a genetic trade-off between growth and reproduction and may be related to the demographic compensation phenomenon frequently observed in marginal natural populations (e.g. Doak and Morris 2010, Sheth and Angert 2018). Finally, not only considering traits related to the viability component of fitness (e.g. survival, growth), but also traits related to the reproductive component seems necessary to get a complete picture of how natural populations will perform over the long run under warmer temperatures.

To conclude, population reaction norms derived from multi-site common gardens have been extremely useful for identifying genetic and plastic effects on phenotypes along environmental gradients and for inferring population responses to future climates. However, they rely on phenotypic data from common gardens, which are expensive and time-consuming to maintain. Therefore, determining how genomic information may be combined with current modelling approaches to facilitate monitoring of population evolution in the face of global change would be extremely valuable and needed.

4.3 Specificities of forest tree genomics

Forest trees represent unique experimental systems to investigate plastic and adaptive variation in quantitative traits because their populations remain nearly undomesticated (i.e. their genetic variation has been little influenced by human-induced selection, even in species with breeding programs), they often have large effective population size, and they are distributed across wide geographical and environmental gradients (Alberto et al. 2013, Neale and Savolainen 2004). Most forest trees are outcrossing, have high lifetime reproductive output, and show important gene flow among populations through long-distance pollen dispersal (Kremer et al. 2012) and slow rates of macroevolution (i.e. low nucleotide substitution rates and low speciation rates; Petit and Hampe 2006). They are particularly challenging to study because of their long generation times and their large and complex genomes. In particular, conifer

genomes, apart from being large, display a rapid decay of linkage disequilibrium and contain numerous repetitive sequences (i.e. transposable elements) and gene duplications (Ahuja and Neale 2005, Kovach et al. 2010, Mackay et al. 2012, Morse et al. 2009, Zonneveld 2012). Therefore, extensive and dense genotyping is required in conifers to identify most of the relevant polymorphisms underlying (highly polygenic) quantitative traits (Jaramillo-Correa et al. 2015, Neale and Savolainen 2004). Since the whole genome sequencing of Norway spruce (19.6 Gbp; Nystedt et al. 2013), an increasing number of conifer reference genomes have been released, e.g. white spruce (20.8 Gbp; Birol et al. 2013, Warren et al. 2015), loblolly pine (20.1 Gbp; Neale et al. 2014, Zimin et al. 2014), sugar pine (31 Gbp; Stevens et al. 2016), coastal Douglas-fir (16 Gbp; Neale et al. 2017) and giant sequoia (8.1 Gbp; Scott et al. 2020). However, as whole genome sequencing remains highly challenging in conifers and concerns only a limited number of species and few individuals, candidate gene approaches are still mainly used to identify adaptive genetic variants and determine the genomic architecture of local adaptation.

As previously mentioned in Section 3.2, genomic information is already broadly used in tree breeding to predict phenotypes, but these predictions strongly depend on the relatedness among individuals and are therefore limited to within-population (or within-family) predictions. Recent pioneering studies in forest trees have explored how to incorporate genomic information in trait-based approaches encompassing multiple populations. Browne et al. (2019) used GEBVs in valley oak to identify individuals with the best performing progeny under future climates and showed that selecting these individuals for assisted gene flow strategies would help to considerably mitigate the predicted negative effects of rising temperatures on growth rates. In loblolly pine, Mahony et al. (2020) and MacLachlan et al. (2021) used counts of phenotype-associated positive-effect alleles (i.e. number of SNPs that each individual has among the 1% top hits from a GWAS on the trait of interest; that they called PEAs) as they were more robust to stochastic SNP sampling effects than PRS. Mahony et al. (2020) showed that local adaptation patterns of several traits (growth cessation and initiation, cold injury and shoot mass) were similarly or slightly better described by the PEAs than by climate or geographical data. MacLachlan et al. (2021) demonstrated the usefulness of PEAs for phenotypic predictions and to rapidly assess the effects of artificial selection on adaptive genetic variation of polygenic traits. These first studies pave the way for combining phenotypic, genomic and environmental data for predictive and monitoring purposes in forest trees, although a robust framework is still needed to make reliable phenotype predictions across species ranges and for many species.

Landscape genomics approaches based on genomic offset are also becoming popular in forest trees, in part because the founding studies of these methods were conducted on trees (poplar in Fitzpatrick and Keller 2015 and white oak species in Rellstab et al. 2016), and because these methods, as they do not require hard-to-collect data from common gardens, are particularly well suited to long-lived and sessile organisms. Populations at risk of short-term maladaptation (in a scenario of persistence in the current location without evolution or migration) have already been identified in many forest tree species, e.g. in balsam poplar (Fitzpatrick et al. 2021, Gougherty et al. 2020b, Keller et al. 2018), loblolly pine (Lu et al. 2019), European aspen (Ingvarsson and Bernhardsson 2020), yellow box (Supple et al. 2018) and cork oak (Vanhove et al. 2021). Interestingly, Gougherty et al. (2020b) differentiated the local (i.e. the commonly used metric, which assumes no gene flow), forward (i.e. minimum genetic distance between the focal population under current climate and all possible locations under future climates, which will be high if the focal population is maladapted to all future climates) and reverse (i.e. minimum genetic distance between the focal population under future climate and all possible locations under current climates, which will be high in regions where we expect changes in the gene-climate associations that are not found elsewhere) genomic offsets to identify populations

that might be maladapted to future climates and cannot benefit from migration or gene flow (natural or artificial). Much-needed experimental support of the genomic offset approach is also underway. For instance, in balsam poplar, Fitzpatrick et al. (2021) validated the predictions based on the genomic offset with data from common gardens and showed that they were more accurate than predictions based on climatic tranfer distances. Therefore, integrating genomic offset into the toolbox of forest tree conservation and management seems very promising, however, caution and further validation are still needed (Rellstab et al. 2021).

5 Objectives

The general objectives of the present PhD are to contribute to (i) accumulating knowledge on the mechanisms underlying population adaptation to the environment, and (ii) improving predictions of population responses to changing environments, such as climate change. Maritime pine (*Pinus pinaster* Ait), a long-lived conifer native to the western part of the Mediterranean Basin and the Atlantic regions of Iberian and southern France, is used as a case study. Chapter 1 aims to understand how quantitative genetic variation is maintained within populations by testing three competing, but not mutually exclusive, hypotheses for several traits: (i) admixed populations have higher quantitative genetic variation due to introgression from other gene pools, (ii) quantitative genetic variation is lower in populations from harsher environments (i.e. experiencing stronger selection; Fisher 1930), and (iii) quantitative genetic variation is higher in populations from spatially heterogeneous environments (McDonald and Yeaman 2018, Yeaman and Jarvis 2006). Importantly, chapter 1 provides insights into which populations may be able to adapt more quickly to climate change since the adaptive potential of populations depends directly on their genetic variation.

In the genomic era where the new sequencing technologies are now at affordable (and steadily decreasing) costs, an important part of this PhD dealt with how genomic data can be combined with environmental and phenotypic data to obtain more robust and fine-scale phenotypic predictions than current approaches, or predictions of the short-term population (mal)adaptation, without having to go through the cumbersome process of setting up common gardens. More specifically, chapter 2 investigates whether models combining climatic, phenotypic and genomic data could capture the underlying drivers of height-growth variation, and thus improve predictions at large geographic scales, especially compared to the predictions from climate-based population response functions that are currently commonly used in forest trees (e.g. Leites et al. 2012a, Rehfeldt et al. 1999). Finally, the first goal of chapter 3 is to identify maritime pine populations whose gene-environment relationships will be the most disrupted under climate change (i.e. populations at risk of short-term climate maladaptation) using, to meet this objective, the genomic offset approach (see Section 3.3; Fitzpatrick and Keller 2015). The second goal is to evaluate a key assumption of the genomic offset approach, namely that populations with the highest predicted genomic offset do show a decrease in absolute fitness or declining demographic trends (Capblancq et al. 2020a). Chapters 2 and 3 provide valuable information for managing maritime pine populations while accounting for adaptive processes: the genetic height-growth response of transferred individuals (populations) to new environments (chapter 2) and the short-term risk of climate maladaptation in local populations (chapter 3). Further, these two chapters investigate how to combine phenotypic, genomic, and environmental data in two very different modelling frameworks (trait-based approaches and landscape genomics, respectively), and thus contribute to the far-reaching goal of predicting how forest tree populations will respond to climate change, and which management

and conservation strategies will be most effective in rescuing declining populations.

III

Materials & methods

1 Model species

Maritime pine (*Pinus pinaster* Ait., Pinaceae) is an ecologically and economically important forest tree species, largely exploited for its wood (Viñas et al. 2016), widely used for stabilising coastal dunes in its Atlantic distribution, and considered as keystone species supporting forest biodiversity in large parts of its range (Fig. III.1). Native to the western part of the Mediterranean Basin, the Atlas mountains in Morocco and the south-west Atlantic coast of Europe, its natural distribution spans from the High Atlas mountains in the south (Morocco) to French Brittany in the north, and from the coast of Portugal in the west to western Italy in the east. It was also introduced for commercial purposes in Australia where it is now considered as a highly invasive species (Viñas et al. 2016).

Figure III.1. Genetic Conservation Unit (GCU) of maritime pine located in Lacanau (southwestern France).

.

Maritime pine is a wind-pollinated, outcrossing and long-lived tree species that can grow on a wide range of substrates, from sandy and acidic soils to more calcareous soils. It can also withstand many different climates: the dry climate along the northern coasts of the Mediterranean Basin (from Portugal to western Italy), the mountainous climates of south-eastern Spain and Morocco, the wetter climate of the Atlantic region (from the Spanish Iberian region to the western part of France) and the continental climate of central Spain.

As many Mediterranean tree species, maritime pine has a strong population genetic structure as well as highly fragmented populations (Alberto et al. 2013). Populations can be grouped into six gene pools (Jaramillo-Correa et al. 2015), that is, genetic clusters that cannot be differentiated on the basis of neutral genetic markers and that probably derive from a common glacial refuge (Bucci et al. 2007, Santos-del-Blanco et al. 2012).

2 CLONAPIN experiment

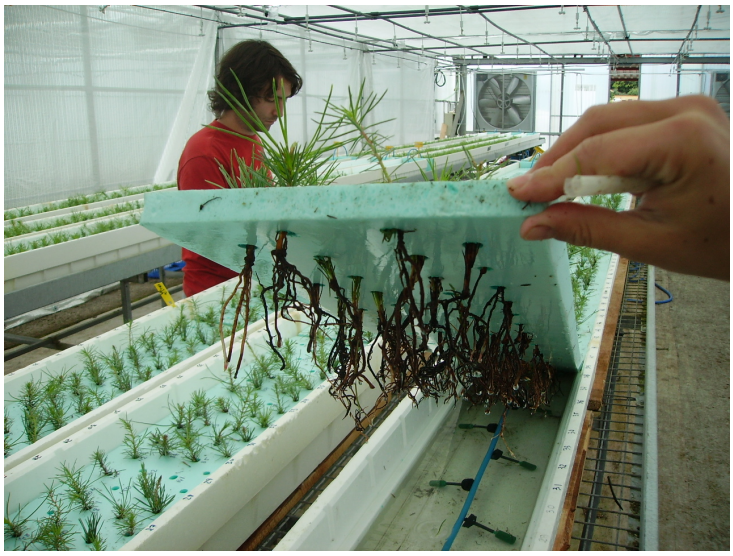
Phenotypic and genomic data used in the present PhD work comes from the clonal common garden network CLONAPIN, consisting of five test sites located in different environments (Fig. VI.1). Three sites are located in the Atlantic region, with mild winters, high annual rainfall and relatively wet summers: Bordeaux in the French part, and Asturias and Portugal in the Iberian part (Figs. III.2a & III.2b), the Portugal site experiencing slightly colder winters and half the summer precipitation than the site in Asturias. The two other sites, Cáceres and Madrid, are located in the Mediterranean region with high temperatures and intense summer drought, as well as large precipitation differences between summer and winter.

(a) Asturias site. **(b)** Portugal site. **(c)** Dead tree in Cáceres.

Figure III.2

In 2010 or 2011, depending on the test site, clonal replicates from 34 provenances were planted in a randomized complete block design with eight blocks (Fig. III.3a). For each population, trees represent between 2 and 28 clones (genotypes), on average about 15. To obtain unrelated clones, trees at least 50 m apart were sampled in natural stands, and one seed per tree was planted in a nursery and vegetatively propagated by cuttings (Fig. III.3b; see Rodríguez-Quilón et al. 2016 for details).

(a) Seedling plantation in Madrid.



(b) Nursery where the trees were vegetatively propagated to obtain clones.

Figure III.3

Tree mortality and height were measured in all common gardens and at different tree ages: 10, 21, and 37 months in Asturias, 25, 37, and 49 months in Bordeaux (plus at 13 months for mortality, and 85 months for height), 8 months in Cáceres, 13 months in Madrid, and 11, 15, 20 and 27 months in Portugal. Noticeably, tree height was not measured on dead trees, which resulted in a strongly unbalanced height measurements in Cáceres and Madrid in which 92% and 75% of the trees died, respectively (partly due to the clay soils and a strong summer drought; Fig. III.2c). Two phenology-related traits, the mean bud burst date and the mean duration of bud burst, were measured in Bordeaux when trees were 2, 3, 4 and 6 years old. Bud burst corresponds to the date of brachyblast emergence in accumulated degree-days (with base temperature 0°C) from the first day of the year to account for between-year variability in temperature. The duration of bud burst corresponds to the number of degree-days between the beginning of bud elongation and the total elongation of the needles (see Hurel et al. 2019). Last, two functional traits, $\delta^{13}C$ and the specific leaf area (SLA), were measured in Portugal.

The 34 populations planted in the five common gardens represent a rangewide sample of natural populations covering all known gene pools in maritime pine (Fig. V.1). A total of 523 clones collected in the Asturias common garden were genotyped with the Illumina Infinium assay, resulting in 5,165 high-quality polymorphic SNPs. There were on average only 3.3 missing values per genotype (ranging between 0 and 142). Details about DNA extraction and genotyping can be found in Plomion et al. (2016b). This first genomic dataset was used in CHAPTERS 1 and 2. In CHAPTER 3, we combined this first genomic dataset with another one developed within the framework of the H2020 EU B4EST project (4Tree; https://b4est.eu). SNPs with MAF < 1% or more than 20% missing data were filtered out, which resulted in 454 clones (i.e. genotypes) and 9,817 high-quality polymorphic SNPs, of which 2,855 were genotyped by both assays to ensure sample identity and estimate genotyping errors. The percentage of missing data per clone was less than 12% for all clones, with an average of 2.5%.

3 Bayesian inference

'The search for the most reliable paths to knowledge is arguably one of the greatest joys of being human.' (Hoang 2020).

Most of the statistical models of this PhD work were implemented within the Bayesian framework. The choice to use Bayesian statistics has several origins: (1) my interest in learning how to use Bayesian statistics that have already been extensively used in quantitative genetics (Beaumont and Rannala 2004, Shoemaker et al. 1999) and are increasingly common in evolutionary biology (Holder and Lewis 2003, Huelsenbeck et al. 2001, O'hara et al. 2008) and ecology (Clark 2005); (2) the greater flexibility, transparency, robustness and intuitiveness of Bayesian statistics (McElreath 2016); (3) the benefits in my daily life experience and reasoning that learning Bayesian philosophy, which some authors have even called a universal philosophy of knowledge (Hoang 2020), has brought to me. I will briefly present some key concepts of Bayesian inference, and some of the main epistemic differences between Bayesian and 'classical' frequentist statistics.

A first epistemic difference between the two approaches is that frequentist inference estimates the probability of the data given a particular hypothesis (or event or parameter) while Bayesian inference estimates the probability of a hypothesis being true (or the occurrence of an event, or the value of a parameter) in light of the available data (Ellison 2004). For example, suppose we have collected some observations y corresponding to random deviations from a normal distribution of mean μ and variance σ^2 (i.e. the parameters to be estimated). In a frequentist analysis, we will estimate $p(y|\mu, \sigma^2)$, i.e. the probability of the observed data y given the parameters μ and σ^2, known as the **likelihood** and estimated with the **maximum likelihood estimation**. In a Bayesian analysis, we will estimate $p(\mu, \sigma^2|y)$, i.e. the conditional probability of the parameters given the observed data. This probability is referred as the **posterior probability distribution** and, according to **Bayes's Theorem**, can be expressed as follows:

$$p(\mu, \sigma^2|y) = \frac{p(y|\mu, \sigma^2)p(\mu, \sigma^2)}{p(y)}$$

The first term of the numerator is the likelihood, identical to its frequentist counterpart. The second term of the numerator is the **prior probability distribution** and reflects a prior belief about the parameters μ and σ^2 expressed as a probability distribution, i.e. what we know before seeing the data. The denominator is often referred as the **marginal likelihood** and is the marginal probability density of the data across all possible parameters (a normalizing constant). It quickly becomes a high-dimensional integral impossible to solve as soon as several parameters have to be estimated. **Markov Chain Monte Carlo** (MCMC) algorithms bypass this problem by directly sampling the posterior distribution via probability ratios. They are both a strength of Bayesian statistics, being extremely flexible and robust (e.g. easily accommodating generalized linear mixed models; de Villemereuil 2019), but also a weakness as they require large computational capacities.

A second essential conceptual difference between Bayesian and frequentist statistics stems from their different interpretations of probability. A frequentist probability is defined as the **long-term frequency** of events in a sequence of trials (often hypothetical), while a Bayesian probability quantifies an individual's **degree of belief** (i.e. uncertainty) in the likelihood of an event (Ellison 2004). The interpretation of a **frequentist confidence interval** directly derives from this interpretation of probability, and corresponds to the range of values including the

true value of the parameter with some minimum probability, e.g. 95%. In other words, in n hypothetical runs of the study and analysis, 95% of the computed confidence intervals will cover the true parameter value. Therefore, a frequentist interval can strictly be interpreted only with respect to a sequence of similar inferences that might be replicated in practice. In contrast, the Bayesian interpretation of probability implies that probabilities are associated to any possible parameter values (i.e. inferences in the form of of a full posterior distribution). It follows that a **Bayesian credibility interval** is interpreted as an individual's belief that there is a 95% probability that the parameter of interest lies within the interval, which corresponds to an intuitive interpretation of the uncertainty around estimates. As applied statistics increasingly emphasizes interval estimation rather than hypothesis testing, Bayesian thinking seems particularly appropriate since it provides a common-sense interpretation of statistical conclusions (Gelman et al. 2020a).

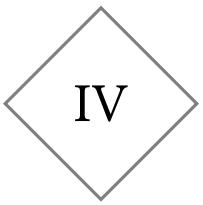
Another major strength of Bayesian statistics comes from the different benefits brought by the prior probability distribution. First, state-of-the-art knowledge, expert opinion or information from previous studies can be incorporated into prior distributions. In particular, as Bayesian inference is iterative, a posterior probability distribution derived from a previous analysis can then be used as prior in a subsequent analyses. Importantly, this allows uncertainty to be propagated between different subsequent analyses/studies (de Villemereuil 2019). In contrast, a frequentist analysis is always a *de novo* exercise, testing a null hypothesis by assuming that there is no relevant information available, even if the null hypothesis had been repeatedly falsified in previous experiments (Ellison 2004).

The prior probability distribution can also be used for **regularization**, a statistical procedure used to reduce overfitting and give more stable posterior estimates (Gelman et al. 2020a, McElreath 2016). For instance, '**weakly informative priors**' are regularizing priors that contain enough information to keep the posterior distributions within plausible bounds without aiming at fully incorporating one's expert knowledge about the underlying parameters (Gelman 2006, Gelman et al. 2008). In other words, weakly informative priors will have only a marginal influence on the posterior distribution if there is a large enough amount of data (in which case the likelihood will dominate), whereas if there is little data, they can have a considerable effect on posterior inference. Hence, such priors can prove highly useful in enabling inference of key parameters that would otherwise be impossible to estimate, especially in very small samples (de Villemereuil 2019, McElreath 2016). In this PhD work, I mostly used weakly informative priors.

Historically, Bayesian inference has been criticized as 'subjective' because of the requirement to make a necessarily arbitrary choice of priors (McElreath 2016). However, non-Bayesian procedures also have subjective choices to make, which are often swept under the rug, such as the choice of estimator or likelihood penalty (McElreath 2016). The inclusion of prior information in a completely transparent way in Bayesian procedures therefore makes them less opaque than likelihood-based approaches. It is now easy to find recommendations for the priors to use depending on the questions being asked, the likelihood and the different types of parameters (e.g. 'Prior choice recommendations' of Andrew Gelman). Moreover, the consistency of the chosen priors with domain knowledge can be assessed with **prior predictive checks**, which simulate predictions from a model using only the prior distribution instead of the posterior distribution (Gabry et al. 2019, McElreath 2016). Last, it is important to keep in mind that as the amount of data increases, the influence of priors on posterior inference decreases and the estimates of Bayesian or frequentist analyses converge towards the same values.

Finally, Bayesian inference is above all characterized by its explicit use of probability for quantifying uncertainty, and can be defined as 'the process of fitting a probability model to a set of data and summarizing the result by a probability distribution on the parameters of the model and on unobserved quantities such as predictions for new observations' (Gelman et al. 2020a). It provides a very intuitive, robust and powerful framework (see the very complete Bayesian workflow recently proposed in Gelman et al. 2020b), with great flexibility and generality to deal with complex problems (Gelman et al. 2020a). This applies to **hierarchical models** (also known as multilevel models), which have been extensively used in the present PhD work. Hierarchical models are used when information is available on several different levels of observation units (e.g. genotype, population, test site) and exchangeable at each level of units (Gelman et al. 2020a).

In the present PhD work, I used the *Stan* probabilistic programming language (Carpenter et al. 2017) and its default Hamiltonian Monte Carlo algorithm, the no-U-turn sampler (NUTS; Hoffman and Gelman 2014). The Bayesian models of CHAPTERS 1 and 3 were implemented through the R package *rstan*, the R interface of *Stan*. Models in CHAPTER 2 were implemented through the R package *brms* (Bürkner 2017), whose functions are passed to *rstan*.



Chapter 1. Extreme climatic events but not environmental heterogeneity shape within-population genetic variation in maritime pine

Juliette Archambeau[1], Marta Benito Garzón[1], Marina de Miguel Vega[1,2], Benjamin Brachi[1], Frédéric Barraquand[3], Santiago C. González-Martínez[1]

1 INRAE, Univ. Bordeaux, BIOGECO, F-33610 Cestas, France
2 EGFV, Univ. Bordeaux, Bordeaux Sciences Agro, INRAE, ISVV, F-33882, Villenave d'Ornon, France
3 CNRS, Institute of Mathematics of Bordeaux, F-33400 Talence, France

Keywords: quantitative genetic variation, adaptive potential, forest tree, natural selection, environmental heterogeneity, severe cold events.

Preregistration: The objectives and methods of this manuscript have been pre-registered at the Center For Open Science and are available here (anonymized repository): https://osf.io/knx6z/?view_only=41bb7b5cbf7241d0856e8b9e393cc795.

1 Abstract

How evolutionary forces interact to maintain quantitative genetic variation within populations has been a matter of extensive theoretical debates. While mutation and migration increase genetic variation, natural selection and genetic drift are expected to deplete it. To date, levels of genetic variation observed in natural populations are hard to predict without accounting for other processes, such as balancing selection in heterogeneous environments. We aimed to empirically test three hypotheses: (i) admixed populations have higher quantitative genetic variation due to introgression from other gene pools, (ii) quantitative genetic variation is lower in populations from harsher environments (i.e. experiencing stronger selection), and (iii) quantitative genetic variation is higher in populations from spatially heterogeneous environments. We used phenotypic measurements of five growth, phenological and functional traits from three clonal common gardens, consisting of 523 clones from 33 populations of maritime pine (*Pinus pinaster* Aiton). Populations from harsher climates (mainly colder areas) showed lower genetic variation for height in the three common gardens. Surprisingly, we did not find any association between within-population genetic variation and environmental heterogeneity or population admixture for any trait. Our results suggest a predominant role of natural selection in driving within-population genetic variation, and therefore indirectly their adaptive potential.

2 Introduction

Most complex traits show substantial heritable variation in natural populations. How evolutionary forces interact to maintain such variation remains a long-standing dilemma in evolutionary biology and quantitative genetics (Johnson and Barton 2005). While mutation and genetic drift have straightforward roles, generating and eliminating variation respectively, the effect of natural selection is more complicated (Walsh and Lynch 2018). Stabilizing selection, i.e. the selection of intermediate phenotypes, is often strong in natural populations (Hereford et al. 2004). This type of selection is expected to deplete genetic variation (Fisher 1930), either directly on the focal trait or indirectly via pleiotropic effects (Johnson and Barton 2005). Theoretical models based on the balance between mutation, drift and stabilizing selection support this idea, but they suggest lower fitness heritability values than those generally observed in empirical studies (Johnson and Barton 2005). Balancing selection encompasses various evolutionary processes that can maintain greater than neutral genetic variation within populations (Mitchell-Olds et al. 2007). The most widely studied of these processes are heterozygote advantage, frequency-dependent selection, e.g. in disease resistance or self-incompatibility systems (Bergelson et al. 2001, Charlesworth et al. 2005), and temporally or spatially fluctuating selection pressures (Felsenstein 1976). The maintenance of stable polymorphism in spatially heterogeneous environments was first theorized by Levene's archetypal model (1953), under the assumptions of random mating within generations and soft selection. Since then, a large corpus of single-locus and polygenic models, most often deterministic, have generally concluded that genetic polymorphisms can only be maintained under restrictive conditions (Byers 2005, Spichtig and Kawecki 2004). In this line, McDonald and Yeaman (2018) showed with stochastic individual-based simulations that substantial within-population genetic variation can be maintained in spatially heterogeneous environments at intermediate migration rates, regardless of population size. However, the relative importance of the different evolutionary forces driving within-population genetic variation remains largely unknown.

Long-dating empirical work has addressed the evolutionary processes underlying the maintenance of genetic and discrete-trait polymorphisms (reviewed in Hedrick 1986, 2006), e.g. plant-pathogen interactions (Karasov et al. 2014), antagonistic pleiotropy (Carter and Nguyen 2011), environmental heterogeneity (Chakraborty and Fry 2016), and temporal fluctuations (Bergland et al. 2014). Genomics have allowed the broad application of genome-wide scans for signatures of selection. Overall these scans suggest that many loci are under adaptive directional selection (Barreiro et al. 2008, Fu and Akey 2013) and that the proportion of genetic polymorphisms maintained by environmental heterogeneity tends to be low (Hedrick 2006). However those scans typically have low power to detect signatures of balancing selection or local adaptation (Fijarczyk and Babik 2015). Far fewer empirical studies have focused on assessing the distribution and extent of the quantitative genetic variation within populations, and its underlying causes (Lynch and Walsh 1998). Traits more closely related to fitness, such as life-history traits, have generally higher additive genetic variance, but lower heritabilities, than morphometric traits (Houle 1992, Kruuk et al. 2000, Price and Schluter 1991). The hypothesis that populations evolving under strong selection pressures display lower levels of genetic variation has been supported in experimentally evolving quail populations under unfavorable vs favorable treatments (Marks 1978), in controlled experiments (Colautti et al. 2010; but see Merilä et al. 2004, Stock et al. 2014), in natural populations of *Drosophila birchii* subject to climatic selection (but see *D. bennata* and *D. serrata*; van Heerwaarden et al. 2009) and in some natural populations of great tits subject to varying levels of food availability (Charmantier et al. 2004). Higher genetic variation in populations evolving under spatially varying selection pressures is supported by experimental evolution of *Drosophila* populations (Huang et al. 2015, Mackay 1981; but not Yeaman et al. 2010) and in forest trees evaluated in common gardens (Yeaman and Jarvis 2006). The lack of general trends from these empirical studies can be explained by method-specific pitfalls to accurately estimate quantitative genetic variation, e.g. the genetic and environmental variances are hard to disentangle in the wild, and when estimated in common gardens, their environment-dependent nature does not allow for wide generalization of estimates (Charmantier et al. 2004, Hoffmann and Parsons 1991, Merilä et al. 2001). In addition, gene flow has been hypothesized to have either a positive effect on the adaptive potential, by increasing standing genetic variation, or a negative effect via gene swamping (Kremer et al. 2012, Tigano and Friesen 2016), which may depend on the spatial scale considered (Bridle et al. 2009).

Forest trees have specific life-history traits and genomic features making them interesting model species in population and quantitative genetic studies (Petit and Hampe 2006, Savolainen et al. 2007). Compared to crop species, they remain largely undomesticated (Neale and Savolainen 2004). Most forest trees are outcrossing, have high lifetime reproductive output and long generation times. They often display important gene flow among populations through long-distance pollen dispersal (Kremer et al. 2012). They show slow rates of macroevolution (i.e. low nucleotide substitution rates and low speciation rates; Petit and Hampe 2006), generally have large effective population sizes, with distributions often covering a wide range of environmental conditions (Alberto et al. 2013). Extensive work has revealed strong clines at large geographical scales in the population-specific mean values of phenotypic traits (reviewed in Benito Garzón et al. 2019, Savolainen et al. 2007), e.g. phenological traits with latitude or altitude (Alberto et al. 2011, Thibault et al. 2020) or height growth with cold hardiness (Leites et al. 2012b, Rehfeldt et al. 1999). Genetic differentiation at microgeographic spatial scales has also been repeatedly observed (reviewed in Jump and Peñuelas 2005, Linhart and Grant 1996, Scotti et al. 2016), suggesting rapid rates of microevolution (Petit and Hampe 2006). Possible explanations include the fact that forest trees have high levels of genetic diversity and that most of their quantitative and neutral genetic variation is within populations (Hamrick 2004).

To our knowledge, only two empirical studies investigated the potential causes underlying the maintenance of quantitative trait variation within forest tree populations. Yeaman and Jarvis (2006) showed that 20% of growth genetic variation in lodgepole pine populations was attributable to regional heterogeneity, suggesting an important role of gene flow and varying selection pressures. In the neotropical oak *Q. oleoides*, Ramírez-Valiente et al. (2019) found lower quantitative genetic variation in harsher environments, but not higher quantitative genetic variation in temporally fluctuating environments. They also suggested only a marginal effect of genetic structure and diversity on the maintenance of within-population genetic variation.

In this study, we aimed to test competing hypotheses regarding the relationship between quantitative genetic variation within maritime pine populations and the potential underlying drivers that maintain this variation. We used phenotypic measurements of growth (height), phenological (bud burst and duration of bud burst) and functional (δ^{13}C and specific leaf area, SLA) traits from three clonal common gardens, consisting of 522 clones (i.e. genotypes) from 33 populations, spanning all known gene pools in the species (Jaramillo-Correa et al. 2015) and genotyped for 5,165 SNPs. For each trait, we compared Bayesian hierarchical models that estimate the relationship between the total genetic variances within populations and some potential drivers, namely climate's harshness at the locations of origin of the populations (i.e. drought intensity and severe cold events), environmental heterogeneity in the forested areas surrounding the populations, and the level and origin of admixture in the populations, as estimated with SNP markers. The competing, but not mutually exclusive, hypotheses tested are: i) the most admixed populations have higher quantitative genetic variation due to introgression from other gene pools, and this relationship is proportional to the divergence between sink and source gene pools; ii) quantitative genetic variation is lower in populations that have evolved in harsher environments, as a result of higher selection pressures in these regions; and iii) quantitative genetic variation is higher in populations that have evolved in spatially heterogeneous environments. Importantly, the last two hypotheses require the action of natural selection, while the first does not. Therefore, we expect the last two hypotheses to be mostly supported for fitness-related traits, while the first hypothesis may apply uniformly to all traits. Determining the patterns of within-population quantitative genetic variation across species' ranges and the relative importance of the evolutionary forces driving the maintenance of such variation is necessary to assess the evolutionary potential of forest tree populations. Empirical studies tackling these questions remain extremely rare in forest trees (but see Ramírez-Valiente et al. 2019, Yeaman and Jarvis 2006), yet they are much needed to anticipate forest tree responses to ongoing global change and therefore develop adaptive management and conservation strategies.

3 Materials & Methods

3.1 Maritime pine, a forest tree growing in heterogeneous environments

Maritime pine (*Pinus pinaster* Ait., Pinaceae) is a wind-pollinated, outcrossing and long-lived tree species with large ecological and economical importance in western Europe and North Africa. Maritime pine is largely appreciated for its wood, for stabilizing coastal and fossil dunes and, as a keystone species, for supporting biodiversity (Viñas et al. 2016). The distribution of maritime pine natural populations is scattered and covers a wide range of environmental

conditions. Several studies have provided evidence of genetic differentiation for adaptive traits in this species, suggesting local adaptation (e.g. de Miguel et al. 2020, González-Martínez et al. 2002). Maritime pine can grow in widely different climates: the dry climate along the northern coasts of the Mediterranean Basin (from Portugal to western Italy), the mountainous climates of south-eastern Spain and Morocco, the wetter climate of the Atlantic region (from the Spanish Iberian region to the western part of France) and the continental climate of central Spain. Maritime pine can also grow on a wide range of substrates, from sandy and acidic soils to more calcareous ones. Maritime pine presents a strong population genetic structure with occasional admixture, suggesting gene flow among gene pools. Six gene pools have been described by previous literature, located in the French Atlantic region, Iberian Atlantic region, central Spain, south-eastern Spain, Corsica and Northern Africa (Fig. 1; Alberto et al. 2013, Jaramillo-Correa et al. 2015). These gene pools probably result from the expansion of different glacial refugia (Bucci et al. 2007).

3.2 Phenotypic data

Phenotypic data was obtained from three clonal common gardens (Table X.1 and Fig. IV.1), planted in 2011 and located in environments considered favorable to maritime pine, as evidenced by the high survival rate at these sites (Table X.1). The common gardens of Asturias (Spain, Iberian Atlantic region) and Bordeaux (France, French Atlantic region) have very similar climates, with mild winters, no severe cold events, high annual rainfall and relatively wet summers (Tables X.3-X.5 and Fig. IV.1). The common garden of Portugal (planted in Fundão) shows slightly colder winters and lower summer precipitation than in Asturias (Table X.4 and Fig. IV.1). In each of these common gardens, trees belonging to 522 clones (i.e. genotypes) from 33 populations, including the six known gene pools in the species, were planted following a randomized complete block design with 8 blocks, 8 trees per clone and from 2 to 28 clones per population (with an average of 15). To obtain the clones, trees at least 50 m apart were sampled in natural stands, and one seed per tree was planted in a nursery and vegetatively propagated by cuttings (see Rodríguez-Quilón et al. 2016 for details). Clones were therefore considered unrelated.

One growth trait, height, was measured in all common gardens and at different tree ages (Table X.1). Two phenology-related traits, the mean bud burst date over four years and the mean duration of bud burst over three years, were measured in Bordeaux and were averaged over several years to suppress differences across years and approximate a normal distribution of their trait values (Table X.1). Bud burst corresponds to the date of brachyblast emergence in accumulated degree-days (with base temperature 0°C) from the first day of the year to account for between-year variability in temperature. The duration of bud burst corresponds to the number of degree-days between the beginning of bud elongation and the total elongation of the needles (see Hurel et al. 2019). Last, two functional traits, δ^{13}C and the specific leaf area (SLA) were measured in Portugal (Table X.1). These traits were selected because they showed broad-sense heritabilities that were mostly low but with credibility intervals not crossing zero (> 0.08 in de Miguel et al. 2020). For each trait, phenotypic means and variances across populations are shown in Section 1.1 of the Supplementary Information. Prior to analyses, some traits were log-transformed to get closer to normality or mean-centered to help model convergence (Table X.1).

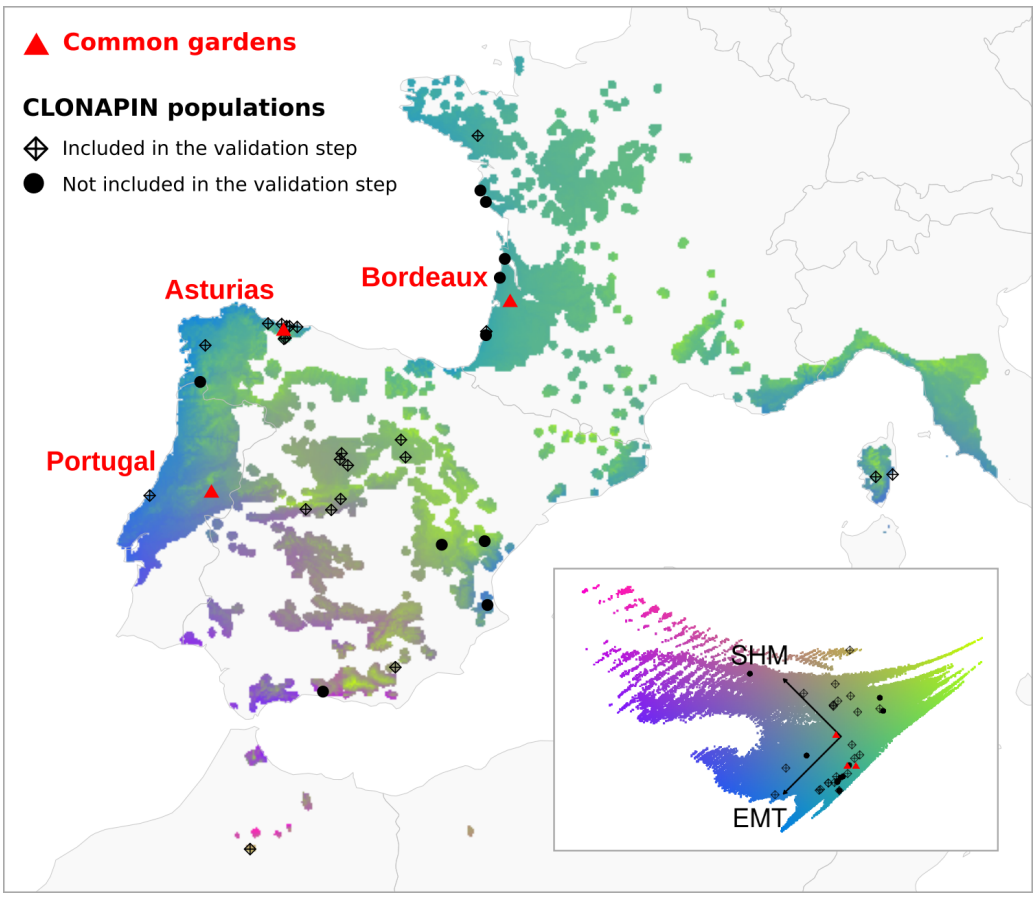


Figure IV.1. Location of the three common gardens and the 33 populations used in the study. The colors represent the gradients of the extreme minimum temperature (EMT) and summer heat moisture index (SHM) over the period 1901-1950 within the maritime pine range. The climatic gradients were obtained by performing a centered and scaled principal component analysis (shown in the inset on the bottom right) based on EMT and SHM values. The maritime pine distribution combines the EUFORGEN distribution (http://www.euforgen.org/) and 10-km radius areas around the National Forest Inventory plots with maritime pines. However, this remains a rough approximation of the actual distribution of maritime pine and therefore probably includes areas experiencing more intense cold or drought episodes than the climatic range of maritime pine.

3.3 SNP genotyping and population admixture

The 522 clones planted in the Asturias common garden were genotyped with the Illumina Infinium assay described in Plomion et al. (2016a), resulting in 5,165 high-quality polymorphic SNPs. There were on average only 3.3 missing values per genotype (ranging between 0 and 142). For each clone, the proportions of ancestry from each of the six known gene pools were estimated in Jaramillo-Correa et al. (2015) using the Bayesian approach available in Structure (Pritchard et al. 2000), and were then averaged by population. Populations were assigned to the gene pool that contributed more than 50% ancestry and the other gene pools were considered as 'foreign' gene pools. First, we calculated a population admixture score A, as the proportion of ancestry from foreign gene pools (Table X.6). Second, we calculated a population admixture score D that considers both the proportion of foreign ancestries and the divergence between the main and foreign gene pools (Table X.6). For that, we weighted the proportions of ancestry from foreign gene pools by the sum of the allele frequency divergence of the main and foreign gene pool from the common ancestral one (F_k, which should be numerically similar to F_{ST}; Falush et al. 2003). We developed D considering that some gene pools are more divergent than others and thus may bring higher genetic diversity to an admixed population at the same level of introgression. A was highly correlated with D (Pearson correlation coefficient of 0.91; Table X.6), and also with a related index that used weights based on pairwise F_{ST} (Table X.6).

3.4 Population-specific environmental heterogeneity and climate harshness indexes

To describe the climate under which the populations have evolved, we used the climatic variables at 1-km resolution and averaged over the period 1901-1950 from the *ClimateEU* database (Marchi et al. 2020). Topographic data were generated from NASA's Shuttle Radar Topography Mission (SRTM) at 90-m resolution and then aggregated at 1-km resolution. We used the SAGA v 2.3.1 (Conrad et al. 2015) to calculate the topographic ruggedness index (TRI) which quantifies the terrain heterogeneity, i.e. differences in elevation between adjacent cells (Riley et al. 1999). Soil variables were extracted from the European Soil Database at 1-km resolution (Hiederer et al. 2013). All environmental variables used are listed in Table X.7 and were mean-centered and divided by their standard deviation prior to analyses.

To calculate the environmental heterogeneity around each population location, we extracted raster cell values of the climatic, topographic and soil variables within a 20-km radius around each population location, and kept only raster cells that fell within forested areas, to avoid including environmental data from non-suitable areas (e.g. lakes, mountain peaks; Section 1.3.2 of the Supplementary Information). We then performed a principal component analysis (PCA) on the raster cell values and extracted the PC1 and PC2 scores of each cell, accounting for 45.2% and 34.1% of the variance, respectively (Fig. X.10). To obtain the four indexes of environmental heterogeneity, we calculated the variances of the PC1 and PC2 scores in a 20-km and 1.6-km radius around each population location. The environmental heterogeneity indexes were only very weakly correlated (Pearson correlation coefficients lower than 0.36) with the number of forested cells (i.e. the area considered to calculate the indexes), ensuring that the estimated effects of environmental heterogeneity in further analyses were not due to the area per se (Stein et al. 2014, Triantis et al. 2003).

To describe the climate harshness at each population location, we used a drought index (the summer heat moisture index averaged over the period 1901-1950, SHM, Table X.7) and an index related to severe cold events (the inverse of the extreme minimum temperature during the period 1901-1950, invEMT, Table X.7). These two indexes were selected as maritime pine shows local adaptation patterns associated with cold tolerance (Grivet et al. 2011) and because detecting changes in the within-population genetic variation along a drought gradient would be key to anticipate tree population responses to ongoing climate change.

3.5 Bayesian statistical modelling

We modeled the eight phenotypic traits with the same Bayesian statistical model, in which we estimate the linear relationship between the within-population genetic variance and each of the potential drivers successively (i.e. one model per driver): the two admixture scores, the four environmental heterogeneity indexes and the two climate harshness indexes. Each trait y followed a normal distribution (Fig. X.1), such as:

$$\begin{aligned} y_{bpcr} &= \mathcal{N}(\mu_{bpc}, \sigma_r^2) \\ \mu_{bpc} &= \beta_0 + B_b + P_p + C_{c(p)} \end{aligned} \tag{3.1}$$

where σ_r^2 is the residual variance, β_0 the global intercept, and B_b, P_p and $C_{c(p)}$ are the block, population and clone (nested within population) varying intercepts, which are drawn from a

common distribution, such as:

$$\begin{bmatrix} B_b \\ P_p \end{bmatrix} \sim \mathcal{N}\left(0, \begin{bmatrix} \sigma_B^2 \\ \sigma_P^2 \end{bmatrix}\right)$$
$$C_{c(p)} \sim \mathcal{N}(0, \sigma_{C_p}^2) \tag{3.2}$$

where σ_B^2 and σ_P^2 are the variance among blocks and populations and $\sigma_{C_p}^2$ are the population-specific variances among clones (i.e. the within-population genetic variation). To estimate the association between $\sigma_{C_p}^2$ and its potential underlying drivers, we expressed σ_{C_p} as follows:

$$\sigma_{C_p} \sim \mathcal{LN}\left(\ln(\overline{\sigma_{C_p}}) - \frac{\sigma_K^2}{2} + \beta_X X_p, \sigma_K^2\right) \tag{3.3}$$

where $\overline{\sigma_{C_p}}$ is the mean of the population-specific standard deviation among clones σ_{C_p} and X_p is the potential driver considered (see Section 2 in the Supplementary Information for more details).

To test the accuracy of the model estimates for σ_K^2 and β_X, we simulated data based on two traits (height in Portugal and Bordeaux at 20 and 25-month old, respectively). For each trait, we ran 100 simulations and extracted the mean standard error and bias error of the estimates and the coverage of the 80% and 95% credible intervals.

Model specification and fit were performed using the Stan probabilistic programming language (Carpenter et al. 2017), based on the no-U-turn sampler algorithm. Models were run with four chains and between 2,500 iterations per chain depending on the models (including 1,250 warm-up samples not used for the inference). All analyses were undertaken in R version 3.6.3 (R Core Team 2020) and scripts are available at https://github.com/JulietteArchambeau/H2Pinpin.

3.6 Validation step on independent data

To validate our results for height, we used an independent dataset provided by Ricardo Alía in which 23 populations shared with the CLONAPIN network were planted in a progeny test near Asturias (thus in a similar environment). As the progeny test is based on families, we were able to estimate the additive genetic variance within populations. We applied the same model as in our study (replacing clones by families) to height measurements when the trees were 3 and 6-year old (see Section 8 of the Supplementary Information for more details).

4 Results

In the data simulation, σ_K^2 (the standard deviation of the logarithm of the within-population genetic variation) and β_X (the coefficient of the potential drivers of the within-population genetic variation) were properly estimated by the models (Table X.9 and X.10). Across 100 simulations, the mean standard error was around 0.066 for σ_K^2 and 0.054 for β_X, the mean bias error was around 0.018 for σ_K^2 and -0.004 for β_X, the coverage of the 80% credible interval was around 93% for σ_K^2 and 80% for β_X, and the coverage of the 95% credible interval was around

98% for σ^2_K and 96% for β_X (Table X.9 and X.10). These simulations therefore showed that, under the assumption that the statistical model reflects the processes at work, our model displayed a satisfactory accuracy to be used in the following analyses.

The proportion of variance explained by the models (i.e. the sum of the among-population, among-clone and among-block variances) and the variance partitioning varied broadly across traits (Fig. X.11 and Section 5.3 in the Supplementary Information). More specifically, the models explained between 40% and 50% of the variance for phenology-related traits, between 30% and 40% for functional traits, and from 20% for height in Portugal to almost 60% for height in Bordeaux at 85-month old (Fig. X.11). Residual variance explained most of the variance for all traits, except for height in Bordeaux at 85-month old, where 40% of the variance came from variation among populations, 40% from residuals and the remaining 20% from variation among clones (Fig. X.18). Variation among populations was higher than variation among clones for height and δ_{13}C (Figs. X.14, X.16, X.18, X.20 and X.28), but not for SLA and phenology-related traits (Figs. X.26, X.22 and X.24).

Environmental heterogeneity indexes and population admixture scores were not associated with within-population genetic variation for any trait (Figs. IV.2 and X.12). In contrast, we found a consistent negative association with the inverse of the extreme minimum temperature across the three common gardens for height, indicating that populations undergoing severe cold events display less genetic variation (Fig. IV.2). Interestingly, in the Bordeaux common garden, this negative relationship was found at 25-month old, but not at 85-month old (Fig. IV.2). A negative association with the summer heat moisture index was also detected for height in Asturias, and less markedly but still with a high probability in Bordeaux at 25-month old (Fig. IV.2). Holding all other parameters constant, a one-standard deviation increase in the inverse of the extreme minimum temperature was associated, on average, with a 32.6%, 21.6% and 17.9% decrease of σ_{C_p} for height in Portugal, Bordeaux at 25-month old and Asturias, respectively. Similarly, a one-standard deviation increase in the summer heat moisture index was associated, on average, with 15.6% and 23.8% decrease of σ_{C_p} for height in Bordeaux at 25-month old and Asturias, respectively (see details of the calculation in Section 4 of the Supplementary Information). Unexpectedly, populations experiencing severe cold events showed higher genetic variation for SLA (Fig. IV.2). Within-population genetic variation was not correlated with the number of clones per population for any trait (maximum Pearson correlation coefficient = 0.57; Table X.11).

Importantly, in the validation analysis, we also found a negative association between the inverse of the extreme minimum temperature and the within-population additive genetic variation for height at 3-year old, but not at 6-year old, and we did not find any association with the other potential drivers (Fig. IV.3).

5 Discussion

How quantitative genetic variation is maintained within populations remains a long-standing open question that has been extensively explored in theoretical work but lacks empirical evidence to date (Johnson and Barton 2005). Our study suggests that genetic variation for height in maritime pine is lower in populations exposed to severe cold events, thus supporting the hypothesis that quantitative genetic variation in fitness-related traits is lower in populations under strong selection (Fisher 1930). Across all traits studied, we did not find higher genetic

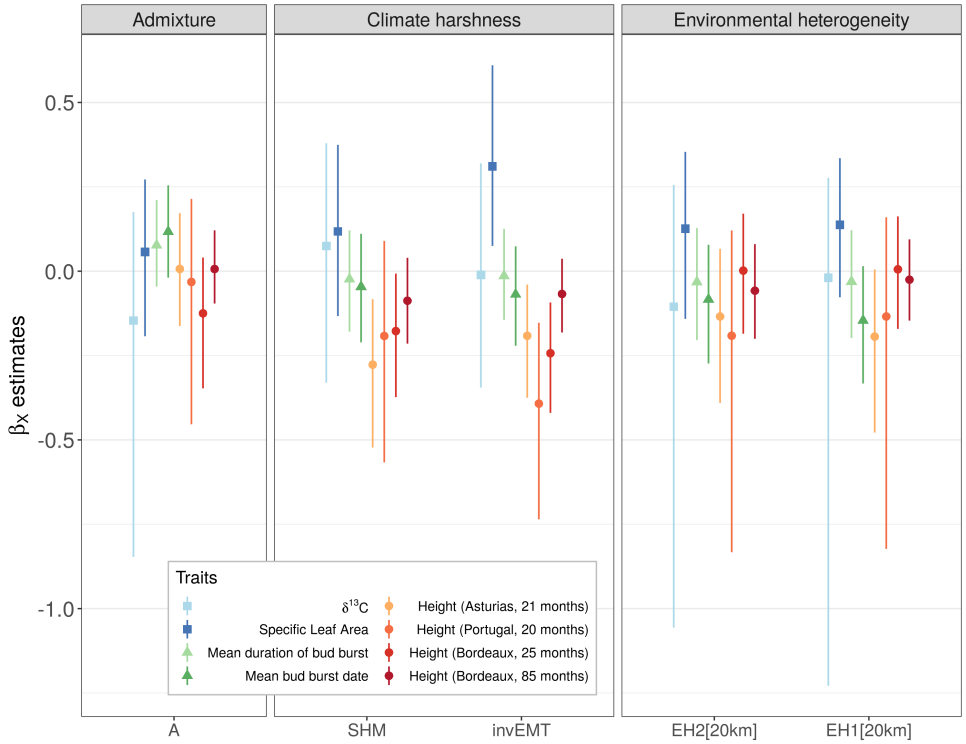


Figure IV.2. Median and 95% credible intervals of the β_X posterior distributions. β_X coefficients stand for the association between the within-population genetic variation and its potential underlying drivers on the x-axis: the inverse of the extreme minimum temperature during the studied period (invEMT), the summer heat moisture index (SHM), an admixture score (A), the environmental heterogeneity in a 20-km radius around the population location (EH1[20km] and EH2[20km]) calculated based on the projection of the PC1 and PC2 scores. Colors stand for the different traits under study and the shapes for the different types of traits, i.e. functional traits (squares), phenology-related traits (triangles) and height (circles).

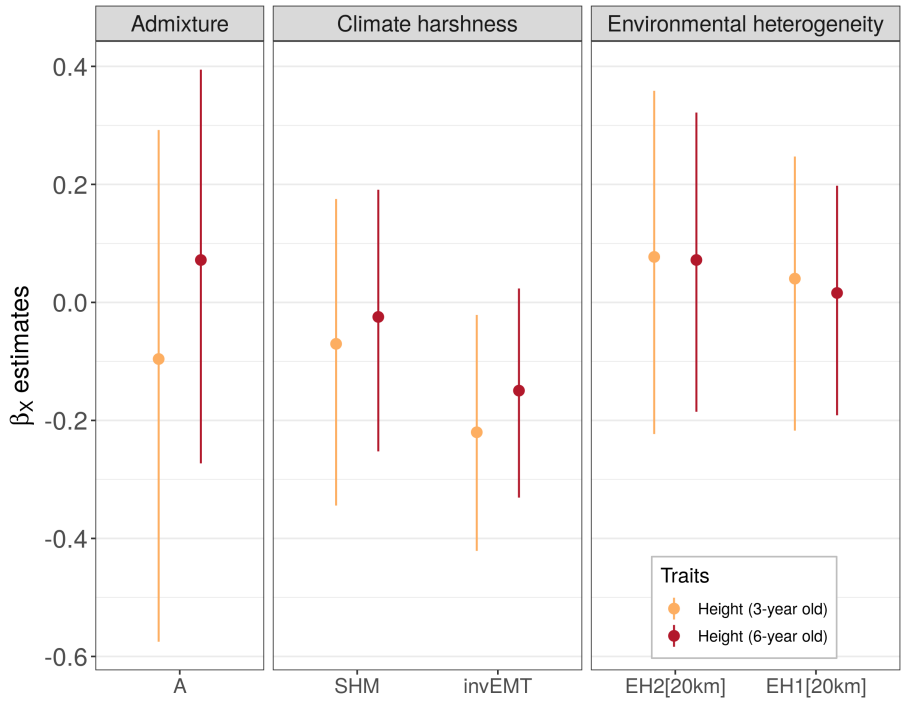


Figure IV.3. Validation step using independent height measurements from a common garden near Asturias. Median and 95% credible intervals of the β_X posterior distributions are shown. In the validation analysis, β_X coefficients stand for the association between the within-population additive genetic variation and its potential underlying drivers on the x-axis. A description of the drivers can be found in the legend of Fig. IV.2.

variation in populations located in heterogeneous landscapes, which goes against the predictions of some theoretical models (McDonald and Yeaman 2018, Walsh and Lynch 2018) and an empirical study in lodgepole pine (Yeaman and Jarvis 2006). Admixed populations did not show higher genetic variation, suggesting that the observed patterns are not confounded by gene flow between distinct gene pools increasing genetic variation. Empirically-based detection of the footprints of natural selection on within-population genetic variation is much needed to understand how populations are adapted to their current environments and will evolve under changing conditions.

5.1 Severe cold events may decrease within-population genetic variation

Height genetic variation was lower in populations experiencing harsher climates, especially severe cold events (invEMT; Fig. IV.2). This result supports the hypothesis that strong stabilizing selection in harsh environments depletes quantitative genetic variation within populations (Fisher 1930) and echoes similar results in another forest tree, *Quercus oleoides*. For this Mesoamerican white oak species, Ramírez-Valiente et al. (2019) found lower genetic variation averaged over functional and growth traits in populations experiencing low precipitation and high temperatures during the dry season. The importance of severe cold events as a driver of height genetic variation in maritime pine is supported by the association between candidate-gene allele frequency and temperature gradients (Grivet et al. 2011, Jaramillo-Correa et al. 2015), suggesting a major role of minimum temperatures in the species adaptive evolution. Indeed, lower genetic variation in areas subject to cold events may enhance adaptation to local conditions, but it may also hamper the adaptive potential of populations under new climates. Noticeably, severe cold events were highly correlated with altitude in our study (Pearson's correlation of 0.9), and adaptation patterns along altitudinal gradients are common in forest trees (e.g. Kurt et al. 2012). Therefore, we cannot exclude that the association between height genetic variation and severe cold events is triggered by more complex environmental factors typical of high altitude conditions (e.g. reduced vapor pressure deficit, higher maximum solar radiation; Körner 1995).

The lower within-population genetic variation for height in populations experiencing harsher climates was unlikely to be the result of demographic factors (i.e. processes that affect the effective population size independently from natural selection; Lawton-Rauh 2008) given that (i) the pattern of reduced genetic variation was only observed for height and not for the other traits, whereas we would have expected demographic factors to impact all traits similarly; (ii) we did not find any association between quantitative genetic variation and genetic diversity estimated with molecular markers (i.e. expected heterozygosity) for any trait (see Section 7 of the Supplementary Information; Rodríguez-Quilón et al. 2015) suggesting negligible effects on trait genetic variation of differences in effective population size among populations; and (iii) our results suggest a negligible impact of gene flow across gene pools (as evaluated by population admixture indexes) on within-population genetic variation for any trait (Fig. IV.2).

The differences in height genetic variation among populations were also unlikely to originate from the expression of hidden genetic variation in novel environments (i.e. 'cryptic genetic variation'; Schlichting 2008) as the lower height genetic variation in populations from harsher climates was consistent across the three common gardens (i.e. independent of their environmental conditions) and thus likely to be intrinsic to the populations. Last, the sampled populations may not fully cover the climatic range of maritime pine (Fig. IV.1), which may

reduce our ability to detect an association between some climatic drivers and within-population genetic variation; this could explain, for example, the lack of association with the summer heat moisture index (SHM), an important climatic factor in Mediterranean environments.

Most importantly, the validation analysis provided independent evidence that additive within-population genetic variation for height was lower in populations experiencing extreme cold events for young trees but not for older trees (Fig. IV.3). This supports the robustness of our study and suggests that our results were unlikely to be biased by considering the total variance instead of the additive one, which was somehow expected as two previous studies in maritime pine found low non-additive effects for growth (Gaspar et al. 2013), and height and diameter (Lepoittevin et al. 2011).

With respect to specific leaf area (SLA), where only a single common garden (i.e. a single environment) was assessed, cryptic genetic variation (as defined above) may indeed underlie the higher genetic variation found in populations experiencing severe cold events. A study in maritime pine suggests that SLA depends strongly on environmental conditions (Alía et al. 2014), which is also supported in our study by its low genetic control (Fig. X.26). Cryptic genetic variation is more likely to be expressed when the differences between original and current environments are large (Paaby and Rockman 2014) , as it may be the case for some of the populations planted in the Portugal common garden. However, this is not a general pattern as we did not find any association between the climatic transfer distances (i.e. the absolute difference between the climate in the population and the climate in the test site) and the within-population genetic variation for SLA (see Section 6 of the Supplementary Information).

5.2 Environmental heterogeneity is not associated with higher genetic variation

Populations from heterogeneous environments did not show higher genetic variation for any trait (Fig. IV.2), which was also the case for the independent height data from the validation analysis (Fig. IV.3). This goes against a previous estimate in lodgepole pine suggesting that up to 20% of the genetic variation in growth within populations is explained by environmental heterogeneity (Yeaman and Jarvis 2006). A potential explanation of this discrepancy is the smaller experiment size in our study compared to that of Yeaman and Jarvis (103 populations with an average of 28 planting sites per population). However, in our study, we obtained reasonable credible intervals for most traits (allowing the detection of associations with other drivers) and data simulations suggested that our models have adequate power, rendering this explanation unlikely.

Another explanation is that genetic variation within populations is not affected by the environmental heterogeneity at the regional scale imposed by the 1×1 km resolution of our climate dataset but at finer spatial scales (also discussed in Yeaman and Jarvis 2006). Indeed, populations can adapt along microgeographic environmental gradients despite the homogenizing effect of gene flow (Richardson et al. 2014), even for forest tree populations with their long-generation times and large effective population sizes (Scotti et al. 2016). However, a correlation between regional and microgeographic environmental heterogeneity across the maritime pine range is very likely: populations showing the highest environmental heterogeneity in our study were located in mountainous areas in which we also expect higher microgeographic variation, e.g. the Cómpeta population (COM) located in the Tejeda and Almijara mountains (southern Spain),

the Arenas de San Pedro population (ARN) located in the Sierra de Gredos (central Spain) or the Pineta population (PIE) located close to the Punta di Forchelli (Corsican mountains), while populations with the lowest environmental heterogeneity were located on flat plateaus, e.g. populations from the Landes plateau and the Atlantic coastal regions in France (HOU, MIM, PET, VER, OLO, STJ, PLE), and populations from the central Spain plateau near to Segovia (CUE, COC, CAR). Thus, even if genetic variation was maintained by migration-selection balance at microgeographic scales, we would have been able to detect the effect of environmental heterogeneity at the regional scale. Nevertheless, more studies characterizing adaptation at microgeographic scales are needed to assess the spatial scale of genetic adaptation in maritime pine.

Another explanation of the discrepancy with Yeaman and Jarvis (2006) could be that we used young trees (between 20 and 85-month old) while they used 20-year old trees. Indeed, the processes generating within-population genetic variation might be age-dependent, as shown for climate harshness in Bordeaux, where the association was present when the trees were 25-month old but not in older trees. In forest trees, genetic parameters often vary with age; e.g. heritability generally increases with age until reaching a plateau, especially for height-related traits (Balocchi et al. 1993, Jansson et al. 2003, Johnson et al. 1997, Kroon et al. 2011, Sierra-Lucero et al. 2002), but may also decrease in some cases (Kroon et al. 2011, Lu and Charrette 2008). In maritime pine, an increase in heritability with age was found in Costa and Durel (2011) but not in Kusnandar et al. (1998). To our knowledge, the drivers of heritability changes with age remain unclear. Competition among trees in common gardens might play a role in the expression of age-dependent heritabilities for diameter growth, but not for height in Pinus radiata (Lin et al. 2013). Replicating our analysis in older trees would be interesting to further assess patterns of association between within-population genetic variation and environmental heterogeneity, and their underlying causes.

Finally, a last explanation is related to the different biological features between lodgepole pine and maritime pine. Lodgepole pine has extensive gene flow and low population structure (F_{ST} = 0.016 in Yeaman et al. 2016) while maritime pine shows restricted gene flow with strong population structure (at least six distinct gene pools and F_{ST} = 0.112; Jaramillo-Correa et al. 2015; our study) and fragmented distribution (Alberto et al. 2013). Pollen dispersal kernels in maritime pine are highly leptokurtic, as for other wind-pollinated pines (Robledo-Arnuncio and Gil 2005, Schuster and Mitton 2000), with estimated mean dispersal distances from 78.4 to 174.4m (de-Lucas et al. 2008). Interestingly, McDonald and Yeaman (2018) showed that high levels of quantitative genetic variance can be maintained when a trait is under stabilizing selection only at intermediate levels of migration. Migration rates in maritime pine may therefore not be strong enough to compensate for the purifying effect of natural selection in heterogeneous environments, especially in mountainous areas which may represent barriers to gene flow and where populations are more isolated (see González-Martínez et al. 2007 for maritime pine). Meanwhile, in the homogeneous plateaus of the Landes forest and central Spain, natural selection may be low because conditions are more favorable, and these populations are less isolated, which may maintain genetic variation at levels similar to those of populations in heterogeneous landscapes. Investigating local adaptation and gene flow at microgeographic scales in natural populations of maritime pine located in both homogeneous and heterogeneous environments would be highly valuable to understand why environmental heterogeneity does not seem to play a major role in maintaining genetic variation in this species. Moreover, conducting similar analyses in sister species such as Scots pine, with low population genetic structure and continuous populations (Alberto et al. 2013), could help to determine whether genetic variation in forest tree populations experiencing higher migration rates are more prone

to be impacted by environmental heterogeneity.

5.3 Link to fitness and genetic constraints may explain the different patterns across traits

Height was the only trait that showed a consistent association between within-population genetic variation and climate harshness. This pattern supports the hypothesis that natural selection mainly depletes genetic variation of traits most directly related to fitness. Indeed, height can be seen as the end-product of multiple ecophysiological processes (Grattapaglia et al. 2009). Taller trees perform better in the competition for light, water and nutrients, and are therefore more likely to have higher fecundity (Aitken and Bemmels 2015, Rehfeldt et al. 1999, Wu and Ying 2004) and lower mortality (Wyckoff and Clark 2002, Zhu et al. 2017). However, taller trees are also more susceptible to spring and fall cold injury (Howe et al. 2003) and to drought (Bennett et al. 2015, McDowell and Allen 2015, Stovall et al. 2019). In maritime pine, effective reproductive success (i.e. the number of successfully established offspring) is related to tree size. Indeed, González-Martínez et al. (2006) found a significant positive female selection gradient for diameter (height was not tested, but diameter and height are strongly correlated in conifers; see, for example, Fig. 1 in Castedo-Dorado et al. 2005 for maritime pine) and suggested that offspring mothered by bigger trees could have a selective advantage due to better quality seeds favouring resilience in the face of severe summer droughts and microsite variation. This evidence also supports the idea of height as a relevant fitness component in maritime pine.

Although less directly related to fitness than height, leaf phenology-related traits exhibit steep adaptation gradients in forest trees and have a relatively high heritability, e.g. 0.15-0.51 for bud burst in pedunculate oak (Scotti-Saintagne et al. 2004), 0.45-1 in Sitka spruce (Alfaro et al. 2000) and 0.54 for bud burst and 0.30 for the duration of bud burst in our study in maritime pine. Gauzere et al. (2020) showed that both the mean and the variance of leaf phenology-related traits varied along an altitudinal gradient in natural oak populations, with populations at high altitude having a narrower fitness peak. We might therefore have expected lower genetic variation for leaf phenology-related traits in populations experiencing severe cold events (and at higher altitude), as found along an altitudinal gradient in sessile oak for bud phenology (Alberto et al. 2011). However, such association may be hidden in common gardens with different climates from those of the populations' location, because of the release of high levels of cryptic genetic variation (Schlichting 2008). Moreover, phenology-related traits can show opposite genetic clines in common gardens and natural populations (e.g. Vitasse et al. 2009). Estimating genetic parameters of phenology-related traits directly in the field, which is now technically possible by using large genomic datasets and advanced statistical methodologies (Gienapp et al. 2017), may therefore be necessary to investigate potential associations between within-population genetic variation and climate harshness, or other selective pressures.

Importantly, theoretical work suggests that much of the genetic variation associated with a trait is likely maintained by pleiotropic effects, which are independent of the selection on that trait, implying that stabilizing selection can only act on a reduced number of independent dimensions in the trait space (Barton 1990, Walsh and Lynch 2018). As we used univariate models, we cannot exclude that the likely associations with height genetic variation originate from genetic correlation with other traits under selection, or that the lack of association with other traits (notably functional traits such as $\delta^{13}C$) does not originate from genetic constraints (Walsh and Blows 2009). For example, in maritime pine, trait canalisation and genetic constraints may explain low quantitative genetic differentiation for hydraulic traits

(e.g. P50, the xylem pressure inducing 50% loss of hydraulic conductance; Lamy et al. 2014), and sapling height was found to be either positively or negatively associated with disease susceptibility depending on the pathogen (e.g. necrosis length caused by *Diplodia sapinea* or *Armillaria ostoyae*, respectively; Hurel et al. 2019). Trade-offs between traits may also explain the unexpected association between minimum temperatures and high genetic variation for SLA, as, for instance, SLA is known to be positively correlated with leaf life span, low assimilation rates and nutrient retention, i.e. traits linked to conservation of acquired resources (Ackerly et al. 2002).

6 Conclusion

Our manuscript contributes to the current debate on the maintenance of quantitative genetic variation within populations by providing empirical support for the role of natural selection in decreasing genetic variation. Indeed, our results consistently showed that genetic variation for height is lower in maritime pine populations experiencing severe cold events (i.e. experiencing stronger selection). Surprisingly, we found no association between environmental heterogeneity at the regional scale and within-population genetic variation for several traits; whether for technical reasons (e.g. sample size, spatial scale considered) or for genuine biological reasons (e.g. too low migration), it would be worth further exploration. Indeed, understanding the evolutionary forces shaping within-population genetic variation could shed light on how populations adapt to their local environment, thereby providing insight into how they may respond to future changes in environmental conditions.

7 Acknowledgments

We thank A. Saldaña, F. del Caño, E. Ballesteros and D. Barba (INIA) and the 'Unité Expérimentale Forêt Pierroton' (UEFP, INRAE; doi:10.15454/1.5483264699193726E12) for field assistance (plantation and measurements). Data used in this research are part of the Spanish Network of Genetic Trials (GENFORED, http://www.genfored.es). We thank all persons and institutions linked to the establishment and maintenance of field trials used in this study. We are very grateful to Ricardo Alía and Juan Majada who initiated and supervised the establishment of the CLONAPIN network and provided the progeny test height data that we used in the validation analysis. We thank Maurizio Marchi for kindly providing the raster files of the climate variables corresponding to the desired time period and spatial extent. JA was funded by the University of Bordeaux (ministerial grant). This study was funded by the 'Initiative d'Excellence (IdEx) de l'Université de Bordeaux - Chaires d'installation 2015' (EcoGenPin) and the European Union's Horizon 2020 research and innovation programme under grant agreements No 773383 (B4EST) and No 862221 (FORGENIUS).

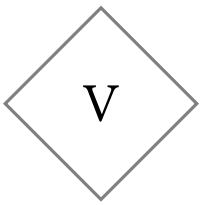
8 Author contributions

SCG-M designed the experiment and supervised the curation of field data. MdM cleaned and formatted the phenotypic data. SCG-M, JA, FB, MBG and BB conceived the paper methodology. JA and FB built the model equations and codes. JA conducted the data and simulation analyses.

All authors interpreted the results. JA led the writing of the manuscript. All authors contributed to the manuscript and gave final approval for publication.

9 Data and script availability

Data are publicly available. SNP data were deposited in the Dryad repository at http://dx.doi.org/10.5061/dryad.8d6k1. Height data have been deposited in GENFORED, the Spanish Network of Genetic Trials (http://www.genfored.es). Scripts are available at https://github.com/JulietteArchambeau/H2Pinpin.



Chapter 2. Combining climatic and genomic data improves range-wide tree height growth prediction in a forest tree

Juliette Archambeau[1], Marta Benito Garzón[1], Frédéric Barraquand[2], Marina de Miguel Vega[1,3], Christophe Plomion[1], Santiago C. González-Martínez[1]

1 INRAE, Univ. Bordeaux, BIOGECO, F-33610 Cestas, France
2 CNRS, Institute of Mathematics of Bordeaux, F-33400 Talence, France
3 EGFV, Univ. Bordeaux, Bordeaux Sciences Agro, INRAE, ISVV, F-33882, Villenave d'Ornon, France

Keywords: Climate change, local adaptation, phenotypic plasticity, population response functions, positive-effect alleles, range-wide predictive models, maritime pine.

1 Abstract

Population response functions based on climatic and phenotypic data from common gardens have long been the gold standard for predicting quantitative trait variation in new environments. However, prediction accuracy might be enhanced by incorporating genomic information that captures the neutral and adaptive processes behind intra-population genetic variation. We used five clonal common gardens containing 34 provenances (523 genotypes) of maritime pine (*Pinus pinaster* Aiton) to determine whether models combining climatic and genomic data capture the underlying drivers of height-growth variation, and thus improve predictions at large geographical scales. The plastic component explained most of the height-growth variation, probably resulting from population responses to multiple environmental factors. The genetic component stemmed mainly from climate adaptation, and the distinct demographic and selective histories of the different maritime pine gene pools. Models combining climate-of-origin and gene pool of the provenances, and positive-effect height-associated alleles (PEAs) captured most of the genetic component of height-growth and better predicted new provenances compared to the climate-based population response functions. Regionally-selected PEAs were better predictors than globally-selected PEAs, showing high predictive ability in some environments, even when included alone in the models. These results are therefore promising for the future use of genome-based prediction of quantitative traits.

2 Introduction

Global change is expected to have a profound impact on forests (Franklin et al. 2016, Seidl et al. 2017), and whether tree populations will be able to migrate or persist across their current range is uncertain (Aitken et al. 2008). Assessing the potential of populations to accommodate future environmental conditions requires a thorough understanding of the origin of variation in quantitative traits subject to natural selection (Alberto et al. 2013, Shaw and Etterson 2012). To this aim, a necessary first step is to quantify the plastic and genetic components of adaptive traits and their interaction in multiple environments (Des Marais et al. 2013, Merilä and Hendry 2014), which has been done extensively in forest trees (Franks et al. 2014). A second step consists of identifying the underlying drivers of these components (Merilä and Hendry 2014). The plastic component corresponds to the ability of one genotype to produce varying phenotypes depending on the environment (Bradshaw 1965). Phenotypic plasticity can help individuals to overcome new conditions up to a certain threshold (Nicotra et al. 2010), and can be to some extent genetically assimilated and therefore involved in the evolutionary process of adaptation (Pigliucci et al. 2006). The genetic component can stem from both neutral (e.g. population demographic history and genetic drift) and adaptive processes (e.g. adaptation to local biotic and abiotic environments), both processes implying changes in allele frequencies. Populations are locally adapted when they have higher fitness in their own environment than populations from other environments (Kawecki and Ebert 2004). In forest trees, a large amount of work highlighted the importance of climate in driving the plastic and genetic responses of quantitative traits to new environmental conditions (Savolainen et al. 2007, Valladares et al. 2014b). However, it is still unclear how multiple and interacting drivers underlying quantitative trait variation could be combined to improve predictions of population responses to global change. The increasing availability of genomic data opens new opportunities to boost prediction accuracy, which is critical for breeding (i.e. genomic selection; Grattapaglia and Resende 2011), to anticipate future distribution of natural populations (e.g. Razgour et al. 2019), or to support

the ongoing development of assisted gene flow strategies aiming to help populations adapt to future environments (Browne et al. 2019, MacLachlan et al. 2021, Mahony et al. 2020).

In forest trees, a long history of common gardens (Langlet 1971) has provided a unique framework to associate population-specific quantitative trait variation with large environmental or geographical gradients, and thus identify populations at risk under climate change (Fréjaville et al. 2020, Pedlar and McKenney 2017, Rehfeldt et al. 2018, 2003, 1999, Savolainen et al. 2007). The development of population response functions was a step forward to evaluate the relative contribution of plasticity -associated to current climatic conditions (i.e. the climate in the common gardens)- and genetic adaptation -associated to the past climatic conditions under which the populations have evolved (i.e. the climate-of-origin of the provenances tested)- in explaining quantitative trait variation (O'Neill et al. 2008, Wang et al. 2010). These models have now been applied to a large variety of traits (Benito Garzón et al. 2019, Leites et al. 2012a,b, Vizcaíno-Palomar et al. 2020) and one of their main conclusions is that trait variation across species ranges is mostly associated with the climate in the common garden (i.e. related to the plastic component) and, only to a much lesser extent, with the climate-of-origin of the provenances (i.e. related to the genetic component) (Benito Garzón et al. 2019, Leites et al. 2012b). Importantly, these models do not allow us to determine to what extent associations between trait variation and provenance climate-of-origin, or the higher trait values of local compared to foreign populations, are caused by adaptive or neutral processes (Franks et al. 2014, Hereford 2009, Leimu and Fischer 2008). This limits our understanding of the genetic processes that led to the current patterns of quantitative trait variation, and therefore our ability to predict trait variation of new (untested in common gardens) populations under new environments.

The advent and generalization of genomic tools have enhanced our understanding of adaptive and neutral genetic processes resulting in trait variation, and their relationship with climatic gradients (Leroy et al. 2020, Savolainen et al. 2013, Sork 2018). Integrating genomic information into quantitative trait prediction would be highly valuable to consider intraspecific variability at a finer scale than in current models (Mahony et al. 2020), thereby probably improving model accuracy, especially for populations not previously planted in commons gardens. More specifically, rapidly growing knowledge on trait-associated alleles identified by Genome-Wide Association Studies (GWAS) is promising for anticipating the genetic response of populations to new environments (Browne et al. 2019, Exposito-Alonso et al. 2018a). For example, Mahony et al. (2020) used counts of alleles positively associated with the traits of interest (PEAs) to describe patterns and identify drivers of local adaptation in lodgepole pine. Recent studies have shown that most quantitative traits are highly polygenic (see reviews in Barghi et al. 2020, Pritchard et al. 2010; and de Miguel et al. 2020 for maritime pine) and that the effect of trait-associated alleles may vary across environments (Anderson et al. 2013, Tiffin and Ross-Ibarra 2014), which complicates the use of genomic information in trait prediction. In addition, patterns in allele frequencies induced by population demographic history are often correlated with environmental gradients (Alberto et al. 2013, Latta 2009, Nadeau et al. 2016), which makes difficult to separate the signature of population structure from that of adaptive processes (Sella and Barton 2019, Sohail et al. 2019). At the species range scale, population structure hinders the use of genomic relationship matrices, which provide more accurate estimates of genetic parameters (e.g. breeding values, additive and non-additive variance) within breeding populations than previously used pedigree-based approaches (Bouvet et al. 2016, El-Dien et al. 2018). Indeed, admixed populations or distinct genetic groups may present different means and variances of their genetic values, which requires new statistical methods to estimate them (e.g. Muff et al. 2019). Thus, integrating genomic information into quantitative trait prediction in

natural populations, while highly valuable, remains challenging.

Forest trees are remarkable models to study the genetic and plastic components of quantitative trait variation. Forest tree populations often have large effective population size and are distributed along a large range of environmental conditions, which makes them especially suitable to study current and future responses to climate (Alberto et al. 2013, Savolainen et al. 2007). Moreover, forest trees remain largely undomesticated (including those species with breeding programs) and, therefore, genetic variation in natural populations has been little influenced by human-induced selection (Neale and Savolainen 2004). However, forest trees have also large and complex genomes (especially conifers; Mackay et al. 2012), that show a rapid decay of linkage disequilibrium (Olson et al. 2010), and extensive genotyping would be needed to identify all (most) relevant polymorphisms underlying (highly polygenic) quantitative traits (Jaramillo-Correa et al. 2015, Neale and Savolainen 2004). In addition, although early results have been convincing in predicting trait variation within tree breeding populations (i.e. using populations with relatively low effective population size; Jarquín et al. 2014, Resende Jr et al. 2012, Resende et al. 2012), predicting the genetic component of trait variation across populations or geographical regions of forest trees remains poorly explored.

In the present study, we aim to identify the potential drivers of the plastic and genetic components of height growth in distinct maritime pine gene pools (i.e. genetic clusters) and investigate how common garden data can be combined with genomics to efficiently predict height-growth variation across the species range. We compared Bayesian hierarchical mixed models that inferred height-growth variation in maritime pine as a function of climatic and genomic-related variables, using a clonal common garden network (CLONAPIN) consisting of five sites and 34 provenances (523 genotypes and 12,841 trees). First, we evaluated the relative importance of potential drivers underlying height-growth variation. We expected that: (i) the plastic component explains most trait variation and is associated with climate in the common gardens, (ii) the genetic component is driven by both adaptive processes, such as adaptation to climate, and neutral processes, such as population demographic history. Second, we compared the out-of-sample predictive ability (on unknown observations or provenances) of models based exclusively on the common garden design and models including (either separately or jointly) potential predictors of the genetic component of trait variation, notably those related to climate and positive-effect height-associated alleles (PEAs). We expected that the distinct demographic history of maritime pine gene pools, the provenance climate-of-origin and the counts of PEAs, either combined or alone, may improve height-growth predictions of unknown provenances. We also expected that height-associated alleles selected regionally, i.e. in particular environments, would have a better predictive ability than globally-selected alleles. Our study is a step towards integrating the recent knowledge brought by large genomic datasets to the modelling of quantitative trait variation in forest trees. Combining common gardens with genomic tools hold great promise for speeding up and improving trait predictions at large scales and for a wide range of species and populations. However, a robust framework is needed to make reliable predictions and to determine when and to what extent genomics can help in making decisions in conservation strategies or in anticipating population responses to climate change.

3 Materials & Methods

3.1 Plant material and phenotypic measurements

Maritime pine (*Pinus pinaster* Ait., Pinaceae) is an economically important forest tree, largely exploited for its wood (Viñas et al. 2016). It has also an important ecological function stabilizing coastal and fossil dunes and as keystone species supporting forest biodiversity. Native to the western part of the Mediterranean Basin, the Atlas mountains in Morocco, and the south-west Atlantic coast of Europe, its natural distribution spans from the High Atlas mountains in the south (Morocco) to French Brittany in the north, and from the coast of Portugal in the west to western Italy in the east. Maritime pine is a wind-pollinated, outcrossing and long-lived tree species that can grow on a wide range of substrates, from sandy and acidic soils to more calcareous ones. It can also withstand many different climates: from the dry climate of the Mediterranean Basin to the highly humid climate of the Atlantic Europe region, and the continental climate of central Spain. Maritime pine populations are highly fragmented and can be grouped into six gene pools (Alberto et al. 2013, Jaramillo-Correa et al. 2015; see Fig. V.1), that is genetic clusters that cannot be differentiated on the basis of neutral genetic markers and that probably derive from a common glacial refuge (Bucci et al. 2007, Santos-del-Blanco et al. 2012).

Height growth is a key adaptive trait in forest trees, including maritime pine. Height can be seen as the end-product of multiple ecophysiological processes that are both genetically regulated and affected by multiple environmental effects (Grattapaglia et al. 2009). As such, taller trees compete more efficiently for light, water and nutrients, and are also more likely to have high fecundity (Aitken and Bemmels 2015, Rehfeldt et al. 1999, Wu and Ying 2004). We obtained height data from the clonal common garden network CLONAPIN, consisting of five common gardens located in different environments (also referred as test sites; Fig. V.1). Three sites are located in the Atlantic Europe region, with mild winters, high annual rainfall and relatively wet summers: Bordeaux in the French part, and Asturias and Portugal in the Iberian part, the Portugal site experiencing slightly colder winters and half the summer precipitation than the site in Asturias. The two other sites, Cáceres and Madrid, are located in the Mediterranean region with high temperatures and intense summer drought, as well as large precipitation differences between summer and winter. In 2010 or 2011 depending on the test site, clonal replicates from 34 provenances were planted in a randomized complete block design with eight blocks. For each provenance, trees represent between 2 and 28 genotypes (clones), on average about 15 (see Rodríguez-Quilón et al. 2016 for details). Genotypes were originally sampled from natural populations, with enough distance among trees (over 50 m) to avoid sampling related individuals. Depending on the site, height was measured from one to four times, when the trees were between 13 and 41 month old (Table XI.1). Only survivors were measured for height, which resulted in a strongly unbalanced design as 92% and 75% of the trees died in Cáceres and Madrid, respectively (partly due to the clay soils and a strong summer drought). After removing genotypes for which we had no genomic information, we analyzed 33,121 height observations from 12,841 trees and 523 genotypes (Table XI.2).

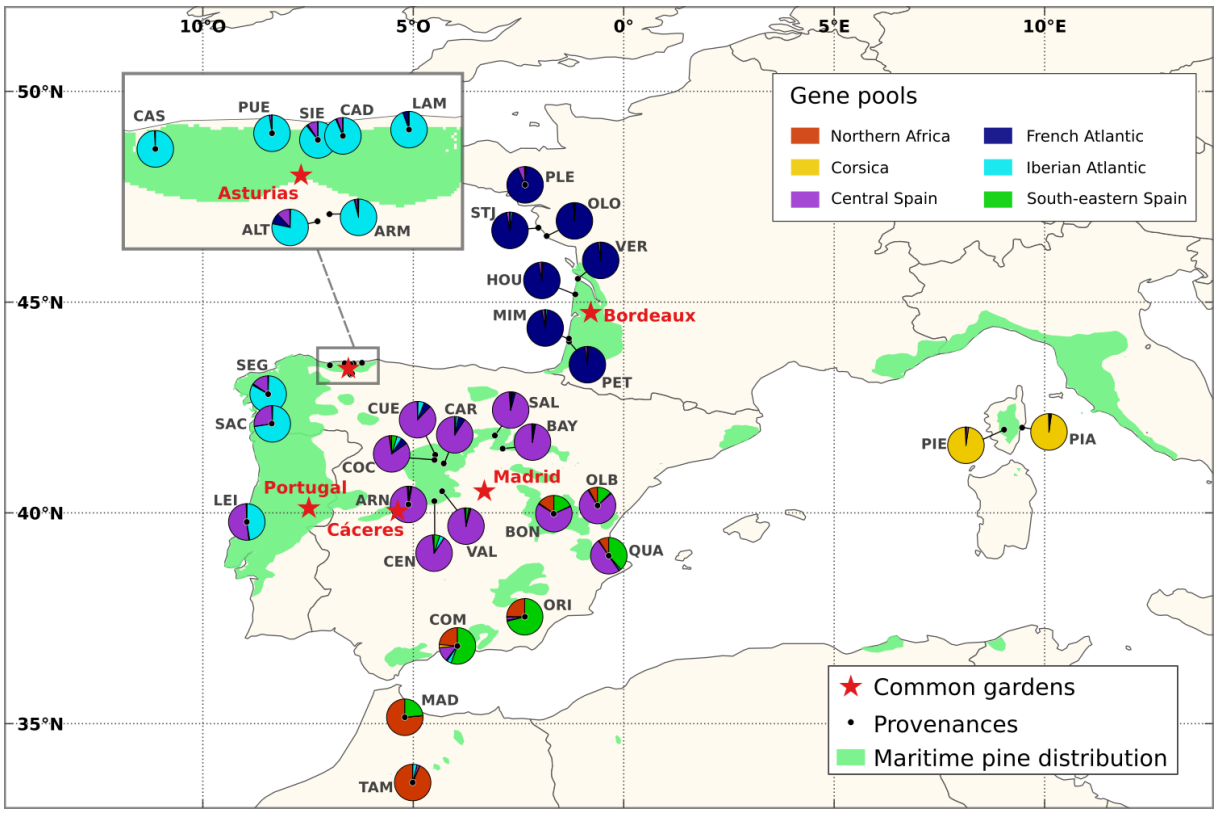


Figure V.1. The five common gardens and 34 provenances of maritime pine (CLONAPIN common garden network) used in this study. The distribution of maritime pine is also shown (based on EUFORGEN map, http://www.euforgen.org/). Pie charts represent the proportions belonging to each gene pool for each provenance (see legend) as estimated in Jaramillo-Correa et al. (2015). Provenance names can be found in Table XI.2.

.

3.2 Gene pool assignment and positive-effect alleles (PEAs)

DNA was extracted from leaves collected in the Asturias common garden and genotyped with a 9k Illumina Infinium SNP assay (described in Plomion et al. (2016a)), resulting in 5,165 high-quality polymorphic SNPs scored on 523 genotypes. There were on average only 3.3 missing values per genotype (ranging between 0 and 142). For each genotype, the proportion belonging to each gene pool was estimated in Jaramillo-Correa et al. (2015), using nine nuSSRs as well as a subset of the same SNPs as in our study (1,745 SNPs) and the Bayesian approach available in STRUCTURE v2.3.3 (Pritchard et al. 2000; Table XI.3). This gene pool assignment aimed at reflecting the neutral genetic structure in maritime pine, which results from population demographic history and genetic drift, but may also arise from different selective histories across gene pools.

Based on the 523 genotypes for which there were both genotypic and phenotypic data, we performed four GWAS following the Bayesian variable selection regression (BVSR) methodology implemented in the piMASS software (Guan and Stephens 2011), correcting for population structure and using the height BLUPs reported in de Miguel et al. (2020), that accounted for site and block effects. First, a global GWAS was performed to identify SNPs that have an association with height at range-wide geographical scales, thus using the combined phenotypic data from the five common gardens. Second, three regional GWAS were performed to identify SNPs that have a local association with height in a particular geographical region *r* (i.e. in a particular environment), thus using separately data from the Iberian Atlantic common gardens (Asturias and Portugal), the French Atlantic common garden (Bordeaux) and the Mediterranean common

gardens (Madrid and Cáceres). For each of the four GWAS, we selected the 350 SNPs (~7% top associations) with the highest absolute Rao-Blackwellized estimates of the posterior effect size, corresponding approximately to the estimated number of SNPs with non-zero effects on height in a previous multi-trait study using the same SNP marker set (de Miguel et al. 2020). These SNPs were used to compute the counts of global and regional positive-effect alleles (gPEAs and rPEAs) for each genotype (see Section 2.1 of the Supplementary Information for more details).

3.3 Climatic data

In forest trees, large-scale patterns of allele frequencies or quantitative trait variation are known to be associated with climatic variables related to mean temperature and precipitation (e.g. Eckert et al. 2010, Fréjaville et al. 2020, Leites et al. 2019, Mahony et al. 2020, McLane et al. 2011), or episodic climatic conditions, such as summer aridity or maximum temperatures (Fréjaville et al. 2020, Grivet et al. 2011, Jaramillo-Correa et al. 2015, McLane et al. 2011, Rehfeldt et al. 2003). As climate change will cause major changes in temperature and precipitation in the near future, particularly in the Mediterranean basin, there is a need to understand the complex influence of climatic variables on quantitative trait variation. We extracted monthly and yearly climatic data from the EuMedClim database with 1 km resolution (Fréjaville and Benito Garzón 2018). The climatic similarity among test sites was described by a covariance matrix Ω including six variables related to both extreme and average temperature and precipitation in the test sites during the year preceding the measurements, and with at most a correlation coefficient of 0.85 among each other (see Section 3.1 in the Supplementary Information for more details). The climatic similarity among provenances was described by a covariance matrix Φ including four variables related to the mean temperature and precipitation in the provenance locations over the period from 1901 to 2009 (i.e. representing the climate under which provenances have evolved), and with at most a correlation coefficient of 0.77 among each other (see Section 3.2 in the Supplementary Information for more details).

3.4 Hierarchical height-growth models

Twelve height-growth models were compared. We first built two baseline models relying exclusively on the common garden design and aimed at quantifying the relative contribution of the genetic and plastic components of height-growth variation (*models M1* and *M2*; Table V.1). Second, we used climatic and genomic data to detect association of height-growth variation with potential underlying drivers related to plasticity, adaptation to climate or gene pool assignment (i.e. a proxy of the population demographic history and genetic drift experienced by the populations), and estimated gene pool-specific total genetic variances (*models M3* to *M6*; Table V.1). Third, we built models either including separately or combining potential drivers of the genetic component of height-growth variation to predict unknown observations and provenances without relying on the common garden design (*models M7* to *M12*; Table V.1). In every model, the logarithm of height ($\log(h)$) was used as a response variable to stabilize the variance. Tree age at the time of measurement i was included as a covariate to account for the average height-growth trajectory. This implies that all models shared the form $\log(h_i) = f(\text{age}_i) + m(\text{covariates})$, where $m(\text{covariates})$ is the rest of the model. Therefore, all models can also be written $h_i = \exp(f(\text{age}_i))\exp(m(\text{covariates}))$, which explains why covariates in our models affect height growth (i.e. modulate the height-growth trajectory) rather than simply height. We used a second-degree polynomial to account for tree age ($f(\text{age}_i + \text{age}_i^2)$) because the logarithm of height first increases linearly with age and then

reaches a threshold (Fig. XI.11). Each tree was measured between one and four times (14% of the trees were measured only once), but we did not include a varying intercept for each tree as it resulted in model miss-specification warnings and strong overfitting. A description of each model specification follows.

Variables	Baseline		Explanatory models				Predictive models					
	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11	M12
Site/Block	×	×	×	×	×	×	×	×	×	×	×	×
Provenance	×	×	×	×	×	×						
Genotype	×	×	×	×		×						
Site × Provenance		×										
Climatic similarity among sites			×	×	×	×						
Proportion belonging to each gene pool				×	×	×	×	×	×			
Gene pool-specific genetic variance					×							
Climatic similarity among provenances						×						
Provenance climate-of-origin							×	×		×		
Global PEAs (gPEAs)							×				×	
Regional PEAs (rPEAs)								×				×

Table V.1. Variables included in the height-growth models. Baseline *models M1* and *M2* separate the genetic and plastic components of height-growth variation via varying intercepts relying exclusively on the common garden design. Explanatory models (*models M3* to *M6*) test different hypotheses regarding the potential drivers underlying height-growth variation. Predictive models (*models M7* to *M12*) are used to compare the predictions on new observations and provenances when combining or including separately genomic and climatic drivers of height-growth variation. The provenance climate-of-origin is evaluated using the precipitation of the driest month, *min.pre*, and the maximum temperature of the warmest month, *max.temp*. gPEAs and rPEAs correspond to the counts of height-associated positive-effect alleles, selected either globally (across all common gardens) or regionally (in specific common gardens). The provenance climate-of-origin and the PEAs were included in the predictive models with site-specific slopes. All models also contained the age effect, not shown in the table.

Baseline *models M1* and *M2*: separating the genetic and plastic components of height-growth variation

In the baseline *model M1*, height h was modeled as a function of tree age, varying intercepts for the sites S_s and blocks nested within sites $B_{b(s)}$ (i.e. the plastic component), and varying intercepts for the provenances P_p and genotypes within provenances $G_{g(p)}$ (i.e. the genetic component):

$$\begin{aligned} \log(h_{isbpg}) &\sim \mathcal{N}(\mathbf{X}\beta + \mu_{sbpg}\,,\, \sigma^2) \\ \mathbf{X}\beta &= \beta_0 + \beta_{age}\text{age}_i + \beta_{age2}\text{age}_i^2 \\ \mu_{sbpg} &= S_s + B_{b(s)} + P_p + G_{g(p)} \end{aligned} \tag{3.1}$$

where $\mathbf{X}$ is the 3-column design matrix and β is a vector including the intercept β_0 and the coefficients β_{age} and β_{age2} of the fixed effect variables (*age* and age^2, respectively). μ_{sbpg} is the vector of varying intercepts. *Model M2* was based on *model M1* but including an interaction term between provenance and site (S_sP_p). We also performed a model without the genetic component (called *M0*) whose outputs are reported in the Supplementary Information.

Explanatory *models M3* to *M6*: potential drivers underlying height-growth variation

In *model M3*, we hypothesized that the plastic component of height growth was influenced by the climatic similarity among test sites during the year preceding the measurements. This model can be expressed with the same likelihood as *M1* but with the vector of varying intercepts equal to:

$$\begin{aligned} \mu_{isbpg} &= S_s + B_{b(s)} + P_p + G_{g(p)} + cs_{is} \\ cs_{is} &\sim \mathcal{N}(0, \Omega\,\sigma^2_{cs_{is}}) \end{aligned} \tag{3.2}$$

where Ω is the covariance matrix describing the climatic similarity between test sites s during the year i preceding the measurements (Fig. XI.6) and cs_{is} are varying intercepts associated with the climatic conditions in each test site s during the year i. In *M3*, the plastic component was partitioned between the regression on the climatic covariates (cs_{is}) and the deviations related to block and site effects due to the local environmental conditions that are not accounted for by the selected climatic covariates.

In *models M4*, *M5* and *M6*, we investigated the drivers of the genetic component of height growth. In *M4*, we hypothesized that the genetic component was influenced by the proportion belonging to each gene pool j. *M5* extends *M4* by estimating different total genetic variances in each gene pool while accounting for admixture among gene pools, following Muff et al. (2019). Equations for *M4* and *M5* can be found in Section 4 of the Supplementary Information. In *M6*, we hypothesized that populations are genetically adapted to the climatic conditions in which they evolved. Thus, we quantified the association between height growth and the climatic similarity among provenances, while still accounting for the gene pool assignment, such as:

$$\begin{aligned} \mu_{ijsbpg} &= S_s + B_{b(s)} + P_p + G_{g(p)} + cs_{is} + cp_p + \sum_{j=1}^{6} q_{gj} g_j \\ cp_p &\sim \mathcal{N}(0, \Phi\, \sigma^2_{cp_p}) \end{aligned} \tag{3.3}$$

where q_{gj} corresponds to the proportion belonging of each genotype g to the gene pool j, g_j is the mean relative contribution of gene pool j to height growth, Φ is the covariance matrix describing the climatic similarity between provenances p (Fig. XI.9) and cp_p are varying intercepts associated with the climate in each provenance p. Therefore, in *M6*, the genetic component was partitioned among the regression on the climatic covariates (cp_p), the gene pool covariates (g_j), and the deviations related to the genotype ($G_{g(p)}$) and provenance (P_p) effects (resulting, for example, from adaptation to environmental variables not measured in our study).

Predictive *models M7* to *M12*: combining climatic and genomic information to improve predictions

In this last set of models, we replaced the provenance and genotype intercepts with different potential drivers of height-growth variation that do not rely directly on the common garden design, namely the gene pool assignment (as in *M4*), two variables describing the climate in the provenance locations (*min.pre* the precipitation of the driest month and *max.temp* the maximum temperature of the warmest month) and either global or regional PEAs. This allowed us to determine whether these potential drivers were able to predict the height-growth genetic component as accurately as the provenance and genotype intercepts (i.e. the variables relying directly on the common garden design). In *models M7* and *M8*, the potential predictors were all included together in the models to quantify their predictive performance conditionally to the other predictors, and were expressed as follows (here for *M7*):

$$\begin{aligned} \mu_{jsbpg} = S_s + B_{b(s)} + \sum_{j=1}^{6} q_{gj} g_j + \beta_{min.pre,s} min.pre_p \\ + \beta_{max.temp,s} max.temp_p + \beta_{gPEA,s} gPEA_g \end{aligned} \tag{3.4}$$

where $min.pre_p$ and $max.temp_p$ are the climatic variables in the provenance locations, $\beta_{min.pre,s}$ and $\beta_{max.temp,s}$ their site-specific slopes, $gPEA_g$ the counts of global PEAs and $\beta_{gPEA,s}$ its site-specific slopes. *M8* is identical to *M7*, except that the counts of gPEAs were replaced by

counts of rPEAs (i.e. regionally-selected alleles, with positive effects in specific geographical regions/environments). We also performed models in which the potential predictors were included individually to determine their specific predictive performance: the gene pool assignment in *M9*, the provenance climate-of-origin in *M10* and the counts of gPEAs and rPEAs, in *M11* and *M12*, respectively.

All models were inferred in a Bayesian framework as this approach better handles unbalanced and multilevel designs (Clark 2005) and also to better propagate sources of uncertainty from data and parameter values into the estimates (de Villemereuil 2019). Priors used in the models were weakly informative and are provided in Section 4.2 of the Supplementary Information. To build the models, we used the R package *brms* (Bürkner 2017), based on the no-U-turn sampler algorithm. Models were run with four chains and between 2,000 and 3,000 iterations per chain depending on the models (including 1,000 warm-up samples not used for the inference). All analyses were undertaken in R version 3.6.3 (R Core Team 2020) and scripts are available at https://github.com/JulietteArchambeau/HeightPinpinClonapin.

3.5 Comparing model goodness-of-fit and predictive ability

Three partitions of the data (P1, P2 and P3) were used to evaluate model goodness-of-fit (i.e. in-sample explanatory power, using training datasets) and predictive ability (out-of-sample predictive power, using test datasets). In P1, we aimed to predict new observations, an observation being a height-growth measurement in a given year on one individual. P1 corresponds to a random split of the data between 75% of observations used to fit the models (the training dataset of 24,840 observations) and 25% of observations used to evaluate model predictions (the test dataset of 8,281 observations). Notice that the test dataset of the P1 partition was not totally independent from the training dataset as it belongs to the same genotypes/provenances and blocks/sites. In P2 and P3, we aimed to predict new provenances. P2 corresponds to a random split between a training dataset of 28 provenances and a test dataset containing the remaining 6 provenances. P3 corresponds to a non-random split between a training dataset of 28 provenances and a test dataset containing 6 provenances with at least one provenance from each under-represented gene pool (i.e. northern Africa, south-eastern Spain and Corsican gene pools; see Section 6.3 of the Supplementary Information for details). Therefore, the test datasets of the P2 and P3 partitions represent fully independent sets of provenances.

To evaluate the model goodness-of-fit, we calculated the in-sample (in the training dataset) proportion of the variance explained by each model m in each common garden s, conditional on the age effect, such as: $\mathcal{R}^2_{ms}|age = (V_{\text{pred}_{ms}} - V_{\text{age}_{2s}})/(V_{\text{y}_s} - V_{\text{age}_{2s}})$, where $V_{\text{pred}_{ms}}$ is the variance of the modeled predictive means from model m in site s, V_{y_s} the phenotypic variance in the site s and $V_{\text{age}_{2s}}$ the variance explained by the age effect in the model *M2* in site s. We used V_{age_2} of model *M2* and not of model m because the variance predicted by the different fixed effects of some of the models (*M7* to *M12*) could not be properly separated. Moreover, as *M2* is the model with the highest predictive ability among the models relying only on the common garden design (Table XI.4), it constitutes an adequate baseline for model comparison. In addition, for baseline models *M1* and *M2*, we also calculated the in-sample proportion of the variance explained by the different model components (i.e. genetic, environment and genetic × environment) conditional on the age effect, e.g. for the genetic component in *M1*: $\mathcal{R}^2_{1,gen}|age = (V_{\text{pred}_{1,gen}} - V_{\text{age}_1})/(V_{\text{y}} - V_{\text{age}_1})$ where $V_{\text{pred}_{1,gen}}$ is the variance explained by the

genetic component (including the provenance and clone effects) in *M1*, V_y the phenotypic variance and V_{age_1} the variance explained by the age effect in *M1*.

Finally, to evaluate the model predictive ability, we calculated the out-of-sample (in the test dataset) proportion of the variance predicted by each model *m* in each common garden *s* conditional on the age effect, that we called *prediction* $\mathcal{R}^2_{ms}|age$. Details about calculating *prediction* $\mathcal{R}^2_{ms}|age$ and some supplementary indexes used for model comparison are presented in Section 5 of the Supplementary Information.

4 Results

4.1 Underlying drivers of height-growth variation

In this part, we disentangled the different components of height-growth variation and provided insights on their underlying drivers. Baseline and explanatory models (i.e. *models M1* to *M6*) explained ~81.5% of height-growth variation, including 57% due to the age effect (Table XI.4). Based on *M1*, ~47% (45-48% CIs) of the variation that was not explained by the age effect (i.e. deviating from the growth trajectory) came from the plastic component, ~11% (11-12% CIs) from the genetic component and ~43% (42-44% CIs) remained unexplained (Fig. V.2A & Table XI.5). In *M2* (same model as *M1* but adding the provenance-by-site interaction), the proportion of variance explained by the provenance-by-site interaction was not different from zero (Table XI.5). Therefore, we mostly interpret parameter estimates of *M1* (Fig. V.3), whose results are very similar to *M2*, but with smaller credible intervals (Tables XI.15 & XI.18). The plastic component was largely driven by the variance among sites (σ^2_S), with very little contribution of the variance among blocks (σ^2_B; XI.15). Trees grew the least in Madrid and the most in Asturias (Fig. V.3 & Table XI.16). The genetic component was equally attributed to the variance among provenances (σ^2_P) and genotypes (σ^2_G; Table XI.15), with the average height of the provenances appearing to be influenced by their belonging to particular gene pools (Fig. V.3; and more details in Section 6.1.1 of the Supplementary Information).

Based on *M3*, the plastic component of height-growth came only marginally from the variance associated with climate similarity among sites, which was more than five times lower than the variance associated with site intercepts (Fig. V.2B & Table XI.19). However, *M3* may be unable to separate the effect of these two components (see Section 6.1.2 in the Supplementary Information). Indeed, when estimating the effect of the climate similarity among sites in a model that did not include varying intercepts for the sites, we found that height growth was positively associated with the climatic conditions in Bordeaux and Asturias, and negatively with those in Madrid and Cáceres, the two Mediterranean sites, and to a lesser extent also in Portugal (Table XI.24).

Based on *M6*, the genetic component of height growth was mostly determined by the climatic similarity among provenances and to a lesser extent by the gene pool assignment (Fig. V.2C & Table XI.29). However, the effects of the gene pools and climatic similarity among provenances were partially confounded, so that the association between height growth and the gene pools was stronger when the climatic similarity among provenances was not included in the models (i.e. *model M4*; Table XI.25). Populations from climatic regions neighboring the Atlantic Ocean, and mainly belonging to the French and Iberian Atlantic gene pools, were generally

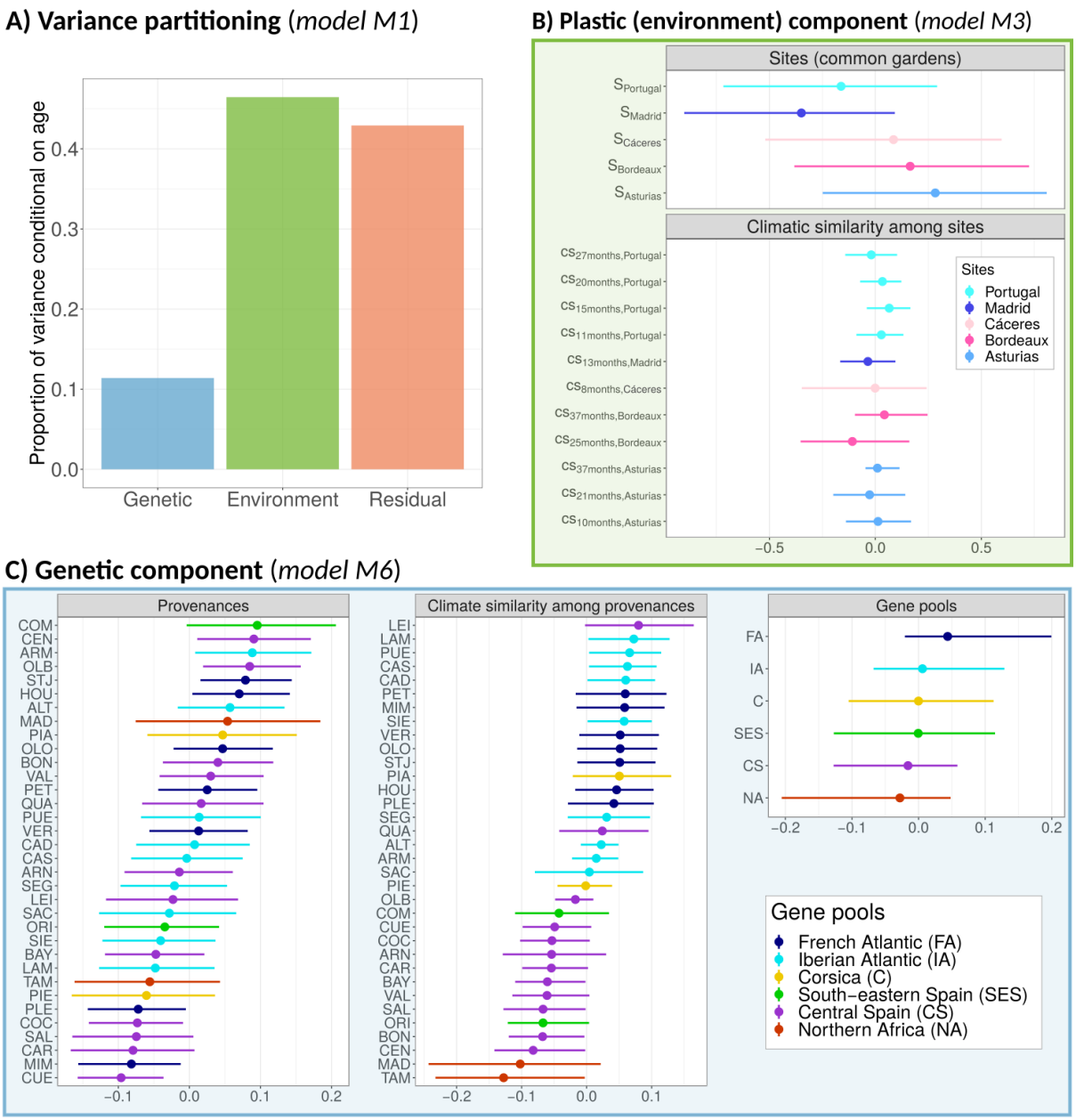


Figure V.2. Understanding the genetic and plastic bases of height-growth variation and their potential underlying drivers. A) shows the variance partitioning conditional on age from *model M1* in the P1 partition. B) displays the partitioning of the plastic (i.e. environment) component in *model M3* among the intercepts of the sites (common gardens) (S_s) and the intercepts associated with the climatic similarity among sites during the year preceding the measurements (cs_{is}). C) displays the partitioning of the genetic component in model *M6* among the intercepts of the provenances (P_p), the intercepts associated with the climatic similarity among provenances (cp_p) and the intercepts of the the gene pools (g_j). The median and 0.95 credible intervals shown in B) and C) were obtained by fitting the *models M3* and *M6* on the P1 partition. Provenance names can be found in Table XI.2.

the tallest (e.g. CAD, SIE, PUE, LAM and CAS in northwestern Spain; all provenances along the French Atlantic coast; Figs. V.2 & V.3). Interestingly, the Leiria (LEI) provenance, which has a strong Iberian Atlantic component (Table XI.3) and had the highest climate intercept estimate (similar to that of the French Atlantic provenances; Fig. V.2C), was not among the tallest provenances (Fig. V.3), probably due to its mixed ancestry with the central Spain gene pool (Table XI.3). Also, the Corsican provenances showed contrasted climate intercepts (Fig. V.2), with a positive influence on height growth for Pinia (PIA) but not for Pineta (PIE), located under more Mediterranean conditions, which could explain their large differences in height growth (Fig. V.3). Finally, the four provenances from south-eastern Spain and northern Africa gene pools, under harsh Mediterranean climates, showed all negative climate intercepts (Fig. V.2). Noticeably, the total genetic variance of the Iberian and French Atlantic gene pools were

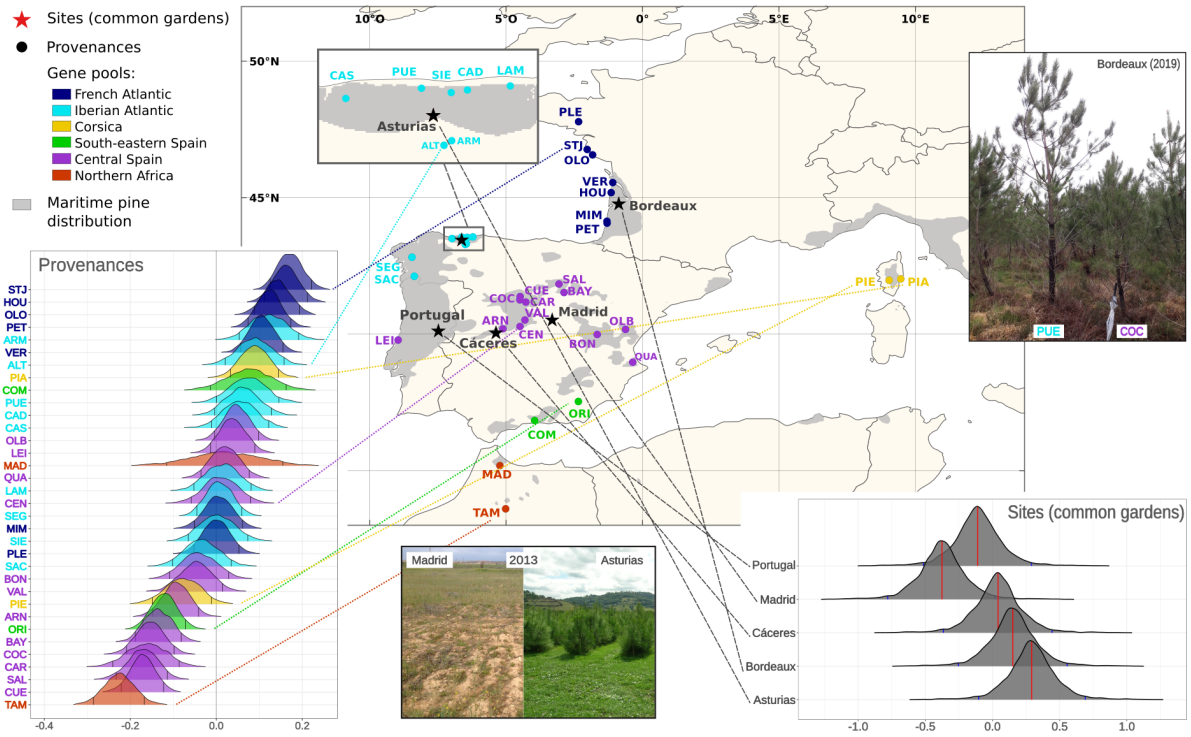


Figure V.3. Posterior distributions of the site and provenance intercepts (S_s and P_p) in *model M1* on a map representation. Provenances are colored according to the main gene pool they belong to. The exact values of the median, standard deviation and 0.95 credible interval of the posterior distributions of the site and provenance intercepts are shown in Tables XI.16 and XI.17, respectively. The top right picture shows the height difference in 2019 between one tree from Coca in central Spain (COC) and another from Puerto de Vega in the Iberian Atlantic region (PUE) growing next to each other in the Bordeaux common garden. The bottom picture shows the height difference between the trees growing in Madrid and Asturias, under highly contrasted environments, three years after plantation (2013). Provenance names can be found in Table XI.2.

likely to be lower than that of the Corsican and south-eastern Spain gene pools, and to a lesser extent the central Spain gene pool, thus resulting in gene pool-specific heritabilities (*model M5*; Table XI.28 and Fig. XI.13A).

4.2 Improved prediction of new observations and provenances by combining climatic and genomic data

In this part, we compared the baseline model *M2* (relying exclusively on the common garden design) to the predictive models that either combine genomic and climatic drivers of height-growth variation (i.e. *models M7* and *M8*) or include each driver separately (i.e. *models M9* to *M12*). Models combining genomic and climatic data generally explained in-sample variation almost as well as *M2*, and sometimes even better; e.g. *model M8* (which includes regional PEAs, rPEAs) in the Mediterranean sites (Madrid and Cáceres) (Fig. XI.10). Models including each driver of height-growth variation separately had a lower goodness-of-fit (for all common gardens) than both *M2* and the models combining the genomic and climatic data, except for *M12* (the model including only rPEAs), which explained in-sample variation almost as well as *M2* and even better than *M7* in Madrid (Fig. XI.10).

Model differences in their predictive ability on new observations (observations not used to fit the models; test dataset of the P1 partition) showed similar patterns than for the goodness-of-fit (Table V.4), which was expected as the new observations were sampled among the same provenances and genotypes. However, importantly, models combining genomic and climatic data provided much better predictions of height-growth on new provenances (provenances not

used to fit the models; test datasets of the P2 and P3 partitions) than did *M2*, with *M8* having a better predictive ability than *M7* in the Mediterranean sites in the P2 partition and in the Atlantic sites in the P3 partition (Fig. V.4). Models including each driver of height-growth variation separately had also a higher predictive ability on new provenances than *M2*, albeit lower than models combining genomic and climatic data, except *model M12* that showed a higher predictive ability than *M7* in the Mediterranean sites in the P2 partition (Fig. V.4). In *model M12*, one standard deviation increase in rPEAs was associated, on average, with 19.0% increase in height in Madrid, 12.7% in Cáceres, 13.0% in Portugal, 10.4% in Asturias and 9.6% in Bordeaux (section 6.4 of the Supplementary Information). More details on model comparisons are given in Section 5 of the Supplementary Information.

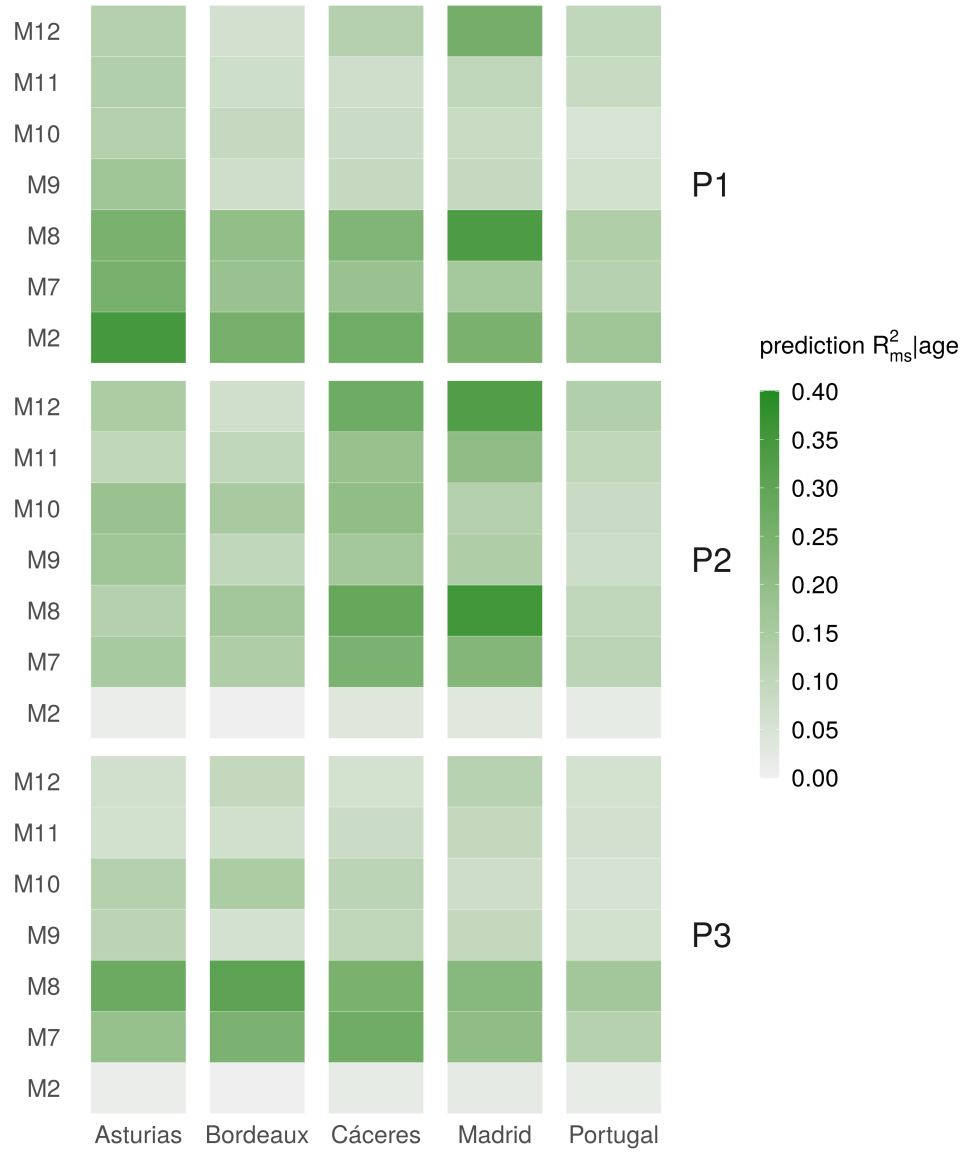


Figure V.4. Model predictive ability on new observations (P1 partition) or new provenances (P2 and P3 partitions) based on the out-of-sample proportion of predicted variance conditional on the age effect (*prediction* $\mathcal{R}^2_{ms}|age$) in the test datasets (data not used to fit the models). In the P1 partition, the training dataset was obtained by randomly sampling 75% of the observations and the test dataset contains the remaining 25% observations. In the P2 partition, the training dataset was obtained by randomly sampling 28 provenances and the test dataset contains the remaining 6 provenances. The P3 partition corresponds to a non-random split between a training dataset of 28 provenances and a test dataset containing 6 provenances with at least one provenance from each under-represented gene pool. The exact values of the *prediction* $\mathcal{R}^2_{ms}|age$ estimates and their associated credible intervals can be found in Tables XI.4 (P1 partition), XI.9 (P2 partition) and XI.12 (P3 partition).

5 Discussion

We combined genomic, climatic and phenotypic data from five common gardens and 34 provenances of maritime pine (over 30,000 observations) to predict range-wide variation in height growth, a key adaptive trait in forest trees. The plastic component explained the largest part of the deviation from the mean height-growth trajectory (~47%), probably due to multiple (confounded) environmental factors, including climate. The genetic component explained ~11% of the deviation from the mean height-growth trajectory and was mainly associated with the provenance climate-of-origin (a proxy of adaptation to climate), whose effect was partially confounded with the proportion belonging to distinct gene pools (a proxy for population demographic history and genetic drift, probably reflecting also the different selective histories of the gene pools). Importantly, we showed that models combining climatic drivers of adaptation, gene pool assignment and counts of height-associated positive-effect alleles (PEAs) captured well the genetic component underlying height-growth variation. They also better predicted height growth of new provenances than models relying exclusively on the common garden design or models including separately climatic and genomic information (e.g. the widely used climate-based population response functions). Interestingly, PEAs that show a regional association with height growth (rPEAs) had a higher predictive ability than PEAs identified globally across the species range (gPEAs). These results pave the way towards integrating genomics into large-scale predictive models of quantitative trait variation.

5.1 Predominant role of height-growth plasticity

Plants are known for their remarkable phenotypic plasticity to changing environments (Bradshaw 1965). In long-lived forest trees, the plastic component of quantitative trait variation estimated based on the common garden design is generally higher than the genetic component (Benito Garzón et al. 2019, Franks et al. 2014), e.g. in maritime pine (Chambel et al. 2007, Corcuera et al. 2010, de la Mata et al. 2012, Vizcaíno-Palomar et al. 2020). This plastic component is also generally associated with the climatic conditions experienced by the trees (Benito Garzón et al. 2019, Franks et al. 2014), allowing them to overcome changing climate up to a certain threshold (Matesanz et al. 2010, Nicotra et al. 2010, Valladares et al. 2014a). In our study, the plastic component of height growth was largely higher than the genetic component (Fig. V.2) and, although climate plays a role, was likely to be driven by multiple and interacting drivers including the biotic environment, soil quality, and other factors not considered in our study.

Plants also present an important genetic variation in plasticity (i.e. the genotype-by-environment interaction, G×E; Des Marais et al. 2013, Sork 2018), often approximated by the family or provenance-by-site interaction in forest tree common gardens, as is the case in our study. G×E is particularly prevalent for growth traits in trees (Li et al. 2017), as already shown in maritime pine (Alía et al. 1997, Corcuera et al. 2010, Correia et al. 2010, de la Mata et al. 2012; but see Chambel et al. (2007) where no provenance-specific responses were observed under two different watering regimes). In our study, provenance-by-site interaction was only weakly associated with height growth and the proportion of variance it explained was not different from zero (*model M2*; Table XI.5). Previous work in the context of tree breeding argued that G×E may hinder model transferability across sites and populations (Resende Jr et al. 2012, Resende et al. 2012). In maritime pine, our results suggest that large-scale predictions of height-growth variation will be only marginally impacted by not accounting for provenance-by-environment interaction. However, further work is necessary to assess the importance of

the genetic variation of plasticity at the genotype level.

5.2 Potential drivers underlying height-growth genetic component

Our study shows that the height-growth genetic component in maritime pine is mostly associated with adaptation to climate, whose effect is partially confounded with the effect of gene pool assignment, reflecting both adaptive (different selective histories) and neutral processes (population demographic history and genetic drift) (Fig. V.2; see also Jaramillo-Correa et al. 2015). For example, the higher growth of most provenances from the French Atlantic gene pool (known for their high growth under a wide range of conditions, including Mediterranean sites in our study; see also Alía et al. 1997, Corcuera et al. 2010, de la Mata et al. 2012) was both associated with the provenance climate-of-origin and the gene pool assignment. As another example, in the northern Africa gene pool, the Madisouka (MAD) provenance was taller than the Tamrabta (TAM) provenance, which could be both explained by its noticeable ancestry proportion (23.3%) from the south-eastern Spanish gene pool (Jaramillo-Correa et al. 2015) or its adaptation to lower elevation (300 m lower than TAM). As a last example, the Leiria (LEI) provenance grew well in Asturias and Bordeaux as was the case for French Atlantic provenances (that share similar climates) but unlike them, it did not maintain growth in drier and warmer sites, probably due to a different genetic background (this provenance has a strong central Spain gene pool component; Table XI.3). Nevertheless, in contrast to the three examples above, for some provenances, the effects of the gene pool assignment and adaptation to climate on height growth could be clearly separated. This was the case, for example, for the Corsican provenances: the higher growth of Pinia (PIA) than Pineta (PIE) can only be explained by adaptation to different environmental conditions (and in particular climate), as both belong to the same gene pool. Indeed PIA is at the sea level under a climate similar to that of provenances from Central and south-eastern Spain whereas PIE is located at an altitude of 750 m a.s.l. in the mountains under a climate similar to that of the Atlantic provenances (Fig. XI.9). These different adaptations within a same gene pool calls for a more targeted investigation of the Corsican gene pool. More generally, a $Q_{ST} - F_{ST}$ analysis supported adaptive differentiation of height growth in maritime pine (see details in Section 7 of the Supplementary Information).

The entanglement of the effect of climate adaptation and gene pool assignment to explain the genetic component of height-growth variation may partly stem from the distinct selective histories experienced in different parts of maritime pine range, despite gene pools being identified using genetic markers considered neutral (Jaramillo-Correa et al. 2015). This is supported by the estimation of gene pool-specific heritabilities in our study (*model M5*): the Corsican gene pool, and to a lesser extent the south-eastern Spain gene pool, have higher heritabilities than the French and Iberian Atlantic gene pools (Fig. XI.13; and see Section 6.1.3 for a potential explanation of this pattern).

Overall, maritime pine proved to be a particularly suitable model species to study the joint influence of genetic neutral (population demographic history, genetic drift) and adaptive (climate adaptation) processes on quantitative traits. Further work on provenances that have different demographic histories but are exposed to similar climates (e.g. the LEI provenance and provenances from the Atlantic gene pools) would be relevant for understanding how a given genetic background guides population adaptation. Conversely, targeting provenances that have a similar demographic history but are found in highly contrasted environments (e.g. the Corsican provenances) would be valuable to identify signatures of adaptation while avoiding common issues due to confounding population structure (Berg et al. 2019, Sella and Barton

2019, Sohail et al. 2019). Likewise, investigating trait genetic architecture will also help better understand how adaptive and neutral processes have shaped the genotype-phenotype map and how this will influence future responses to selection (e.g. Kardos and Luikart 2021; see de Miguel et al. 2020 for maritime pine). Finally, it would also be critical to consider drivers of adaptation other than climate, such as resistance to pathogens or other biotic-related traits.

5.3 Towards integrating genomics into population response functions

Anticipating how provenances will grow in new environments is key to guide forest conservation strategies and population translocations to compensate for rapid climate change (Aitken and Whitlock 2013). To date, population response functions based on the climate in the provenance location have been the most widely used method for anticipating trait values when transplanting provenances in new environments (Fréjaville et al. 2020, O'Neill et al. 2008, Pedlar and McKenney 2017, Rehfeldt et al. 2018, 2003, 1999, Wang et al. 2010). Genome-informed predictive modelling of key adaptive traits is highly promising as it may provide a means to further integrate adaptive or neutral genetic variation in the predictions, and to consider intraspecific variability at a finer scale than current models, thus gaining in prediction accuracy (Holliday et al. 2017). In valley oak, Browne et al. (2019) used genomic estimated breeding values (GEBVs; sum of the marker predicted effects, also known as polygenic scores) to identify genotypes that will grow faster under future climates. In lodgepole pine, Mahony et al. (2020) showed that phenotype-associated positive-effect alleles (PEAs, as used in our study) can predict phenotypic traits (e.g. cold injury) as well as climatic or geographical variables. In our study, we investigated whether including genomic information related to past demographic and selective processes resulting in distinct gene pools and counts of trait-associated alleles could improve range-wide height-growth predictions in maritime pine. Models combining climatic conditions in the provenance location, gene pool assignment, and PEAs captured most of the genetic component of height-growth variation (see Fig. XI.10) and better predicted height growth of new provenances, compared to models relying exclusively on the common garden design or models including separately climatic or genomic information (see Fig. V.4). This suggests that range-wide trait prediction would benefit from jointly considering different sources of information (i.e. climatic and genomic), even though they may have overlapping effects (e.g. confounded effects of provenance climate-of-origin and gene pool assignment), as it may help to embrace the complexity and multidimensionality of the genetic component underlying quantitative traits. Noticeably, regional PEAs were generally better predictors of height growth in new provenances than gene pool assignment or provenance climate-of-origin as, when they were included alone in the models, they made better predictions in the driest common gardens (Madrid, Cáceres and Portugal) and similar ones to models combining multiple drivers of height growth variation in all common gardens except Bordeaux (P2 partition in Fig. V.4). Although this highlights the major role that trait-associated alleles identified using GWAS may play in predictive modelling, predicting traits of new provenances depends also on the number of provenances used to fit the models and the strength of the genetic relationship among them (Hidalgo et al. 2016, Jarquín et al. 2014, Moghaddar et al. 2014, Resende et al. 2012). This was reflected in our study by better predictive ability on new provenances in the P2 partition (random) compared to the P3 partition (containing provenances from underrepresented gene pools) for models including climatic and genomic information separately but not for models considering both jointly (Fig. V.4). Thus combining multiple sources of information may also be particularly relevant for predicting traits in marginal or difficult-to-access populations, as they normally belong to underrepresented geographical areas/gene pools in ecological and genetic studies.

The high predictive ability of PEAs, both alone and combined with climatic and gene pool information, was somehow unexpected given the sparse genomic sampling in our study: 5,165 SNPs to cover the 28 Gbp maritime pine genome (Zonneveld 2012). Indeed, conifers have particularly huge genomes, generally ranging from 18 to 35 Gbp (Mackay et al. 2012) and thus rendering the current cost of whole-genome resequencing prohibitive (Holliday et al. 2017). Targeted genotyping approaches, such as the one used in the present study, select candidate genes based on previous population and functional studies, thus allowing to include potential targets of selection and climate adaptation, but probably inducing an ascertainment bias (Jaramillo-Correa et al. 2015). However, as height is a particularly polygenic trait (degree of polygenicity estimated at ~7% in de Miguel et al. 2020), we were able to identify a considerable number of PEAs despite the weak genome coverage of our study. Further genomic sampling would be highly valuable to capture the polygenic architecture of height more broadly, turning PEAs into much better predictors than the provenance climate-of-origin or the gene pool assignment, and ultimately making climatic data redundant, at least for main range populations (see above for marginal populations). This would also allow to characterize the genetic variation within provenances more precisely, thereby increasing the estimation accuracy and reducing the residual variance. Similar to Mahony et al. (2020) and MacLachlan et al. (2021) who selected the positive-effect alleles as the 1% of SNPs that showed the strongest association with phenotypes (estimated via a GWAS performed on 18,525 SNPs), we used PEA counts instead of the more commonly used polygenic scores (Browne et al. 2019, Fuller et al. 2020, Pritchard et al. 2010). Unlike polygenic scores, PEAs do not account for allele effect sizes, thus minimizing the circularity of the analysis (i.e. effect sizes that are estimated based on the same dataset as the one used for the models, only serve for PEAs identification) and potentially enhancing the prediction accuracy across genetic groups compared to polygenic scores. Indeed, low observed transferability of polygenic scores across genetic groups (Barton et al. 2019, Martin et al. 2017, 2019) may stem from varying effect sizes of "peripheral" alleles (i.e. alleles indirectly affecting the phenotype), as suggested in Mathieson 2021).

Although combining climatic and genomic information allowed us to capture most of the genetic component of height-growth variation (Fig. XI.10), the residual variance remained high in our study. As already mentioned, this may be partly related to the models' difficulty in accounting for genetic variation within provenances, which might be improved by denser genomic sampling. However, this unexplained variance may also originate from developmental stochasticity, which can play an important role in explaining differences between individuals with the same genotype (Ballouz et al. 2019, Vogt 2015). Height growth may also be influenced by the correlative effects of other traits. For example, Stern et al. (2020) recently showed that variation in some human traits (hair color and educational attainment), previously thought to be under selection, can instead be explained by indirect selection via a correlated response to other traits. Therefore, multi-trait models may be the next necessary step to improve our understanding and predictive ability of quantitative trait variation at large geographical scales (e.g. Csilléry et al. 2020).

A last noticeable results was that rPEAs (positive-effect alleles identified in specific geographical regions, i.e. particular environments) had generally a higher predictive ability than gPEAs (positive-effect alleles identified range-wide) (Fig. V.4). Interestingly, only a small proportion of rPEAs were shared among geographical regions in our study (20% shared between the Iberian and French Atlantic regions, 12% between the French Atlantic and Mediterranean regions, and 24% between the Iberian Atlantic and Mediterranean regions; Fig. XI.2), although we cannot exclude that the proportion of shared rPEAs among regions is a function of the sample size (see details in Section 2.2 of the Supplementary Information). Moreover, those that were shared

among different regions showed consistently similar effects across regions (e.g. positive effects in two or more regions rather than antagonist effects). This supports the predominance of conditional neutrality, i.e. alleles that are advantageous in some environments and neutral in others, over antagonistic pleiotropy, i.e. alleles that are advantageous in some environments and disadvantageous in others (Tiffin and Ross-Ibarra 2014). Such pattern has already been reported in plants (Anderson et al. 2013, Prunier et al. 2012). Our results show that, despite a high stability in the level of polygenicity for height between the Atlantic and Mediterranean regions (de Miguel et al. 2020), height-growth variation in Mediterranean sites is unlikely to be affected by the same loci as in the other regions, probably as a result of genetic divergence in separated southern refugia during the last glaciation. Overall, identifying positive-effect alleles for different geographical regions separately has the potential to greatly improve the predictive ability of the models, but at the cost of reducing GWAS power (due to lower sample size than in global, wide-range analyses).

Finally, caution has to be taken when generalizing our results to older trees as the drivers of height growth in young trees may differ from that of adult trees. For example, G×E on tree height can be age-dependant (Gwaze et al. 2001, Rehfeldt et al. 2018, Zas et al. 2003) and the plastic component may be higher in younger trees, especially in maritime pine (Vizcaíno-Palomar et al. 2020). Nevertheless, a recent measurement in the Bordeaux common garden (2018) showed a high correlation between young saplings and 10-year old trees for height (Pearson's correlation coefficient of 0.893 based on height BLUPs; see de Miguel et al. 2020 for details on BLUP estimation). Moreover, our study remains indicative of how trees respond to varying environmental conditions during establishment and early-growing stages, a critical phase where most mortality (i.e. selection) is expected to take place (Postma and Ågren 2016). In addition to ontogenic effects, high mortality in the Mediterranean common gardens (Cáceres and Madrid), after a marked summer drought, may have biased estimates of some parameters of interest. Indeed, if this environmental filtering was not independent of tree height, it could have resulted in an underestimation of the genetic variance. Nonetheless, height distributions in Cáceres and Madrid were only slightly right-skewed, suggesting uniform selection across height classes (Fig. XI.21), and thus no bias due to high mortality in these common gardens.

6 Conclusion

The present study connects climate-based population response functions that have been extensively used in predictive models for forest trees (Leites et al. 2012a, Rehfeldt et al. 2003, 1999, Wang et al. 2010) with recent genomic approaches to investigate the potential drivers behind the genetic and plastic components of height-growth variation and predict how provenances will grow when transplanted into new climates. The integration of genomic data into range-wide predictive models is in its infancy and still lacks a well-established framework, especially for non-model species such as forest trees. We showed that combining climatic and genomic information (i.e. provenance climate-of-origin, gene pool assignment and trait-associated positive-effect allele counts) can improve model predictions for a highly polygenic adaptive trait such as height growth, despite sparse genomic sampling. Further genomic sampling may help to improve the accuracy of the estimates, notably through improved characterization of within-provenance genetic variation. Moreover, comparative studies between maritime pine and more continuously distributed species (e.g. Scots pine; Alberto et al. 2013) and/or living under stronger climatic limitations, would be highly valuable to determine whether our findings can be generalized to species with contrasted population demographic and selective history.

Finally, our study focuses specifically on the height-growth genetic component of standing populations, but considering evolutionary processes (e.g. genetic drift in small populations, extreme selection events, etc.) into the predictions would be necessary to anticipate the response of future forest tree generations to changing climatic conditions and thus provide a much-needed longer-term vision (Waldvogel et al. 2020).

7 Acknowledgements

We thank A. Saldaña, F. del Caño, E. Ballesteros and D. Barba (INIA) and the 'Unité Expérimentale Forêt Pierroton' (UEFP, INRAE; doi:10.15454/1.5483264699193726E12) for field assistance (plantation and measurements). Data used in this research are part of the Spanish Network of Genetic Trials (GENFORED, http://www.genfored.es). We thank all persons and institutions linked to the establishment and maintenance of field trials used in this study. We are also very grateful to Ricardo Alía who contributed to the design and establishment of the CLONAPIN network and provided comments on the manuscript, and to Benjamin Brachi, Thibault Poiret, Andrew J. Eckert and one anonymous reviewer who provided constructive and valuable comments on a previous version of the manuscript. Thanks are extended to Juan Majada for initiating and supervising the establishment of the CLONAPIN network. JA was funded by the University of Bordeaux (ministerial grant). This study was funded by the Spanish Ministry of Economy and Competitiveness through projects RTA2010-00120-C02-02 (CLONAPIN), CGL2011-30182-C02-01 (AdapCon) and AGL2012-40151-C03-02 (FENOPIN). The study was also supported by the 'Initiative d'Excellence (IdEx) de l'Université de Bordeaux - Chaires d'installation 2015' (EcoGenPin) and the European Union's Horizon 2020 research and innovation programme under grant agreements No 773383 (B4EST) and 862221 (FORGENIUS).

8 Author contributions

SCG-M and CP designed the experiment and supervised the curation of field data. MdM cleaned and formatted the phenotypic and genomic data, and produced the BLUPs used in GWAS. SCG-M and MdM ran the GWAS to identify the positive-effect alleles. SCG-M, MBG, JA and FB conceived the paper methodology. JA conducted the data analyses. SCG-M, MBG, JA and FB interpreted the results. JA led the writing of the manuscript. All authors contributed to the manuscript and gave final approval for publication.

9 Data and script availability

Data are publicly available. SNP data were deposited in the Dryad repository at http://dx.doi.org/10.5061/dryad.8d6k1. Height data have been deposited in GENFORED, the Spanish Network of Genetic Trials (http://www.genfored.es). Scripts are available at https://github.com/JulietteArchambeau/HeightPinpinClonapin.

VI

Chapter 3. Short-term climate maladaptation in a keystone forest tree: genomic prediction and validation

Juliette Archambeau[1], Marta Benito Garzón[1], Marina de Miguel Vega[1,2], Alexandre Changenet[1], Camilla Avanzi[3], Francesca Bagnoli[3], Frédéric Barraquand[4], Giovanni G. Vendramin[3], Santiago C. González-Martínez[1]

1 INRAE, Univ. Bordeaux, BIOGECO, F-33610 Cestas, France
2 EGFV, Univ. Bordeaux, Bordeaux Sciences Agro, INRAE, ISVV, F-33882, Villenave d'Ornon, France
3 Institute of Biosciences and BioResources, National Research Council, 50019 Sesto Fiorentino, Italy
4 CNRS, Institute of Mathematics of Bordeaux, F-33400 Talence, France

Keywords: Climate change, risk of maladaptation, genomic offset, maritime pine, landscape genomics, local adaptation.

1 Abstract

A major goal in evolutionary biology is to understand how populations adapt to their environment and predict how they will respond to future conditions, in particular climate change. The genomic offset approach is an increasingly popular metric aiming at identifying populations for which the gene-environment relationships will be the most disrupted under new climatic conditions, i.e. populations at risk of short-term climate maladaptation. However, this approach relies on key assumptions in need of empirical validation. We used 9,817 SNPs from 34 populations of maritime pine, a keystone forest tree from southwestern Europe and north Africa. Our results suggest that populations already experiencing mild-winter conditions (i.e. most Atlantic populations and Mediterranean populations in southeastern France and northwestern Italy) are at higher maladaptation risk under climate change. In a validation analysis based on independent height and mortality data from common gardens and natural populations (National Forest Inventories, NFI), genomic offset predictions were generally negatively associated with population performance. This confirms the genomic offset assumption that its predictions are indicative of (future) fitness declines. However, genomic offset predictions were highly sensitive to how strictly the candidate SNPs were selected and the modelling approach used. Moreover, among the different ways tested to predict genomic offset, none was consistently accurate in the different validation steps. This highlights the need for further research evaluating different genomic offset estimation and validation approaches to produce robust and accurate predictions that can confidently be used to guide conservation and management strategies.

2 Introduction

There is growing evidence that biodiversity is already being affected by anthropogenic climate change (Parmesan and Yohe 2003). Populations and species are experiencing range shifts or losses (Chen et al. 2011), and changes in their genomic composition (Bradshaw and Holzapfel 2006, Jump and Peñuelas 2005), which may result in widespread local population decline and, potentially, species extinctions (Urban 2015, Wiens 2016). Forecasting the short-term impacts of climate change on species and populations has recently been improved by shifting from traditional distribution-based approaches, which assume that populations within species respond uniformly to climate (Guisan and Thuiller 2005), to approaches integrating intraspecific genetic variation (Banta et al. 2012, Benito Garzón et al. 2019, Jay et al. 2012, Razgour et al. 2019, Valladares et al. 2014a). In particular, adaptive genetic variation is a key component of population response to changing environmental conditions. Optimal phenotypes to a given environment result from the evolutionary process of local adaptation through natural selection (Davis and Shaw 2001, Hereford 2009, Jump and Peñuelas 2005, Kawecki and Ebert 2004, Leimu and Fischer 2008, Savolainen et al. 2007). Malaptation is the flip side of adaptation: the process of producing suboptimal phenotypes (Brady et al. 2019b, Crespi 2000). Maladaptation occurs within populations as, for selection to act on a population, not all individuals can have the optimal phenotype (Brady et al. 2019a, Crespi 2000). Maladaptation is also common at the population level, with some populations deviating from the adaptive peak in their native environment, which may be due to changes in population trait distributions or in the environment, or arising from eco-evolutionary or eco-plasticity feedbacks (Angert et al. 2020, Brady et al. 2019b). Climate change is potentially one major cause of future maladaptation as it disrupts the current adaptation patterns by shifting the optimal phenotype away from current phenotypes (i.e. scenario of the moving target in Brady et al. 2019a). Populations with

sufficient adaptive genetic variation are likely to cope with moderate environmental changes by evolving towards the new optimum (Brady et al. 2019b). However, identifying the adaptive component of individual and population genetic variation is highly challenging (Oetting et al. 2017), and whether populations will be able to adapt to the predicted rapid climate change remains uncertain (Chevin et al. 2010).

In forest trees, adaptation to abiotic and biotic gradients shapes the geographical patterns of phenotypic and genetic variation (Aitken et al. 2008, Alberto et al. 2013, Langlet 1971, Savolainen et al. 2007, Sork et al. 1993). Maladaptation is also common, one of the best documented patterns across boreal, temperate and Mediterranean tree species being that populations at the leading edge would benefit from a temperature increase, at least in the short term (Fréjaville et al. 2020, Pedlar and McKenney 2017, Rehfeldt et al. 2018, 2003, 2002, Savolainen et al. 2007). A potential explanation is that these populations may not be at equilibrium with the current climate due to adaptation lags along their post-glacial colonization routes (García-Valdés et al. 2013, Johnstone and Chapin 2003). Populations in the Mediterranean area are particularly prone to maladaptation as they are often small and fragmented and therefore more subject to genetic drift (Alberto et al. 2013). Moveover, the Mediterranean area is expected to experience strong changes in precipitation and temperature in the coming decades (Hoegh-Guldberg et al. 2018), rendering the tree populations in this region particularly vulnerable. This high vulnerability is accentuated by the specific features of forest trees: they are sessile and have long generation times, so that they may not be able to adapt or migrate fast enough to track rapid climate change (Aitken et al. 2008, Browne et al. 2019, Dauphin et al. 2021, Jump and Peñuelas 2005). They also show high phenotypic plasticity that could help to survive climate change up to a certain threshold (Benito Garzón et al. 2019, Nicotra et al. 2010), or could even promote evolutionary changes via genetic accommodation (Wund 2012). For all these reasons, to what extent the short-term maladaptations caused by climate change will trigger widespread local extinctions or be offset by rapid adaptations or large plastic responses remains largely unknown in forest trees. This question requires particular attention in forest trees as they have great economic and ecological importance. Indeed, forest trees play a major role in terrestrial ecosystems as foundation species, provide wood and fibre to satisfy an increasing demand of wood-based products, and are a main source of carbon sequestration (Bonan 2008, Brockerhoff et al. 2017, Gibson et al. 2011, Hooper et al. 2005).

In this study, we used maritime pine (*Pinus pinaster* Ait.), a tree species with fragmented and highly structured populations in the Mediterranean and Atlantic regions of western Europe and North Africa (Jaramillo-Correa et al. 2015), to identify the populations whose gene-environment relationships will be the most disrupted under climate change, i.e. populations at risk of short-term climate maladaptation (Capblancq et al. 2020a). Based on 9,817 common (MAF> 1%) single nucleotide polymorphisms (SNPs), we modeled the turnover in allele frequencies along current environmental gradients and estimated a metric of change of the genomic composition required to maintain the current gene-environment relationships under future climates across the species range (referred as 'genomic offset' in Fitzpatrick and Keller 2015). Then we put special emphasis in validating the key assumption that populations with high predicted genomic offset should experience a decrease in absolute fitness or declining demographic trends (Capblancq et al. 2020a). A first validation step consisted in detecting a negative association between population performance (i.e. height and mortality rates) and genomic offset predictions in common gardens (rather than in future climates), and compared it with the association between population performance and five climatic transfer distances. A second validation step involved searching for an association between recent mortality rates in natural populations (based on National Forest Inventories data) and genomic offset predictions under future climates, assuming that

populations predicted to suffer from short-term climate maladaptation already experience higher mortality rates. Genomic offset predictions in common gardens and natural populations were derived for all possible combinations of four sets of SNPs (i.e. reference SNPs and three sets of more or less strictly selected candidate SNPs for adaptation to climate using two gene-environment association analyses, GEAs), two modelling approaches (i.e. Gradient Forests, GF, and Generalised Dissimilarity Modelling, GDM) and two more or less alarming future climate scenarios (only for predictions in natural populations). As the genomic offset approach does not rely on phenotypes, such validation of genomic offset predictions with independent phenotypic data is of major importance to assess whether populations with high genomic offset are truly at risk of fitness and demographic decline, and can therefore be confidently considered as vulnerable to climate change (Capblancq et al. 2020a, Rellstab et al. 2021).

3 Materials & Methods

3.1 Model species

Maritime pine (*Pinus pinaster* Ait., Pinaceae) is a wind-pollinated, outcrossing and long-lived tree species with large economic and ecological importance in Western Europe and North Africa: largely exploited for its wood, stabilizing coastal and fossil dunes and, as a foundation species, supporting biodiversity (Viñas et al. 2016). Natural populations of maritime pine are scattered over a large range of environmental conditions, which makes the species a relevant case study for studying local adaptation. In addition, several studies have provided evidence of genetic differentiation for adaptive traits in maritime pine (e.g. de Miguel et al. 2020, González-Martínez et al. 2002). Maritime pine can grow in widely different climates: the dry climate along the northern coasts of the Mediterranean Basin (from Portugal to western Italy), the (Mediterranean) mountainous climates of south-eastern Spain and Morocco, the wetter climate of the Atlantic region (from north-western Spain and Portugal to the western part of France), and the continental climate of central Spain. Maritime pine can also grow on a wide range of substrates, from sandy and acidic soils to more calcareous ones, and can live in fire-prone regions, showing intraspecific variability in fire-related traits such as early flowering and serotiny (Budde et al. 2014, Tapias et al. 2004). Studying local adaptation in maritime pine is challenged by its strong population genetic structure: six distinct gene pools have been identified (Alberto et al. 2013, Jaramillo-Correa et al. 2015), presumably resulting from the expansion of as many glacial refugia (Bucci et al. 2007, Santos-del-Blanco et al. 2012).

3.2 Single-nucleotide polymorphism (SNP) genotyping

A rangewide sample of natural populations covering all known gene pools in maritime pine (454 trees from 34 populations; see Table XII.1 for the number of trees in each population and Fig. VI.1 for their location) were genotyped with the Illumina Infinium assay described in Plomion et al. (2016b) and with an Affymetrix assay developed in the framework of the H2020 EU B4EST project (4Tree; https://b4est.eu). We filtered out SNPs with MAF $< 1\%$ and more than 20% missing data, which resulted in 9,817 high-quality polymorphic SNPs, of which 2,855 were genotyped by both assays to ensure sample identity and estimate genotyping errors. The percentage of missing data per tree was less than 12% for all trees, with an average of 2.5%.

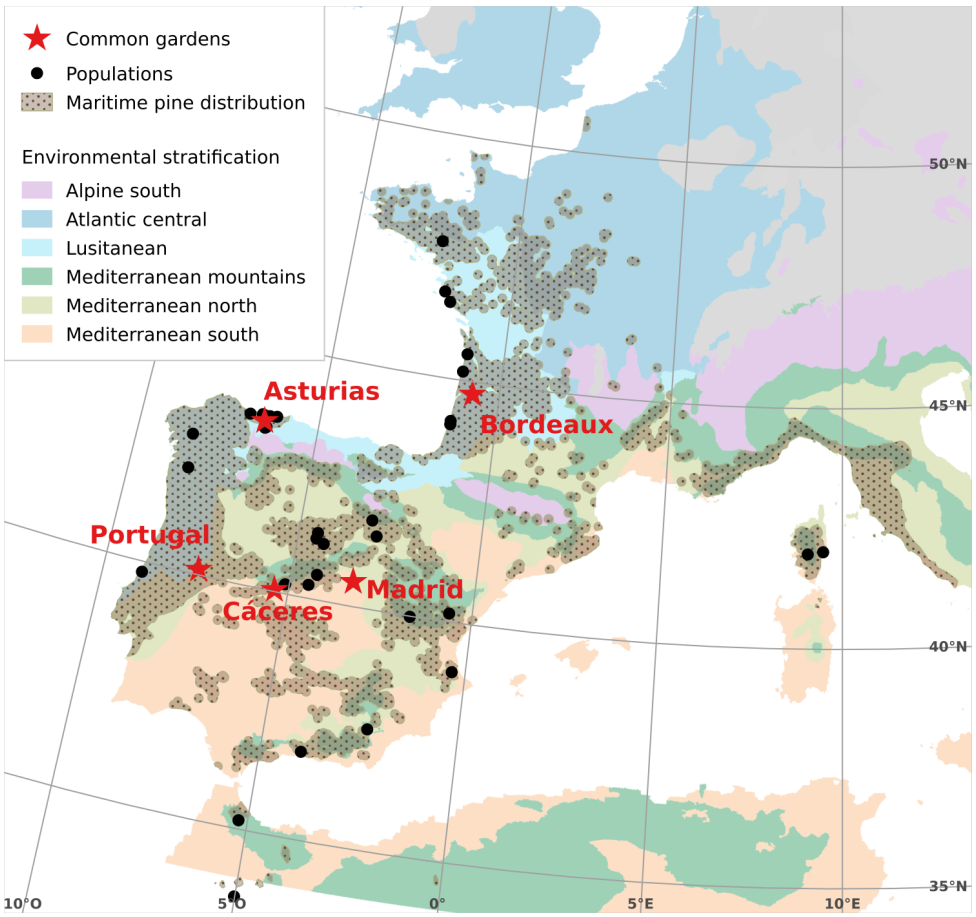


Figure VI.1. Location of the 34 genotyped populations and the five common gardens used in the validation steps and in which the same 34 populations were planted (CLONAPIN network). The environment stratification from Metzger (2018) shows the variety of environments inhabited by maritime pine. The maritime pine distribution combines the EUFORGEN distribution, based on natural populations (http://www.euforgen.org/), and 10-km radius areas around the French and Spanish National Forest Inventory plots with maritime pine occurrence.

3.3 Climatic, soil, topographic and fire-related data

Climatic data were extracted from the WorldClim database at 30 arc-seconds spatial resolution (Fick and Hijmans 2017). As a proxy of the climate under which the populations have evolved, we used the average of four climatic variables over the period 1970-2000, namely the maximum temperature of the warmest month (°C), the minimum temperature of the coldest month (°C), the annual precipitation (mm) and the precipitation seasonality (coefficient of variation, %). The potential future climate was described by two shared socio-economic pathways (the moderately alarming SSP3-7.0 and the strongly alarming SSP5-8.5) averaged over the period 2041-2060 at 2.5 arc-minutes spatial resolution. For each of these SSPs, the predictions of the nine global climate models available in WorldClim were averaged to account for scenarios' uncertainty (see Section 1.2 in the Supplementary Information). Topographic data were generated from NASA's Shuttle Radar Topography Mission (SRTM) at 90 m resolution. We used the SAGA v2.3.1 (Conrad et al. 2015) to calculate the topographic ruggedness index (TRI; m) which quantifies the terrain heterogeneity, i.e. differences in elevation between adjacent cells (Riley et al. 1999). Two soil variables were extracted from the European Soil Database at 1 km resolution: the total water content in the topsoil (0-30 cm; mm) and the depth available to roots (cm) (Hiederer et al. 2013). The average of the monthly burned area (in hectares) from June 1995 to December 2014 was extracted from the GFED4 database at 0.25 degrees resolution (~28 km resolution) (Giglio et al. 2013). At the population level (i.e. one value per population location), these eight selected environmental variables had at most a correlation coefficient of 0.7 (Fig. XII.1).

3.4 Identification of candidate SNPs potentially involved in local adaptation

We identified candidate SNPs potentially involved in local adaptation using both univariate and multivariate genotype–environment association (GEA) methods following Forester et al. (2018). Univariate GEA methods perform well for loci under moderate and strong selection but not under weak selection, whereas multivariate GEA methods perform well across all levels of selection. Missing allelic values were imputed using the most common allele within the main gene pool of the genotype of concern (although we acknowledge that some genotypes had high admixture rates).

For univariate GEA analysis, we used the standard covariate model with the Importance sampling approximation implemented in the BayPass software (Gautier 2015). This model explicitly accounts for the population genetic structure via a covariance matrix based on the population allele frequencies. The association between each environmental covariate and the SNPs was assessed according to the median Bayes Factor (in deciban units, db) calculated over five independent runs. We used a threshold of 5 db (corresponding to 3:1 odds) to identify the candidate SNPs.

For multivariate GEA analysis, we used Redundancy Analysis (RDA). Forester et al. (2018) showed RDA to have the best ratio between a low rate of false positives and a high rate of true positives across all levels of selection. We performed a partial RDA conditioned on the ancestry coefficients obtained with nine nuSSRs and 1,745 SNPs common to our study using the Structure software (see details in Jaramillo-Correa et al. 2015). We extracted from the significant constrained axes the SNP loadings in the ordination space. The SNPs that are more likely to be under selection (i.e. the candidate SNPs) are in the tail of the SNP loading distribution, thus we used a three standard deviation cutoff to identify them.

We grouped the candidate SNPs identified by BayPass and the RDA as follows (to represent different levels of confidence): 1) the common candidates between the two GEA methods, 2) the candidates under expected strong selection: RDA candidates that show a strong association with at least one covariate, i.e. with $\beta_{RDA} > 0.3$, and all the BayPass candidates, and 3) the merged candidates identified by at least one of the methods, either BayPass or the RDA, this group being the less conservative.

3.5 Genomic offset estimation

We had four sets of SNPs: the three sets of candidate SNPs defined above and a set of reference SNPs including all 9,817 SNPs. First, independence of loci within each set of SNPs was assessed by calculating pairwise linkage disequilibrium (LD) using the 'LD' function of the R package *genetics* v1.3.8.1.2. Then, for each set of SNPs, the relationship between current genomic composition and environmental conditions was estimated using two approaches from community-level modelling that were extended by Fitzpatrick and Keller (2015) to model genomic variation across the landscapes: Generalised Dissimilarity Modelling (GDM; Ferrier et al. 2007) and Gradient Forests (GF; Ellis et al. 2012). These two approaches model the compositional turnover in genomic variation as a function of environmental covariates (via allele turnover functions) and efficiently accommodate nonlinear gene–environment relationships (Fitzpatrick and Keller 2015). Fitzpatrick and Keller (2015) advise to use them in tandem as

they present complementary pros and cons: GDM provides a direct means of considering the geographical distance among populations (i.e. proxy of the population genetic structure) while GF better handles both correlations and interactions among predictors.

In the GDM analysis, the response variable is a matrix of genetic distances/differentiation between populations (a 34 × 34 matrix in our case). We estimated a pairwise F_{ST} matrix according to Weir and Cockerham (1984) using the R package *hierfstat* v0.04-22 (Goudet 2005). To facilitate model convergence and enable comparisons between reference and candidate SNPs (that displayed different ranges of observed F_{ST} values), the F_{ST} matrix was scaled by subtracting the minimum value and then dividing by the maximum minus the minimum value (resulting in F_{ST} values lying between 0 and 1). The GDM models were performed with the R package *gdm* v1.4.2 (Fitzpatrick et al. 2020). The model fit was assessed by the percentage of deviance explained and the model generalization ability by cross-validation (100 iterations of 2, 6 and 9 folds cross-validation) using the R package *sgdm* v1.0 (Leitão et al. 2017). In the GF analysis, the response variable is the population allele frequencies. To ensure regression robustness, we filtered away the SNPs that were polymorphic in only five or less than five populations (see Table XII.8 for the number of SNPs left). The GF models were performed with the R package *gradientForest* (Ellis et al. 2012) and using the same parameters as in Fitzpatrick and Keller (2015): 2000 regression trees per SNP, default values for the number of covariates randomly sampled as candidates at each split and the proportion of samples used for training and testing (~0.63 and ~0.37, respectively), maxLevel = log2(0.368*34/2) and a covariate correlation threshold of 0.5 (accounting for correlation between covariates by calculating conditional covariate importance). No imputation of the missing values was carried out when estimating the pairwise F_{ST} matrix (GDM) and the provenance allele frequencies (GF). The population genetic structure, resulting from the population demographic history and other spatial processes, was partially accounting for with the geographical distance between populations in the GDM analysis and with Moran's eigenvector map variables (MEM; Borcard and Legendre 2002, Dray et al. 2006) in the GF analysis, as calculated with the R package *adespatial* v0.3-8 (Dray et al. 2020).

In the GDM and GF analyses, the maximum height of each turnover function indicates the total amount of turnover in allele frequencies associated with that covariate and thereby corresponds to the covariate's relative importance in explaining changes in allele frequency while holding all other covariates constant (i.e. a partial genetic distance). In the GDM analysis, to obtain an estimate of the relative importance of the covariates between 0 and 1, we rescaled the splines by dividing by the maximum value across all covariates. In the GF models, the relative importance of the covariates is given by a weighted $\mathcal{R}^2$ across all SNPs. The shape of the GDM and GF turnover functions was used to identify the regions along the covariate gradient associated with a high rate of change in allele frequencies, irrespective of the underlying allele frequencies.

The GDM and GF turnover functions were then used to transform the current and future values of the environmental covariates across the maritime pine range into genetic importance values, which allows to project the current and future genomic composition across the landscape under the two scenarios of future climates considered (i.e. SSP3-7.0 and SSP5-8.5). To aid visualization, a principal component analysis (PCA) was used to reduce the dimensionality of the transformed environmental covariates and a RGB colour palette was assigned to the first three PCs, with resulting similar colour for similar expected patterns of genomic composition. We calculated the Euclidean distance between the current and future genetic importance values, which can be seen as the change in genomic composition required to maintain the current gene-environment relationships (referred as the 'genomic offset' in Fitzpatrick and Keller 2015).

It can be noted that the estimated changes in genomic composition between current and future environmental conditions reflect only differences in climatic conditions as soil, topography and burned area were considered fixed over time (although we acknowledge that burned area is expected to increase with climate change; Pausas and Fernández-Muñoz 2012).

3.6 Genomic offset validation

A key preliminary step before translating genomic offset values into potential population (mal)adaptation is to establish the correlation between genomic offset-based predictions and absolute fitness or demographic trends. For this, we used two complementary validation methods.

First, we evaluated whether populations that perform poorly in common gardens were those with the highest predicted genomic offset in the common gardens (calculated based on the environmental differences between the populations and the common gardens, rather than future climates), or those with the highest climatic transfer distance (i.e. absolute difference between the climate of the population locations and the common gardens). It can be noted that genomic offset was not predicted based on climate differences only between the population locations and the common gardens, but also accounted for soil and topographic differences, and therefore did not only represent climate (mal)adaptation (see Section 1.3 in the Supplementary Information for more details). The performance of the 34 genotyped populations (see Section 3.2) was estimated from height and mortality measurements on trees aged 8 to 85 month-old in a network of five common gardens planted in different environments: three under the favorable conditions of the Atlantic European region with mild winters, no severe cold events, high annual rainfall and relatively wet summers (trial sites of Asturias, Bordeaux and Portugal) and two in the harsh environments of the Mediterranean region with high temperatures and an intense summer drought (trial sites of Cáceres and Madrid) (Fig. VI.1). We used height measurements from all five common gardens while mortality measurements were taken from the two Mediterranean common gardens, Cáceres and Madrid, in which a severe summer drought exacerbated by clay soils killed 92% and 72% of the trees, respectively (Table XII.2). For each of the eight combinations of SNP sets (reference SNPs and the three candidate SNP sets) and models (i.e. GDM and GF), we estimated the association between BLUPs for tree height and mortality and the predicted genomic offset (see models in Section 1.3 in the Supplementary Information). Using the same models, we also estimated the association between tree performance and five climate transfer distances obtained independently from the four climate variables used to calculate genomic offset (see Section 3.3) and mean annual temperature (in °C), which is often used to estimate population response functions in forest trees (e.g. Pedlar and McKenney 2017, Rehfeldt et al. 2002). Last, we compared the model goodness-of-fit based on the proportion of variance explained by each model and the model predictive ability based on the leave-one-out cross-validation (LOOCV) procedure from the R package *loo* v2.2.0.

Second, we evaluated whether natural maritime pine populations in geographical regions where genomic offset was predicted to be high showed higher mortality rates. We used mortality data from the French and Spanish National Forest Inventories (NFI) estimated in Changenet et al. (2021) and covering the mortality observations between 2000 and 2014 for the French inventory and between 1986 and 2008 for the Spanish inventory. Trees diameter at breast height ranged from 10 to 263 cm, hence including sapling and adult trees (Changenet et al. 2021). We acknowledge that the predicted genomic offset, which measures a potential future

climate maladaptation, cannot causally explain the recent mortality patterns. Nevertheless, a positive association between the genomic offset and recent mortality patterns could indicate that populations predicted to suffer from climate maladaptation are already showing higher mortality rates. We independently estimated the association between mortality rates and the different genomic offset predictions obtained from the sixteen combinations of SNP sets (i.e. reference SNPs and the three candidate SNP sets), models (i.e. GDM and GF) and scenarios of future climates (i.e. SSP3-7.0 and SSP5-8.5) using the statistical model described in Section 1.4 in the Supplementary Information.

All models of the validation part were implemented in a Bayesian framework using the *Stan* probabilistic programming language (Carpenter et al. 2017), based on the no-U-turn sampler algorithm. Models were run with four chains and 2,000 iterations per chain (including 1,000 warm-up samples not used for the inference). All analyses were undertaken in R version 3.6.3 (R Core Team 2020) and scripts are available at https://github.com/JulietteArchambeau/GenomicOffsetPinPin.

4 Results

4.1 Genomic offset estimation

Eight candidate SNPs were identified by the two GEA methods (i.e. common candidates), 79 candidate SNPs showed a strong association with at least one covariate (i.e. candidates under expected strong selection) and 370 candidate SNPs were identified by at least one GEA method (i.e. merged candidates) (Table XII.4). LD was low for both reference SNPs and the three candidate SNP sets (Table XII.6). In the univariate GEA (BayPass), the vast majority of candidate SNPs were associated with the minimum temperature of the coldest month, including seven of the eight common candidate SNPs (Table XII.4). These seven common candidate SNPs were all also associated with the minimum temperature of the coldest month in the multivariate GEA (RDA) and five had $\beta_{RDA} > 0.3$.

The percentage of deviance explained by the GDM models ranged from 41% for the common candidates to 64% for the reference SNPs (Table XII.7; see also the predicted vs. observed genomic distance in Fig. XII.6), while their mean predictive ability quantified via 100 independent samples of 9-fold cross validation ranged from 23% to 50%, respectively (Table XII.7; see also mean predictive ability for 6-fold and 2-fold cross-validations in this table). In the GF models, individual SNP $\mathcal{R}^2$ averaged across all SNPs ranged from 23% for the common candidates to 37% for the candidates under expected strong selection (Table XII.8).

In both the GDM and GF analyses, covariates related to population structure (i.e. the geographical distance in the GDM models and the MEM in the GF models) were the most important predictors contributing to the genomic turnover, i.e. fitted I-splines and turnover functions with the highest maximum height in the GDM (Fig. VI.2 & Table XII.7) and GF models (Figs. XII.7-XII.10), respectively. A noteworthy exception was the set of common candidate SNPs in the GF model for which the most important covariate was by far the minimum temperature of the coldest month (Fig. XII.7). In the GDM analyses, the fitted I-splines of the reference SNPs and merged candidates were similar and indicated a very strong contribution of the geographical distance, a mid-contribution of the topographic ruggedness index (and

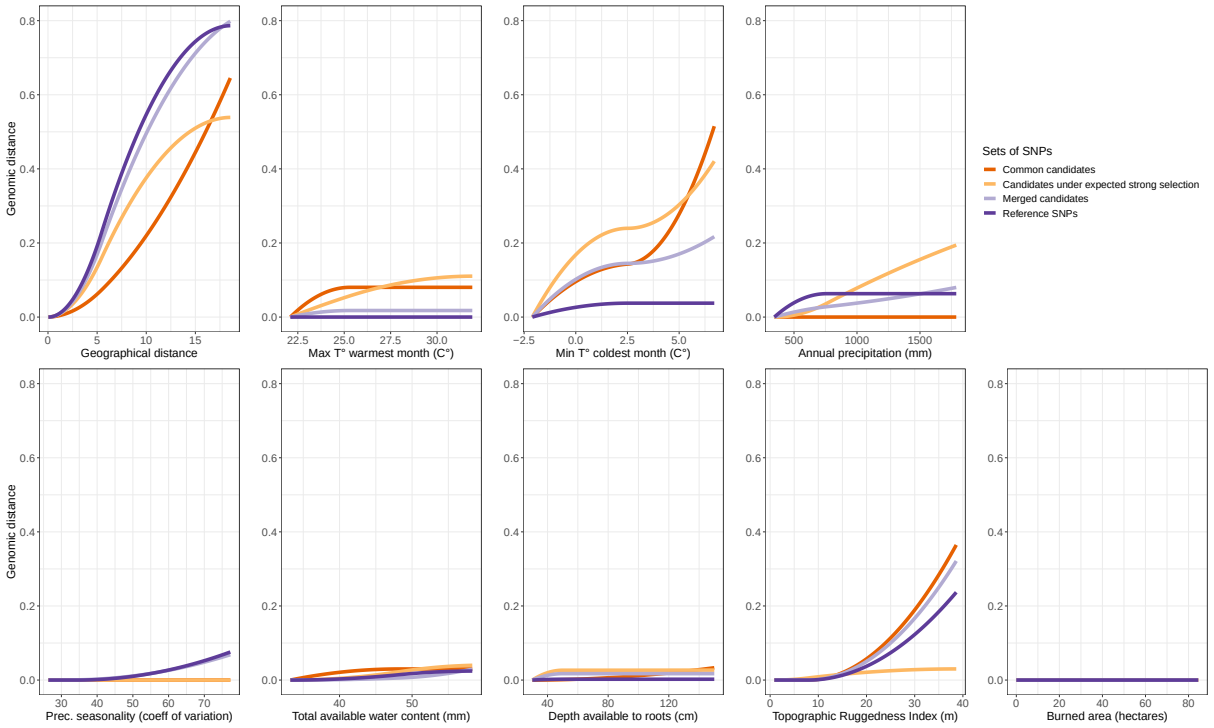


Figure VI.2. Fitted I-splines for the geographical distance and each environmental covariate in the GDM analyses performed with the four sets of SNPs.

the minimum temperature of the coldest month for the merged candidates only), a small contribution of the annual precipitation and precipitation seasonality, and a null (or almost null) contribution of the other covariates (Fig. VI.2 & Table XII.7). In comparison, for the common candidates and the candidates under expected strong selection, the importance of the geographical distance was smaller, while the importance of the minimum temperature of the coldest month was about 80% of the geographical distance's importance and that of the maximum temperature of the warmest month was between 12% and 20% (Fig. VI.2 & Table XII.7). Fitted I-splines of the common candidates also indicated a mid-contribution of the topographic ruggedness index, while those of the candidates under expected strong selection indicated a mid-contribution of the annual precipitation (Fig. VI.2). In the GF models, the minimum temperature of the coldest month was the most important environmental covariate for the three sets of candidate SNPs, but not for the reference SNPs, for which the environmental covariates had a minor relative importance compared to the MEM covariates (Figs. XII.7-XII.10). Burned area had also a slight importance in the genomic turnover of the set of common candidates (Fig. XII.7), and the annual precipitation and the maximum temperature of the warmest month noticeably contributed to the turnover of the set of candidates under expected strong selection (Fig. XII.8). Noticeably, for the common candidate SNPs, both the GDM fitted I-spline and the GF turnover function of the minimum temperature of the coldest month show a steep slope between 4 and 6°C (Figs. VI.2 & XII.7), indicative of a rapid turnover in allele frequency.

The predicted spatial variation in current and future genomic composition varied greatly among the models used to estimate the current gene-environment relationships (i.e. GDM vs GF models) and the sets of SNPs considered (see Section 2.3 in the Supplementary Information). As a result, the predicted spatial variation in genomic offset across the species range also showed strong differences among models and sets of SNPs but not among the two scenarios of future climate (Figs. VI.3 and XII.27). In both the GDM and GF approaches, higher predicted values of genomic offset were found for the more stringently selected sets of candidate SNPs, with the highest values being for the set of common candidate SNPs and with relatively small

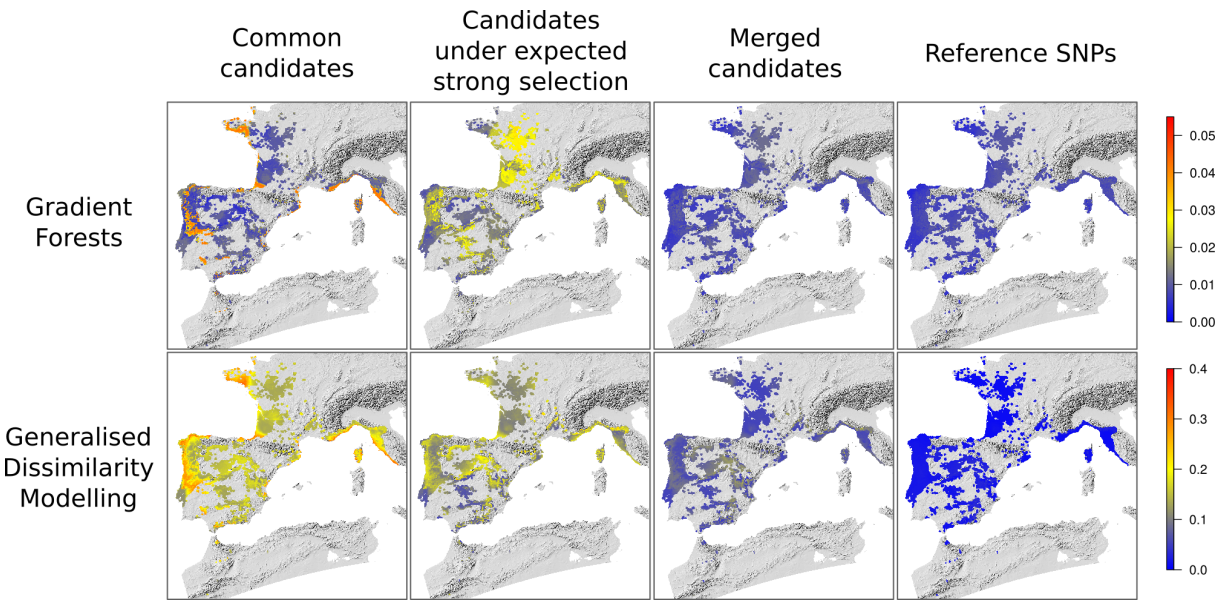


Figure VI.3. Predicted spatial variation in genomic offset for each combination of modelling approaches (i.e. Gradient Forests or Generalised Dissimilarity modelling) and sets of SNPs (i.e. three sets of candidate SNPs and the reference SNPs) under the future climate scenario SSP3-7.0 (moderately alarming). See Fig. XII.27 for predictions under the future climate scenario SSP5-8.5 (strongly alarming), and see Figure XII.28 for the same predictions but visualized with different scales so that the spatial variation in genomic offset for the merged candidates and reference SNPs are visible.

values for the sets of merged candidates and reference SNPs (Figs. VI.3 and XII.27). For the set of common candidates, GDM and GF-based predictions were very similar and the highest genomic offset values were found in French Brittany, the coasts of north-western Italy and south-eastern France, the southern part of the Landes forest in south-western France, and some inland parts of Galicia and Portugal (Figs. VI.3 and XII.27). For the set of candidates under expected strong selection, both GDM and GF-based genomic offset predictions were higher in the mountain foothills of south-eastern France and north-western Italy, and some parts of Galicia and northern Portugal, but only GF-based predictions were higher in a large part of western France (except Brittany) and the southern part of Central Spain, while only GDM-based predictions were higher in French Brittany, the south of the Landes forest and the southern part of Central Spain (Figs VI.3 and XII.27).

4.2 Genomic offset validation

Regression coefficients accounting for the linear association between the eight different predicted genomic offsets in common gardens and tree height or mortality were nearly all in the expected direction: negative in height models (Fig. XII.30), thereby suggesting that populations with higher genomic offset grow less, and positive in mortality models (Fig. XII.32), thereby suggesting that populations with higher genomic offset show higher mortality rates.

The proportion of explained variance varied widely among height models and no clear pattern emerged regarding which modelling approach (GDM or GF) or set of SNPs (reference SNPs or the three sets of candidate SNPs) provided the best performing genomic offset across all common gardens (Fig. VI.4). GF-based genomic offsets were better predictors than GDM-based genomic offsets in two of the Atlantic common gardens (Portugal and Asturias, except for reference SNPs) while the opposite was true in Bordeaux (except for the common candidate SNPs; Fig. VI.4). Indeed, for Atlantic common gardens, genomic offset predictions based on the set of common candidates with the GDM approach were not associated with height differences

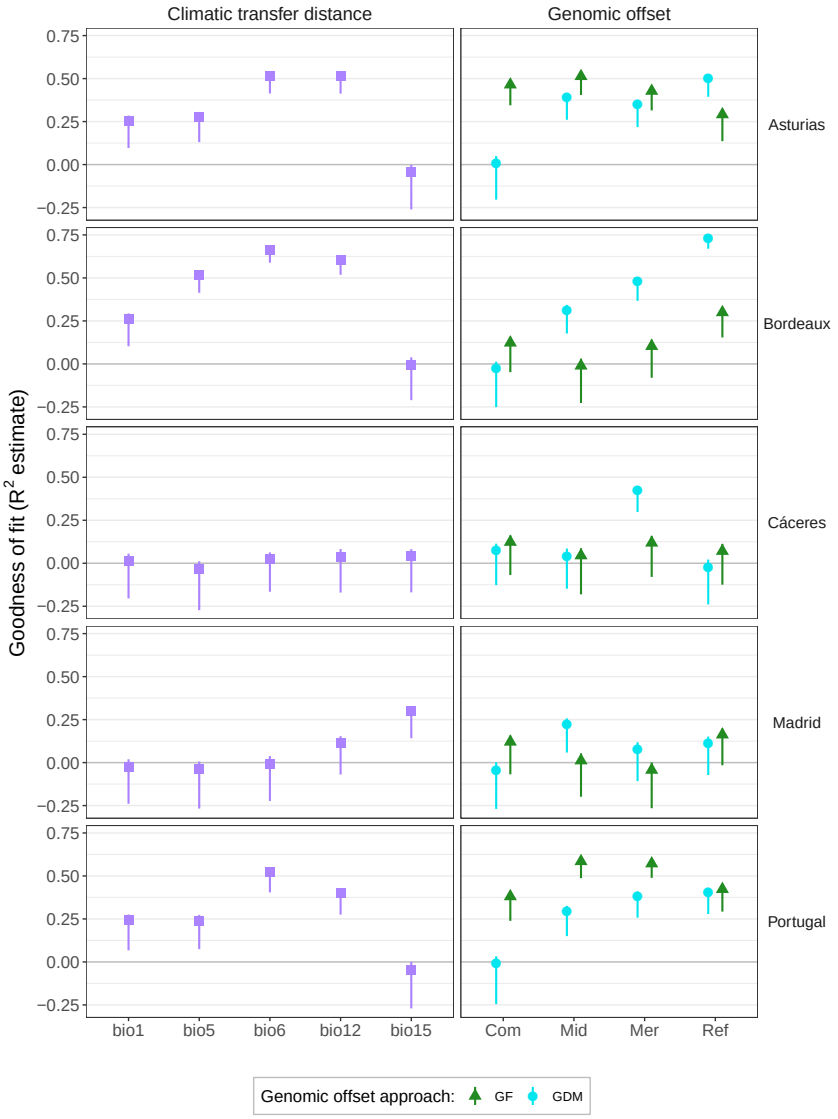


Figure VI.4. Proportion of variance explained ($\mathcal{R}^2$ estimate) of the models estimating the association between BLUPs for height in five common gardens and the climatic transfer distances (left panels) or the predicted genomic offset (right panels). The climatic transfer distances were calculated based on five climatic variables: the annual daily mean temperature (*bio1*; °C), the maximum temperature of the warmest month (*bio5*; C°), the minimum temperature of the coldest month (*bio6*; °C), the annual precipitation (*bio12*; mm) and the precipitation seasonality (*bio15*; %). Genomic offset predictions were obtained for each combination of two modelling approaches (GDM or GF) and four sets of SNPs, i.e. the common candidates (Com), the candidates under expected strong selection (Mid), the merged candidates (Mer) and the reference SNPs (Ref). Tree height in the common gardens was measured at 37-month old in Asturias, 85-month old in Bordeaux, 8-month old in Cáceres, 13-month old in Madrid and 27-month old in Portugal (see Figure XII.31 for Asturias at 10-month old, Bordeaux at 25-month old and Portugal at 11-month old).

in any common garden while those based on the other sets of SNPs explained between 25% and 50% of the height-BLUPs variance. Conversely, with the GF approach, more than 25% of the height-BLUPs variance was explained when using any of the three sets of candidate SNPs (including the common candidate SNPs) in Asturias, the reference SNPs in Bordeaux, and all SNP sets in Portugal (Fig. VI.4). Height differences in the Mediterranean commons gardens were poorly explained by the genomic offset predictions, except for the GDM-based predictions in Cáceres using the set of merged candidates, which explained about 40% of the height-BLUPs variance, and the GDM-based predictions in Madrid for the set of candidates under expected strong selection, which explained about 25% (Fig. VI.4). Last, in all common gardens, at least one of the genomic offset predictions showed similar, or better (e.g. in Cáceres), explanatory ability than the best performing climatic transfer distance (which were most often the minimum temperature of the coldest month and the annual precipitation; Figure VI.4). Model differences in predictive ability evaluated with the Bayesian leave-one-out cross validation supported the

model ranking obtained with the $\mathcal{R}^2$ estimates, and are therefore not shown here (see Tables XII.9 to XII.16).

With respect to the mortality models in the two Mediterranean common gardens, GF-based genomic offset predictions were generally better predictors than those based on GDM, and were as good predictors as the best performing climatic transfer distances (Tables XII.17-XII.18). In Cáceres, the strength of the association between mortality rates and GF-based genomic offset predictions decreased with the stringency in the selection of SNP sets, i.e. no association for common candidates, and maximum association for reference SNPs (Fig. XII.32). In contrast, GDM and GF-based genomic offset predictions in Madrid were better predictors of the mortality rates when based on the set of candidates under expected strong selection (Fig. XII.32).

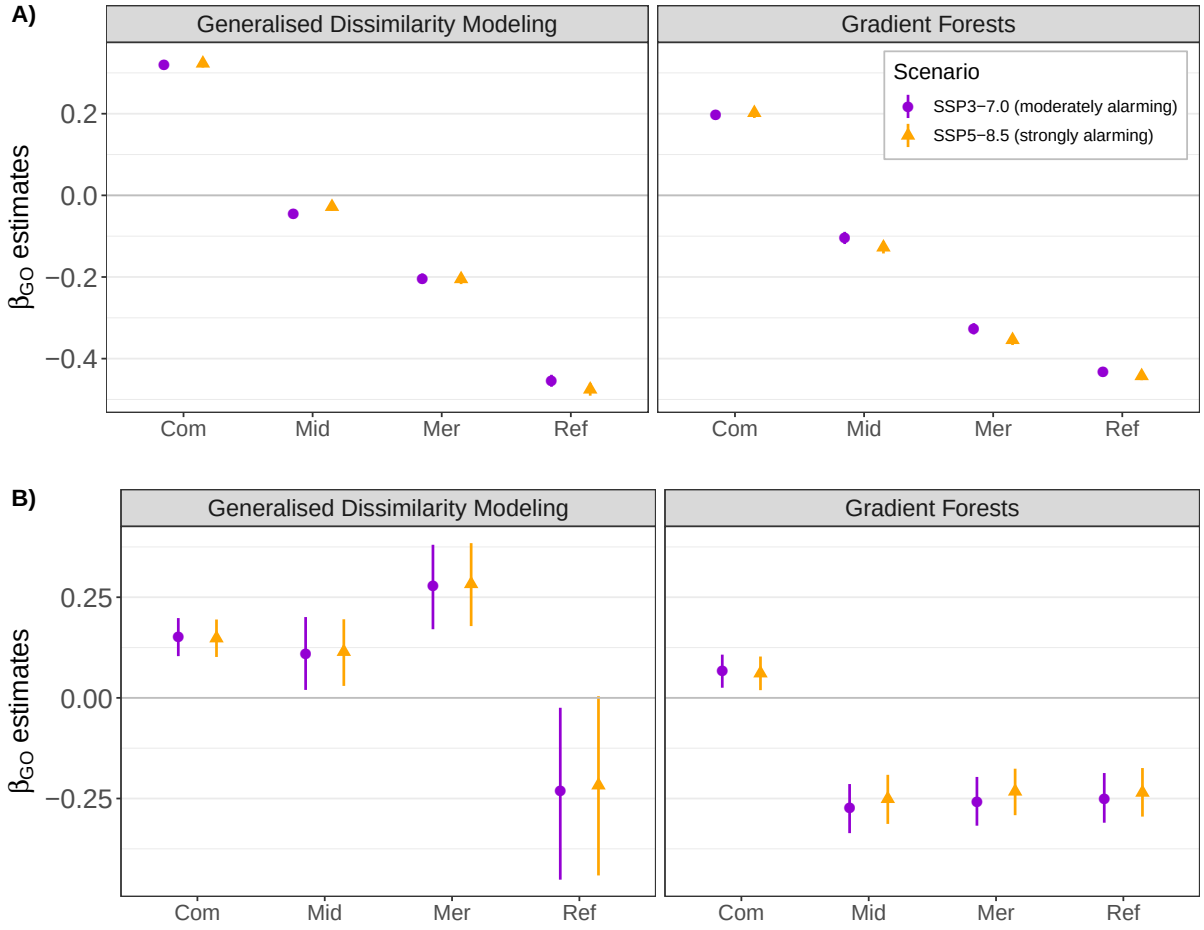


Figure VI.5. Regression coefficients describing the linear association between tree mortality rates in natural populations of maritime pine extracted from the Spanish (A) and French (B) National Forest Inventories and the genomic offset predictions under two scenarios of future climate (see equation 1.7 in the Supplementary Information).

Finally, when considering natural populations of maritime pine (National Forest Inventory data from France and Spain), we found a positive association between both GDM and GF-based genomic offset predictions and mortality rates for the set of common candidates (Fig. VI.5). This was also true for the sets of candidates under expected strong selection and merged candidates in the GDM approach in France (Fig. VI.5B). However, in all other cases, the association between genomic offset predictions and mortality rates was negative, suggesting lower mortality rates in populations subject to higher genomic mismatch between current and future climates (Fig. VI.5). The strength of the relationship between genomic offset predictions and mortality rates was very similar between the two scenarios of future climate (Fig. VI.5).

5 Discussion

Since Fitzpatrick and Keller (2015) and Rellstab et al. (2015), the genomic offset approach has become popular for identifying populations at risk of short-term climate maladaptation. It appears as a promising method to guide conservation and management strategies, in particular for long-lived and sessile species such as forest trees, for which *in situ* adaptation or migration of adaptive alleles may not be fast enough to track climate change (Fitzpatrick and Keller 2015). However, this approach relies on a number of key assumptions in need of robust empirical validation (discussed in Rellstab et al. 2021 and Capblancq et al. 2020b). The main objective of this study was twofold: (1) identify maritime populations at risk of short-term climate maladaptation, (2) verify the assumption that populations with the highest predicted genomic offset do show a decrease in absolute fitness or declining demographic trends (Capblancq et al. 2020b). About (1), genomic offset predictions based on the set of common candidate SNPs (i.e. the most confidently selected) point towards a higher maladaptation risk for populations currently experiencing mild-winter conditions (i.e. most Atlantic populations and populations in southeastern France and northwestern Italy), which may be related to the past cold adaptation of these populations preventing optimal growth and survival under warming temperatures. About (2), we showed that the genomic offset predictions were generally negatively associated with population performance in common gardens and natural populations (i.e. height or mortality rates). This suggests that genomic offset predictions may be indicative of (future) fitness declines, and thereby validate our results in maritime pine. Nevertheless, genomic offset predictions were highly sensitive to the set of SNPs considered (i.e. how strictly they were selected) and the modelling approach used (i.e. GDM or GF), while they were very similar under two different future climate scenarios. Noticeably, none of the different ways tested to predict genomic offset had better predictive ability across all the validation steps. Our results therefore confirm that the genomic offset approach is promising, but also suggest that further validation, based on both experimental and observational independent data, is needed to determine which modelling methods and ways of selecting the adaptive genetic component, are the most robust and accurate.

5.1 Past cold adaptation may trigger short-term adaptation mismatch in mild-winter regions

Our results highlighted the key role of cold temperatures in the past adaptive history of maritime pine populations, as previously suggested by studies evaluating the association between candidate-gene allele frequency and temperature gradients (Grivet et al. 2011, Jaramillo-Correa et al. 2015). Interestingly, the selection pressures resulting from severe cold events might even reduce genetic variation within maritime pine populations for some fitness-related phenotypic traits (e.g. for height; Archambeau et al. 2021b). In this line, we found that the minimum temperature of the coldest month was the environmental covariate associated with the largest number of SNPs in the univariate GEA (Table XII.4) and the most important in explaining allele frequency turnover in the GF and GDM models (once accounting for geographic distance; Table XII.7 & Figs. VI.2 & XII.7-XII.10), in particular for the common candidate SNPs. We base our interpretation on the genomic offset predictions for the set of common candidate SNPs, as they were selected most stringently (and thus with the most confidence) and were the only SNP set to show consistent spatial patterns of genomic offset between GF- and GDM-based predictions (Fig. VI.3). Populations with the highest predicted genomic offset were those at the warm edge of the cold temperature gradient: mainly Atlantic populations in

the northwestern part of the Iberian Peninsula, southwestern France and Brittany, but also Mediterranean populations of Corsica, southeastern France and northwestern Italy (Figs. VI.3 & XII.3). Importantly, the warm edge of the cold temperature gradient (i.e. between 4 and 6°C) is characterized by a strong turnover in the frequency of common candidate SNPs (Figs. VI.2 and XII.7), which may be indicative of a loss of cold hardiness alleles in populations from warmer regions. Consequently, in these populations, a slight increase in minimum winter temperatures due to climate change will result in a large disruption of current gene-environment relationships. Whether this mismatch will translate into a maladaptation risk associated with declines in absolute fitness remains to be confirmed, but there are good reasons to think so.

Based on height growth data from common gardens, Fréjaville et al. (2020) predicted that some maritime pine populations in southwestern France, northwestern Iberia, Corsica, and coastal areas of southeastern France and northwestern Italy will grow less under future temperatures and precipitation. These spatial patterns of potential adaptation lags in terms of growth are very similar to those obtained in our study with the genomic offset approach. This may be partly explained by the trade-off between growth and cold hardiness in forest trees (Aitken and Hannerz 2001, Howe et al. 2003 and Prada et al. 2016 in maritime pine), which may constrain cold-hardy tree populations from growing optimally in warming climates, at least in the short-term. The higher susceptibility of populations from mild-winter regions to increasing winter minimum temperatures may also be explained by reduced winter hardening of the trees, and thus greater exposure to late frosts (e.g. Hänninen 2006). Last, the genomic offset predictions obtained with the common candidate SNPs were positively associated with mortality rates in natural forest inventory populations calculated over the last four decades, which may suggest that populations that will be at risk of climate maladaptation are already experiencing fitness declines (although this conclusion has to be tempered, as discussed in the next section).

In conclusion, the populations most at risk of short-term climate maladaptation might be those already living in relatively mild-winter conditions but for which the transition to slightly higher temperatures would entail a large evolutionary step. Importantly, these populations, being at the warm margin of the cold winter temperature gradient, will not benefit from the migration of adaptive alleles from other populations. It therefore seems crucial to monitor their demographic dynamics and adaptive trajectories in the coming years, especially knowing that southwestern France and northeastern Iberia populations are the ones with the highest commercial value.

5.2 Associations between genomic offset predictions and fitness declines

A key assumption of the genomic offset approach is that its predictions have to be correlated with a decrease in absolute fitness (Capblancq et al. 2020b, Rellstab et al. 2021). This assumption was first successfully demonstrated in experimental conditions by Fitzpatrick et al. (2021), in which the authors used the GF approach to predict genomic offsets resulting from transplanting populations of balsam poplar from their home environment to a common garden environment (more recently, see a similar validation analysis in Capblancq and Forester 2021). They showed that genomic offset predictions were negatively associated with height growth in the common gardens and were better predictors of height growth variation than climate transfer distances. Interestingly, Borrell et al. (2020) found a negative association between the estimated risk of nonadaptedness (a similar measure to genomic offset; Rellstab et al. 2015) and catkin production in dwarf birch, thus suggesting that the reproduction-related component of fitness may also

be captured by the genomic offset approach. In our study, we showed that the genomic offset predictions were in most cases negatively associated with population performance: negatively associated with height in the Atlantic common gardens (but almost no association in the Mediterranean common gardens; Fig. XII.30) and positively associated with mortality for the GF approach in the Mediterranean common gardens (Fig. XII.32). However, we did not find that genomic offset predictions outperformed predictions based on climatic transfer distances, as in each common garden at least one climatic transfer distance (of the five tested) predicted population performance as well as the best genomic offset prediction (Figs. XII.30 & XII.32). A potential explanation of the discrepancy between Fitzpatrick et al. (2021) and our study is that we evaluated the association between population performance and the climatic transfer distance for each climatic variable independently, whereas Fitzpatrick et al. (2021) calculated a climatic transfer distance based on a PCA of the climatic variables (thus merging in a single index climatic variables more or less associated with adaptation gradients). The good performance of GF-based genomic offset predictions to predict mortality rates in the Mediterranean common gardens was particularly encouraging as genetic differentiation in mortality patterns is not a general pattern in forest tree common gardens (e.g. Oddou-Muratorio et al. 2011, Vizcaíno-Palomar et al. 2014). Moreover, genomic offset predicted mortality rates as well as a climatic variable related to severe cold events, which are known to be an important driver of adaptation in maritime pine (Archambeau et al. 2021b, Grivet et al. 2011, Jaramillo-Correa et al. 2015).

The assumption that populations with higher genomic offset show decreased absolute fitness was also evaluated in natural populations of a North American migratory bird by Bay et al. (2018), in which the authors found a negative relationship between GF-based genomic offset predictions under future climates and spatial extrapolations of historical population trends. The authors therefore suggested that some populations may already suffer from reduced fitness due to an inability to adapt. However, this study was criticized because (Fitzpatrick et al. 2018): (1) historical and future climate shifts are assumed to be correlated whereas it is likely that this is not the case; (2) the relationship between historical climate and population trends was not evaluated, and (3) the genomic offset predictions were based on all SNPs with $\mathcal{R}^2 > 0$ in the GF analysis (~8,000 SNPs), which therefore probably represented mainly neutral genetic variation. In our study, for all tested combinations (i.e. combination of the two modelling approaches, the two future climate scenarios and the two countries in which the NFI plots were located, Spain and France), genomic offset predictions for the set of reference SNPs were in the opposite direction to that expected, i.e. higher genomic offset in plots experiencing lower mortality rates, probably reflecting also neutral genetic variation (Fig. VI.5). In stark contrast, as mentioned in the previous section, genomic offset predictions for the set of common candidates (i.e. the most strictly selected; Fig. VI.5) were positively correlated with mortality rates, which may indicate that populations in northwestern Iberian Peninsula, southwestern France, French Brittany and northwestern Italy are already experiencing higher mortality rates as a result of failure to adapt to climate change. However, this result has to be taken with caution for several reasons. First, it can be criticized by some of the same arguments, (1) and (2), made against Bay et al. (2018) (see above). Second, high mortality rates are not necessarily associated with fitness declines as, if they are compensated by high recruitment rates, they can accelerate evolutionary processes by speeding up the generation turnover, allowing populations to escape demographic collapse (Kuparinen et al. 2010). Last, NFI data (like observational data in general) remains highly noisy; although we corrected for some potential confounding factors (competition among trees, different sampling schemes across countries), we cannot assert that the observed mortality patterns are only attributed to climate (e.g. dieback events caused by storms but not recorded as such).

All in all, the different validation steps gave diverging results regarding which combination of SNP sets and modelling approaches best capture true patterns of decreasing fitness. A potential explanation may stem from the different fitness proxies used in the validation parts: height and mortality rates of seedlings in controlled environments vs mortality rates of saplings and adult trees in natural conditions. Indeed, these fitness proxies may be under different selection pressures, thus shaping differently their adaptive genetic component and thereby their gene-environment relationships. Estimating fitness remains highly challenging in forest trees, which have high reproductive outputs, and long lifetime and generation times. For this reason, assessing genomic offset predictions on the basis of different fitness proxies is particularly relevant, as they may convey a different picture of the spatial patterns of adaptive genetic variation. The different conclusions drawn by the two validation approaches may also originate from the inclusion of the topographic and soil-related variables in the genomic offset estimation in the common gardens (i.e. reflecting a wider environmental maladaptation) while only climatic variables were used to estimate the genomic offset at the NFI plots.

5.3 Limitations and promises of the genomic offset approach

A key assumption of the genomic offset approach that we did not directly tackle here is that the set of genomic markers used for genomic offset predictions does capture the adaptive genetic component required to adapt to future climates (Capblancq et al. 2020b). In our study, the adaptive genetic component was identified with two GEA methods, which are inherently correlative. Confirming that the selected candidate SNPs are involved in climate adaptation would therefore require further validation (Ioannidis et al. 2009, Oetting et al. 2017). For instance, this can be done by repeating the GEA analyses on independent samples (Bay et al. 2018) or along parallel climate gradients (van Boheemen and Hodgins 2020), comparing genotype fitness with or without the putative adaptive alleles under experimental conditions that control for non-genetic effects (e.g. in reciprocal transplant experiments or common gardens; Exposito-Alonso et al. 2019), or using functional validation experiments (e.g. through gene expression knockdown; Rohde et al. 2018). However, adaptation to climate probably involves multiple polygenic traits whose genetic component is determined by complex gene networks (Boyle et al. 2017, Liu et al. 2019). Thus, similar adaptive phenotypic changes may result from different combinations of alleles (i.e. 'genetic redundancy'; Barghi et al. 2020). Moreover, single alleles may be associated with multiple phenotypes (i.e. pleiotropy) or could be adaptive only in particular environments (i.e. conditionally neutral). These factors render the identification of the genetic component of climate adaptation particularly tricky, especially in conifers with their huge genome size (Mackay et al. 2012; 28 Gbp for maritime pine, Zonneveld 2012) and complexity (e.g. large number of transposable elements; De La Torre et al. 2014).

Our results show that genomic offset predictions are highly sensitive to the way candidate SNPs are selected, i.e. how strict is the selection process (Fig. VI.3). The estimated gene-environment relationships, the projections of the genomic composition and the genomic offset predictions were the most different between the common candidate and reference SNP sets, as similarly observed in Fitzpatrick et al. (2021). Importantly, results obtained for the sets of merged candidates and candidates under expected strong selection remained highly different from those obtained for the set of common candidates. This implies that the selection process of candidate SNPs is crucial in the genomic offset approach and further research is undoubtedly needed to determine which method best captures the genetic component of climate adaptation and thus produces the most robust genomic offset predictions.

Finally, a source of bias important to consider at each step of the genomic offset approach is the population structure, which is particularly strong in maritime pine (Jaramillo-Correa et al. 2015), and partially covaries with environmental gradients (Archambeau et al. 2021a). First, population structure remains challenging to account for in GEA analyses and can result in high rates of false positives, i.e. alleles mistakenly identified as adaptive or linked to an adaptive allele (Hoban et al. 2016, Hoffmann et al. 2021). We tackled this issue by using two GEA methods that account for population structure in very contrasted ways: by partialling out the ancestry coefficients in the RDA (following Forester et al. 2018) or with a covariance matrix based on the population allele frequencies in BayPass (Gautier 2015). Second, to date, methods to calculate the genomic offset incorporate population structure coarsely by relying only on the geographical location of the populations: geographical distance among populations in the GDM models and Moran's eigenvector map (MEM) variables in the GF models (Fitzpatrick and Keller 2015). However, population structure might show more complex patterns than those resulting from isolation-by-distance, in particular in species with complex demographic histories or in fragmented landscapes (Rellstab et al. 2021). Interestingly, Gain and François (2021) recently proposed a new approach to calculate the genomic offset that directly accounts for the neutral population genetic structure and, importantly, does not rely on the pre-selection of candidate SNPs, which seems particularly relevant given the biases that may derive from this step (see above). Third, population structure can vary in space and time, which warns against extrapolating too widely across the landscape and projecting too far into the future (Rellstab et al. 2021), although for species with long generation times such as forest trees, population structure is unlikely to change significantly over the next few decades. Last, population structure may have also influenced the validation step in the common gardens: BLUPs were estimated without accounting for population structure, and therefore capture height or mortality differences among populations that both arise from adaptive and neutral genetic processes (e.g. demographic history, gene flow).

6 Conclusion

Our study adds to the accumulating evidence on the power of the genomic offset method to predict short-term climate maladaptation. In maritime pine, populations in regions with increasingly mild-winter conditions may be the most likely to experience adaptation lags and should therefore be kept under close scrutiny, e.g. by monitoring the joint dynamics of mortality and recruitment. Such monitoring may be particularly crucial for populations in northwestern Iberia and southwestern France, which have a major economic importance; climate maladaptation affecting their phenotypic traits of interest might have a substantial impact on the local economy. Importantly, the validation steps highlighted the need of combining different genomic offset estimation and validation approaches to be confident about the robustness of genomic offset predictions. It might also be worth remembering that the genomic offset approach is based on static gene-environment associations, assuming that current allele frequencies reflect the adaptive optimum of the populations (which is often not true in forest trees based on phenotype-environment associations; e.g. Fréjaville et al. 2020, Pedlar and McKenney 2017, Rehfeldt et al. 2018, 2003, 2002, Savolainen et al. 2007). Importantly, incorporating processes such as gene flow among populations and selection is necessary to evaluate whether a population will be able to adapt in the long run in face of changing climatic conditions (Waldvogel et al. 2020). Recent studies have investigated ways of integrating migration processes in the genomic offset approach (e.g. Gougherty et al. 2020b), while integrating selection processes remains largely unexplored, likely due to the high complexity and multifactorial nature of the

selection response. For instance, pleiotropy can either boost or slow down adaptive processes depending on whether genetic correlations go in the same or opposite direction as selection, respectively (Hoffmann et al. 2021). Nevertheless, although genomic offset predictions have to be interpreted with caution, they remain a major step towards integrating adaptive processes into management and conservation strategies (Waldvogel et al. 2020).

7 Author contributions

SCG-M designed the experiment and supervised the curation of field data. MdM cleaned and formatted the phenotypic data. FB, SCG-M and GGV did the DNA extractions, cleaning and checking of the genomic data. AC cleaned and formatted mortality data from the National Forest Inventories. SCG-M, JA and MBG conceived the paper methodology. JA and FB built the Bayesian model equations and codes of the validation part. JA conducted the data analyses. All authors interpreted the results. JA led the writing of the manuscript. All authors contributed to the manuscript and gave final approval for publication.

8 Acknowledgments

We thank A. Saldaña, F. del Caño, E. Ballesteros and D. Barba (INIA) and the 'Unité Expérimentale Forêt Pierroton' (UEFP, INRAE; doi:10.15454/1.5483264699193726E12) for field assistance (plantation and measurements). Data used in this research are part of the Spanish Network of Genetic Trials (GENFORED, http://www.genfored.es). We thank all persons and institutions linked to the establishment and maintenance of field trials used in this study. We are very grateful to Ricardo Alía and Juan Majada who initiated and supervised the establishment of the CLONAPIN network (i.e. the five common gardens used in the validation part). We thank Paloma Ruiz-Benito and Sylvain Schmitt for their constructive comments, and Charlie Pauvert for help on database extraction. JA was funded by the University of Bordeaux (ministerial grant). The authors also thank the MITECO ('Ministerio para la Transición Ecológica y Reto Demográfico'), MAPA ('Ministerio de Agricultura, Pesca y Alimentación'), and the French Forest Inventory (IGN) for making NFI data from Spain and France available. This study was funded by the European Union's Horizon 2020 research and innovation programme under grant agreements No 773383 (B4EST) and No 862221 (FORGENIUS).

VII

Synthesis & discussion

Forest trees are keystone species that are essential to ecosystem functioning, maintaining biodiversity and sustaining multiple ecosystem services. Predicting how forest tree populations will adapt *in situ* to future environmental conditions, especially those caused by climate change, is becoming a critical and urgent issue, which requires a deep understanding of the evolutionary processes at stake. Moreover, some populations will not be able to adapt fast enough to keep pace with climate change and will need to be accurately identified and prioritized to implement relevant conservation and management strategies. In strategies accounting for adaptive processes, the successful moving of individuals into environments in which they will be assumed to be more adapted (i.e. assisted gene flow) or into threatened populations in need of additional genetic variation (i.e. evolutionary rescue) will require predicting the response of transplanted individuals in the new environments, e.g. the mean absolute fitness of the transplanted sample. Importantly, the increasing availability of genomic data for non-model species provides new opportunities to understand adaptive processes, identify populations at risk of maladaptation or improve individual-level predictions of quantitative traits. Genomics-based predictive modelling approaches are developing rapidly, but most are still lacking robust validation against independent data and have considerable room for improvement.

Here I used maritime pine as a case study to investigate key questions in the field of evolutionary biology through the use of innovative modelling approaches combining phenotypic, environmental and genomic data: do populations show differentiation patterns along environmental gradients in terms of phenotypic and genomic variation across the species range? If yes, can the phenotypic and genomic differentiation be attributable to adaptive processes, and more specifically climate-driven adaptation? How does adaptation to harsh climatic conditions or heterogeneous landscapes impact the maintenance of within-population quantitative genetic variation, and therefore indirectly the adaptive potential of the populations? Answering these questions has provided a better understanding of how past adaptations have shaped current maritime pine populations and has offered a detailed picture of the current adaptive state of the populations, a necessary first step in predicting their future states.

The field of evolutionary biology is indeed currently experiencing a growing interest in predicting the future evolutionary trajectories of populations. A consequential part of the present PhD work was therefore dedicated to investigate, through quantitative genetics and landscape genomics methods, the adaptive potential and the risk of short-term climate maladaptation of maritime pine populations across the species range.

Finally, the present PhD work provides an original attempt to combine phenotype, environmental and genomic data using various modelling approaches in a non-model species. To what extent the conclusions drawn from maritime pine, a Mediterranean forest tree species with highly differentiated and fragmented populations, may be extended to other forest trees would require further investigation, but the approaches used here are undoubtedly applicable to other forest trees. Ultimately, by advancing knowledge of the evolutionary mechanisms underlying adaptive responses of populations to changing conditions and providing key population-level metrics related to population adaptive capacity, the results of this PhD work add building blocks to the development of more mechanistic modelling approaches aimed at predicting the evolutionary response of populations on medium- to long-term time scales.

1 What have we learned in maritime pine?

'Without deep biological understanding of the system under study, predictive models are not likely to offer much insight into either the past or future.' (Reznick and Travis 2018).

1.1 High quantitative and molecular genetic differentiation

Due to its high economical importance, maritime pine has long been planted in provenance trials. Strong genetic differentiation has commonly been found among populations for a wide variety of phenotypic traits, such as traits related to growth, survival, wood quality, tree form, drought resistance or reproduction (Alía et al. 2014, 1997, Chambel et al. 2007, Correia et al. 2010, de la Mata et al. 2012, Gaspar et al. 2013, Guyon and Kremer 1982, Lamy et al. 2011, Santos del Blanco et al. 2010, Santos-del-Blanco et al. 2012, Sierra-de-Grado et al. 2008). Mediterranean populations generally recover better after drought and have a more conservative strategy than Atlantic populations, e.g. they have more stable yield across different environments (Alía et al. 1997). In contrast, Atlantic populations grow well under favorable conditions, show larger size at first reproduction (Santos-del-Blanco et al. 2012) but often recover poorly after drought, thus being more prone to drought-related mortality (Alía et al. 1997, Zas et al. 2020). This strong interactive behavior of Atlantic populations may explain why most trait-based studies found significant levels of genotype-by-environment interaction in maritime pine (e.g. Alía et al. 1997, Corcuera et al. 2010, Correia et al. 2010, de la Mata et al. 2012). As counter-examples, Chambel et al. (2007) did not find phenotypic differentiation among populations under two different watering regimes and Santos del Blanco et al. (2010) reported weak genotype-by-environment both at the family and population level for traits related to reproduction strategies (i.e. threshold size for reproduction and reproductive allocation).

High levels of genetic differentiation among maritime pine populations have also been repeatedly found with molecular data, from different types of markers, e.g. polymorphic allozyme loci, micro-satellites and SNPs, and different genomes, i.e. mitochondrial, chloroplast and nuclear (Bucci et al. 2007, Burban and Petit 2003, González-Martínez et al. 2002, Jaramillo-Correa et al. 2015, Santos-del-Blanco et al. 2012, Serra-Varela et al. 2015). Because of this strong quantitative and molecular genetic differentiation, some authors argued to define two subspecies of maritime pine (splitting the Atlantic and Mediterranean populations) and five varieties based on the geographical location: Algeria and Tunisia, Morocco, Corsica, the Mediterranean region and the Atlantic region (Barbéro et al. 1998).

In the present PhD work, spatial patterns of genetic differentiation were investigated through both trait-based approaches (CHAPTER 2) and landscape genomics approaches (CHAPTER 3). In CHAPTER 2, clear differences in height growth among populations were identified, with populations belonging to the French and Iberian Atlantic gene pools growing faster on average than populations from the central Spain gene pool (Fig. VII.1). The genetic component of height growth was strongly associated with the population structure resulting from demographic history and gene flow among populations. This was expected as population structure is known to be particularly strong in maritime pine (Alberto et al. 2013, Jaramillo-Correa et al. 2015), which may be explained by its fragmented distribution and the historical isolation of different glacial refugia (Bucci et al. 2007, Burban and Petit 2003, Naydenov et al. 2014, Serra-Varela et al. 2015). Interestingly, most gene pools involved population pairs with contrasted mean height growth (Fig. VII.1), thus suggesting that height growth variation among populations does not result only from neutral evolutionary processes, but also from adaptive processes. In the case of the northern African gene pool, other possible explanations include the lack of survivors from the Madisouka (MAD) population in the Cáceres common garden, which may bias the height growth estimation for that population, or that the highest growth of this population compared to the Tamrabta (TAM) population originates from its ancestry proportion (23.3%) from the south-eastern Spanish gene pool (Jaramillo-Correa et al. 2015). A last noteworthy contribution of CHAPTER 2 to describing patterns of genetic differentiation in maritime pine is the identification of an association between height growth and provenance-by-site interaction, although its proportion of explained variance was weak. This may suggest that most genotype-by-environment interaction is within populations or that genotype-by-environment does not explain an important part of height-growth variation in maritime pine, most genetic differences among populations being stable across varying environments.

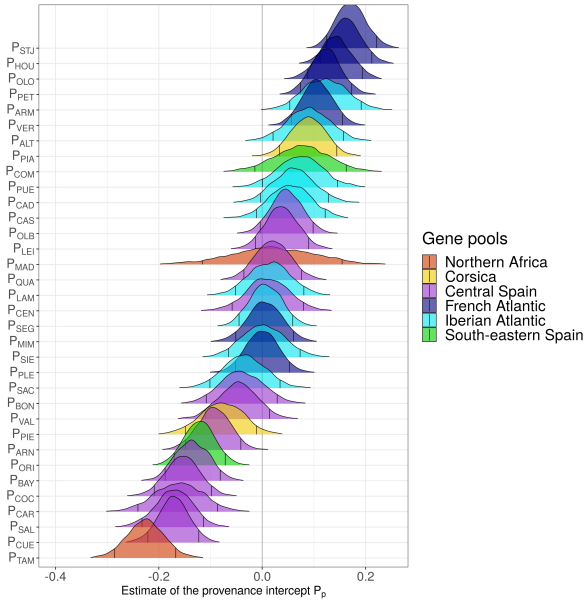


Figure VII.1. Population (i.e. provenance) varying intercepts in the height-growth model *M1* from CHAPTER 2, in which only the population and clone (i.e. genotype) intercepts were included in the model to account for the genetic component of height-growth variation.

Part of CHAPTER 3 was dedicated to predict spatial variation in genomic composition across the range of maritime pine based on a set of reference SNPs. Results from this chapter confirm that covariates related to the geographical distance among populations explain most of the turnover in allelic frequencies, as expected for a species with a strong population structure such as maritime pine. Therefore the projected spatial variation in genomic composition when both environmental and geographical covariates are considered was almost entirely driven by

the geographical distance among populations (Fig. VII.2A). Interestingly, the most important environmental covariates to explain the turnover in reference SNPs frequency were not the same in the GDM and GF approaches, thus leading to very contrasted projections of the genomic composition when only environmental covariates were considered (Figs. VII.2B and VII.2C). This is not surprising under the assumption that the turnover in reference SNPs frequency results primarily from neutral evolutionary processes and therefore the gene-environment associations remain very marginal and hard to estimate. Nevertheless, we can note that, once the effect of the geographical distance among populations was removed, GDM-based projections identified a specific genomic composition in the mountainous areas (mainly explained by the topographic ruggedness) and in central Spain (mainly explained by the annual precipitation) while GF-based projections identified a specific genomic composition in Portugal (mainly explained by the fire intensity).

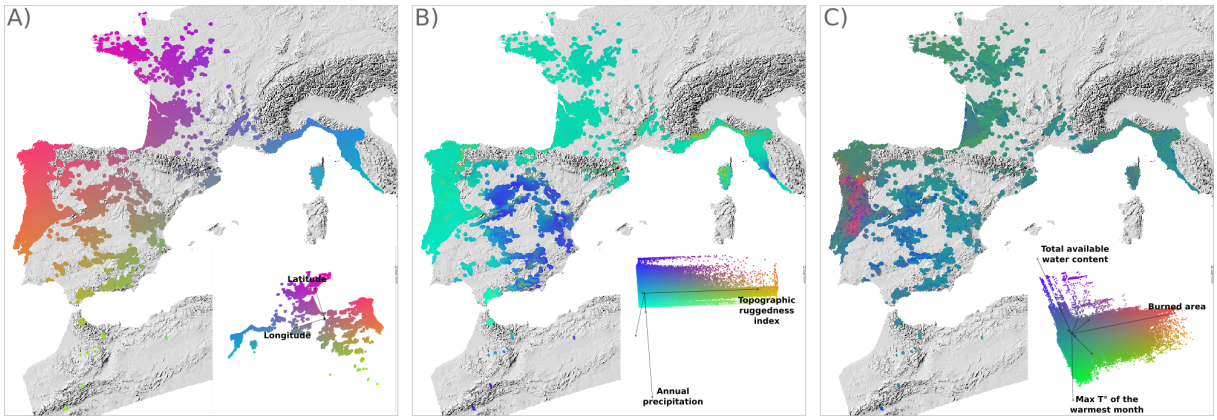


Figure VII.2. Predicted spatial variation in genomic composition under current climates for the set of reference SNPs. A) Projection based on the GDM modelling approach and using both environmental and geographical covariates. B) Projection based on the GDM modelling approach and using only the environmental covariates. C) Projection based on the GF modelling approach and using only the environmental covariates. Similar colors correspond to expected similar genomic composition. In the bottom right corner of each figure are the associated principal coordinate analyses, which inform on the contribution of the geographical and environmental covariates to the predicted variation in genomic composition, with the arrows indicating the direction and magnitude of each covariate. Only labels of the most important covariates to explain the turnover in allele frequency are shown. Figures from CHAPTER 3.

1.2 Multiple evidence of climate adaptation in maritime pine

Adaptive signatures have been repeatedly observed in maritime pine, although it remains challenging to distinguish the relative importance of adaptive and neutral evolutionary processes in the genetic differentiation observed among populations. Among studies specifically aiming at detecting footprints of natural selection on phenotypic traits, González-Martínez et al. (2002) and de Miguel et al. (2020) showed that quantitative (Q_{ST}) differentiation was higher than allozyme or SNPs (F_{ST}) differentiation for a variety of traits (e.g. stem form, total height growth, survival, phenology-related traits, functional traits), thus suggesting the action of natural selection. In addition, the negative correlation between SNP effect-size and minor allele frequency in de Miguel et al. (2020) for tree height, bud burst and SLA may be an indicator of the action of negative selection, i.e. the purging of deleterious alleles (O'Connor et al. 2019). As an example of phenotype-genotype association studies, Budde et al. (2014) found 18 candidate SNPs associated with resistance to fire and explaining ~ 29% of a fire-related trait in the eastern Iberian Peninsula. Serra-Varela et al. (2015) highlighted the likely major contribution of environmental adaptation in the observed patterns of genetic differentiation across the range of maritime pine by showing that similar groups of populations are obtained based on either genetics or environment, and that environment explained a greater proportion of the variation in phylogeographic distance than geography. In gene-environment association

studies, Jaramillo-Correa et al. (2015) identified 18 candidate SNPs associated with climate, whose adaptive patterns varied between the Iberian Mediterranean and Atlantic regions, suggesting contrasted selection pressures across gene pools. Noticeably, the frequency of the 18 candidate SNPs was correlated with survival in a dry and hot environment, thus providing an experimental validation of the set of candidate SNPs. Grivet et al. (2011) identified two loci associated with climate, one of which was also detected in another pine (*Pinus halepensis*). Interestingly, both Jaramillo-Correa et al. (2015) and Grivet et al. (2011) found that most gene-environment associations were related to temperature indices, thus supporting the major role of temperature in the adaptive history of maritime pine.

In line with the above mentioned work, results from CHAPTER 2 showed that the genetic component of height growth in maritime pine was associated with the climatic similarity among populations, which underline the potential major role of adaptation to climate for this trait. However, the population genetic structure (included by accounting for the gene pool assignment of each clone) had a confounded association with climatic variables, so that separating their relative importance to explain the genetic component of height growth was not completely possible. Such confounded association between population structure and adaptation patterns is common in forest trees (Alberto et al. 2013, Latta 2009, Nadeau et al. 2016). However, it may also partly stem from the gene pool assignment that reflects both adaptive (different selective histories) and neutral processes (population demographic history and genetic drift), despite gene pools being identified using genetic markers considered neutral (Jaramillo-Correa et al. 2015). This was notably supported by the different heritabilities estimated between some gene pools in CHAPTER 2 (using a method that accounts for admixture among gene pools; Muff et al. 2019), i.e. the Corsican gene pool, and to a lesser extent the south-eastern Spain gene pool, had higher heritabilities than the French and Iberian Atlantic gene pools.

Further, CHAPTER 3 contributed significantly to describing and understanding patterns of climate (mal)adaptation in maritime pine, using in combination two GEAs (BAYPASS and RDA) and two recent landscape genomics approaches (GDM and GF). Different sets of candidate SNPs were defined based on the stringency of the selection threshold above which they can be considered as potential candidates for environment adaptation: the common candidates (i.e. 8 candidate SNPs identified by the two GEA methods), the candidates under expected strong selection (79 candidate SNPs with a strong association with at least one covariate) and the merged candidates (370 candidate SNPs identified by at least one GEA method). Then, we compared the relative importance of environmental covariates in explaining the turnover in allele frequencies across the range of maritime pine for each SNP set and landscape genomics approach. A first key finding was the consistent influence of minimum temperatures on the adaptive genetic component: minimum temperature of the coldest month was involved in 25 of the 26 gene-environment associations identified in BAYPASS and was generally the most important environmental covariate in explaining the turnover in allele frequencies in the GDM and GF models (being even a better predictor than covariates related to population structure for the set of common candidates in the GF models; Fig. VII.3A & B). This result is well in line with previous studies discussed above that also supported a strong influence of temperatures (and especially cold temperatures) in the selective history of maritime pine (Grivet et al. 2011, Jaramillo-Correa et al. 2015). A second key message of CHAPTER 3 concerns the marked differences among the different sets of candidate SNPs in the relative importance of environmental covariates in explaining allelic frequency turnover. This is of concern because most current studies using the genomic offset approach do not first validate the adaptive genetic component (sometimes even using all SNPs directly, which is certainly biased by population structure; Fitzpatrick et al. 2018) and therefore may identify different maladapted populations

depending on the initial set of candidate SNPs.

Further, CHAPTER 3 goes one step further than the work of Jaramillo-Correa et al. (2015) by using recent landscape genomics approaches to project the predicted genomic composition of maritime pine populations across the species range. These projections were based on the gene-environment relationships after accounting for the relationship between allele frequencies and the geographical distance among populations (a proxy of the population structure), and they may thus capture the adaptive genetic component. For example, the genomic composition was relatively uniform in France, whereas it showed contrasted patterns in Galicia and Portugal for the set of common candidate SNPs (Fig. VII.3C). Nevertheless, considering the geographical distance among populations is likely insufficient to account for population structure, which might show more complex patterns than those resulting from isolation-by-distance, in particular in species with complex demographic histories or in fragmented landscapes (Rellstab et al. 2021). Further method development is undoubtedly needed to better account for the confounded effect of population structure, e.g. Gain and François (2021).

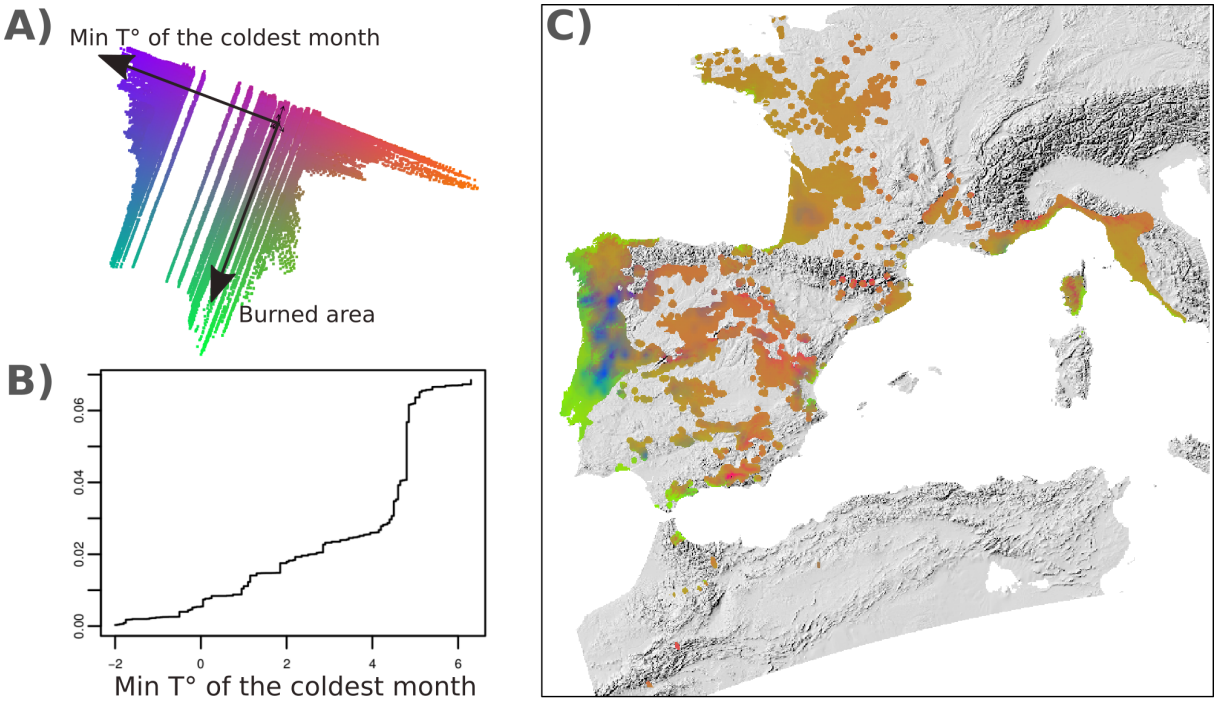


Figure VII.3. Figures from CHAPTER 3 generated based on the Gradient Forest estimations of the gene-environment relationships for the set of common candidate SNPs (i.e. SNPs identified by both BAYPASS and RDA). A) Principal coordinate analysis of the predicted variation in genomic composition across the maritime pine range. Only labels of the two most important environmental covariates explaining the turnover in allele frequency are shown: the minimum temperature of the coldest month (C°) and the burned area (hectares). B) Turnover function of the minimum temperature of the coldest month (C°). C) Predicted spatial variation in genomic composition under current climates.

In CHAPTER 1, we showed that not only do populations vary in their mean genetic values and allelic frequencies along environmental gradients, but they also exhibit different levels of quantitative genetic variation for height in three different common gardens (e.g. see Fig. VII.4 for the common garden in Bordeaux when the trees were 25 month-old). Remarkably, our results again confirmed the strong influence of extreme minimum temperatures on adaptation gradients in maritime pine, as we showed that genetic variation in height was lower in populations subjected to severe cold events. This supports the hypothesis that quantitative genetic variation in fitness-related traits is lower in populations under strong selection. Estimating within-population quantitative genetic variation is particularly valuable in describing adaptation patterns because, in theory, low genetic variation may indicate a high degree of adaptation

to harsh environmental conditions, at the expense of the adaptive potential under changing environmental conditions.

Another key finding of CHAPTER 1 was that, contrary to our expectations, quantitative genetic variation of all traits studied (i.e. growth, phenological and functional traits) was not associated with environmental heterogeneity, which goes against the predictions of some theoretical models (McDonald and Yeaman 2018, Walsh and Lynch 2018) and an empirical study in lodgepole pine (Yeaman and Jarvis 2006). Last, we did not find any association between within-population genetic variation for height (and other traits) and population admixture, which suggests that, despite the strong population structure in maritime pine, within-population quantitative genetic variation was unlikely to be influenced by gene flow across gene pools. In conclusion, CHAPTER 1 consists of an original attempt to study the imprints of natural selection on quantitative genetic variation and informs how genetic variation is maintained within populations, and therefore how populations adapt along environmental gradients.

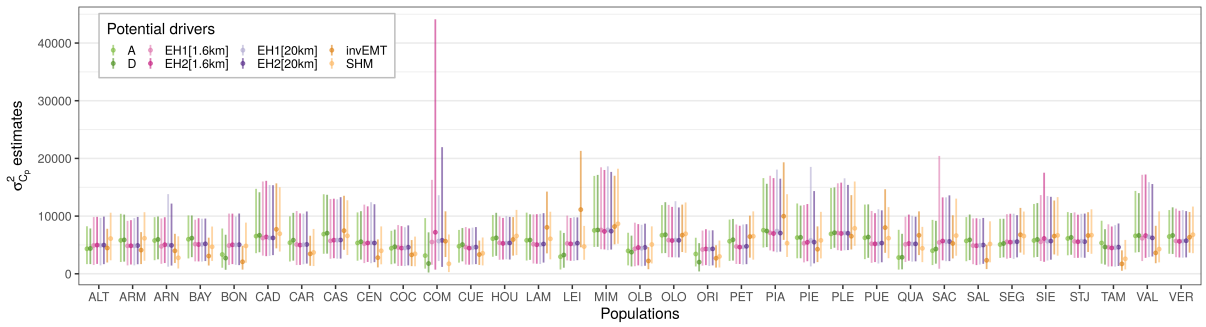


Figure VII.4. Median and 95% intervals of the posterior distributions of $\sigma^2_{C_p}$ (i.e. standing for the within-population genetic variation) in the model for height at 25 month-old in the Bordeaux common garden. The potential drivers are: two admixture scores (A and D), four indexes representing the environmental heterogeneity in a 1.6-km and 20-km radius around the population location (EH1[1.6km], EH2[1.6km], EH1[20km] and EH2[20km]), the inverse of the extreme minimum temperature during the studied period (invEMT) and the summer heat moisture index (SHM). Figure from the Supplementary Information of CHAPTER 1.

1.3 Genomics may help predicting short-term population responses to changing conditions

A few large-scale studies in maritime pine demonstrated the benefits of including intraspecific variability to predict changes in the occurrence or performance of populations under new environmental conditions. For instance, species distribution models (SDMs) accounting for intraspecific variability yielded more realistic projections of suitable habitat under future climate scenarios, generally leading to an increase in the predicted suitable area for the species (Benito Garzón et al. 2011, Serra-Varela et al. 2015). As another example, Fréjaville et al. (2020) estimated the association between height-growth performance in common gardens and the climate-of-origin of the populations and reported that maritime pine populations in the driest or hottest parts of the species range may grow less under future climates, while populations in the coldest and wettest regions may grow more. However, these studies integrate intraspecific variability within large-scale predictions at a coarse-grained level, i.e. variability at the population level (Benito Garzón et al. 2011, Fréjaville et al. 2020), or among genetically similar clusters defined based on a combination of mitochondrial, chloroplast and nuclear molecular markers (Serra-Varela et al. 2015). Moreover, they do not distinguish the relative contributions of adaptive and neutral processes to genetic differentiation among populations. Therefore, there is potentially considerable room for improvement in incorporating intraspecific variability at a finer grain

level (which might be enabled by individual-level genomic information) and differentiating the role of local adaptation, population demographic history and gene flow in predictions.

In CHAPTERS 2 and 3, genomic information from new genotyping technologies was incorporated within innovative statistical methods to evaluate the short-term performance of populations under new conditions. In CHAPTER 2, we showed that models combining genomic and climatic information (i.e. climate-of-origin of the populations, gene pool assignment and counts of trait-associated positive-effect alleles, PEAs) predict height growth of populations not included during the model fit better than models based only on the common garden design or incorporating climatic information, such as the climate-based population response functions which are currently commonly used to predict phenotypic variation of forest tree populations along climatic gradients (e.g. Leites et al. 2012a, Rehfeldt et al. 2018, 2002). A noteworthy finding of this chapter was that the regional PEAs (identified in specific geographical regions; Fig. VII.5) had a better predictive ability than global PEAs (identified range-wide; Fig. VII.5). It may be noted that the PEA effect on height growth was specific to each genotype (i.e. clone), and in the case of regional PEAs, specific also to the geographic region in which each tree was planted, thus making the genome-based predictions both individual-based and environment-dependent. Such integration of individual-based genomic information into trait-based approaches thus holds great promise for predicting how populations, and genotypes within populations, will perform when transplanted into new environments.

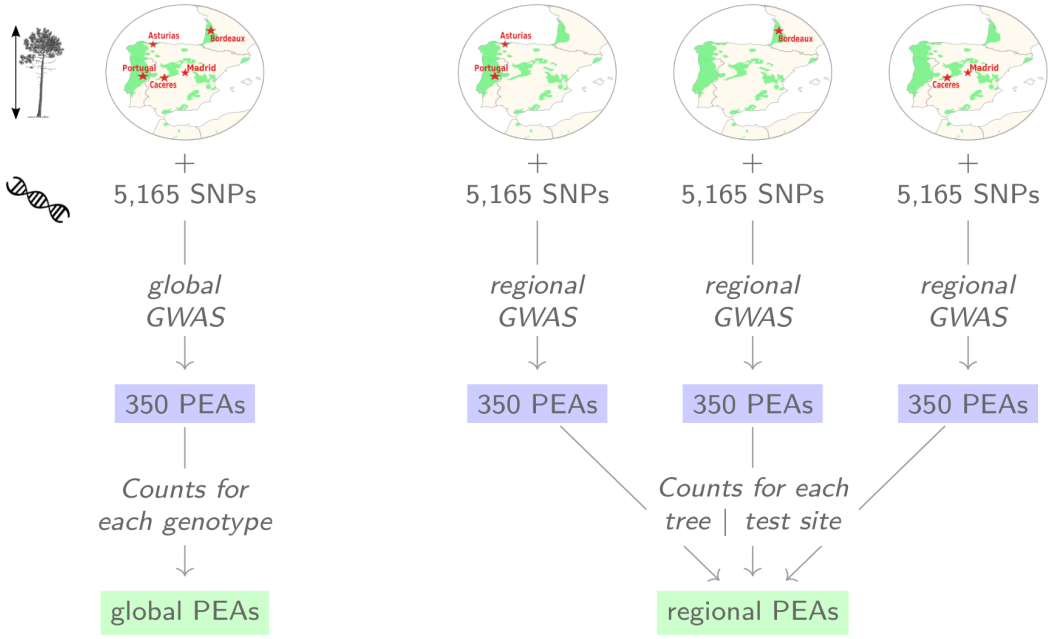


Figure VII.5. Schematic representation of the calculation of the global PEA counts (i.e. SNPs selected range-wide) and regional PEA counts (i.e. SNPs selected in a specific geographical region). Figure from Section 2.1 in the Supplementary Information of CHAPTER 2, in which more details are given on PEA calculation.

In CHAPTER 3, we found that genomic offset predictions were promising to predict short-term climate maladaptation of maritime pine populations. In particular, predictions based on the set of common SNP candidates (identified by the two GEA methods) were consistent between the two landscape genomic approaches (GDM and GF) and were associated with recent demographic trends in natural populations (i.e. mortality rates from National Forest Inventories). Nevertheless, it has be noted that the accuracy of genomic offset predictions in the validation analyses were highly sensitive to the set of SNPs and the modelling approach used

to estimate the gene-environment relationships. Moreover, contrary to Fitzpatrick et al. (2021), the genomic offset predictions in common gardens did not outperform predictions based on climatic transfer distances. Finally, results from CHAPTER 3 imply that before confidently using the genomic offset approach in the toolkit of conservation and management strategies, the effects of different choices on predictions must be clearly understood, e.g. choice of candidate SNPs, landscape genomics approach or method to control for population structure (Capblancq et al. 2020a, Rellstab et al. 2021).

2 Limitations, challenges and perspectives

'From the outset, it is important to acknowledge that predictions may often be imprecise even when they are accurate. Characterizing the magnitude of uncertainty is itself worthwhile.' (Shaw 2018)

2.1 Combining phenotypic, environment and genomic information

One of the major strengths of this PhD work is the original way in which genomic, phenotypic and environmental data have been combined to (i) provide a broad picture of adaptation patterns in maritime pine and the evolutionary processes underlying them, (ii) evaluate the short-term maladaptation risk of maritime pine populations in the face of climate change.

Combining phenotypic, environmental and genomic information have been repeatedly advised to study adaptation patterns in forest trees (de Villemereuil et al. 2016, Lepais and Bacles 2014, Sork et al. 2013). For instance, Kort et al. (2014) combined gene-environment and gene-phenotype association analyses to demonstrate the major role of temperature in driving adaptation patterns in *Alnus glutinosa.* As another example, Mahony et al. (2020) used phenotypic data from two common gardens of seedlings and 20-year trees in *Pinus contorta* to validate the climatic variables identified by gene-environment associations. Their results suggest that genomic information could be a reliable option to identify the climatic drivers of local adaptation when no phenotypic data is available.

In the present PhD work, I combined trait-based approaches (CHAPTERS 1 and 2) with landscape genomics (CHAPTER 3). In forest trees, trait-based approaches relying on common gardens have long been the gold standard to separate the genetic and environment component of quantitative trait variation and to determine which genotypes (populations) perform best in different environments through the use of genotype (population) reaction norms. This strong background in quantitative genetics in forest tree research has largely benefited from a long history of common gardens (often called provenance trials; Langlet 1971). However, common gardens are costly and time-consuming to maintain, especially in forest trees, and therefore the number of species and populations that can be measured in common gardens is bound to be limited. In addition, landscape genomics approaches are recent methods that still require development and validation steps, but are gaining considerable popularity in forest trees (e.g. Gougherty et al. 2020b, Ingvarsson and Bernhardsson 2020, Lu et al. 2019, Martins et al. 2018, Supple et al. 2018, Vanhove et al. 2021). Notably, they require only environmental and genomic information, which is becoming more and more available due to the steadily decreasing costs of new genotyping technologies. A strong assumption of these phenotype-free approaches is that populations are currently at their phenotypic optimum and that a change in environmental

conditions will break the optimal gene-environment relationships, thereby reducing relative fitness (Brady et al. 2019b). However, disrupted gene-environment relationships would not necessarily be associated with declining absolute fitness (i.e. decreasing demographic trends) and therefore the use of landscape genomics approaches to predict future population declines is currently in need of validation (Láruson et al. 2021).

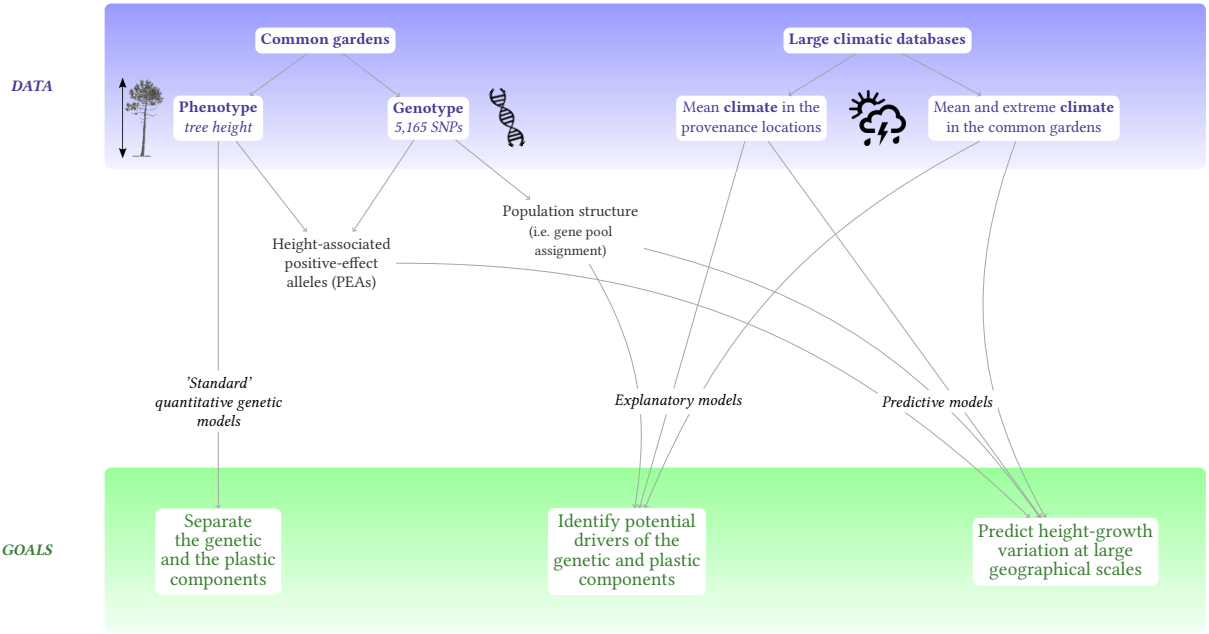


Figure VII.6. Schematic representation of the modelling framework used in CHAPTER 2. This figure is intended to illustrate how the different climatic, phenotypic, and genomic data were combined (and transformed into variables that could be incorporated into the models) to meet the objectives of CHAPTER 2.

Further, the novelty of the present PhD work lies in incorporating genomic information within trait-based approaches (CHAPTER 2), and in validating landscape genomics approaches with phenotypic information (CHAPTER 3). In CHAPTER 2, we showed that models incorporating genomic information (i.e. gene pool assignment and PEA counts) performed better (i.e. better predicted height-growth of populations not included during the model fit) than those using only phenotypic or climatic information, highlighting the potential of combining all data sources in large-scale trait predictions. In forest trees, the promising predictive ability of PEA counts was first supported by Mahony et al. (2020). More recently, MacLachlan et al. (2021) showed that PEA counts were more effective that single locus approaches to identify the adaptive climatic gradients for complex traits and were more robust to stochastic SNP sampling effects than polygenic scores. CHAPTER 2 thus follows the line of these previous works by confirming the superiority of PEAs compared to climate variables, but also goes further by suggesting the higher predictive ability of regional PEAs compared to global PEAs (Fig. VII.5). However, this improvement of the model predictive ability for regional PEAs is at the cost of reducing GWAS power (due to lower sample size than in global, wide-range analyses). Further comparison of the predictive ability of PEAs selected range-wide or regionally with denser genomic sampling will undoubtedly be of great interest to determine whether the higher predictive ability of regional PEAs observed under limited genomic sampling can be verified.

In CHAPTER 3, we combined phenotypic, climatic and genomic information in the following way: (1) we identified the adaptive genetic component through gene-environment associations; (2) we conducted landscape genomics analyses relying only on environment and genomic information; and (3) we validated the outputs from landscape genomics analyses with independent phenotypic data (i.e. height and mortality rates) from both common gardens and

natural populations (i.e. National Forest Inventories). The validation steps of CHAPTER 3 were inspired by the validation analysis based on demographic trends in natural populations from Bay et al. (2018) (but see critics in Fitzpatrick et al. 2018), and by the experimental validation based on height measurements in common gardens from Fitzpatrick et al. (2021). Importantly, we demonstrated that combining the validation steps in natural populations and common gardens can be particularly relevant, as they may point to very different conclusions about which genomic offset calculation methods work best. Finally, our work is consistent with several recent reviews calling for caution in the use of the genomic offset metric (Capblancq et al. 2020a, Hoffmann et al. 2021, Rellstab et al. 2021), since, among other concerns, it relies only on genomic and climatic data, and thus confronting genomic offset-based predictions with phenotypic information seems necessary.

Last, a particularly attractive avenue of research is the combination of gene-environment and gene-phenotype associations to identify potential candidate SNPs for local adaptation with gene expression analyses (i.e. transcriptomics) to validate the candidate SNPs (DeBiasse and Kelly 2016, Franks and Hoffmann 2012, Sork 2018). In a case study in *Arabidopsis*, Lasky et al. (2014) identified genes with variable expression response to environmental change, thus corresponding to a genotype-by-environment interaction in the expression patterns, reflecting local adaptation. In this line, Depardieu et al. (2021) first selected 285 candidate SNPs with gene-environment or gene-phenotype associations and then identified 110 high-confidence candidate SNPs that were differentially expressed under different drought treatments in white spruce, a widespread boreal conifer. Other examples in forest trees include the validation of genes involved in disease resistance in western balsam-poplar (Muchero et al. 2018) or in seedling water-stress response in valley oak (Gugger et al. 2017). Combining transcriptomic studies with population genetics remains rare in forest trees, although transcriptomics provide phenotypes that directly translate the response to the environment and thus have the potential to be a valuable tool for understanding local adaptation and for adding further evidence to the involvement of genes identified in association studies.

2.2 To what extent can the results for maritime pine be generalized?

In the present PhD work, I used maritime pine as a case study but the trends and patterns observed are likely to be informative about forest tree adaptation more generally. Indeed, in conifers, convergent patterns of local adaptation have been repeatedly detected, usually through the identification of common signatures of selection among related species in a set of orthologous genes (i.e. genes descending from a common ancestral gene by speciation). Perhaps the most striking example is the identification of 47 common genes involved in local adaptation between two distantly related conifers, interior spruce and lodgepole pine, separated by 140 million years of independent evolution (Yeaman et al. 2016). Similarly, Mosca et al. (2012) found seven SNPs associated with an environmental gradient that were located in genes common to four cohabiting conifers of the European Alps, i.e. silver fir, Swiss mountain pine, Swiss stone pine and European larch. Other evidence of convergent adaptation were found in more closely related species, e.g. two pine species from southeast China in which signals of recent selection were species-specific while most signals of ancient selection were common between the two species (Zhou et al. 2014), two spruce species (Siberian spruce and Norway spruce) showing footprints of convergent adaptation in the control of growth cessation (Chen et al. 2014), two other spruce species (white spruce and black spruce) showing common adaptive patterns in nine gene families -a number higher than expected by chance- (Prunier et al. 2011) and two white pine species (eastern white pine and western white pine) in which three orthologous

genes showed signatures of selection (Nadeau et al. 2016). In the Mediterranean pines, which have diverged recently (about 10 million years ago), Grivet et al. (2013) found two genes related to defense and stress response that showed adaptation patterns in the different taxon studied. In contrast, the Mediterranean pines exhibited divergent evolution in their life-history traits and different genetic correlations among growth-development, reproduction and fire-related traits, thus reflecting diverging adaptive histories (Grivet et al. 2013).

A consistent pattern emerging from studies comparing the adaptive evolution of different conifer species, and more generally forest species, in the northern hemisphere is the predominant role of cold temperatures in shaping the observed adaptation gradients. Indeed, in most of the studies mentioned above, some genes suggested to be under recent or ancient selection were already known to be associated with cold response in other plants or forest trees (Nadeau et al. 2016, Prunier et al. 2011, Zhou et al. 2014). Noticeably, the strongest phenotypic signatures of adaptation to climate found in Yeaman et al. (2016) were related to fall and winter cold injury traits and low-temperature stress–related environmental factors. These findings are consistent with the strong phenotypic and genetic clines along latitudinal and temperature gradients long known in forest trees (e.g. Aitken and Hannerz 2001, Joyce and Rehfeldt 2013, Langlet 1971, Morgenstern 1996, O'Neill et al. 2008, Rehfeldt et al. 1999) and involving the synchronised response of multiple complex traits (Howe et al. 2003).

Results from CHAPTER 3 are in line with the aforementioned body of work on the prevalence of cold adaptation in conifers since we showed the major contribution of cold temperature variables in explaining allelic turnover along environmental gradients. More remarkably, the work presented in CHAPTER 1 is, to my knowledge, the first to demonstrate an association between the levels of quantitative genetic variation within populations (for height in this chapter) and extreme cold events in a long-lived forest tree, thus supporting the key role of climate-induced selection in reducing genetic variation within populations. Interestingly, Ramírez-Valiente et al. (2019) also found a decrease in within-population genetic variation for functional and growth traits in drier and hotter conditions during the dry season in a Mesoamerican white oak, *Quercus oleoides*. In contrast, Anderegg et al. (2021) found no association between within-population phenotypic variance for several functional traits and aridity for eight *Acacia* species in Western Australia and Tasmania, regions where we would expect strong drought-related selection potentially leading to a decrease in within-population variance (however, these results would need to be confirmed by examining the additive genetic -and not phenotypic- variance). To conclude, assessing whether the association between within-population genetic variation for height and severe cold temperatures in maritime pine populations is a general pattern in conifers -and more generally in forest trees- would be particularly informative to further investigate the convergence in adaptive evolution among closely or more distantly related species. Moreover, determining whether such a decrease in within-population genetic variation can be found along other environmental gradients (e.g. Ramírez-Valiente et al. 2019) or for other traits would inform about how this pattern is general or specific to tree height along cold gradients.

Studying forest tree adaptation to cold temperatures is particularly important in management and conservation strategies. Indeed, within the assisted migration framework, moving trees to latitudes too far from their current location (outside their breeding zone in particular) may lead to maladaptation to new cold conditions (e.g. Grady et al. 2015). Predicting the phenotypes of translocated individuals can be done at the population level using using climatic data (i.e. climate-based population response functions; e.g. Leites et al. 2012a, Rehfeldt et al. 1999) or at the individual level using genomic data, or a combination of genomic and climatic data

(CHAPTER 2). Ultimately, landscape genomics approaches such as the genomic offset may also be appropriate to predict the risk of environmental maladaptation of the populations when transplanted in new environments (CHAPTER 3). Furthermore, careful thought must be taken when selecting for certain traits in breeding programs (e.g. growth) since genetic correlations among traits may impede cold hardiness (MacLachlan et al. 2021).

Besides their convergent patterns of adaptive evolution, conifers show highly different patterns of neutral genetic variation across their ranges. In Europe for instance, a clear distinction can be made between populations in the Mediterranean region, which are often fragmented and have a high population genetic structure (e.g. populations of black pine and maritime pine), and populations in central and northern Europe, which are often continuously distributed and have a low population structure (e.g. populations of silver fir, Norway spruce and Scots pine) (Alberto et al. 2011). Population genetic structure, induced by population demographic history and gene flow among populations, may be considered as a nuisance parameter to be controlled when searching for selection footprints, but also as a key factor shaping the nature, direction and efficiency of natural selection (Siol et al. 2010). Therefore, to what extent the results from the present PhD work in maritime pine, a Mediterranean pine species with fragmented populations and a strong population genetic structure (Jaramillo-Correa et al. 2015), can be transferred to other conifer species (with different population structure) remains to be tested. For example, similar modelling approaches as the one used in CHAPTER 1 could be applied to Scots pine populations, which show low levels of genetic differentiation and are distributed almost continuously from southern Spain to eastern Asia. Similarly to the results obtained in maritime pine, we may hypothesize to detect a decrease in within-population genetic variation subjected to severe cold events, because the distribution of Scots pine covers a broad latitudinal gradient and populations have been repeatedly shown to follow steep adaptive clines along temperature gradients (Savolainen et al. 2007). However, unlike maritime pine, which has reduced gene flow among fragmented populations, strong gene flow among Scots pine populations might counteract the negative effect of natural selection on within-population genetic variation. Thus, the extension of the work in this PhD to other forest tree species remains highly desirable.

Another interesting perspective would be to focus on maritime pine populations with very low levels of differentiation but located in highly different environments, and assess whether the results obtained in this PhD work are robust in these populations. Corsican populations of maritime pine may be used for this purpose as they show almost no genetic structure and are located in highly contrasted environments, ranging from Mediterranean beaches on the seashore to populations at high altitude a few kilometers away. Extension of the work in this PhD in that direction is also desirable.

2.3 Towards predicting the response of populations to future environmental conditions

CHAPTERS 1 and 3 present a first step towards predicting the response of maritime pine populations to future environmental conditions (e.g. climate change) by assessing which populations may be able to adapt quickly (CHAPTER 1) and which populations may have to evolve the most to maintain current gene-environment relationships (CHAPTER 3). In CHAPTER 1, quantitative genetic variation estimated within populations informs about their adaptive potential. According to the breeder's equation, $R = h^2S$ (univariate form, Falconer and Mackay 1996, Lush 1937; see Lande and Arnold 1983 for the multivariate form), the per generation

response (i.e. evolutionary change) of a quantitative trait under selection (R) is the product of the trait heritability (h^2) and the selection differential (S), which is the difference between the population mean before and after selection, and more generally, the phenotypic covariance between the relative fitness and the trait. Therefore, CHAPTER 1 provides half the puzzle for determining which populations may adapt more rapidly to changing conditions on short time scales, as we showed that genetic variation for height varies among populations, in particular being lower in populations experiencing severe cold events. CHAPTER 3 revealed that the risk of short-term climate maladaptation, i.e. the genomic offset, also varies among populations, with populations currently experiencing mild-winter conditions (i.e. most Atlantic populations and populations in southeastern France and northwestern Italy) being at higher risk. Importantly, the genomic offset informs about the magnitude of genetic change required to maintain the current gene-environment relationships, and thus in theory to maintain the optimal phenotype under changing environmental conditions (i.e. scenario of the 'moving target' in Brady et al. 2019a). However, it does not account for the ability of populations to evolve towards the new optimum, either through shifts in allele frequencies caused by selection, the migration of adaptive alleles, or at a lesser extent through the onset of new mutations (but see Gougherty et al. 2020b). Considering these evolutionary processes may lower the estimates of maladaptation risk (Exposito-Alonso et al. 2018b), but for that, other approaches are needed as the genomic offset approach is limited to estimating maladaptation of populations before evolution happens.

Predicting the fate of populations in the face of climate change involves predicting their short-, medium-, and if possible long-term evolutionary trajectories. Here I use 'predict' in the sense of forecasting future attributes of a population (e.g. trait values, mean fitness or allele frequencies) based on theory and the current attributes of the population. But, is evolution predictable? This question has been the subject of long-standing debate attempting to determine whether evolution acts primarily via deterministic or stochastic processes (Grant and Grant 2002, Lässig et al. 2017, Reznick and Travis 2018). Stochastic processes include the individual-level processes of mutation and stochastic developmental variation and the population-level processes of recombination and genetic drift (Fig. VII.7). Although the influence of new mutations on the fitness of a given individual may be negligible most of the time, long-term experimental evolution studies of short-lived microorganisms (i.e. over thousands of generations) have revealed the importance of rare but large-effect random mutations on evolutionary trajectories (Blount et al. 2008). In forest trees with long generation times and in which most adaptations derive from standing genetic variation, the contribution of new mutations to the evolutionary trajectories of populations in the face of climate change is likely to be very marginal, if not absent. Gene flow may also be considered as a random process acting at the metapopulation level, although it has also a deterministic component (Edelaar and Bolnick 2012, Rice and Papadopoulos 2009). Importantly, the relative importance of random processes on evolutionary trajectories fixes an upper bound to the degree to which evolution is in theory predictable, i.e. the 'random limit' hypothesis in Nosil et al. (2020) (see Fig. VII.7).

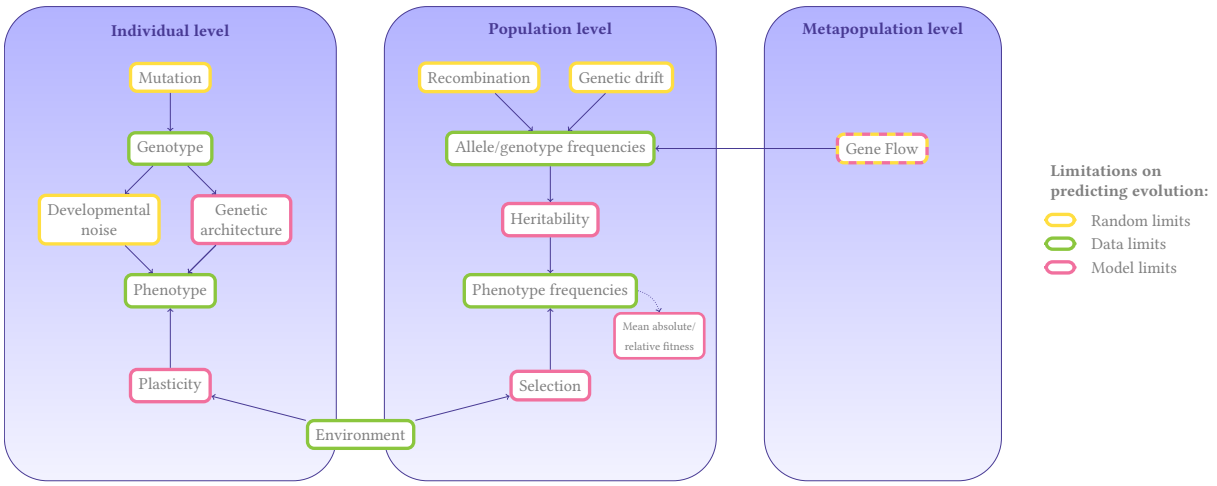


Figure VII.7. Diagram representing different sources of uncertainty in predicting evolution, originating from stochastic evolutionary processes (i.e. random limits), limited understanding (and thereby accurate modelling) of the deterministic evolutionary processes (i.e. model limits) and the lack of genomic, phenotypic, and environmental data (i.e. data limits). The underlying rationale of the model and data limits (gathered under the term 'data limits' in Nosil et al. 2020) is that with sufficient data, and theoretical and mathematical models accurately capturing the processes at play, we would be able to predict the evolutionary trajectories of populations. Modified from Ovaskainen et al. (2016) and Nosil et al. (2020).

Our ability to predict evolution may also, if not more, be constrained by our understanding of the deterministic processes underlying the evolutionary trajectories of the populations (referred as 'model limits' in Fig. VII.7). At the individual level, deterministic processes encompass the molecular mechanisms generating the phenotypes based on the environment-mediated expression of the genotypes, e.g. epigenetic mechanisms behind phenotypic plasticity, *cis*/*trans* effects or pleiotropic effects that underlie the genetic architecture of most quantitative traits. At the population level, phenotype frequencies are determined by the action of natural selection and inheritance rules based on the initial allele and genotype frequencies. A deep understanding of such processes is required to build robust theoretical and mathematical models aiming at predicting the future adaptive states of natural populations. I will briefly present three major modelling approaches that have been developed so far. First, quantitative genetics theory offers key models for predicting the adaptive potential of populations over short time scales. The breeder's equation (see above) or the Robertson's secondary theorem of selection (Price 1970, Robertson 1966) can be used to predict the adaptive evolution of phenotypic traits under selection (Walsh and Lynch 2018). However, a crucial limitation is that these equations require choosing which phenotypic traits to consider among the large number of traits potentially subject to selection (Shaw 2018). According to Shaw (2018), Fisher's Fundamental Theorem provides a direct and potentially more robust alternative to predict the rate of adaptation, i.e. the change in mean absolute fitness, of populations (Fisher 1930). Second, prediction of population evolutionary trajectories over medium time scales can be achieved using long-term series of observational data from natural populations, i.e. by determining whether trends at the beginning of the time series can predict subsequent temporal trends (e.g. Grant and Grant 2002, Nosil et al. 2018). Third, predicting evolution over long time scales (i.e. across many generations) requires to directly incorporate the evolutionary processes within the models, which can be achieved via mechanistic individual-based evolutionary dynamics models. Such models are particularly relevant for predicting the evolutionary trajectories of populations as they can simulate how the different deterministic and stochastic evolutionary forces (selection, gene flow, mutation and genetic drift) interact by explicitly integrating the underlying mechanisms, i.e. inheritance rules (Oddou-Muratorio et al. 2020). In particular, they can be calibrated with quantitative genetic models which provide an efficient and robust framework to infer the genetic architecture of polygenic traits (e.g. Coulson et al. 2017). Ultimately, evolutionary dynamics models may be combined with ecophysiological and forest dynamics models to

improve our understanding and the predictions of the eco-evolutionary dynamics of forests under changing environmental conditions (Oddou-Muratorio et al. 2020).

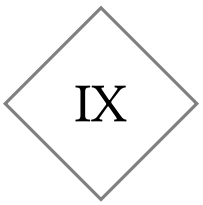
A common limitation of the modelling approaches presented above stems from the insufficient empirical data to both calibrate and validate the models, i.e. 'data limits' in Fig. VII.7. For example, the evolutionary trajectories of beak size in Darwin's finches on Daphne Major Island were not predictable, primarily due to rare and large-effect climate events that were not predicted by climate models (Grant and Grant 2002). Similarly, large databases of quantitative genetic parameters for performance (e.g. survival or reproduction) or functional traits are lacking to robustly calibrating evolutionary dynamics models (Oddou-Muratorio et al. 2020). Moreover, in forest trees, most quantitative genetic parameters come from controlled experiments (e.g. common gardens) of young trees, thus covering only a limited range of environments and under growth conditions far different from those experienced *in situ* (e.g. relaxed competition among trees).

Our ability to make accurate predictions of future evolutionary trajectories of populations, in particular under climate change, is therefore constrained by a combination of random limits caused by unpredictable stochastic processes, and model and data limits, originating from insufficient knowledge of the deterministic processes at stake (Nosil et al. 2020). Noticeably, our current insufficient knowledge on how the phenotypes are generated based on the genotypes at the individual level (i.e. the genotype-phenotype map) explains why using genomic data to predict multi-generational evolution is at present out of reach (Shaw 2018). Indeed, a robust conceptual framework on how genes interact among them and with the environment to generate the phenotype is still lacking. The omnigenic model is so far the most complete theoretical model of genetic architecture, suggesting that phenotypes are determined by a minority of direct-effect genes and a majority of indirect-effect genes. Interestingly, Mathieson (2021) argues that direct-effect genes are likely to have a stable effect across populations while indirect-effect genes are more likely to show population-dependent effects, and that the omnigenic model can be extended to the 'omni-environmental' model in which some 'core' environmental effects would be consistent across populations while 'peripheral' environmental effects would be unpredictable. Importantly, this would imply that genomic-based predictions of individual phenotypes will be inherently environment and population-specific, and therefore inappropriate for predicting adaptation of populations across many generations under changing environments (Shaw 2018). Other arguments in this direction include (i) the highly challenging (if not impossible) task of capturing most genetic variants under weak selection with genome scans, (ii) the mismatch between the temporal scale of adaptation signals detectable in genomic studies and that of the necessary short-term adaptation to climate change and, more importantly, (iii) the inability of genomic studies to estimate the mean absolute fitness of populations, which corresponds to their degree of adaptation (Shaw 2018). Therefore, predicting the evolution of populations under future conditions using quantitative genetic theory, whose predictions are based on the aggregate effects of multiple polymorphic loci, seems much more promising than using genomic and molecular tools, which necessitate understanding the effect of each adaptive loci in isolation. Conversely, although genomic data cannot be used directly in the prediction of evolutionary trajectories, they are particularly useful for estimating genetic parameters of interest in natural populations *in situ*, which is of great value for the calibration of mechanistic evolutionary dynamics models.

VIII

Conclusion

Using maritime pine as a case study, the present PhD work originally combined trait-based approaches with landscape genomics, thereby bringing together phenotypic, environmental and genomic data from a large network of five common gardens and 34 populations. Results from the different chapters converge on the key role of cold temperatures in the adaptive history of maritime pine, impacting both the current adaptive states of populations, but potentially also the within-population variance of some quantitative traits (e.g. tree height). Genomic data showed particular promise for improving predictions of short-term population responses to environmental changes. Indeed, quantitative trait predictions for individuals translocated to new environments may be improved by incorporating genomic information within the models, which would be of great interest in conservation or management strategies (e.g. assisted migration or evolutionary rescue). Moreover, genome-based predictions of the extent to which gene-environment relationships will be disrupted by climate change may prove to be a highly relevant tool for rapidly identifying populations at risk of short-term climate maladaptation. More broadly, this work contributes to a better understanding of adaptive processes, and to the accumulation of knowledge and robust statistical methods necessary for the implementation of evolution-based management strategies, and to progress towards the ambitious but urgent goal of predicting the response of populations to climate change.



Bibliography

Ackerly D., Knight C., Weiss S., Barton K., and Starmer K. (2002). Leaf size, specific leaf area and microhabitat distribution of chaparral woody plants: contrasting patterns in species level and community level analyses. *Oecologia* **130**.3: 449– 457. DOI: 10.1007/s004420100805.

Aguirre-Liguori J. A., Ramírez-Barahona S., Tiffin P., and Eguiarte L. E. (2019). Climate change is predicted to disrupt patterns of local adaptation in wild and cultivated maize. *Proceedings of the Royal Society B: Biological Sciences* **286**.1906: 20190486. DOI: 10.1098/rspb.2019.0486.

Ahuja M. R. and Neale D. B. (2005). Evolution of Genome Size in Conifers. *Silvae Genetica* **54**.1-6: 126– 137. DOI: 10.1515/sg-2005-0020.

Aitken S. N. and Bemmels J. B. (2015). Time to get moving: assisted gene flow of forest trees. *Evolutionary Applications* **9**.1: 271– 290. DOI: 10.1111/eva.12293.

Aitken S. N. and Hannerz M. (2001). Genecology and Gene Resource Management Strategies for Conifer Cold Hardiness. *Conifer Cold Hardiness*. Ed. by Bigras F. J. and Colombo S. J. Dordrecht: Springer Netherlands. Page 23– 53. ISBN: 978-94-015-9650-3. DOI: 10.1007/978-94-015-9650-3_2.

Aitken S. N. and Whitlock M. C. (2013). Assisted Gene Flow to Facilitate Local Adaptation to Climate Change. *Annual Review of Ecology, Evolution, and Systematics* **44**.1: 367– 388. DOI: 10.1146/annurev-ecolsys-110512-135747.

Aitken S. N., Yeaman S., Holliday J. A., Wang T., and Curtis-McLane S. (2008). Adaptation, migration or extirpation: climate change outcomes for tree populations. *Evolutionary Applications* **1**.1: 95– 111. DOI: 10.1111/j.1752-4571.2007.00013.x.

Alberto F., Aitken S. N., Alía R., González-Martínez S. C., Hänninen H., Kremer A., Lefèvre F., Lenormand T., Yeaman S., Whetten R., and Savolainen O. (2013). Potential for evolutionary responses to climate change – evidence from tree populations. *Global Change Biology* **19**.6: 1645– 1661. DOI: 10.1111/gcb.12181.

Alberto F., Bouffier L., Louvet J.-M., Lamy J.-B., Delzon S., and Kremer A. (2011). Adaptive responses for seed and leaf phenology in natural populations of sessile oak along an altitudinal gradient. *Journal of Evolutionary Biology* **24**.7: 1442– 1454. DOI: 10.1111/j.1420-9101.2011.02277.x.

Alfaro R. I., Lewis K. G., King J. N., El-Kassaby Y. A., Brown G., and Smith L. D. (2000). Budburst phenology of sitka spruce and its relationship to white pine weevil attack. *Forest Ecology and Management* **127**.1: 19– 29. DOI: 10.1016/S0378-1127(99)00115-2.

Alía R., Chambel R., Notivol E., Climent J., and González-Martínez S. C. (2014). Environment-dependent microevolution in a Mediterranean pine (Pinus pinasterAiton). *BMC Evolutionary Biology* **14**: 200. DOI: 10.1186/s12862-014-0200-5.

Alía R., Moro J., and Denis J. B. (1997). Performance of Pinus pinaster provenances in Spain: interpretation of the genotype by environment interaction. *Canadian Journal of Forest Research* **27**.10: 1548– 1559.

Allen C. D., Breshears D. D., and McDowell N. G. (2015). On underestimation of global vulnerability to tree mortality and forest die-off from hotter drought in the Anthropocene. *Ecosphere* **6**.8: 1– 55. DOI: 10.1890/ES15-00203.1.

Allen C. D., Macalady A. K., Chenchouni H., Bachelet D., McDowell N., Vennetier M., Kitzberger T., Rigling A., Breshears D. D., Hogg E. H. (, Gonzalez P., Fensham R., Zhang Z., Castro J., Demidova N., Lim J.-H., Allard G., Running S. W., Semerci A., and Cobb N. (2010). A global overview of drought and heat-induced tree mortality reveals emerging climate change risks for forests. *Forest Ecology and Management* **259**.4: 660– 684. DOI: 10.1016/j.foreco.2009.09.001.

Anderegg L. D. L., Loy X., Markham I. P., Elmer C. M., Hovenden M. J., HilleRisLambers J., and Mayfield M. M. (2021). Aridity drives coordinated trait shifts but not decreased trait variance across the geographic range of eight Australian trees. *New Phytologist* **229**.3: 1375–1387. DOI: 10.1111/nph.16795.

Anderegg W. R. L., Hicke J. A., Fisher R. A., Allen C. D., Aukema J., Bentz B., Hood S., Lichstein J. W., Macalady A. K., McDowell N., Pan Y., Raffa K., Sala A., Shaw J. D., Stephenson N. L., Tague C., and Zeppel M. (2015). Tree mortality from drought, insects, and their interactions in a changing climate. *New Phytologist* **208**.3: 674– 683. DOI: 10.1111/nph.13477.

Anderson J. T., Inouye D. W., McKinney A. M., Colautti R. I., and Mitchell-Olds T. (2012). Phenotypic plasticity and adaptive evolution contribute to advancing flowering phenology in response to climate change. *Proceedings of the Royal Society B: Biological Sciences* **279**.1743: 3843– 3852. DOI: 10.1098/rspb.2012.1051.

Anderson J. T., Lee C.-R., Rushworth C. A., Colautti R. I., and Mitchell-Olds T. (2013). Genetic trade-offs and conditional neutrality contribute to local adaptation. *Molecular Ecology* **22**.3: 699– 708. DOI: 10.1111/j.1365-294X.2012.05522.x.

Anderson R. G., Canadell J. G., Randerson J. T., Jackson R. B., Hungate B. A., Baldocchi D. D., Ban-Weiss G. A., Bonan G. B., Caldeira K., Cao L., Diffenbaugh N. S., Gurney K. R., Kueppers L. M., Law B. E., Luyssaert S., and O'Halloran T. L. (2011). Biophysical considerations in forestry for climate protection. *Frontiers in Ecology and the Environment* **9**.3: 174– 182. DOI: 10.1890/090179.

Andrews K. R., Good J. M., Miller M. R., Luikart G., and Hohenlohe P. A. (2016). Harnessing the power of RADseq for ecological and evolutionary genomics. *Nature Reviews Genetics* **17**.2: 81– 92. DOI: 10.1038/nrg.2015.28.

Angert A. L., Bontrager M. G., and Ågren J. (2020). What Do We Really Know About Adaptation at Range Edges? *Annual Review of Ecology, Evolution, and Systematics* **51**.1: 341– 361. DOI: 10.1146/annurev-ecolsys-012120-091002.

Antonovics J. (2006). Evolution in closely adjacent plant populations X: long-term persistence of prereproductive isolation at a mine boundary. *Heredity* **97**.1: 33– 37. DOI: 10.1038/sj.hdy.6800835.

Antonovics J. and Bradshaw A. D. (1970). Evolution in closely adjacent plant populations VIII. Clinal patterns at a mine boundary. *Heredity* **25**.3: 349– 362. DOI: 10.1038/hdy.1970.36.

Archambeau J., Garzón M. B., Barraquand F., Vega M. d. M., Plomion C., and González-Martínez S. C. (2021a). Combining climatic and genomic data improves range-wide tree height growth prediction in a forest tree: 2020.11.13.382515. DOI: 10.1101/2020.11.13.382515.

Archambeau J., Garzón M. B., Vega M. d. M., Brachi B., Barraquand F., and González-Martínez S. C. (2021b). Extreme climatic events but not environmental heterogeneity shape within-population genetic variation in maritime pine: 2021.08.17.456636.

Arnold P. A., Kruuk L. E. B., and Nicotra A. B. (2019). How to analyse plant phenotypic plasticity in response to a changing climate. *New Phytologist* **0**.0. DOI: 10.1111/nph.15656.

Arnold S. J., Pfrender M. E., and Jones A. G. (2001). The adaptive landscape as a conceptual bridge between micro- and macroevolution. *Microevolution Rate, Pattern, Process*. Ed. by

Hendry A. P. and Kinnison M. T. Dordrecht: Springer Netherlands. Page 9– 32. ISBN: 978-94-010-0585-2. DOI: 10.1007/978-94-010-0585-2_2.

Aschard H., Chen J., Cornelis M. C., Chibnik L. B., Karlson E. W., and Kraft P. (2012). Inclusion of Gene-Gene and Gene-Environment Interactions Unlikely to Dramatically Improve Risk Prediction for Complex Diseases. *The American Journal of Human Genetics* **90**.6: 962– 972. DOI: 10.1016/j.ajhg.2012.04.017.

Aubret F. and Shine R. (2010). Fitness costs may explain the post-colonisation erosion of phenotypic plasticity. *Journal of Experimental Biology* **213**.5: 735– 739. DOI: 10.1242/jeb.040576.

Baird N. A., Etter P. D., Atwood T. S., Currey M. C., Shiver A. L., Lewis Z. A., Selker E. U., Cresko W. A., and Johnson E. A. (2008). Rapid SNP Discovery and Genetic Mapping Using Sequenced RAD Markers. *PLOS ONE* **3**.10: e3376. DOI: 10.1371/journal.pone.0003376.

Baldwin J. M. (1896). A New Factor in Evolution (Continued). *The American Naturalist* **30**.355: 536– 553. DOI: 10.1086/276428.

Ballouz S., Pena M. T., Knight F. M., Adams L. B., and Gillis J. A. (2019). The transcriptional legacy of developmental stochasticity. *bioRxiv*. DOI: 10.1101/2019.12.11.873265.

Balocchi C. E., Bridgwater F. E., Zobel B. J., and Jahromi S. (1993). Age Trends in Genetic Parameters for Tree Height in a Nonselected Population of Loblolly Pine. *Forest Science* **39**.2: 231– 251. DOI: 10.1093/forestscience/39.2.231.

Balvanera P., Pfisterer A. B., Buchmann N., He J.-S., Nakashizuka T., Raffaelli D., and Schmid B. (2006). Quantifying the evidence for biodiversity effects on ecosystem functioning and services. *Ecology Letters* **9**.10: 1146– 1156. DOI: 10.1111/j.1461-0248.2006.00963.x.

Banta J. A., Ehrenreich I. M., Gerard S., Chou L., Wilczek A., Schmitt J., Kover P. X., and Purugganan M. D. (2012). Climate envelope modelling reveals intraspecific relationships among flowering phenology, niche breadth and potential range size in Arabidopsis thaliana. *Ecology Letters* **15**.8: 769– 777. DOI: 10.1111/j.1461-0248.2012.01796.x.

Barbéro M., Loisel R., Quézel P., Richardson D. M., and Romane F. (1998). *Ecology and biogeography of Pinus*. 153-170. Cambridge; New York, NY, USA: Cambridge University Press. ISBN: 978-0-521-55176-2.

Barghi N., Hermisson J., and Schlötterer C. (2020). Polygenic adaptation: a unifying framework to understand positive selection. *Nature Reviews Genetics* **21**.12: 769– 781. DOI: 10.1038/s41576-020-0250-z.

Barreiro L. B., Laval G., Quach H., Patin E., and Quintana-Murci L. (2008). Natural selection has driven population differentiation in modern humans. *Nature Genetics* **40**.3: 340– 345. DOI: 10.1038/ng.78.

Barrett R. D. H. and Schluter D. (2008). Adaptation from standing genetic variation. *Trends in Ecology & Evolution* **23**.1: 38– 44. DOI: 10.1016/j.tree.2007.09.008.

Barton N. H. (1990). Pleiotropic models of quantitative variation. *Genetics* **124**.3: 773– 782. DOI: 10.1093/genetics/124.3.773.

Barton N. H. and Partridge L. (2000). Limits to natural selection. *BioEssays* **22**.12: 1075– 1084. DOI: 10.1002/1521-1878(200012)22:12<1075::AID-BIES5>3.0.CO;2-M.

Barton N. H. and Turelli M. (1987). Adaptive landscapes, genetic distance and the evolution of quantitative characters. *Genetics Research* **49**.2: 157– 173. DOI: 10.1017/S0016672300026951.

Barton N., Hermisson J., and Nordborg M. (2019). Why structure matters. *eLife* **8**: e45380. DOI: 10.7554/eLife.45380.

Bateson W. (1909). *Mendel's principles of heredity*. Cambridge University Press, Cambridge.

Bay R. A., Harrigan R. J., Underwood V. L., Gibbs H. L., Smith T. B., and Ruegg K. (2018). Genomic signals of selection predict climate-driven population declines in a migratory bird. *Science* **359**.6371: 83– 86. DOI: 10.1126/science.aan4380.

Beaulieu J., Doerksen T., Clément S., MacKay J., and Bousquet J. (2014). Accuracy of genomic selection models in a large population of open-pollinated families in white spruce. *Heredity* **113**.4: 343– 352. DOI: 10.1038/hdy.2014.36.

Beaumont M. A. and Rannala B. (2004). The Bayesian revolution in genetics. *Nature Reviews Genetics* **5**.4: 251– 261. DOI: 10.1038/nrg1318.

Bell G. (2013). Evolutionary rescue and the limits of adaptation. *Philosophical Transactions of the Royal Society B: Biological Sciences* **368**.1610: 20120080. DOI: 10.1098/rstb.2012.0080.

Bellard C., Leclerc C., Leroy B., Bakkenes M., Veloz S., Thuiller W., and Courchamp F. (2014). Vulnerability of biodiversity hotspots to global change. *Global Ecology and Biogeography* **23**.12: 1376– 1386. DOI: 10.1111/geb.12228.

Benito Garzón M., Alía R., Robson T. M., and Zavala M. A. (2011). Intra-specific variability and plasticity influence potential tree species distributions under climate change. *Global Ecology and Biogeography* **20**.5: 766– 778. DOI: 10.1111/j.1466-8238.2010.00646.x.

Benito Garzón M., Robson T. M., and Hampe A. (2019). ΔTraitSDMs: species distribution models that account for local adaptation and phenotypic plasticity. *New Phytologist* **222**.4: 1757– 1765. DOI: 10.1111/nph.15716.

Bennett A. C., McDowell N. G., Allen C. D., and Anderson-Teixeira K. J. (2015). Larger trees suffer most during drought in forests worldwide. *Nature Plants* **1**.10: 15139. DOI: 10.1038/nplants.2015.139.

Bérénos C., Ellis P. A., Pilkington J. G., and Pemberton J. M. (2014). Estimating quantitative genetic parameters in wild populations: a comparison of pedigree and genomic approaches. *Molecular Ecology* **23**.14: 3434– 3451. DOI: 10.1111/mec.12827.

Berg J. J., Harpak A., Sinnott-Armstrong N., Joergensen A. M., Mostafavi H., Field Y., Boyle E. A., Zhang X., Racimo F., Pritchard J. K., and Coop G. (2019). Reduced signal for polygenic adaptation of height in UK Biobank. *eLife* **8**: e39725. DOI: 10.7554/eLife.39725.

Bergelson J., Kreitman M., Stahl E. A., and Tian D. (2001). Evolutionary Dynamics of Plant R-Genes. *Science* **292**.5525: 2281– 2285. DOI: 10.1126/science.1061337.

Bergland A. O., Behrman E. L., O'Brien K. R., Schmidt P. S., and Petrov D. A. (2014). Genomic Evidence of Rapid and Stable Adaptive Oscillations over Seasonal Time Scales in Drosophila. *PLOS Genetics* **10**.11: e1004775. DOI: 10.1371/journal.pgen.1004775.

Birol I., Raymond A., Jackman S. D., Pleasance S., Coope R., Taylor G. A., Yuen M. M. S., Keeling C. I., Brand D., Vandervalk B. P., Kirk H., Pandoh P., Moore R. A., Zhao Y., Mungall A. J., Jaquish B., Yanchuk A., Ritland C., Boyle B., Bousquet J., Ritland K., MacKay J., Bohlmann J., and Jones S. J. (2013). Assembling the 20 Gb white spruce (Picea glauca) genome from whole-genome shotgun sequencing data. *Bioinformatics* **29**.12: 1492– 1497. DOI: 10.1093/bioinformatics/btt178.

Blount Z. D., Borland C. Z., and Lenski R. E. (2008). Historical contingency and the evolution of a key innovation in an experimental population of Escherichia coli. *Proceedings of the National Academy of Sciences* **105**.23: 7899– 7906. DOI: 10.1073/pnas.0803151105.

Boisvert-Marsh Laura, Périé Catherine, and de Blois Sylvie (2014). Shifting with climate? Evidence for recent changes in tree species distribution at high latitudes. *Ecosphere* **5**.7: art83. DOI: 10.1890/ES14-00111.1.

Bonamour S., Chevin L.-M., Charmantier A., and Teplitsky C. (2019). Phenotypic plasticity in response to climate change: the importance of cue variation. *Philosophical Transactions of the Royal Society B: Biological Sciences* **374**.1768: 20180178. DOI: 10.1098/rstb.2018.0178.

Bonan G. B. (2008). Forests and Climate Change: Forcings, Feedbacks, and the Climate Benefits of Forests. *Science* **320**.5882: 1444– 1449. DOI: 10.1126/science.1155121.

Borcard D. and Legendre P. (2002). All-scale spatial analysis of ecological data by means of principal coordinates of neighbour matrices. *Ecological Modelling* **153**.1: 51– 68. DOI: 10.1016/S0304-3800(01)00501-4.

Borrell J. S., Zohren J., Nichols R. A., and Buggs R. J. A. (2020). Genomic assessment of local adaptation in dwarf birch to inform assisted gene flow. *Evolutionary Applications* **13**.1: 161– 175. DOI: 10.1111/eva.12883.

Bouvet J.-M., Makouanzi G., Cros D., and Vigneron P. (2016). Modeling additive and non-additive effects in a hybrid population using genome-wide genotyping: prediction accuracy implications. *Heredity* **116**.2: 146– 157. DOI: 10.1038/hdy.2015.78.

Boyle E. A., Li Y. I., and Pritchard J. K. (2017). An Expanded View of Complex Traits: From Polygenic to Omnigenic. *Cell* **169**.7: 1177– 1186. DOI: 10.1016/j.cell.2017.05.038.

Brachi B., Morris G. P., and Borevitz J. O. (2011). Genome-wide association studies in plants: the missing heritability is in the field. *Genome Biology* **12**.10: 232. DOI: 10.1186/gb-2011-12-10-232.

Bradshaw A. D. (1965). Evolutionary Significance of Phenotypic Plasticity in Plants. *Advances in Genetics*. Ed. by Caspari E. W. and Thoday J. M. 13. Academic Press. Page 115– 155. DOI: 10.1016/S0065-2660(08)60048-6.

Bradshaw A. D. (2006). Unravelling phenotypic plasticity – why should we bother? *New Phytologist* **170**.4: 644– 648. DOI: 10.1111/j.1469-8137.2006.01761.x.

Bradshaw W. E. and Holzapfel C. M. (2006). Evolutionary Response to Rapid Climate Change. *Science* **312**.5779: 1477– 1478. DOI: 10.1126/science.1127000.

Brady S. P., Bolnick D. I., Angert A. L., Gonzalez A., Barrett R. D. H., Crispo E., Derry A. M., Eckert C. G., Fraser D. J., Fussmann G. F., Guichard F., Lamy T., McAdam A. G., Newman A. E. M., Paccard A., Rolshausen G., Simons A. M., and Hendry A. P. (2019a). Causes of maladaptation. *Evolutionary Applications* **12**.7: 1229– 1242. DOI: 10.1111/eva.12844.

Brady S. P., Bolnick D. I., Barrett R. D. H., Chapman L., Crispo E., Derry A. M., Eckert C. G., Fraser D. J., Fussmann G. F., Gonzalez A., Guichard F., Lamy T., Lane J., McAdam A. G., Newman A. E. M., Paccard A., Robertson B., Rolshausen G., Schulte P. M., Simons A. M., Vellend M., and Hendry A. (2019b). Understanding Maladaptation by Uniting Ecological and Evolutionary Perspectives. *The American Naturalist* **194**.4: 495– 515. DOI: 10.1086/705020.

Bridle J. R., Gavaz S., and Kennington W. J. (2009). Testing limits to adaptation along altitudinal gradients in rainforest Drosophila. *Proceedings of the Royal Society B: Biological Sciences*. DOI: 10.1098/rspb.2008.1601.

Brockerhoff E. G., Barbaro L., Castagneyrol B., Forrester D. I., Gardiner B., González-Olabarria J. R., Lyver P. O., Meurisse N., Oxbrough A., Taki H., Thompson I. D., van der Plas F., and Jactel H. (2017). Forest biodiversity, ecosystem functioning and the provision of ecosystem services. *Biodiversity and Conservation* **26**.13: 3005– 3035. DOI: 10.1007/s10531-017-1453-2.

Browne L., Wright J. W., Fitz-Gibbon S., Gugger P. F., and Sork V. L. (2019). Adaptational lag to temperature in valley oak (Quercus lobata) can be mitigated by genome-informed assisted gene flow. *Proceedings of the National Academy of Sciences* **116**.50: 25179– 25185. DOI: 10.1073/pnas.1908771116.

Bucci G., González-Martínez S. C., Provost G. L., Plomion C., Ribeiro M. M., Sebastiani F., Alía R., and Vendramin G. G. (2007). Range-wide phylogeography and gene zones in Pinus pinaster Ait. revealed by chloroplast microsatellite markers. *Molecular Ecology* **16**.10: 2137– 2153. DOI: 10.1111/j.1365-294X.2007.03275.x.

Budde K. B., Heuertz M., Hernández-Serrano A., Pausas J. G., Vendramin G. G., Verdú M., and González-Martínez S. C. (2014). In situ genetic association for serotiny, a fire-related trait, in Mediterranean maritime pine (Pinus pinaster). *New Phytologist* **201**.1: 230– 241. DOI: 10.1111/nph.12483.

Burban C. and Petit R. J. (2003). Phylogeography of maritime pine inferred with organelle markers having contrasted inheritance. *Molecular Ecology* **12**.6: 1487– 1495. DOI: 10.1046/j.1365-294X.2003.01817.x.

Bürkner P.-C. (2017). brms: An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software* **80**.1: 1– 28. DOI: 10.18637/jss.v080.i01.

Burrows M. T., Schoeman D. S., Richardson A. J., Molinos J. G., Hoffmann A., Buckley L. B., Moore P. J., Brown C. J., Bruno J. F., Duarte C. M., Halpern B. S., Hoegh-Guldberg O., Kappel C. V., Kiessling W., O'Connor M. I., Pandolfi J. M., Parmesan C., Sydeman W. J., Ferrier S., Williams K. J., and Poloczanska E. S. (2014). Geographical limits to species-range shifts are suggested by climate velocity. *Nature* **507**.7493: 492– 495. DOI: 10.1038/nature12976.

Burt A. (1995). Perspective: The Evolution of Fitness. *Evolution* **49**.1: 1– 8. DOI: 10.2307/2410288.

Butchart S. H. M., Walpole M., Collen B., van Strien A., Scharlemann J. P. W., Almond R. E. A., Baillie J. E. M., Bomhard B., Brown C., Bruno J., Carpenter K. E., Carr G. M., Chanson J., Chenery A. M., Csirke J., Davidson N. C., Dentener F., Foster M., Galli A., Galloway J. N., Genovesi P., Gregory R. D., Hockings M., Kapos V., Lamarque J.-F., Leverington F., Loh J., McGeoch M. A., McRae L., Minasyan A., Morcillo M. H., Oldfield T. E. E., Pauly D., Quader S., Revenga C., Sauer J. R., Skolnik B., Spear D., Stanwell-Smith D., Stuart S. N., Symes A., Tierney M., Tyrrell T. D., Vié J.-C., and Watson R. (2010). Global Biodiversity: Indicators of Recent Declines. *Science* **328**.5982: 1164– 1168. DOI: 10.1126/science.1187512.

Byers D. L. (2005). Evolution in heterogeneous environments and the potential of maintenance of genetic variation in traits of adaptive significance. *Genetica* **123**.1: 107. DOI: 10.1007/s10709-003-2721-5.

Caignard T., Delzon S., Bodénès C., Dencausse B., and Kremer A. (2018). Heritability and genetic architecture of reproduction-related traits in a temperate oak species. *Tree Genetics & Genomes* **15**.1: 1. DOI: 10.1007/s11295-018-1309-2.

Caignard T., Kremer A., Bouteiller X. P., Parmentier J., Louvet J.-M., Venner S., and Delzon S. (2021). Counter-gradient variation of reproductive effort in a widely distributed temperate oak. *Functional Ecology* **35**.8: 1745– 1755. DOI: 10.1111/1365-2435.13830.

Capblancq T., Fitzpatrick M. C., Bay R. A., Exposito-Alonso M., and Keller S. R. (2020a). Genomic Prediction of (Mal)Adaptation Across Current and Future Climatic Landscapes. *Annual Review of Ecology, Evolution, and Systematics* **51**.1: null. DOI: 10.1146/annurev-ecolsys-020720-042553.

Capblancq T. and Forester B. R. (2021). Redundancy analysis: A Swiss Army Knife for landscape genomics. *Methods in Ecology and Evolution* **n/a**.n/a. DOI: 10.1111/2041-210X.13722.

Capblancq T., Morin X., Gueguen M., Renaud J., Lobreaux S., and Bazin E. (2020b). Climate-associated genetic variation in Fagus sylvatica and potential responses to climate change in the French Alps. *Journal of Evolutionary Biology* **33**.6: 783– 796. DOI: 10.1111/jeb.13610.

Cardinale B. J., Duffy J. E., Gonzalez A., Hooper D. U., Perrings C., Venail P., Narwani A., Mace G. M., Tilman D., Wardle D. A., Kinzig A. P., Daily G. C., Loreau M., Grace J. B., Larigauderie A., Srivastava D. S., and Naeem S. (2012). Biodiversity loss and its impact on humanity. *Nature* **486**.7401: 59. DOI: 10.1038/nature11148.

Carlborg Ö. and Haley C. S. (2004). Epistasis: too often neglected in complex trait studies? *Nature Reviews Genetics* **5**.8: 618– 625. DOI: 10.1038/nrg1407.

Carlson S. M., Cunningham C. J., and Westley P. A. H. (2014). Evolutionary rescue in a changing world. *Trends in Ecology & Evolution* **29**.9: 521– 530. DOI: 10.1016/j.tree.2014.06.005.

Carpenter B., Gelman A., Hoffman M. D., Lee D., Goodrich B., Betancourt M., Brubaker M., Guo J., Li P., and Riddell A. (2017). Stan : A Probabilistic Programming Language. *Journal of Statistical Software* **76**.1. DOI: 10.18637/jss.v076.i01.

Carter A. J. and Nguyen A. Q. (2011). Antagonistic pleiotropy as a widespread mechanism for the maintenance of polymorphic disease alleles. *BMC Medical Genetics* **12**.1: 160. DOI: 10.1186/1471-2350-12-160.

Castedo-Dorado F., Anta M. B., Parresol B. R., and González J. G. Á. (2005). A stochastic height-diameter model for maritime pine ecoregions in Galicia (northwestern Spain). *Annals of Forest Science* **62**.5: 455– 465. DOI: 10.1051/forest:2005042.

Ceballos G., Ehrlich P. R., and Dirzo R. (2017). Biological annihilation via the ongoing sixth mass extinction signaled by vertebrate population losses and declines. *Proceedings of the National Academy of Sciences* **114**.30: E6089– E6096. DOI: 10.1073/pnas.1704949114.

Chakraborty M. and Fry J. D. (2016). Evidence that Environmental Heterogeneity Maintains a Detoxifying Enzyme Polymorphism in Drosophila melanogaster. *Current Biology* **26**.2: 219– 223. DOI: 10.1016/j.cub.2015.11.049.

Chambel M. R., Climent J., and Alía R. (2007). Divergence among species and populations of Mediterranean pines in biomass allocation of seedlings grown under two watering regimes. *Annals of Forest Science* **64**.1: 87– 97. DOI: 10.1051/forest:2006092.

Changenet A., Ruiz-Benito P., Ratcliffe S., Fréjaville T., Archambeau J., Porte A. J., Zavala M. A., Dahlgren J., Lehtonen A., and Benito Garzón M. (2021). Occurrence but not intensity of mortality rises towards the climatic trailing edge of tree species ranges in European forests. *Global Ecology and Biogeography* **30**.7: 1356– 1374. DOI: 10.1111/geb.13301.

Charlesworth D., Vekemans X., Castric V., and Glémin S. (2005). Plant self-incompatibility systems: a molecular evolutionary perspective. *New Phytologist* **168**.1: 61– 69. DOI: 10.1111/j.1469-8137.2005.01443.x.

Charmantier A., Kruuk L. E. B., Blondel J., and Lambrechts M. M. (2004). Testing for microevolution in body size in three blue tit populations. *Journal of Evolutionary Biology* **17**.4: 732–743. DOI: 10.1111/j.1420-9101.2004.00734.x.

Charmantier A. and Garant D. (2005). Environmental quality and evolutionary potential: lessons from wild populations. *Proceedings of the Royal Society B: Biological Sciences* **272**.1571: 1415–1425. DOI: 10.1098/rspb.2005.3117.

Chen I.-C., Hill J. K., Ohlemüller R., Roy D. B., and Thomas C. D. (2011). Rapid Range Shifts of Species Associated with High Levels of Climate Warming. *Science* **333**.6045: 1024– 1026. DOI: 10.1126/science.1206432.

Chen J., Tsuda Y., Stocks M., Källman T., Xu N., Kärkkäinen K., Huotari T., Semerikov V. L., Vendramin G. G., and Lascoux M. (2014). Clinal Variation at Phenology-Related Genes in Spruce: Parallel Evolution in FTL2 and Gigantea? *Genetics* **197**.3: 1025– 1038. DOI: 10.1534/genetics.114.163063.

Cheptou P.-O., Carrue O., Rouifed S., and Cantarel A. (2008). Rapid evolution of seed dispersal in an urban environment in the weed Crepis sancta. *Proceedings of the National Academy of Sciences* **105**.10: 3796– 3799. DOI: 10.1073/pnas.0708446105.

Chevalet C. (1994). An approximate theory of selection assuming a finite number of quantitative trait loci: 22.

Chevin L.-M. (2013). Genetic Constraints on Adaptation to a Changing Environment. *Evolution* **67**.3: 708– 721. DOI: 10.1111/j.1558-5646.2012.01809.x.

Chevin L.-M. and Hoffmann A. A. (2017). Evolution of phenotypic plasticity in extreme environments. *Philosophical Transactions of the Royal Society B: Biological Sciences* **372**.1723: 20160138. DOI: 10.1098/rstb.2016.0138.

Chevin L.-M., Lande R., and Mace G. M. (2010). Adaptation, Plasticity, and Extinction in a Changing Environment: Towards a Predictive Theory. *PLOS Biology* **8**.4: e1000357. DOI: 10.1371/journal.pbio.1000357.

Clark J. S. (2005). Why environmental scientists are becoming Bayesians. *Ecology Letters* **8**.1: 2–14. DOI: 10.1111/j.1461-0248.2004.00702.x.

Colautti R. I. and Barrett S. C. H. (2013). Rapid Adaptation to Climate Facilitates Range Expansion of an Invasive Plant. *Science* **342**.6156: 364– 366. DOI: 10.1126/science.1242121.

Colautti R. I., Eckert C. G., and Barrett S. C. H. (2010). Evolutionary constraints on adaptive evolution during range expansion in an invasive plant. *Proceedings of the Royal Society B: Biological Sciences* **277**.1689: 1799– 1806. DOI: 10.1098/rspb.2009.2231.

Conover D. O., Duffy T. A., and Hice L. A. (2009). The Covariance between Genetic and Environmental Influences across Ecological Gradients: Reassessing the Evolutionary Significance of Countergradient and Cogradient Variation. *Annals of the New York Academy of Sciences* **1168**.1: 100– 129. DOI: 10.1111/j.1749-6632.2009.04575.x.

Conover D. O. and Schultz E. T. (1995). Phenotypic similarity and the evolutionary significance of countergradient variation. *Trends in Ecology & Evolution* **10**.6: 248– 252. DOI: 10.1016/S0169-5347(00)89081-3.

Conrad O., Bechtel B., Bock M., Dietrich H., Fischer E., Gerlitz L., Wehberg J., Wichmann V., and Böhner J. (2015). System for Automated Geoscientific Analyses (SAGA) v. 2.1.4. *Geoscientific Model Development* **8**.7: 1991– 2007. DOI: 10.5194/gmd-8-1991-2015.

Corcuera L., Gil-Pelegrin E., and Notivol E. (2010). Phenotypic plasticity in Pinus pinaster δ13C: environment modulates genetic variation. *Annals of Forest Science* **67**.8: 812– 812. DOI: 10.1051/forest/2010048.

Cordell H. J. (2002). Epistasis: what it means, what it doesn't mean, and statistical methods to detect it in humans. *Human Molecular Genetics* **11**.20: 2463– 2468. DOI: 10.1093/hmg/11.20.2463.

Cornelius J. (1994). Heritabilities and additive genetic coefficients of variation in forest trees. *Canadian Journal of Forest Research* **24**.2: 372– 379. DOI: 10.1139/x94-050.

Correia I., Alía R., Yan W., David T., Aguiar A., and Almeida M. H. (2010). Genotype × Environment interactions in Pinus pinaster at age 10 in a multienvironment trial in Portugal: a maximum likelihood approach. *Annals of Forest Science* **67**.6: 612– 612. DOI: 10.1051/forest/2010025.

Costa P. and Durel C. E. (2011). Time trends in genetic control over height and diameter in maritime pine. *Canadian Journal of Forest Research*. DOI: 10.1139/x26-135.

Coulson T., Kendall B. E., Barthold J., Plard F., Schindler S., Ozgul A., and Gaillard J.-M. (2017). Modeling Adaptive and Nonadaptive Responses of Populations to Environmental Change. *The American Naturalist* **190**.3: 313– 336. DOI: 10.1086/692542.

Crespi B. J. (2000). The evolution of maladaptation. *Heredity* **84**.6: 623– 629. DOI: 10.1046/j.1365-2540.2000.00746.x.

Crispo E. (2008). Modifying effects of phenotypic plasticity on interactions among natural selection, adaptation and gene flow. *Journal of Evolutionary Biology* **21**.6: 1460– 1469. DOI: 10.1111/j.1420-9101.2008.01592.x.

Crispo E. (2007). The Baldwin Effect and Genetic Assimilation: Revisiting Two Mechanisms of Evolutionary Change Mediated by Phenotypic Plasticity. *Evolution* **61**.11: 2469– 2479. DOI: 10.1111/j.1558-5646.2007.00203.x.

Crow J. F. and Kimura M. (1970). An introduction to population genetics theory. *An introduction to population genetics theory.*

Csilléry K., Ovaskainen O., Sperisen C., Buchmann N., Widmer A., and Gugerli F. (2020). Adaptation to local climate in multi-trait space: evidence from silver fir (Abies alba Mill.) populations across a heterogeneous environment. *Heredity* **124**.1: 77– 92. DOI: 10.1038/s41437-019-0240-0.

Curtin S. J., Tiffin P., Guhlin J., Trujillo D. I., Burghardt L. T., Atkins P., Baltes N. J., Denny R., Voytas D. F., Stupar R. M., and Young N. D. (2017). Validating Genome-Wide Association Candidates Controlling Quantitative Variation in Nodulation. *Plant Physiology* **173**.2: 921–931. DOI: 10.1104/pp.16.01923.

Dai Z., Long N., and Huang W. (2020). Influence of Genetic Interactions on Polygenic Prediction. *G3 Genes|Genomes|Genetics* **10**.1: 109– 115. DOI: 10.1534/g3.119.400812.

Darwin C. R. (1859). *On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life.* John Murray, London, UK.

Darwin C. R. and Wallace A. R. (1858). On the tendency of species to form varieties; and on the perpetuation of varieties and species by natural means of selection. *Journal of the Proceedings of the Linnean Society of London, Zoology* **3**.9: 45– 62.

Dauphin B., Rellstab C., Schmid M., Zoller S., Karger D. N., Brodbeck S., Guillaume F., and Gugerli F. (2021). Genomic vulnerability to rapid climate warming in a tree species with a long generation time. *Global Change Biology* **27**.6: 1181– 1195. DOI: 10.1111/gcb.15469.

Davey J. W., Hohenlohe P. A., Etter P. D., Boone J. Q., Catchen J. M., and Blaxter M. L. (2011). Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nature Reviews Genetics* **12**.7: 499– 510. DOI: 10.1038/nrg3012.

Davis M. B. and Shaw R. G. (2001). Range Shifts and Adaptive Responses to Quaternary Climate Change. *Science* **292**.5517: 673– 679. DOI: 10.1126/science.292.5517.673.

De La Torre A. R., Birol I., Bousquet J., Ingvarsson P. K., Jansson S., Jones S. J., Keeling C. I., MacKay J., Nilsson O., Ritland K., Street N., Yanchuk A., Zerbe P., and Bohlmann J. (2014). Insights into Conifer Giga-Genomes. *Plant Physiology* **166**.4: 1724– 1732. DOI: 10.1104/pp.114.248708.

De la Mata R., Voltas J., and Zas R. (2012). Phenotypic plasticity and climatic adaptation in an Atlantic maritime pine breeding population. *Annals of Forest Science* **69**.4: 477– 487. DOI: 10.1007/s13595-011-0173-0.

De Rubeis S., He X., Goldberg A. P., Poultney C. S., Samocha K., Ercument Cicek A., Kou Y., Liu L., Fromer M., Walker S., Singh T., Klei L., Kosmicki J., Fu S.-C., Aleksic B., Biscaldi M., Bolton P. F., Brownfeld J. M., Cai J., Campbell N. G., Carracedo A., Chahrour M. H., Chiocchetti A. G., Coon H., Crawford E. L., Crooks L., Curran S. R., Dawson G., Duketis E., Fernandez B. A., Gallagher L., Geller E., Guter S. J., Sean Hill R., Ionita-Laza I., Jimenez Gonzalez P., Kilpinen H., Klauck S. M., Kolevzon A., Lee I., Lei J., Lehtimäki T., Lin C.-F., Ma'ayan A., Marshall C. R., McInnes A. L., Neale B., Owen M. J., Ozaki N., Parellada M., Parr J. R., Purcell S., Puura K., Rajagopalan D., Rehnström K., Reichenberg A., Sabo A., Sachse M., Sanders S. J., Schafer C., Schulte-Rüther M., Skuse D., Stevens C., Szatmari P., Tammimies K., Valladares O., Voran A., Wang L.-S., Weiss L. A., Jeremy Willsey A., Yu T. W., Yuen R. K. C., Cook E. H., Freitag C. M., Gill M., Hultman C. M., Lehner T., Palotie A., Schellenberg G. D., Sklar P., State M. W., Sutcliffe J. S., Walsh C. A., Scherer S. W., Zwick M. E., Barrett J. C., Cutler D. J., Roeder K., Devlin B., Daly M. J., and Buxbaum J. D. (2014). Synaptic, transcriptional and chromatin genes disrupted in autism. *Nature* **515**.7526: 209– 215. DOI: 10.1038/nature13772.

De Jong G. (1990). Genotype-by-environment interaction and the genetic covariance between environments: multilocus genetics. *Genetica* **81**.3: 171– 177. DOI: 10.1007/BF00360862.

De Miguel M., Rodríguez-Quilón I., Heuertz M., Hurel A., Grivet D., Jaramillo-Correa J.-P., Vendramin G. G., Plomion C., Majada J., Alía R., Eckert A. J., and González-Martínez S. C. (2020). Polygenic adaptation and negative selection across traits, years and environments in a long-lived plant species (Pinus pinaster Ait.) *bioRxiv*. DOI: 10.1101/2020.03.02.974113.

De Villemereuil P., Gaggiotti O. E., Mouterde M., and Till-Bottraud I. (2016). Common garden experiments in the genomic era: new perspectives and opportunities. *Heredity* **116**.3: 249– 254. DOI: 10.1038/hdy.2015.93.

De Villemereuil P. (2019). On the relevance of Bayesian statistics and MCMC for animal models. *Journal of Animal Breeding and Genetics* **136**.5: 339– 340. DOI: 10.1111/jbg.12426.

De Villemereuil P., Frichot É., Bazin É., François O., and Gaggiotti O. E. (2014). Genome scan methods against more complex models: when and how much should we trust them? *Molecular Ecology* **23**.8: 2006– 2019. DOI: 10.1111/mec.12705.

DeBiasse M. B. and Kelly M. W. (2016). Plastic and Evolved Responses to Global Change: What Can We Learn from Comparative Transcriptomics? *Journal of Heredity* **107**.1: 71– 81. DOI: 10.1093/jhered/esv073.

Delph L. F. and Kelly J. K. (2014). On the importance of balancing selection in plants. *New Phytologist* **201**.1: 45– 56. DOI: 10.1111/nph.12441.

Depardieu C., Gérardi S., Nadeau S., Parent G. J., Mackay J., Lenz P., Lamothe M., Girardin M. P., Bousquet J., and Isabel N. (2021). Connecting tree-ring phenotypes, genetic associations and transcriptomics to decipher the genomic architecture of drought adaptation in a widespread conifer. *Molecular Ecology* **30**.16: 3898– 3917. DOI: 10.1111/mec.15846.

Derry A. M., Fraser D. J., Brady S. P., Astorg L., Lawrence E. R., Martin G. K., Matte J.-M., Dastis J. O. N., Paccard A., Barrett R. D. H., Chapman L. J., Lane J. E., Ballas C. G., Close M., and Crispo E. (2019). Conservation through the lens of (mal)adaptation: Concepts and meta-analysis. *Evolutionary Applications* **12**.7: 1287– 1304. DOI: 10.1111/eva.12791.

Des Marais D. L., Hernandez K. M., and Juenger T. E. (2013). Genotype-by-Environment Interaction and Plasticity: Exploring Genomic Responses of Plants to the Abiotic Environment. *Annual Review of Ecology, Evolution, and Systematics* **44**.1: 5– 29. DOI: 10.1146/annurev-ecolsys-110512-135806.

DeWitt T. J. and Scheiner S. M. (2004). *Phenotypic Plasticity: Functional and Conceptual Approaches*. Oxford University Press. 262 pp. ISBN: 978-0-19-513896-2.

DeWitt T. J., Sih A., and Wilson D. S. (1998). Costs and limits of phenotypic plasticity. *Trends in Ecology & Evolution* **13**.2: 77– 81. DOI: 10.1016/S0169-5347(97)01274-3.

El-Dien O. G., Ratcliffe B., Klápště J., Porth I., Chen C., and El-Kassaby Y. A. (2016). Implementation of the Realized Genomic Relationship Matrix to Open-Pollinated White Spruce Family Testing for Disentangling Additive from Nonadditive Genetic Effects. *G3: Genes, Genomes, Genetics* **6**.3: 743– 753. DOI: 10.1534/g3.115.025957.

— (2018). Multienvironment genomic variance decomposition analysis of open-pollinated Interior spruce (Picea glauca x engelmannii). *Molecular Breeding* **38**.3: 26. DOI: 10.1007/s11032-018-0784-3.

Dirzo R., Young H. S., Galetti M., Ceballos G., Isaac N. J. B., and Collen B. (2014). Defaunation in the Anthropocene. *Science* **345**.6195: 401– 406. DOI: 10.1126/science.1251817.

Doak D. F. and Morris W. F. (2010). Demographic compensation and tipping points in climate-induced range shifts. *Nature* **467**.7318: 959– 962. DOI: 10.1038/nature09439.

Dobrowski S. Z., Thorne J. H., Greenberg J. A., Safford H. D., Mynsberge A. R., Crimmins S. M., and Swanson A. K. (2011). Modeling plant ranges over 75 years of climate change in California, USA: temporal transferability and species traits. *Ecological Monographs* **81**.2: 241– 257. DOI: 10.1890/10-1325.1.

Dray S., Bauman D., Blanchet G., Borcard D., Clappe S., Guenard G., Jombart T., Larocque G., Legendre P., Madi N., and Wagner H. H. (2020). adespatial: Multivariate Multiscale Spatial Analysis. R package version 0.3-8.

Dray S., Legendre P., and Peres-Neto P. R. (2006). Spatial modelling: a comprehensive framework for principal coordinate analysis of neighbour matrices (PCNM). *Ecological Modelling* **196**.3: 483– 493. DOI: 10.1016/j.ecolmodel.2006.02.015.

Drouineau H., Durif C., Castonguay M., Mateo M., Rochard E., Verreault G., Yokouchi K., and Lambert P. (2018). Freshwater eels: A symbol of the effects of global change. *Fish and Fisheries* **19**.5: 903– 930. DOI: 10.1111/faf.12300.

Eckert A. J., Heerwaarden J. van, Wegrzyn J. L., Nelson C. D., Ross-Ibarra J., González-Martínez S. C., and Neale D. B. (2010). Patterns of Population Structure and Environmental Associations to Aridity Across the Range of Loblolly Pine (Pinus taeda L., Pinaceae). *Genetics* **185**.3: 969– 982. DOI: 10.1534/genetics.110.115543.

Edelaar P. and Bolnick D. I. (2012). Non-random gene flow: an underappreciated force in evolution and ecology. *Trends in Ecology & Evolution* **27**.12: 659– 665. DOI: 10.1016/j.tree.2012.07.009.

— (2019). Appreciating the Multiple Processes Increasing Individual or Population Fitness. *Trends in Ecology & Evolution* **34**.5: 435– 446. DOI: 10.1016/j.tree.2019.02.001.

Ehrenreich I. M. and Pfennig D. W. (2016). Genetic assimilation: a review of its potential proximate causes and evolutionary consequences. *Annals of Botany* **117**.5: 769– 779. DOI: 10.1093/aob/mcv130.

Ellis N., Smith S. J., and Pitcher C. R. (2012). Gradient forests: calculating importance gradients on physical predictors. *Ecology* **93**.1: 156– 168. DOI: 10.1890/11-0252.1.

Ellison A. M. (2004). Bayesian inference in ecology. *Ecology Letters* **7**.6: 509– 520. DOI: 10.1111/j.1461-0248.2004.00603.x.

Elshire R. J., Glaubitz J. C., Sun Q., Poland J. A., Kawamoto K., Buckler E. S., and Mitchell S. E. (2011). A Robust, Simple Genotyping-by-Sequencing (GBS) Approach for High Diversity Species. *PLOS ONE* **6**.5: e19379. DOI: 10.1371/journal.pone.0019379.

Ensing D. J. and Eckert C. G. (2019). Interannual variation in season length is linked to strong co-gradient plasticity of phenology in a montane annual plant. *New Phytologist* **224**.3: 1184–1200. DOI: 10.1111/nph.16009.

Estes S. and Arnold S. J. (2007). Resolving the Paradox of Stasis: Models with Stabilizing Selection Explain Evolutionary Divergence on All Timescales. *The American Naturalist* **169**.2: 227–244. DOI: 10.1086/510633.

Exposito-Alonso M., Burbano H. A., Bossdorf O., Nielsen R., and Weigel D. (2018a). A map of climate change-driven natural selection in Arabidopsis thaliana. *bioRxiv*. DOI: 10.1101/321133.

Exposito-Alonso M., Burbano H. A., Bossdorf O., Nielsen R., and Weigel D. (2019). Natural selection on the Arabidopsis thaliana genome in present and future climates. *Nature* **573**.7772: 126– 129. DOI: 10.1038/s41586-019-1520-9.

Exposito-Alonso M., Vasseur F., Ding W., Wang G., Burbano H. A., and Weigel D. (2018b). Genomic basis and evolutionary potential for extreme drought adaptation in Arabidopsis thaliana. *Nature Ecology & Evolution* **2**.2: 352– 358. DOI: 10.1038/s41559-017-0423-0.

Falconer D. S. and Mackay T. F. (1996). *Introduction to quantitative genetics*. Vol. Fourth Edition. Pearson Education Limited.

Falush D., Stephens M., and Pritchard J. K. (2003). Inference of Population Structure Using Multilocus Genotype Data: Linked Loci and Correlated Allele Frequencies. *Genetics* **164**.4: 1567– 1587. DOI: 10.1093/genetics/164.4.1567.

FAO (2020). *Global Forest Resources Assessment 2020: Main report*. Rome, Italy: FAO. 184 pp. ISBN: 978-92-5-132974-0. DOI: 10.4060/ca9825en.

Felsenstein J. (1976). The theoretical population genetics of variable selection and migration. *Annual review of genetics* **10**.1: 253– 280.

Ferrier S., Manion G., Elith J., and Richardson K. (2007). Using generalized dissimilarity modelling to analyse and predict patterns of beta diversity in regional biodiversity assessment. *Diversity and Distributions* **13**.3: 252– 264. DOI: 10.1111/j.1472-4642.2007.00341.x.

Fick S. E. and Hijmans R. J. (2017). WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology* **37**.12: 4302– 4315. DOI: 10.1002/joc.5086.

Fijarczyk A. and Babik W. (2015). Detecting balancing selection in genomes: limits and prospects. *Molecular Ecology* **24**.14: 3529– 3545. DOI: 10.1111/mec.13226.

Fisher R. A. (1918). The Correlation between Relatives on the Supposition of Mendelian Inheritance. *Transactions of The Royal Society of Edinburgh* **52**.2: 399– 433. DOI: 10.1017/S0080456800012163.

Fisher R. A. (1930). *The genetical theory of natural selection*. Oxford: Clarendon Press. 289 pp. ISBN: 978-1-176-62502-0.

Fitzpatrick M. C., Chhatre V. E., Soolanayakanahally R. Y., and Keller S. R. (2021). Experimental support for genomic prediction of climate maladaptation using the machine learning approach Gradient Forests. *Molecular Ecology Resources* **n/a**.n/a. DOI: 10.1111/1755-0998.13374.

Fitzpatrick M. C. and Keller S. R. (2015). Ecological genomics meets community-level modelling of biodiversity: mapping the genomic landscape of current and future environmental adaptation. *Ecology Letters* **18**.1: 1– 16. DOI: 10.1111/ele.12376.

Fitzpatrick M. C., Keller S. R., and Lotterhos K. E. (2018). Comment on "Genomic signals of selection predict climate-driven population declines in a migratory bird". *Science* **361**.6401. DOI: 10.1126/science.aat7279.

Fitzpatrick M. C., Mokany K., Manion G., Lisk M., Ferrier S., and Nieto-Lugilde D. (2020). gdm: Generalized Dissimilarity Modeling. R package version 1.4.2.

Foden W. B., Young B. E., Akçakaya H. R., Garcia R. A., Hoffmann A. A., Stein B. A., Thomas C. D., Wheatley C. J., Bickford D., Carr J. A., Hole D. G., Martin T. G., Pacifici M., Pearce-Higgins J. W., Platts P. J., Visconti P., Watson J. E. M., and Huntley B. (2019). Climate change vulnerability assessment of species. *WIREs Climate Change* **10**.1: e551. DOI: 10.1002/wcc.551.

Forester B. R., Lasky J. R., Wagner H. H., and Urban D. L. (2018). Comparing methods for detecting multilocus adaptation with multivariate genotype–environment associations. *Molecular Ecology* **27**.9: 2215– 2233. DOI: 10.1111/mec.14584.

Fox R. J., Donelson J. M., Schunter C., Ravasi T., and Gaitán-Espitia J. D. (2019). Beyond buying time: the role of plasticity in phenotypic adaptation to rapid environmental change. *Philosophical Transactions of the Royal Society B: Biological Sciences* **374**.1768: 20180174. DOI: 10.1098/rstb.2018.0174.

Franklin J., Serra-Diaz J. M., Syphard A. D., and Regan H. M. (2016). Global change and terrestrial plant community dynamics. *Proceedings of the National Academy of Sciences* **113**.14: 3725–3734. DOI: 10.1073/pnas.1519911113.

Franks S. J. and Hoffmann A. A. (2012). Genetics of Climate Change Adaptation. *Annual Review of Genetics* **46**.1: 185– 208. DOI: 10.1146/annurev-genet-110711-155511.

Franks S. J., Sim S., and Weis A. E. (2007). Rapid evolution of flowering time by an annual plant in response to a climate fluctuation. *Proceedings of the National Academy of Sciences* **104**.4: 1278– 1282. DOI: 10.1073/pnas.0608379104.

Franks S. J., Weber J. J., and Aitken S. N. (2014). Evolutionary and plastic responses to climate change in terrestrial plant populations. *Evolutionary Applications* **7**.1: 123– 139. DOI: 10.1111/eva.12112.

Frazer K. A., Murray S. S., Schork N. J., and Topol E. J. (2009). Human genetic variation and its contribution to complex traits. *Nature Reviews Genetics* **10**.4: 241– 251. DOI: 10.1038/nrg2554.

Fréjaville T. and Benito Garzón M. (2018). The EuMedClim Database: Yearly Climate Data (1901–2014) of 1 km Resolution Grids for Europe and the Mediterranean Basin. *Frontiers in Ecology and Evolution* **6**: 31. DOI: 10.3389/fevo.2018.00031.

Fréjaville T., Vizcaíno-Palomar N., Fady B., Kremer A., and Garzón M. B. (2020). Range margin populations show high climate adaptation lags in European trees. *Global Change Biology* **26**.2: 484– 495. DOI: 10.1111/gcb.14881.

Frichot E., Schoville S. D., Bouchard G., and François O. (2013). Testing for Associations between Loci and Environmental Gradients Using Latent Factor Mixed Models. *Molecular Biology and Evolution* **30**.7: 1687– 1699. DOI: 10.1093/molbev/mst063.

Fu W. and Akey J. M. (2013). Selection and Adaptation in the Human Genome. *Annual Review of Genomics and Human Genetics* **14**.1: 467– 489. DOI: 10.1146/annurev-genom-091212-153509.

Fuller Z. L., Mocellin V. J. L., Morris L. A., Cantin N., Shepherd J., Sarre L., Peng J., Liao Y., Pickrell J., Andolfatto P., Matz M., Bay L. K., and Przeworski M. (2020). Population genetics

of the coral Acropora millepora: Toward genomic prediction of bleaching. *Science* **369**.6501. DOI: 10.1126/science.aba4674.

Gabry J., Simpson D., Vehtari A., Betancourt M., and Gelman A. (2019). Visualization in Bayesian workflow. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* **182**.2: 389–402. DOI: 10.1111/rssa.12378.

Gain C. and François O. (2021). LEA 3: Factor models in population genetics and ecological genomics with R. *Molecular Ecology Resources* **21**.8: 2738– 2748. DOI: 10.1111/1755-0998.13366.

García-Ramos G. and Kirkpatrick M. (1997). Genetic Models of Adaptation and Gene Flow in Peripheral Populations. *Evolution* **51**.1: 21– 28. DOI: 10.1111/j.1558-5646.1997.tb02384.x.

García-Ruiz A., Cole J. B., VanRaden P. M., Wiggans G. R., Ruiz-López F. J., and Tassell C. P. V. (2016). Changes in genetic selection differentials and generation intervals in US Holstein dairy cattle as a result of genomic selection. *Proceedings of the National Academy of Sciences* **113**.28: E3995– E4004. DOI: 10.1073/pnas.1519061113.

García-Valdés R., Zavala M. A., Araújo M. B., and Purves D. W. (2013). Chasing a moving target: projecting climate change-induced shifts in non-equilibrial tree species distributions. *Journal of Ecology* **101**.2: 441– 453. DOI: 10.1111/1365-2745.12049.

Gaspar M. J., Velasco T., Feito I., Alía R., and Majada J. (2013). Genetic Variation of Drought Tolerance in Pinus pinaster at Three Hierarchical Levels: A Comparison of Induced Osmotic Stress and Field Testing. *PLOS ONE* **8**.11: e79094. DOI: 10.1371/journal.pone.0079094.

Gatti L. V., Basso L. S., Miller J. B., Gloor M., Gatti Domingues L., Cassol H. L. G., Tejada G., Aragão L. E. O. C., Nobre C., Peters W., Marani L., Arai E., Sanches A. H., Corrêa S. M., Anderson L., Von Randow C., Correia C. S. C., Crispim S. P., and Neves R. A. L. (2021). Amazonia as a carbon source linked to deforestation and climate change. *Nature* **595**.7867: 388– 393. DOI: 10.1038/s41586-021-03629-6.

Gattuso J.-P., Magnan A., Billé R., Cheung W. W. L., Howes E. L., Joos F., Allemand D., Bopp L., Cooley S. R., Eakin C. M., Hoegh-Guldberg O., Kelly R. P., Pörtner H.-O., Rogers A. D., Baxter J. M., Laffoley D., Osborn D., Rankovic A., Rochette J., Sumaila U. R., Treyer S., and Turley C. (2015). Contrasting futures for ocean and society from different anthropogenic CO2 emissions scenarios. *Science* **349**.6243. DOI: 10.1126/science.aac4722.

Gautier M. (2015). BayPass Genome-Wide Scan for Adaptive Differentiation and Association Analysis with population-specific covariables (en lien avec la publication Gautier M. 2015. Genome-wide scan for adaptive divergence and association with population-specific covariates. Genetics 201 (4): 1555-1579 (10.1534/genetics.115.181453).

Gauzere J., Teuf B., Davi H., Chevin L.-M., Caignard T., Leys B., Delzon S., Ronce O., and Chuine I. (2020). Where is the optimum? Predicting the variation of selection along climatic gradients and the adaptive value of plasticity. A case study on tree phenology. *Evolution Letters* **4**.2: 109– 123. DOI: 10.1002/evl3.160.

Gelber J. (2015). Aristotle on Essence and Habitat. *Oxford Studies in Ancient Philosophy* **48**: 267–293.

Gelman A. (2006). Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper). *Bayesian Analysis* **1**.3: 515– 534. DOI: 10.1214/06-BA117A.

Gelman A., Carlin J. B., Stern H. S., Dunson D. B., Vehtari A., and Rubin D. B. (2020a). *Bayesian Data Analysis, Third Edition*. Chapman & Hall.

Gelman A., Jakulin A., Pittau M. G., and Su Y.-S. (2008). A weakly informative default prior distribution for logistic and other regression models. *The Annals of Applied Statistics* **2**.4: 1360– 1383. DOI: 10.1214/08-AOAS191.

Gelman A., Vehtari A., Simpson D., Margossian C. C., Carpenter B., Yao Y., Kennedy L., Gabry J., Bürkner P.-C., and Modrák M. (2020b). *Bayesian Workflow*. URL: http://arxiv.org/abs/2011.01808 (visited on 11/15/2021).

Ghalambor C. K., McKAY J. K., Carroll S. P., and Reznick D. N. (2007). Adaptive versus non-adaptive phenotypic plasticity and the potential for contemporary adaptation in new environments. *Functional Ecology* **21**.3: 394– 407. DOI: 10.1111/j.1365-2435.2007.01283.x.

Ghalambor C. K., Hoke K. L., Ruell E. W., Fischer E. K., Reznick D. N., and Hughes K. A. (2015). Non-adaptive plasticity potentiates rapid adaptive evolution of gene expression in nature. *Nature* **525**.7569: 372– 375. DOI: 10.1038/nature15256.

Gibbard S., Caldeira K., Bala G., Phillips T. J., and Wickett M. (2005). Climate effects of global land cover change. *Geophysical Research Letters* **32**.23. DOI: 10.1029/2005GL024550.

Gibson G. (2012). Rare and common variants: twenty arguments. *Nature Reviews Genetics* **13**.2: 135– 145. DOI: 10.1038/nrg3118.

Gibson L., Lee T. M., Koh L. P., Brook B. W., Gardner T. A., Barlow J., Peres C. A., Bradshaw C. J. A., Laurance W. F., Lovejoy T. E., and Sodhi N. S. (2011). Primary forests are irreplaceable for sustaining tropical biodiversity. *Nature* **478**.7369: 378– 381. DOI: 10.1038/nature10425.

Gienapp P., Calus M. P. L., Laine V. N., and Visser M. E. (2019). Genomic selection on breeding time in a wild bird population. *Evolution Letters* **3**.2: 142– 151. DOI: 10.1002/evl3.103.

Gienapp P., Fior S., Guillaume F., Lasky J. R., Sork V. L., and Csilléry K. (2017). Genomic Quantitative Genetics to Study Evolution in the Wild. *Trends in Ecology & Evolution* **32**.12: 897– 908. DOI: 10.1016/j.tree.2017.09.004.

Giglio L., Randerson J. T., and Werf G. R. van der (2013). Analysis of daily, monthly, and annual burned area using the fourth-generation global fire emissions database (GFED4). *Journal of Geophysical Research: Biogeosciences* **118**.1: 317– 328. DOI: 10.1002/jgrg.20042.

Gingerich P. D. (2009). Rates of Evolution. *Annual Review of Ecology, Evolution, and Systematics* **40**.1: 657– 675. DOI: 10.1146/annurev.ecolsys.39.110707.173457.

Gömöry D., Longauer R., Hlásny T., Pacalaj M., Strmeň S., and Krajmerová D. (2012). Adaptation to common optimum in different populations of Norway spruce (Picea abies Karst.) *European Journal of Forest Research* **131**.2: 401– 411. DOI: 10.1007/s10342-011-0512-6.

Gomulkiewicz R. and Holt R. D. (1995). When does Evolution by Natural Selection Prevent Extinction? *Evolution* **49**.1: 201– 207. DOI: 10.2307/2410305.

Gomulkiewicz R., Holt R. D., Barfield M., and Nuismer S. L. (2010). Genetics, adaptation, and invasion in harsh environments. *Evolutionary Applications* **3**.2: 97– 108. DOI: 10.1111/j.1752-4571.2009.00117.x.

Gonzalez A., Ronce O., Ferriere R., and Hochberg M. E. (2013). Evolutionary rescue: an emerging focus at the intersection between ecology and evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences* **368**.1610: 20120404. DOI: 10.1098/rstb.2012.0404.

González-Martínez S. C., Alía R., and Gil L. (2002). Population genetic structure in a Mediterranean pine (Pinus pinaster Ait.): a comparison of allozyme markers and quantitative traits. *Heredity* **89**.3: 199– 206. DOI: 10.1038/sj.hdy.6800114.

González-Martínez S. C., Burczyk J., Nathan R., Nanos N., Gil L., and Alía R. (2006). Effective gene dispersal and female reproductive success in Mediterranean maritime pine (Pinus pinaster Aiton). *Molecular Ecology* **15**.14: 4577– 4588. DOI: 10.1111/j.1365-294X.2006.03118.x.

González-Martínez S. C., Gómez A., Carrión J. S., Agúndez D., Alía R., and Gil L. (2007). Spatial genetic structure of an explicit glacial refugium of maritime pine (Pinus pinaster Aiton) insoutheastern Spain. *Phylogeography of Southern European Refugia: Evolutionary perspectives on the origins and conservation of European biodiversity*. Ed. by Weiss S. and Ferrand N. Dordrecht: Springer Netherlands. Page 257– 269. ISBN: 978-1-4020-4904-0. DOI: 10.1007/1-4020-4904-8_9.

Goudet J. (2005). hierfstat, a package for r to compute and test hierarchical F-statistics. *Molecular Ecology Notes* **5**.1: 184– 186. DOI: 10.1111/j.1471-8286.2004.00828.x.

Gougherty A. V., Chhatre V. E., Keller S. R., and Fitzpatrick M. C. (2020a). Contemporary range position predicts the range-wide pattern of genetic diversity in balsam poplar (Populus balsamifera L.) *Journal of Biogeography* **47**.6: 1246– 1257. DOI: 10.1111/jbi.13811.

Gougherty A. V., Keller S. R., Chhatre V. E., and Fitzpatrick M. C. (2020b). Future climate change promotes novel gene-climate associations in balsam poplar (Populus balsamifera L.), a forest tree species. *bioRxiv*: 2020.02.28.961060. DOI: 10.1101/2020.02.28.961060.

Grady K. C., Kolb T. E., Ikeda D. H., and Whitham T. G. (2015). A bridge too far: cold and pathogen constraints to assisted migration of riparian forests. *Restoration Ecology* **23**.6: 811– 820. DOI: 10.1111/rec.12245.

Grant P. R. and Grant B. R. (1995). Predicting Microevolutionary Responses to Directional Selection on Heritable Variation. *Evolution* **49**.2: 241– 251. DOI: 10.1111/j.1558-5646.1995.tb02236.x.

— (2002). Unpredictable Evolution in a 30-Year Study of Darwin's Finches. *Science* **296**.5568: 707– 711. DOI: 10.1126/science.1070315.

Grattapaglia D., Plomion C., Kirst M., and Sederoff R. R. (2009). Genomics of growth traits in forest trees. *Current Opinion in Plant Biology* **12**.2: 148– 156. DOI: 10.1016/j.pbi.2008.12.008.

Grattapaglia D. and Resende M. D. V. (2011). Genomic selection in forest tree breeding. *Tree Genetics & Genomes* **7**.2: 241– 255. DOI: 10.1007/s11295-010-0328-4.

Gratten J. and Visscher P. M. (2016). Genetic pleiotropy in complex traits and diseases: implications for genomic medicine. *Genome Medicine* **8**.1: 78. DOI: 10.1186/s13073-016-0332-x.

Grether G. F. (2005). Environmental Change, Phenotypic Plasticity, and Genetic Compensation. *The American Naturalist* **166**.4: E115– E123. DOI: 10.1086/432023.

Grivet D., Climent J., Zabal-Aguirre M., Neale D. B., Vendramin G. G., and González-Martínez S. C. (2013). Adaptive evolution of Mediterranean pines. *Molecular Phylogenetics and Evolution* **68**.3: 555– 566. DOI: 10.1016/j.ympev.2013.03.032.

Grivet D., Sebastiani F., Alía R., Bataillon T., Torre S., Zabal-Aguirre M., Vendramin G. G., and González-Martínez S. C. (2011). Molecular Footprints of Local Adaptation in Two Mediterranean Conifers. *Molecular Biology and Evolution* **28**.1: 101– 116. DOI: 10.1093/molbev/msq190.

Guan Y. and Stephens M. (2011). Bayesian variable selection regression for genome-wide association studies and other large-scale problems. *The Annals of Applied Statistics* **5**.3: 1780– 1815. DOI: 10.1214/11-AOAS455.

Gugger P. F., Peñaloza-Ramírez J. M., Wright J. W., and Sork V. L. (2017). Whole-transcriptome response to water stress in a California endemic oak, Quercus lobata. *Tree Physiology* **37**.5: 632– 644. DOI: 10.1093/treephys/tpw122.

Guisan A. and Thuiller W. (2005). Predicting species distribution: offering more than simple habitat models. *Ecology Letters* **8**.9: 993– 1009. DOI: 10.1111/j.1461-0248.2005.00792.x.

Gupta A. P. and Lewontin R. C. (1982). A Study of Reaction Norms in Natural Populations of Drosophila Pseudoobscura. *Evolution* **36**.5: 934– 948. DOI: 10.1111/j.1558-5646.1982.tb05464.x.

Guyon J. P. and Kremer A. (1982). Stabilité phénotypique de la croissance en hauteur et cinétique journalière de la pression de sève et de la transpiration chez le pin maritime (Pinuspinaster Ait.) *Canadian Journal of Forest Research* **12**.4: 936– 946. DOI: 10.1139/x82-136.

Gwaze D. P., Wolliams J. A., Kanowski P. J., and Bridgwater F. E. (2001). Interactions of genotype with site for height and stem straightness in Pinus taeda in Zimbabwe. *Silvae Genetica* **50**.3-4: 135– 139.

Haller B. C. and Hendry A. P. (2014). Solving the Paradox of Stasis: Squashed Stabilizing Selection and the Limits of Detection. *Evolution* **68**.2: 483– 500. DOI: 10.1111/evo.12275.

Hallingbäck H. R., Burton V., Vizcaíno-Palomar N., Trotter F., Liziniewicz M., Marchi M., Berlin M., Ray D., and Benito Garzón M. (2021). Managing Uncertainty in Scots Pine Range-

Wide Adaptation Under Climate Change. *Frontiers in Ecology and Evolution* **9**: 807. DOI: 10.3389/fevo.2021.724051.

Hamrick J. L. (2004). Response of forest trees to global environmental changes. *Forest Ecology and Management* **197**.1: 323– 335. DOI: 10.1016/j.foreco.2004.05.023.

Hänninen H. (2006). Climate warming and the risk of frost damage to boreal forest trees: identification of critical ecophysiological traits. *Tree Physiology* **26**.7: 889– 898. DOI: 10.1093/treephys/26.7.889.

Hansen A. J., Neilson R. P., Dale V. H., Flather C. H., Iverson L. R., Currie D. J., Shafer S., Cook R., and Bartlein P. J. (2001). Global Change in Forests: Responses of Species, Communities, and Biomes: Interactions between climate change and land use are projected to cause large shifts in biodiversity. *BioScience* **51**.9: 765– 779. DOI: 10.1641/0006-3568(2001)051[0765:GCIFRO]2.0.CO;2.

Hansen T. F., ÁLvarez-CASTRO J. M., Carter A. J. R., Hermisson J., and Wagner G. P. (2006). Evolution of Genetic Architecture Under Directional Selection. *Evolution* **60**.8: 1523– 1536. DOI: 10.1111/j.0014-3820.2006.tb00498.x.

Hayes B. J., Bowman P. J., Chamberlain A. C., Verbyla K., and Goddard M. E. (2009). Accuracy of genomic breeding values in multi-breed dairy cattle populations. *Genetics Selection Evolution* **41**.1: 51. DOI: 10.1186/1297-9686-41-51.

Hedrick P. W. (1986). Genetic Polymorphism in Heterogeneous Environments: A Decade Later. *Annual Review of Ecology and Systematics* **17**: 535– 566.

— (2006). Genetic Polymorphism in Heterogeneous Environments: The Age of Genomics. *Annual Review of Ecology, Evolution, and Systematics* **37**.1: 67– 93. DOI: 10.1146/annurev.ecolsys.37.091305.110132.

Henderson C. R. (1975). Best Linear Unbiased Estimation and Prediction under a Selection Model. *Biometrics* **31**.2: 423– 447. DOI: 10.2307/2529430.

Hendry A. P. (2017). *Eco-evolutionary Dynamics*. Princeton University Press. ISBN: 978-1-4008-8308-0.

Hendry A. P., Farrugia T. J., and Kinnison M. T. (2008). Human influences on rates of phenotypic change in wild animal populations. *Molecular Ecology* **17**.1: 20– 29. DOI: 10.1111/j.1365-294X.2007.03428.x.

Hendry A. P. and Gonzalez A. (2008). Whither adaptation? *Biology & Philosophy* **23**.5: 673– 699. DOI: 10.1007/s10539-008-9126-x.

Hendry A. P., Gotanda K. M., and Svensson E. I. (2017). Human influences on evolution, and the ecological and societal consequences. *Philosophical Transactions of the Royal Society B: Biological Sciences* **372**.1712: 20160028. DOI: 10.1098/rstb.2016.0028.

Hendry A. P., Grant P. R., Rosemary Grant B., Ford H. A., Brewer M. J., and Podos J. (2006). Possible human impacts on adaptive radiation: beak size bimodality in Darwin's finches. *Proceedings of the Royal Society B: Biological Sciences* **273**.1596: 1887– 1894. DOI: 10.1098/rspb.2006.3534.

Hendry A. P. and Kinnison M. T. (1999). Perspective: The Pace of Modern Life: Measuring Rates of Contemporary Microevolution. *Evolution* **53**.6: 1637– 1653. DOI: 10.1111/j.1558-5646.1999.tb04550.x.

Hereford J. (2009). A Quantitative Survey of Local Adaptation and Fitness Trade-Offs. *The American Naturalist* **173**.5: 579– 588. DOI: 10.1086/597611.

Hereford J., Hansen T. F., and Houle D. (2004). Comparing Strengths of Directional Selection: How Strong Is Strong? *Evolution* **58**.10: 2133– 2143. DOI: 10.1111/j.0014-3820.2004.tb01592.x.

Hidalgo A. M., Bastiaansen J. W. M., Lopes M. S., Calus M. P. L., and Koning D. J. de (2016). Accuracy of genomic prediction of purebreds for cross bred performance in pigs. *Journal of Animal Breeding and Genetics* **133**.6: 443– 451. DOI: 10.1111/jbg.12214.

Hiederer R., European Commission, Joint Research Centre, and Institute for Environment and Sustainability (2013). *Mapping soil properties for Europe spatial representation of soil database attributes*. Luxembourg: Publications Office.

Hill W. G., Goddard M. E., and Visscher P. M. (2008). Data and Theory Point to Mainly Additive Genetic Variance for Complex Traits. *PLOS Genetics* **4**.2: e1000008. DOI: 10.1371/journal.pgen.1000008.

Hivert V., Sidorenko J., Rohart F., Goddard M. E., Yang J., Wray N. R., Yengo L., and Visscher P. M. (2021a). Estimation of non-additive genetic variance in human complex traits from a large sample of unrelated individuals. *The American Journal of Human Genetics* **108**.5: 786– 798. DOI: 10.1016/j.ajhg.2021.02.014.

Hivert V., Wray N. R., and Visscher P. M. (2021b). Gene action, genetic variation, and GWAS: A user-friendly web tool. *PLOS Genetics* **17**.5: e1009548. DOI: 10.1371/journal.pgen.1009548.

Hoang L. N. (2020). *The Equation of Knowledge: From Bayes' Rule to a Unified Philosophy of Science*. 1st edition. Boca Raton: Chapman and Hall/CRC. 438 pp. ISBN: 978-0-367-42815-0.

Hoban S., Kelley J. L., Lotterhos K. E., Antolin M. F., Bradburd G., Lowry D. B., Poss M. L., Reed L. K., Storfer A., and Whitlock M. C. (2016). Finding the Genomic Basis of Local Adaptation: Pitfalls, Practical Solutions, and Future Directions. *The American Naturalist* **188**.4: 379– 397. DOI: 10.1086/688018.

Hoegh-Guldberg O., Jacob D., Bindi M., Brown S., Camilloni I., Diedhiou A., Djalante R., Ebi K., Engelbrecht F., Guiot J., Hijioka Y., Mehrotra S., Payne A., Seneviratne S. I., Thomas A., Warren R., Zhou G., Halim S. A., Achlatis M., Alexander L. V., Allen M., Berry P., Boyer C., Byers E., Brilli L., Buckeridge M., Cheung W., Craig M., Ellis N., Evans J., Fischer H., Fraedrich K., Fuss S., Ganase A., Gattuso J. P., Greve P., Bolaños T. G., Hanasaki N., Hasegawa T., Hayes K., Hirsch A., Jones C., Jung T., Kanninen M., Krinner G., Lawrence D., Lenton T., Ley D., Liverman D., Mahowald N., McInnes K., Meissner K. J., Millar R., Mintenbeck K., Mitchell D., Mix A. C., Notz D., Nurse L., Okem A., Olsson L., Oppenheimer M., Paz S., Petersen J., Petzold J., Preuschmann S., Rahman M. F., Rogelj J., Scheuffele H., Schleussner C.-F., Scott D., Séférian R., Sillmann J., Singh C., Slade R., Stephenson K., Stephenson T., Sylla M. B., Tebboth M., Tschakert P., Vautard R., Wartenburger R., Wehner M., Weyer N. M., Whyte F., Yohe G., Zhang X., and Zougmoré R. B. (2018). Impacts of 1.5°C Global Warming on Natural and Human Systems. *Global warming of 1.5 C. An IPCC Special Report*.

Hoffman M. D. and Gelman A. (2014). The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo. *Journal of Machine Learning Research*: 1593– 1623.

Hoffmann A. A. and Parsons P. A. (1991). *Evolutionary genetics and environmental stress*. Oxford University Press. ISBN: 978-0-19-857732-4.

Hoffmann A. A., Weeks A. R., and Sgrò C. M. (2021). Opportunities and challenges in assessing climate change vulnerability through genomics. *Cell* **184**.6: 1420– 1425. DOI: 10.1016/j.cell.2021.02.006.

Holder M. and Lewis P. O. (2003). Phylogeny estimation: traditional and Bayesian approaches. *Nature Reviews Genetics* **4**.4: 275– 284. DOI: 10.1038/nrg1044.

Holliday J. A., Aitken S. N., Cooke J. E. K., Fady B., González-Martínez S. C., Heuertz M., Jaramillo-Correa J.-P., Lexer C., Staton M., Whetten R. W., and Plomion C. (2017). Advances in ecological genomics in forest trees and applications to genetic resources conservation and breeding. *Molecular Ecology* **26**.3: 706– 717. DOI: 10.1111/mec.13963.

Hooper D. U., Chapin F. S., Ewel J. J., Hector A., Inchausti P., Lavorel S., Lawton J. H., Lodge D. M., Loreau M., Naeem S., Schmid B., Setälä H., Symstad A. J., Vandermeer J., and Wardle D. A. (2005). Effects of Biodiversity on Ecosystem Functioning: A Consensus of Current Knowledge. *Ecological Monographs* **75**.1: 3– 35. DOI: 10.1890/04-0922.

Houle D. (1992). Comparing evolvability and variability of quantitative traits. *Genetics* **130**.1: 195– 204.

Houle D. (1998). How should we explain variation in the genetic variance of traits? *Genetica* **102**.0: 241. DOI: 10.1023/A:1017034925212.

Howe G. T., Aitken S. N., Neale D. B., Jermstad K. D., Wheeler N. C., and Chen T. H. (2003). From genotype to phenotype: unraveling the complexities of cold adaptation in forest trees. *Canadian Journal of Botany* **81**.12: 1247– 1266. DOI: 10.1139/b03-141.

Huang Y., Stinchcombe J. R., and Agrawal A. F. (2015). Quantitative genetic variance in experimental fly populations evolving with or without environmental heterogeneity. *Evolution* **69**.10: 2735– 2746. DOI: 10.1111/evo.12771.

Huelsenbeck J. P., Ronquist F., Nielsen R., and Bollback J. P. (2001). Bayesian Inference of Phylogeny and Its Impact on Evolutionary Biology. *Science* **294**.5550: 2310– 2314. DOI: 10.1126/science.1065889.

Huey R. B., Hertz P. E., and Sinervo B. (2003). Behavioral Drive versus Behavioral Inertia in Evolution: A Null Model Approach. *The American Naturalist* **161**.3: 357– 366. DOI: 10.1086/346135.

Hughes L. (2000). Biological consequences of global warming: is the signal already apparent? *Trends in Ecology & Evolution* **15**.2: 56– 61. DOI: 10.1016/S0169-5347(99)01764-4.

Hurel A., Miguel M. de, Dutech C., Desprez-Loustau M.-L., Plomion C., Rodríguez-Quilón I., Cyrille A., Guzman T., Alía R., González-Martínez S. C., and Budde K. B. (2019). Genetic basis of susceptibility to Diplodia sapinea and Armillaria ostoyae in maritime pine. *bioRxiv*: 699389. DOI: 10.1101/699389.

Ingvarsson P. K. and Bernhardsson C. (2020). Genome-wide signatures of environmental adaptation in European aspen (Populus tremula) under current and future climate conditions. *Evolutionary Applications* **13**.1: 132– 142. DOI: 10.1111/eva.12792.

Ioannidis J. P. A., Thomas G., and Daly M. J. (2009). Validating, augmenting and refining genome-wide association signals. *Nature Reviews Genetics* **10**.5: 318– 329. DOI: 10.1038/nrg2544.

IPCC (2015). Climate change 2014: Synthesis report, contribution of working groups I, II and III to the fifth assessment report of the Intergovernmental Panel on Climate Change.

— (2018). Global warming of 1.5° C: an IPCC special report on the impacts of global warming of 1.5° C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty. *Intergovernmental Panel on Climate Change*.

— (2021). Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. *Cambridge University Press*.

Ishizuka W. and Goto S. (2012). Modeling intraspecific adaptation of Abies sachalinensis to local altitude and responses to global warming, based on a 36-year reciprocal transplant experiment. *Evolutionary Applications* **5**.3: 229– 244. DOI: 10.1111/j.1752-4571.2011.00216.x.

Jannink J.-L., Lorenz A. J., and Iwata H. (2010). Genomic selection in plant breeding: from theory to practice. *Briefings in Functional Genomics* **9**.2: 166– 177. DOI: 10.1093/bfgp/elq001.

Jansson G., Li B., and Hannrup B. (2003). Time Trends in Genetic Parameters for Height and Optimal Age for Parental Selection in Scots Pine. *Forest Science* **49**.5: 696– 705. DOI: 10.1093/forestscience/49.5.696.

Jaramillo-Correa J.-P., Rodríguez-Quilón I., Grivet D., Lepoittevin C., Sebastiani F., Heuertz M., Garnier-Géré P. H., Alía R., Plomion C., Vendramin G. G., and González-Martínez S. C. (2015). Molecular Proxies for Climate Maladaptation in a Long-Lived Tree (Pinus pinaster Aiton, Pinaceae). *Genetics* **199**.3: 793– 807. DOI: 10.1534/genetics.114.173252.

Jarquín D., Crossa J., Lacaze X., Du Cheyron P., Daucourt J., Lorgeou J., Piraux F., Guerreiro L., Pérez P., Calus M., Burgueño J., and de los Campos G. (2014). A reaction norm model for genomic selection using high-dimensional genomic and environmental data. *Theoretical and Applied Genetics* **127**.3: 595– 607. DOI: 10.1007/s00122-013-2243-1.

Jay F., Manel S., Alvarez N., Durand E. Y., Thuiller W., Holderegger R., Taberlet P., and François O. (2012). Forecasting changes in population genetic structure of alpine plants in response to global warming. *Molecular Ecology* **21**.10: 2354– 2368. DOI: 10.1111/j.1365-294X.2012.05541.x.

Jensen J. D., Payseur B. A., Stephan W., Aquadro C. F., Lynch M., Charlesworth D., and Charlesworth B. (2019). The importance of the Neutral Theory in 1968 and 50 years on: A response to Kern and Hahn 2018. *Evolution* **73**.1: 111– 114. DOI: 10.1111/evo.13650.

Johnson G. R., Sniezko R. A., and Mandel N. L. (1997). Age trends in Douglas-fir genetic parameters and implications for optimum selection age. *Silvae Genetica. 46: 349-358.*

Johnson T. and Barton N. H. (2005). Theoretical models of selection and mutation on quantitative traits. *Philosophical Transactions of the Royal Society B: Biological Sciences* **360**.1459: 1411– 1425. DOI: 10.1098/rstb.2005.1667.

Johnstone J. F. and Chapin F. S. (2003). Non-equilibrium succession dynamics indicate continued northern migration of lodgepole pine. *Global Change Biology* **9**.10: 1401– 1409. DOI: 10.1046/j.1365-2486.2003.00661.x.

Josephs E. B., Berg J. J., Ross-Ibarra J., and Coop G. (2019). Detecting Adaptive Differentiation in Structured Populations with Genomic Data and Common Gardens. *Genetics* **211**.3: 989– 1004. DOI: 10.1534/genetics.118.301786.

Joyce D. G. and Rehfeldt G. E. (2013). Climatic niche, ecological genetics, and impact of climate change on eastern white pine (Pinus strobus L.): Guidelines for land managers. *Forest Ecology and Management* **295**: 173– 192. DOI: 10.1016/j.foreco.2012.12.024.

Jump A. S. and Peñuelas J. (2005). Running to stand still: adaptation and the response of plants to rapid climate change. *Ecology Letters* **8**.9: 1010– 1020. DOI: 10.1111/j.1461-0248.2005.00796.x.

Karasov T. L., Kniskern J. M., Gao L., DeYoung B. J., Ding J., Dubiella U., Lastra R. O., Nallu S., Roux F., Innes R. W., Barrett L. G., Hudson R. R., and Bergelson J. (2014). The long-term maintenance of a resistance polymorphism through diffuse interactions. *Nature* **512**.7515: 436– 440. DOI: 10.1038/nature13439.

Kardos M. and Luikart G. (2021). The Genetic Architecture of Fitness Drives Population Viability during Rapid Environmental Change. *The American Naturalist* **197**.5: 511– 525. DOI: 10.1086/713469.

Kawecki T. J. (2000). The Evolution of Genetic Canalization Under Fluctuating Selection. *Evolution* **54**.1: 1– 12. DOI: 10.1111/j.0014-3820.2000.tb00001.x.

Kawecki T. J. and Ebert D. (2004). Conceptual issues in local adaptation. *Ecology Letters* **7**.12: 1225– 1241. DOI: 10.1111/j.1461-0248.2004.00684.x.

Keller S. R., Chhatre V. E., and Fitzpatrick M. C. (2018). Influence of Range Position on Locally Adaptive Gene–Environment Associations in Populus Flowering Time Genes. *Journal of Heredity* **109**.1: 47– 58. DOI: 10.1093/jhered/esx098.

Kelly M. (2019). Adaptation to climate change through genetic accommodation and assimilation of plastic phenotypes. *Philosophical Transactions of the Royal Society B: Biological Sciences* **374**.1768: 20180176. DOI: 10.1098/rstb.2018.0176.

Kern A. D. and Hahn M. W. (2018). The Neutral Theory in Light of Natural Selection. *Molecular Biology and Evolution* **35**.6: 1366– 1371. DOI: 10.1093/molbev/msy092.

Khera A. V., Chaffin M., Aragam K. G., Haas M. E., Roselli C., Choi S. H., Natarajan P., Lander E. S., Lubitz S. A., Ellinor P. T., and Kathiresan S. (2018). Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. *Nature Genetics* **50**.9: 1219– 1224. DOI: 10.1038/s41588-018-0183-z.

Kimura M. (1968). Evolutionary Rate at the Molecular Level. *Nature* **217**.5129: 624– 626. DOI: 10.1038/217624a0.

King J. G. and Hadfield J. D. (2019). The evolution of phenotypic plasticity when environments fluctuate in time and space. *Evolution Letters* **3**.1: 15– 27. DOI: 10.1002/evl3.100.

Kinnison M. T. and Hairston N. G. (2007). Eco-evolutionary conservation biology: contemporary evolution and the dynamics of persistence. *Functional Ecology* **21**.3: 444– 454. DOI: 10.1111/j.1365-2435.2007.01278.x.

Kinnison M. T. and Hendry A. P. (2001). The pace of modern life II: From rates of contemporary microevolution to pattern and process. *Microevolution Rate, Pattern, Process*. Ed. by Hendry A. P. and Kinnison M. T. Dordrecht: Springer Netherlands. Page 145– 164. ISBN: 978-94-010-0585-2. DOI: 10.1007/978-94-010-0585-2_10.

Kirkpatrick M. and Barton N. H. (1997). Evolution of a Species' Range. *The American Naturalist* **150**.1: 1– 23. DOI: 10.1086/286054.

Kleunen M. V. and Fischer M. (2005). Constraints on the evolution of adaptive phenotypic plasticity in plants. *New Phytologist* **166**.1: 49– 60. DOI: 10.1111/j.1469-8137.2004.01296.x.

Kooyers N. J., Colicchio J. M., Greenlee A. B., Patterson E., Handloser N. T., and Blackman B. K. (2019). Lagging Adaptation to Climate Supersedes Local Adaptation to Herbivory in an Annual Monkeyflower. *The American Naturalist* **194**.4: 541– 557. DOI: 10.1086/702312.

Körner C. (1995). Alpine Plant Diversity: A Global Survey and Functional Interpretations. *Arctic and Alpine Biodiversity: Patterns, Causes and Ecosystem Consequences*. Ed. by Chapin F. S. and Körner C. Berlin, Heidelberg: Springer. Page 45– 62. ISBN: 978-3-642-78966-3. DOI: 10.1007/978-3-642-78966-3_4.

Kort H. D., Vandepitte K., Bruun H. H., Closset-Kopp D., Honnay O., and Mergeay J. (2014). Landscape genomics and a common garden trial reveal adaptive differentiation to temperature across Europe in the tree species Alnus glutinosa. *Molecular Ecology* **23**.19: 4709– 4721. DOI: 10.1111/mec.12813.

Kovach A., Wegrzyn J. L., Parra G., Holt C., Bruening G. E., Loopstra C. A., Hartigan J., Yandell M., Langley C. H., Korf I., and Neale D. B. (2010). The Pinus taeda genome is characterized by diverse and highly diverged repetitive sequences. *BMC Genomics* **11**.1: 420. DOI: 10.1186/1471-2164-11-420.

Kremer A., Potts B. M., and Delzon S. (2014). Genetic divergence in forest trees: understanding the consequences of climate change. *Functional Ecology* **28**.1: 22– 36. DOI: 10.1111/1365-2435.12169.

Kremer A., Ronce O., Robledo-Arnuncio J. J., Guillaume F., Bohrer G., Nathan R., Bridle J. R., Gomulkiewicz R., Klein E. K., Ritland K., Kuparinen A., Gerber S., and Schueler S. (2012). Long-distance gene flow and adaptation of forest trees to rapid climate change. *Ecology Letters* **15**.4: 378– 392. DOI: 10.1111/j.1461-0248.2012.01746.x.

Kroon J., Ericsson T., Jansson G., and Andersson B. (2011). Patterns of genetic parameters for height in field genetic tests of Picea abies and Pinus sylvestris in Sweden. *Tree Genetics & Genomes* **7**.6: 1099– 1111. DOI: 10.1007/s11295-011-0398-y.

Kruuk L. E. B., Clutton-Brock T. H., Slate J., Pemberton J. M., Brotherstone S., and Guinness F. E. (2000). Heritability of fitness in a wild mammal population. *Proceedings of the National Academy of Sciences* **97**.2: 698– 703. DOI: 10.1073/pnas.97.2.698.

Kuparinen A., Savolainen O., and Schurr F. M. (2010). Increased mortality can promote evolutionary adaptation of forest trees to climate change. *Forest Ecology and Management* **259**.5: 1003– 1008. DOI: 10.1016/j.foreco.2009.12.006.

Kurt Y., González-Martínez S. C., Alía R., and Isik K. (2012). Genetic differentiation in Pinus brutia Ten. using molecular markers and quantitative traits: the role of altitude. *Annals of Forest Science* **69**.3: 345– 351. DOI: 10.1007/s13595-011-0169-9.

Kusnandar D., Galwey N. W., Hertzler G. L., and Butcher T. B. (1998). Age Trends in Variances and Heritabilities for Diameter and Height in Maritime Pine (Pinus pinaster AIT.) in Western Australia. *Silvae Genetica* **47**.2: 136– 141.

Lamy J.-B., Bouffier L., Burlett R., Plomion C., Cochard H., and Delzon S. (2011). Uniform Selection as a Primary Force Reducing Population Genetic Differentiation of Cavitation Resistance across a Species Range. *PLOS ONE* **6**.8: e23476. DOI: 10.1371/journal.pone.0023476.

Lamy J.-B., Delzon S., Bouche P. S., Alia R., Vendramin G. G., Cochard H., and Plomion C. (2014). Limited genetic variability and phenotypic plasticity detected for cavitation resistance in a Mediterranean pine. *New Phytologist* **201**.3: 874– 886. DOI: 10.1111/nph.12556.

Lande R. (1976). Natural Selection and Random Genetic Drift in Phenotypic Evolution. *Evolution* **30**.2: 314– 334. DOI: 10.2307/2407703.

— (1979). Quantitative Genetic Analysis of Multivariate Evolution, Applied to Brain: Body Size Allometry. *Evolution* **33**.1: 402– 416. DOI: 10.2307/2407630.

Lande R. and Arnold S. J. (1983). The Measurement of Selection on Correlated Characters. *Evolution* **37**.6: 1210– 1226. DOI: 10.2307/2408842.

Langlet O. (1971). Two Hundred Years Genecology. *Taxon* **20**.5/6: 653– 721. DOI: 10.2307/1218596.

Láruson Á. J., Fitzpatrick M. C., Keller S. R., Haller B. C., and Lotterhos K. E. (2021). Seeing the Forest for the trees: Assessing genetic offset predictions with Gradient Forest: 2021.09.20.461151. DOI: 10.1101/2021.09.20.461151.

Lasky J. R., Des Marais D. L., Lowry D. B., Povolotskaya I., McKay J. K., Richards J. H., Keitt T. H., and Juenger T. E. (2014). Natural Variation in Abiotic Stress Responsive Gene Expression and Local Adaptation to Climate in Arabidopsis thaliana. *Molecular Biology and Evolution* **31**.9: 2283– 2296. DOI: 10.1093/molbev/msu170.

Lässig M., Mustonen V., and Walczak A. M. (2017). Predicting evolution. *Nature Ecology & Evolution* **1**.3: 1– 9. DOI: 10.1038/s41559-017-0077.

Latta R. G. (2009). Testing for local adaptation in Avena barbata: a classic example of ecotypic divergence. *Molecular Ecology* **18**.18: 3781– 3791. DOI: 10.1111/j.1365-294X.2009.04302.x.

Lawton-Rauh A. (2008). Demographic processes shaping genetic variation. *Current Opinion in Plant Biology* **11**.2: 103– 109. DOI: 10.1016/j.pbi.2008.02.009.

Layzer D. (1980). Genetic Variation and Progressive Evolution. *The American Naturalist* **115**.6: 809– 826. DOI: 10.1086/283602.

Lee S. H., Weerasinghe W. M. S. P., Wray N. R., Goddard M. E., and van der Werf J. H. J. (2017). Using information of relatives in genomic prediction to apply effective stratified medicine. *Scientific Reports* **7**.1: 42091. DOI: 10.1038/srep42091.

Leimu R. and Fischer M. (2008). A Meta-Analysis of Local Adaptation in Plants. *PLOS ONE* **3**.12: e4010. DOI: 10.1371/journal.pone.0004010.

Leitão P. J., Schwieder M., and Senf C. (2017). sgdm: An R Package for Performing Sparse Generalized Dissimilarity Modelling with Tools for gdm. *ISPRS International Journal of Geo-Information* **6**.1: 23. DOI: 10.3390/ijgi6010023.

Leites L. P., Rehfeldt G. E., Robinson A. P., Crookston N. L., and Jaquish B. (2012a). Possibilities and Limitations of Using Historic Provenance Tests to Infer Forest Species Growth Responses to Climate Change. *Natural Resource Modeling* **25**.3: 409– 433. DOI: 10.1111/j.1939-7445.2012.00129.x.

Leites L. P., Rehfeldt G. E., and Steiner K. C. (2019). Adaptation to climate in five eastern North America broadleaf deciduous species: Growth clines and evidence of the growth-cold tolerance trade-off. *Perspectives in Plant Ecology, Evolution and Systematics* **37**: 64– 72. DOI: 10.1016/j.ppees.2019.02.002.

Leites L. P., Robinson A. P., Rehfeldt G. E., Marshall J. D., and Crookston N. L. (2012b). Height-growth response to climatic changes differs among populations of Douglas-fir: a novel analysis of historic data. *Ecological Applications* **22**.1: 154– 165. DOI: 10.1890/11-0150.1.

Lepais O. and Bacles C. F. E. (2014). Two are better than one: combining landscape genomics and common gardens for detecting local adaptation in forest trees. *Molecular Ecology* **23**.19: 4671– 4673. DOI: 10.1111/mec.12906.

Lepoittevin C., Rousseau J.-P., Guillemin A., Gauvrit C., Besson F., Hubert F., da Silva Perez D., Harvengt L., and Plomion C. (2011). Genetic parameters of growth, straightness and wood chemistry traits in Pinus pinaster. *Annals of Forest Science* **68**.4: 873– 884. DOI: 10.1007/s13595-011-0084-0.

Leroy T., Louvet J.-M., Lalanne C., Provost G. L., Labadie K., Aury J.-M., Delzon S., Plomion C., and Kremer A. (2020). Adaptive introgression as a driver of local adaptation to climate in European white oaks. *New Phytologist* **226**.4: 1171– 1182. DOI: 10.1111/nph.16095.

Lewin K. (1943). Psychology and the Process of Group Living. *The Journal of Social Psychology* **17**.1: 113– 131. DOI: 10.1080/00224545.1943.9712269.

Li Y., Zhao M., Motesharrei S., Mu Q., Kalnay E., and Li S. (2015). Local cooling and warming effects of forests based on satellite observations. *Nature Communications* **6**.1: 6603. DOI: 10.1038/ncomms7603.

Li Y. I., Geijn B. van de, Raj A., Knowles D. A., Petti A. A., Golan D., Gilad Y., and Pritchard J. K. (2016). RNA splicing is a primary link between genetic variation and disease. *Science* **352**.6285: 600– 604. DOI: 10.1126/science.aad9417.

Li Y., Suontama M., Burdon R. D., and Dungey H. S. (2017). Genotype by environment interactions in forest tree breeding: review of methodology and perspectives on research and application. *Tree Genetics & Genomes* **13**.3: 60. DOI: 10.1007/s11295-017-1144-x.

Liang Y., Duveneck M. J., Gustafson E. J., Serra-Diaz J. M., and Thompson J. R. (2018). How disturbance, competition, and dispersal interact to prevent tree range boundaries from keeping pace with climate change. *Global Change Biology* **24**.1: e335– e351. DOI: 10.1111/gcb.13847.

Lin Y., Yang H., Ivković M., Gapare W. J., Colin Matheson A., and Wu H. X. (2013). Effect of genotype by spacing interaction on radiata pine genetic parameters for height and diameter growth. *Forest Ecology and Management* **304**: 204– 211. DOI: 10.1016/j.foreco.2013.05.015.

Lind B. M., Menon M., Bolte C. E., Faske T. M., and Eckert A. J. (2018). The genomics of local adaptation in trees: are we out of the woods yet? *Tree Genetics & Genomes* **14**.2: 29. DOI: 10.1007/s11295-017-1224-y.

Linhart Y. B. and Grant M. C. (1996). Evolutionary Significance of Local Genetic Differentiation in Plants. *Annual Review of Ecology and Systematics* **27**.1: 237– 277. DOI: 10.1146/annurev.ecolsys.27.1.237.

Liu X., Li Y. I., and Pritchard J. K. (2019). Trans Effects on Gene Expression Can Drive Omnigenic Inheritance. *Cell* **177**.4: 1022– 1034.e6. DOI: 10.1016/j.cell.2019.04.014.

Losos J. B., Warheitt K. I., and Schoener T. W. (1997). Adaptive differentiation following experimental island colonization in Anolis lizards. *Nature* **387**.6628: 70– 73. DOI: 10.1038/387070a0.

Lotterhos K. E. and Whitlock M. C. (2014). Evaluation of demographic history and neutral parameterization on the performance of FST outlier tests. *Molecular Ecology* **23**.9: 2178–2192. DOI: 10.1111/mec.12725.

Lu M., Krutovsky K. V., and Loopstra C. A. (2019). Predicting Adaptive Genetic Variation of Loblolly Pine (Pinus taeda L.) Populations Under Projected Future Climates Based on Multivariate Models. *Journal of Heredity* **110**.7: 857– 865. DOI: 10.1093/jhered/esz065.

Lu P. and Charrette P. (2008). Genetic parameter estimates for growth traits of black spruce in northwestern Ontario. *Canadian Journal of Forest Research*. DOI: 10.1139/X08-133.

de-Lucas A. I., Robledo-Arnuncio J. J., Hidalgo E., and González-Martínez S. C. (2008). Mating system and pollen gene flow in Mediterranean maritime pine. *Heredity* **100**.4: 390– 399. DOI: 10.1038/sj.hdy.6801090.

Luikart G., England P. R., Tallmon D., Jordan S., and Taberlet P. (2003). The power and promise of population genomics: from genotyping to genome typing. *Nature Reviews Genetics* **4**.12: 981– 994. DOI: 10.1038/nrg1226.

Lush J. (1937). *Animal breeding plans*. Iowa State Univ Press Ames.

Lustenhouwer N., Wilschut R. A., Williams J. L., Putten W. H. van der, and Levine J. M. (2018). Rapid evolution of phenology during range expansion with recent climate change. *Global Change Biology* **24**.2: e534– e544. DOI: 10.1111/gcb.13947.

Lynch M. and Walsh B. (1998). "Genetics and Analysis of Quantitative Traits". DOI: 10.1046/j.1439-0388.2002.00356.x.

Mackay J., Dean J. F. D., Plomion C., Peterson D. G., Cánovas F. M., Pavy N., Ingvarsson P. K., Savolainen O., Guevara M. Á., Fluch S., Vinceti B., Abarca D., Díaz-Sala C., and Cervera M.-T. (2012). Towards decoding the conifer giga-genome. *Plant Molecular Biology* **80**.6: 555– 569. DOI: 10.1007/s11103-012-9961-7.

Mackay T. F. C. (1981). Genetic variation in varying environments. *Genetics Research* **37**.1: 79–93. DOI: 10.1017/S0016672300020036.

MacLachlan I. R., McDonald T. K., Lind B. M., Rieseberg L. H., Yeaman S., and Aitken S. N. (2021). Genome-wide shifts in climate-related variation underpin responses to selective breeding in a widespread conifer. *Proceedings of the National Academy of Sciences* **118**.10. DOI: 10.1073/pnas.2016900118.

Mahony C. R., MacLachlan I. R., Lind B. M., Yoder J. B., Wang T., and Aitken S. N. (2020). Evaluating genomic data for management of local adaptation in a changing climate: A lodgepole pine case study. *Evolutionary Applications* **13**.1: 116– 131. DOI: 10.1111/eva.12871.

Manel S., Schwartz M. K., Luikart G., and Taberlet P. (2003). Landscape genetics: combining landscape ecology and population genetics. *Trends in Ecology & Evolution* **18**.4: 189– 197. DOI: 10.1016/S0169-5347(03)00008-9.

Manolio T. A., Collins F. S., Cox N. J., Goldstein D. B., Hindorff L. A., Hunter D. J., McCarthy M. I., Ramos E. M., Cardon L. R., Chakravarti A., Cho J. H., Guttmacher A. E., Kong A., Kruglyak L., Mardis E., Rotimi C. N., Slatkin M., Valle D., Whittemore A. S., Boehnke M., Clark A. G., Eichler E. E., Gibson G., Haines J. L., Mackay T. F. C., McCarroll S. A., and Visscher P. M. (2009). Finding the missing heritability of complex diseases. *Nature* **461**.7265: 747– 753. DOI: 10.1038/nature08494.

Mantgem P. J. van, Stephenson N. L., Byrne J. C., Daniels L. D., Franklin J. F., Fulé P. Z., Harmon M. E., Larson A. J., Smith J. M., Taylor A. H., and Veblen T. T. (2009). Widespread Increase of Tree Mortality Rates in the Western United States. *Science* **323**.5913: 521– 524. DOI: 10.1126/science.1165000.

Marchi M., Castellanos-Acuña D., Hamann A., Wang T., Ray D., and Menzel A. (2020). ClimateEU, scale-free climate normals, historical time series, and future projections for Europe. *Scientific Data* **7**.1: 428. DOI: 10.1038/s41597-020-00763-0.

Marks H. L. (1978). Long term selection for four-week body weight in Japanese quail under different nutritional environments. *Theoretical and Applied Genetics* **52**.3: 105– 111. DOI: 10.1007/BF00264742.

Marland G., Pielke R. A., Apps M., Avissar R., Betts R. A., Davis K. J., Frumhoff P. C., Jackson S. T., Joyce L. A., Kauppi P., Katzenberger J., MacDicken K. G., Neilson R. P., Niles J. O., Niyogi D. d. S., Norby R. J., Pena N., Sampson N., and Xue Y. (2003). The climatic impacts of land surface change and carbon management, and the implications for climate-change mitigation policy. *Climate Policy* **3**.2: 149– 157. DOI: 10.3763/cpol.2003.0318.

Martin A. R., Gignoux C. R., Walters R. K., Wojcik G. L., Neale B. M., Gravel S., Daly M. J., Bustamante C. D., and Kenny E. E. (2017). Human Demographic History Impacts Genetic Risk Prediction across Diverse Populations. *The American Journal of Human Genetics* **100**.4: 635– 649. DOI: 10.1016/j.ajhg.2017.03.004.

Martin A. R., Kanai M., Kamatani Y., Okada Y., Neale B. M., and Daly M. J. (2019). Clinical use of current polygenic risk scores may exacerbate health disparities. *Nature Genetics* **51**.4: 584– 591. DOI: 10.1038/s41588-019-0379-x.

Martínez-Berdeja A., Hamilton J. A., Bontemps A., Schmitt J., and Wright J. W. (2019). Evidence for population differentiation among Jeffrey and Ponderosa pines in survival, growth and phenology. *Forest Ecology and Management* **434**: 40– 48. DOI: 10.1016/j.foreco.2018.12.009.

Martins K., Gugger P. F., Llanderal-Mendoza J., González-Rodríguez A., Fitz-Gibbon S. T., Zhao J.-L., Rodríguez-Correa H., Oyama K., and Sork V. L. (2018). Landscape genomics provides evidence of climate-associated genetic variation in Mexican populations of Quercus rugosa. *Evolutionary Applications* **11**.10: 1842– 1858. DOI: 10.1111/eva.12684.

Matesanz S., Gianoli E., and Valladares F. (2010). Global change and the evolution of phenotypic plasticity in plants. *Annals of the New York Academy of Sciences* **1206**.1: 35– 55. DOI: 10.1111/j.1749-6632.2010.05704.x.

Mathieson I. (2021). The omnigenic model and polygenic prediction of complex traits. *The American Journal of Human Genetics*. DOI: 10.1016/j.ajhg.2021.07.003.

Mavaddat N. et al. (2019). Polygenic Risk Scores for Prediction of Breast Cancer and Breast Cancer Subtypes. *The American Journal of Human Genetics* **104**.1: 21– 34. DOI: 10.1016/j.ajhg.2018.11.002.

Mayhew A. J. and Meyre D. (2017). Assessing the Heritability of Complex Traits in Humans: Methodological Challenges and Opportunities. *Current Genomics* **18**.4: 332– 340. DOI: 10.2174/1389202918666170307161450.

McCain K. W. (2015). "Nothing as practical as a good theory" Does Lewin's Maxim still have salience in the applied social sciences? *Proceedings of the Association for Information Science and Technology* **52**.1: 1– 4. DOI: 10.1002/pra2.2015.145052010077.

McDonald T. K. and Yeaman S. (2018). Effect of migration and environmental heterogeneity on the maintenance of quantitative genetic variation: a simulation study. *Journal of Evolutionary Biology* **31**.9: 1386– 1399. DOI: 10.1111/jeb.13341.

McDowell N. G. and Allen C. D. (2015). Darcy's law predicts widespread forest mortality under climate warming. *Nature Climate Change* **5**.7: 669– 672. DOI: 10.1038/nclimate2641.

McElreath R. (2016). *Statistical Rethinking: A Bayesian Course with Examples in R and Stan*. Boca Raton: Chapman and Hall/CRC. 505 pp. ISBN: 978-1-315-37249-5. DOI: 10.1201/9781315372495.

McLachlan J. S., Clark J. S., and Manos P. S. (2005). Molecular Indicators of Tree Migration Capacity Under Rapid Climate Change. *Ecology* **86**.8: 2088– 2098. DOI: 10.1890/04-1036.

McLane S. C., Daniels L. D., and Aitken S. N. (2011). Climate impacts on lodgepole pine (Pinus contorta) radial growth in a provenance experiment. *Forest Ecology and Management* **262**.2: 115– 123. DOI: 10.1016/j.foreco.2011.03.007.

Merilä J. and Sheldon B. C. (2000). Lifetime Reproductive Success and Heritability in Nature. *The American Naturalist* **155**.3: 301– 310. DOI: 10.1086/303330.

Merilä J., Sheldon B., and Kruuk L. (2001). Explaining stasis: microevolutionary studies in natural populations. *Genetica* **112**.1: 199– 222. DOI: 10.1023/A:1013391806317.

Merilä J. and Hendry A. P. (2014). Climate change, adaptation, and phenotypic plasticity: the problem and the evidence. *Evolutionary Applications* **7**.1: 1– 14. DOI: 10.1111/eva.12137.

Merilä J., Söderman F., O'Hara R., Räsänen K., and Laurila A. (2004). Local Adaptation and Genetics of Acid-Stress Tolerance in the Moor Frog, Rana arvalis. *Conservation Genetics* **5**.4: 513– 527. DOI: 10.1023/B:COGE.0000041026.71104.0a.

Metzger M. J. (2018). *The Environmental Stratification of Europe*. DOI: 10.7488/ds/2356.

Meuwissen T. H. E., Hayes B. J., and Goddard M. E. (2001). Prediction of Total Genetic Value Using Genome-Wide Dense Marker Maps. *Genetics* **157**.4: 1819– 1829. DOI: 10.1093/genetics/157.4.1819.

Millennium Ecosystem Assessment (2005). *Ecosystems and human well-being*. Vol. 5. Island press United States of America.

Miller M. R., Dunham J. P., Amores A., Cresko W. A., and Johnson E. A. (2007). Rapid and cost-effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers. *Genome Research* **17**.2: 240– 248. DOI: 10.1101/gr.5681207.

Mitchell-Olds T., Willis J. H., and Goldstein D. B. (2007). Which evolutionary processes influence natural genetic variation for phenotypic traits? *Nature Reviews Genetics* **8**.11: 845– 856. DOI: 10.1038/nrg2207.

Moghaddar N., Swan A. A., and van der Werf J. H. (2014). Comparing genomic prediction accuracy from purebred, crossbred and combined purebred and crossbred reference populations in sheep. *Genetics Selection Evolution* **46**.1: 58. DOI: 10.1186/s12711-014-0058-4.

Monroe J. G., Powell T., Price N., Mullen J. L., Howard A., Evans K., Lovell J. T., and McKay J. K. (2018). Drought adaptation in Arabidopsis thaliana by extensive genetic loss-of-function. *eLife* **7**. Ed. by Kliebenstein D. J., Hardtke C. S., and Korte A.: e41038. DOI: 10.7554/eLife.41038.

Morelli T. L., Smith A. B., Kastely C. R., Mastroserio I., Moritz C., and Beissinger S. R. (2012). Anthropogenic refugia ameliorate the severe climate-related decline of a montane mammal along its trailing edge. *Proceedings of the Royal Society B: Biological Sciences* **279**.1745: 4279–4286. DOI: 10.1098/rspb.2012.1301.

Morgenstern M. (1996). *Geographic Variation in Forest Trees: Genetic Basis and Application of Knowledge in Silviculture*. Vancouver: Univ. B.C. Press. 224 pp. ISBN: 978-0-7748-4177-1.

Mori A. S., Lertzman K. P., and Gustafsson L. (2017). Biodiversity and ecosystem services in forest ecosystems: a research agenda for applied forest ecology. *Journal of Applied Ecology* **54**.1: 12– 27. DOI: 10.1111/1365-2664.12669.

Morse A. M., Peterson D. G., Islam-Faridi M. N., Smith K. E., Magbanua Z., Garcia S. A., Kubisiak T. L., Amerson H. V., Carlson J. E., Nelson C. D., and Davis J. M. (2009). Evolution of Genome Size and Complexity in Pinus. *PLOS ONE* **4**.2: e4332. DOI: 10.1371/journal.pone.0004332.

Mosca E., Eckert A. J., Di Pierro E. A., Rocchini D., La Porta N., Belletti P., and Neale D. B. (2012). The geographical and environmental determinants of genetic diversity for four alpine conifers of the European Alps. *Molecular Ecology* **21**.22: 5530– 5545. DOI: 10.1111/mec.12043.

Mostafavi H., Harpak A., Agarwal I., Conley D., Pritchard J. K., and Przeworski M. (2020). Variable prediction accuracy of polygenic scores within an ancestry group. *eLife* **9**. Ed. by Loos R., Eisen M. B., and O'Reilly P.: e48376. DOI: 10.7554/eLife.48376.

Mousseau T. A. and Roff D. A. (1987). Natural selection and the heritability of fitness components. *Heredity* **59**.2: 181– 197. DOI: 10.1038/hdy.1987.113.

Muchero W., Sondreli K. L., Chen J.-G., Urbanowicz B. R., Zhang J., Singan V., Yang Y., Brueggeman R. S., Franco-Coronado J., Abraham N., Yang J.-Y., Moremen K. W., Weisberg A. J., Chang J. H., Lindquist E., Barry K., Ranjan P., Jawdy S., Schmutz J., Tuskan G. A., and LeBoldus J. M. (2018). Association mapping, transcriptomics, and transient expression identify candidate genes mediating plant–pathogen interactions in a tree. *Proceedings of the National Academy of Sciences* **115**.45: 11573– 11578. DOI: 10.1073/pnas.1804428115.

Muff S., Niskanen A. K., Saatoglu D., Keller L. F., and Jensen H. (2019). Animal models with group-specific additive genetic variances: extending genetic group models. *Genetics Selection Evolution* **51**.1: 7. DOI: 10.1186/s12711-019-0449-7.

Muñoz P. R., Resende M. F. R., Gezan S. A., Resende M. D. V., Campos G. de los, Kirst M., Huber D., and Peter G. F. (2014). Unraveling Additive from Nonadditive Effects Using Genomic Relationship Matrices. *Genetics* **198**.4: 1759– 1768. DOI: 10.1534/genetics.114.171322.

Murren C. J., Maclean H. J., Diamond S. E., Steiner U. K., Heskel M. A., Handelsman C. A., Ghalambor C. K., Auld J. R., Callahan H. S., Pfennig D. W., Relyea R. A., Schlichting C. D., and Kingsolver J. (2014). Evolutionary Change in Continuous Reaction Norms. *The American Naturalist* **183**.4: 453– 467. DOI: 10.1086/675302.

Muscente A. D., Prabhu A., Zhong H., Eleish A., Meyer M. B., Fox P., Hazen R. M., and Knoll A. H. (2018). Quantifying ecological impacts of mass extinctions with network analysis of

fossil communities. *Proceedings of the National Academy of Sciences* **115**.20: 5217– 5222. DOI: 10.1073/pnas.1719976115.

Nadeau S., Meirmans P. G., Aitken S. N., Ritland K., and Isabel N. (2016). The challenge of separating signatures of local adaptation from those of isolation by distance and colonization history: The case of two white pines. *Ecology and Evolution* **6**.24: 8649– 8664. DOI: 10.1002/ece3.2550.

Narum S. R., Buerkle C. A., Davey J. W., Miller M. R., and Hohenlohe P. A. (2013). Genotyping-by-sequencing in ecological and conservation genomics. *Molecular Ecology* **22**.11: 2841–2847. DOI: 10.1111/mec.12350.

Naydenov K. D., Alexandrov A., Matevski V., Vasilevski K., Naydenov M. K., Gyuleva V., Carcaillet C., Wahid N., and Kamary S. (2014). Range-wide genetic structure of maritime pine predates the last glacial maximum: evidence from nuclear DNA. *Hereditas* **151**.1: 1–13. DOI: 10.1111/j.1601-5223.2013.00027.x.

Neale D. B., McGuire P. E., Wheeler N. C., Stevens K. A., Crepeau M. W., Cardeno C., Zimin A. V., Puiu D., Pertea G. M., Sezen U. U., Casola C., Koralewski T. E., Paul R., Gonzalez-Ibeas D., Zaman S., Cronn R., Yandell M., Holt C., Langley C. H., Yorke J. A., Salzberg S. L., and Wegrzyn J. L. (2017). The Douglas-Fir Genome Sequence Reveals Specialization of the Photosynthetic Apparatus in Pinaceae. *G3 Genes|Genomes|Genetics* **7**.9: 3157– 3167. DOI: 10.1534/g3.117.300078.

Neale D. B. and Savolainen O. (2004). Association genetics of complex traits in conifers. *Trends in Plant Science* **9**.7: 325– 330. DOI: 10.1016/j.tplants.2004.05.006.

Neale D. B., Wegrzyn J. L., Stevens K. A., Zimin A. V., Puiu D., Crepeau M. W., Cardeno C., Koriabine M., Holtz-Morris A. E., Liechty J. D., Martínez-García P. J., Vasquez-Gross H. A., Lin B. Y., Zieve J. J., Dougherty W. M., Fuentes-Soriano S., Wu L.-S., Gilbert D., Marçais G., Roberts M., Holt C., Yandell M., Davis J. M., Smith K. E., Dean J. F., Lorenz W. W., Whetten R. W., Sederoff R., Wheeler N., McGuire P. E., Main D., Loopstra C. A., Mockaitis K., deJong P. J., Yorke J. A., Salzberg S. L., and Langley C. H. (2014). Decoding the massive genome of loblolly pine using haploid DNA and novel assembly strategies. *Genome Biology* **15**.3: R59. DOI: 10.1186/gb-2014-15-3-r59.

Newman J. R. S., Ghaemmaghami S., Ihmels J., Breslow D. K., Noble M., DeRisi J. L., and Weissman J. S. (2006). Single-cell proteomic analysis of S. cerevisiae reveals the architecture of biological noise. *Nature* **441**.7095: 840– 846. DOI: 10.1038/nature04785.

Nicholson G., Smith A. V., Jónsson F., Gústafsson Ó., Stefánsson K., and Donnelly P. (2002). Assessing population differentiation and isolation from single-nucleotide polymorphism data. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **64**.4: 695– 715. DOI: 10.1111/1467-9868.00357.

Nicotra A. B., Atkin O. K., Bonser S. P., Davidson A. M., Finnegan E. J., Mathesius U., Poot P., Purugganan M. D., Richards C. L., Valladares F., and van Kleunen M. (2010). Plant phenotypic plasticity in a changing climate. *Trends in Plant Science* **15**.12: 684– 692. DOI: 10.1016/j.tplants.2010.09.008.

Nosil P., Flaxman S. M., Feder J. L., and Gompert Z. (2020). Increasing our ability to predict contemporary evolution. *Nature Communications* **11**.1: 5592. DOI: 10.1038/s41467-020-19437-x.

Nosil P., Villoutreix R., de Carvalho C. F., Farkas T. E., Soria-Carrasco V., Feder J. L., Crespi B. J., and Gompert Z. (2018). Natural selection and the predictability of evolution in Timema stick insects. *Science* **359**.6377: 765– 770. DOI: 10.1126/science.aap9125.

Nosil P., Vines T. H., and Funk D. J. (2005). Perspective: Reproductive isolation caused by natural selection against immigrants from divergent habitats. *Evolution; International Journal of Organic Evolution* **59**.4: 705– 719.

Novacek M. J. and Wheeler Q. D. (1992). *Extinction and phylogeny*. Columbia University Press.

Nussey D. H., Postma E., Gienapp P., and Visser M. E. (2005). Selection on Heritable Phenotypic Plasticity in a Wild Bird Population. *Science* **310**.5746: 304– 306. DOI: 10.1126/science.1117004.

Nystedt B., Street N. R., Wetterbom A., Zuccolo A., Lin Y.-C., Scofield D. G., Vezzi F., Delhomme N., Giacomello S., Alexeyenko A., Vicedomini R., Sahlin K., Sherwood E., Elfstrand M., Gramzow L., Holmberg K., Hällman J., Keech O., Klasson L., Koriabine M., Kucukoglu M., Käller M., Luthman J., Lysholm F., Niittylä T., Olson Å., Rilakovic N., Ritland C., Rosselló J. A., Sena J., Svensson T., Talavera-López C., Theißen G., Tuominen H., Vanneste K., Wu Z.-Q., Zhang B., Zerbe P., Arvestad L., Bhalerao R., Bohlmann J., Bousquet J., Garcia Gil R., Hvidsten T. R., de Jong P., MacKay J., Morgante M., Ritland K., Sundberg B., Lee Thompson S., Van de Peer Y., Andersson B., Nilsson O., Ingvarsson P. K., Lundeberg J., and Jansson S. (2013). The Norway spruce genome sequence and conifer genome evolution. *Nature* **497**.7451: 579– 584. DOI: 10.1038/nature12211.

Obeso J. R. (2002). The costs of reproduction in plants. *New Phytologist* **155**.3: 321– 348. DOI: 10.1046/j.1469-8137.2002.00477.x.

Ockendon N., Baker D. J., Carr J. A., White E. C., Almond R. E. A., Amano T., Bertram E., Bradbury R. B., Bradley C., Butchart S. H. M., Doswald N., Foden W., Gill D. J. C., Green R. E., Sutherland W. J., Tanner E. V. J., and Pearce-Higgins J. W. (2014). Mechanisms underpinning climatic impacts on natural populations: altered species interactions are more important than direct effects. *Global Change Biology* **20**.7: 2221– 2229. DOI: 10.1111/gcb.12559.

O'Connor L. J., Schoech A. P., Hormozdiari F., Gazal S., Patterson N., and Price A. L. (2019). Extreme Polygenicity of Complex Traits Is Explained by Negative Selection. *The American Journal of Human Genetics* **105**.3: 456– 476. DOI: 10.1016/j.ajhg.2019.07.003.

Oddou-Muratorio S., Davi H., and Lefèvre F. (2020). Integrating evolutionary, demographic and ecophysiologicalprocesses to predict the adaptive dynamics of forest tree populations under globalchange. *Tree Genetics & Genomes* **16**.5: 67. DOI: 10.1007/s11295-020-01451-1.

Oddou-Muratorio S., Klein E. K., Vendramin G. G., and Fady B. (2011). Spatial vs. temporal effects on demographic and genetic structures: the roles of dispersal, masting and differential mortality on patterns of recruitment in Fagus sylvatica. *Molecular Ecology* **20**.9: 1997– 2010. DOI: 10.1111/j.1365-294X.2011.05039.x.

Oetting W. S., Jacobson P. A., and Israni A. K. (2017). Validation Is Critical for Genome-Wide Association Study–Based Associations. *American Journal of Transplantation* **17**.2: 318– 319. DOI: 10.1111/ajt.14051.

O'hara R. B., Cano J. M., Ovaskainen O., Teplitsky C., and Alho J. S. (2008). Bayesian approaches in evolutionary quantitative genetics. *Journal of Evolutionary Biology* **21**.4: 949– 957. DOI: 10.1111/j.1420-9101.2008.01529.x.

Ohta T. and Gillespie J. H. (1996). Development of Neutral and Nearly Neutral Theories. *Theoretical Population Biology* **49**.2: 128– 142. DOI: 10.1006/tpbi.1996.0007.

Olson M. S., Robertson A. L., Takebayashi N., Silim S., Schroeder W. R., and Tiffin P. (2010). Nucleotide diversity and linkage disequilibrium in balsam poplar (Populus balsamifera). *New Phytologist* **186**.2: 526– 536. DOI: 10.1111/j.1469-8137.2009.03174.x.

O'Neill G. A., Hamann A., and Wang T. (2008). Accounting for population variation improves estimates of the impact of climate change on species' growth and distribution. *Journal of Applied Ecology*: 1040– 1049. DOI: 10.1111/j.1365-2664.2008.01472.x@10.1111/(ISSN)1365-2664.CLIMATE_JPE.

Orr H. A. and Unckless R. L. (2008). Population Extinction and the Genetics of Adaptation. *The American Naturalist* **172**.2: 160– 169. DOI: 10.1086/589460.

Ovaskainen O., Knegt H. J. de, and Delgado M. d. M. (2016). *Quantitative Ecology and Evolutionary Biology: Integrating models with data*. Oxford University Press. 301 pp. ISBN: 978-0-19-102422-1.

Paaby A. B. and Rockman M. V. (2014). Cryptic genetic variation: evolution's hidden substrate. *Nature Reviews Genetics* **15**.4: 247– 258. DOI: 10.1038/nrg3688.

Parmesan C. (2006). Ecological and Evolutionary Responses to Recent Climate Change. *Annual Review of Ecology, Evolution, and Systematics* **37**.1: 637– 669. DOI: 10.1146/annurev.ecolsys.37.091305.110100.

Parmesan C. and Yohe G. (2003). A globally coherent fingerprint of climate change impacts across natural systems. *Nature* **421**.6918: 37– 42. DOI: 10.1038/nature01286.

Pausas J. G. and Fernández-Muñoz S. (2012). Fire regime changes in the Western Mediterranean Basin: from fuel-limited to drought-driven fire regime. *Climatic Change* **110**.1: 215– 226. DOI: 10.1007/s10584-011-0060-6.

Pedlar J. H. and McKenney D. W. (2017). Assessing the anticipated growth response of northern conifer populations to a warming climate. *Scientific Reports* **7**: 43881. DOI: 10.1038/srep43881.

Pélabon C., Hansen T. F., Carter A. J. R., and Houle D. (2010). Evolution of Variation and Variability Under Fluctuating, Stabilizing, and Disruptive Selection. *Evolution* **64**.7: 1912–1925. DOI: 10.1111/j.1558-5646.2010.00979.x.

Pennisi E. (2018). Buying time. *Science* **362**.6418: 988– 991. DOI: 10.1126/science.362.6418.988.

Peñuelas J., Sardans J., Estiarte M., Ogaya R., Carnicer J., Coll M., Barbeta A., Rivas-Ubach A., Llusià J., Garbulsky M., Filella I., and Jump A. S. (2013). Evidence of current impact of climate change on life: a walk from genes to the biosphere. *Global Change Biology* **19**.8: 2303– 2338. DOI: 10.1111/gcb.12143.

Petit R. J. and Hampe A. (2006). Some Evolutionary Consequences of Being a Tree. *Annual Review of Ecology, Evolution, and Systematics* **37**.1: 187– 214. DOI: 10.1146/annurev.ecolsys.37.091305.110215.

Phillips B. L., Brown G. P., Webb J. K., and Shine R. (2006). Invasion and the evolution of speed in toads. *Nature* **439**.7078: 803– 803. DOI: 10.1038/439803a.

Pigliucci M. (2005). Evolution of phenotypic plasticity: where are we going now? *Trends in Ecology & Evolution* **20**.9: 481– 486. DOI: 10.1016/j.tree.2005.06.001.

Pigliucci M., Murren C. J., and Schlichting C. D. (2006). Phenotypic plasticity and evolution by genetic assimilation. *Journal of Experimental Biology* **209**.12: 2362– 2367. DOI: 10.1242/jeb.02070.

Plomion C., Bartholomé J., Lesur I., Boury C., Rodríguez-Quilón I., Lagraulet H., Ehrenmann F., Bouffier L., Gion J. M., Grivet D., Miguel M. de, María N. de, Cervera M. T., Bagnoli F., Isik F., Vendramin G. G., and González-Martínez S. C. (2016a). High-density SNP assay development for genetic analysis in maritime pine (Pinus pinaster). *Molecular Ecology Resources* **16**.2: 574– 587. DOI: 10.1111/1755-0998.12464.

Plomion C., Bastien C., Bogeat-Triboulot M.-B., Bouffier L., Déjardin A., Duplessis S., Fady B., Heuertz M., Le Gac A.-L., Le Provost G., Legué V., Lelu-Walter M.-A., Leplé J.-C., Maury S., Morel A., Oddou-Muratorio S., Pilate G., Sanchez L., Scotti I., Scotti-Saintagne C., Segura V., Trontin J.-F., and Vacher C. (2016b). Forest tree genomics: 10 achievements from the past 10 years and future prospects. *Annals of Forest Science* **73**.1: 77– 103. DOI: 10.1007/s13595-015-0488-3.

Postma F. M. and Ågren J. (2016). Early life stages contribute strongly to local adaptation in Arabidopsis thaliana. *Proceedings of the National Academy of Sciences* **113**.27: 7590– 7595. DOI: 10.1073/pnas.1606303113.

Prada E., Climent J., Alía R., and Díaz R. (2016). Life-history correlations with seasonal cold hardiness in maritime pine. *American Journal of Botany* **103**.12: 2126– 2135. DOI: 10.3732/ajb.1600286.

Price G. R. (1970). Selection and Covariance. *Nature* **227**.5257: 520– 521. DOI: 10.1038/227520a0.

Price T. and Schluter D. (1991). On the Low Heritability of Life-History Traits. *Evolution* **45**.4: 853– 861. DOI: 10.1111/j.1558-5646.1991.tb04354.x.

Pritchard J. K., Pickrell J. K., and Coop G. (2010). The Genetics of Human Adaptation: Hard Sweeps, Soft Sweeps, and Polygenic Adaptation. *Current Biology* **20**.4: R208– R215. DOI: 10.1016/j.cub.2009.11.055.

Pritchard J. K., Stephens M., and Donnelly P. (2000). Inference of Population Structure Using Multilocus Genotype Data. *Genetics* **155**.2: 945– 959.

Prober S. M., Hilbert D. W., Ferrier S., Dunlop M., and Gobbett D. (2012). Combining community-level spatial modelling and expert knowledge to inform climate adaptation in temperate grassy eucalypt woodlands and related grasslands. *Biodiversity and Conservation* **21**.7: 1627– 1650. DOI: 10.1007/s10531-012-0268-4.

Prunier J., Gérardi S., Laroche J., Beaulieu J., and Bousquet J. (2012). Parallel and lineage-specific molecular adaptation to climate in boreal black spruce. *Molecular Ecology* **21**.17: 4270– 4286. DOI: 10.1111/j.1365-294X.2012.05691.x.

Prunier J., Laroche J., Beaulieu J., and Bousquet J. (2011). Scanning the genome for gene SNPs related to climate adaptation and estimating selection at the molecular level in boreal black spruce. *Molecular Ecology* **20**.8: 1702– 1716. DOI: 10.1111/j.1365-294X.2011.05045.x.

Purcell S. M., Moran J. L., Fromer M., Ruderfer D., Solovieff N., Roussos P., O'Dushlaine C., Chambert K., Bergen S. E., Kähler A., Duncan L., Stahl E., Genovese G., Fernández E., Collins M. O., Komiyama N. H., Choudhary J. S., Magnusson P. K. E., Banks E., Shakir K., Garimella K., Fennell T., DePristo M., Grant S. G. N., Haggarty S. J., Gabriel S., Scolnick E. M., Lander E. S., Hultman C. M., Sullivan P. F., McCarroll S. A., and Sklar P. (2014). A polygenic burden of rare disruptive mutations in schizophrenia. *Nature* **506**.7487: 185– 190. DOI: 10.1038/nature12975.

R Core Team (2020). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Radersma R., Noble D. W. A., and Uller T. (2020). Plasticity leaves a phenotypic signature during local adaptation. *Evolution Letters* **4**.4: 360– 370. DOI: 10.1002/evl3.185.

Ramírez-Valiente J. A., Etterson J. R., Deacon N. J., and Cavender-Bares J. (2019). Evolutionary potential varies across populations and traits in the neotropical oak Quercus oleoides. *Tree Physiology* **39**.3: 427– 439. DOI: 10.1093/treephys/tpy108.

Razgour O., Forester B., Taggart J. B., Bekaert M., Juste J., Ibáñez C., Puechmaille S. J., Novella-Fernandez R., Alberdi A., and Manel S. (2019). Considering adaptive genetic variation in climate change vulnerability assessment reduces species range loss projections. *Proceedings of the National Academy of Sciences* **116**.21: 10418– 10423. DOI: 10.1073/pnas.1820663116.

Rehfeldt G. E., Leites L. P., Joyce D. G., and Weiskittel A. R. (2018). Role of population genetics in guiding ecological responses to climate. *Global Change Biology* **24**.2: 858– 868. DOI: 10.1111/gcb.13883.

Rehfeldt G. E., Tchebakova N. M., Milyutin L. I., Parfenova E. I., Wykoff W. R., and Kouzmina N. A. (2003). Assessing population responses to climate in Pinus sylvestris and Larix spp. of Eurasia with climate-transfer models. *Eurasian Journal of Forest Research - Hokkaido University (Japan)*.

Rehfeldt G. E., Tchebakova N. M., Parfenova Y. I., Wykoff W. R., Kuzmina N. A., and Milyutin L. I. (2002). Intraspecific responses to climate in Pinus sylvestris. *Global Change Biology* **8**.9: 912– 929. DOI: 10.1046/j.1365-2486.2002.00516.x.

Rehfeldt G. E., Ying C. C., Spittlehouse D. L., and Hamilton D. A. (1999). Genetic Responses to Climate in Pinus Contorta: Niche Breadth, Climate Change, and Reforestation. *Ecological Monographs* **69**.3: 375– 407. DOI: 10.1890/0012-9615(1999)069[0375:GRTCIP]2.0.CO;2.

Rehm E. M. and Feeley K. J. (2015). The inability of tropical cloud forest species to invade grasslands above treeline during climate change: potential explanations and consequences. *Ecography* **38**.12: 1167– 1175. DOI: 10.1111/ecog.01050.

Reich P. B. and Oleksyn J. (2008). Climate warming will reduce growth and survival of Scots pine except in the far north. *Ecology Letters* **11**.6: 588– 597. DOI: 10.1111/j.1461-0248.2008.01172.x.

Rellstab C., Dauphin B., and Exposito-Alonso M. (2021). Prospects and limitations of genomic offset in conservation management. *Evolutionary Applications* **14**.5: 1202– 1212. DOI: 10.1111/eva.13205.

Rellstab C., Gugerli F., Eckert A. J., Hancock A. M., and Holderegger R. (2015). A practical guide to environmental association analysis in landscape genomics. *Molecular Ecology* **24**.17: 4348– 4370. DOI: 10.1111/mec.13322.

Rellstab C., Zoller S., Walthert L., Lesur I., Pluess A. R., Graf R., Bodénès C., Sperisen C., Kremer A., and Gugerli F. (2016). Signatures of local adaptation in candidate genes of oaks (Quercus spp.) with respect to present and future climatic conditions. *Molecular Ecology* **25**.23: 5907– 5924. DOI: 10.1111/mec.13889.

Resende Jr M. F. R., Muñoz P., Acosta J. J., Peter G. F., Davis J. M., Grattapaglia D., Resende M. D. V., and Kirst M. (2012). Accelerating the domestication of trees using genomic selection: accuracy of prediction models across ages and environments. *New Phytologist* **193**.3: 617– 624. DOI: 10.1111/j.1469-8137.2011.03895.x.

Resende M. D. V., Resende M. F. R., Sansaloni C. P., Petroli C. D., Missiaggia A. A., Aguiar A. M., Abad J. M., Takahashi E. K., Rosado A. M., Faria D. A., Pappas G. J., Kilian A., and Grattapaglia D. (2012). Genomic selection for growth and wood quality in Eucalyptus: capturing the missing heritability and accelerating breeding for complex traits in forest trees. *New Phytologist*: 116– 128. DOI: 10.1111/j.1469-8137.2011.04038.x@10.1002/(ISSN)1469-8137(CAT)FeatureIssues(VI)Bioenergytrees.

Reznick D. N. and Ghalambor C. K. (2001). The population ecology of contemporary adaptations: what empirical studies reveal about the conditions that promote adaptive evolution. *Genetica* **112**.1: 183– 198. DOI: 10.1023/A:1013352109042.

Reznick D. N., Shaw F. H., Rodd F. H., and Shaw R. G. (1997). Evaluation of the Rate of Evolution in Natural Populations of Guppies (Poecilia reticulata). *Science* **275**.5308: 1934– 1937. DOI: 10.1126/science.275.5308.1934.

Reznick D. and Travis J. (2018). Is evolution predictable? *Science* **359**.6377: 738– 739. DOI: 10.1126/science.aas9043.

Rice S. H. and Papadopoulos A. (2009). Evolution with Stochastic Fitness and Stochastic Migration. *PLOS ONE* **4**.10: e7130. DOI: 10.1371/journal.pone.0007130.

Richardson J. L., Urban M. C., Bolnick D. I., and Skelly D. K. (2014). Microgeographic adaptation and the spatial scale of evolution. *Trends in Ecology & Evolution* **29**.3: 165– 176. DOI: 10.1016/j.tree.2014.01.002.

Riley S. J., DeGloria S. D., and Elliot R. (1999). Index that quantifies topographic heterogeneity. *Intermountain Journal of Sciences* **5**.1-4: 23– 27.

Robertson A. (1966). A mathematical model of the culling process in dairy cattle. *Animal Science* **8**.1: 95– 108. DOI: 10.1017/S0003356100037752.

Robinson M. R., Santure A. W., DeCauwer I., Sheldon B. C., and Slate J. (2013). Partitioning of genetic variation across the genome using multimarker methods in a wild bird population. *Molecular Ecology* **22**.15: 3963– 3980. DOI: 10.1111/mec.12375.

Robledo-Arnuncio J. J. and Gil L. (2005). Patterns of pollen dispersal in a small population of Pinus sylvestris L. revealed by total-exclusion paternity analysis. *Heredity* **94**.1: 13– 22. DOI: 10.1038/sj.hdy.6800542.

Rodríguez-Quilón I., Santos-del-Blanco L., Grivet D., Jaramillo-Correa J. P., Majada J., Vendramin G. G., Alía R., and González-Martínez S. C. (2015). Local effects drive heterozygosity–fitness correlations in an outcrossing long-lived tree. *Proceedings of the Royal Society B: Biological Sciences* **282**.1820: 20152230. DOI: 10.1098/rspb.2015.2230.

Rodríguez-Quilón I., Santos-del-Blanco L., Serra-Varela M. J., Koskela J., González-Martínez S. C., and Alía R. (2016). Capturing neutral and adaptive genetic diversity for conservation in a highly structured tree species. *Ecological Applications* **26**.7: 2254– 2266. DOI: 10.1002/eap.1361.

Rogalski M. A. (2017). Maladaptation to Acute Metal Exposure in Resurrected Daphnia ambigua Clones after Decades of Increasing Contamination. *The American Naturalist*. DOI: 10.1086/691077.

Rohde P. D., Østergaard S., Kristensen T. N., Sørensen P., Loeschcke V., Mackay T. F. C., and Sarup P. (2018). Functional Validation of Candidate Genes Detected by Genomic Feature Models. *G3 Genes|Genomes|Genetics* **8**.5: 1659– 1668. DOI: 10.1534/g3.118.200082.

Ruegg K., Bay R. A., Anderson E. C., Saracco J. F., Harrigan R. J., Whitfield M., Paxton E. H., and Smith T. B. (2018). Ecological genomics predicts climate vulnerability in an endangered southwestern songbird. *Ecology Letters* **21**.7: 1085– 1096. DOI: 10.1111/ele.12977.

Sala O. E., Chapin F. S., Iii, Armesto J. J., Berlow E., Bloomfield J., Dirzo R., Huber-Sanwald E., Huenneke L. F., Jackson R. B., Kinzig A., Leemans R., Lodge D. M., Mooney H. A., Oesterheld M., Poff N. L., Sykes M. T., Walker B. H., Walker M., and Wall D. H. (2000). Global Biodiversity Scenarios for the Year 2100. *Science* **287**.5459: 1770– 1774. DOI: 10.1126/science.287.5459.1770.

Samis K. E., López-Villalobos A., and Eckert C. G. (2016). Strong genetic differentiation but not local adaptation toward the range limit of a coastal dune plant. *Evolution* **70**.11: 2520– 2536. DOI: 10.1111/evo.13047.

Santos del Blanco L., Zas R., Notivol Paíno E., Chambel M. R., Majada J., and Climent J. (2010). Variation of early reproductive allocation in multi-site genetic trials of Maritime pine and Aleppo pine.

Santos-del-Blanco L., Climent J., González-Martínez S. C., and Pannell J. R. (2012). Genetic differentiation for size at first reproduction through male versus female functions in the widespread Mediterranean tree Pinus pinaster. *Annals of Botany* **110**.7: 1449– 1460. DOI: 10.1093/aob/mcs210.

Savolainen O., Lascoux M., and Merilä J. (2013). Ecological genomics of local adaptation. *Nature Reviews Genetics* **14**.11: 807– 820. DOI: 10.1038/nrg3522.

Savolainen O., Pyhäjärvi T., and Knürr T. (2007). Gene Flow and Local Adaptation in Trees. *Annual Review of Ecology, Evolution, and Systematics* **38**.1: 595– 619. DOI: 10.1146/annurev.ecolsys.38.091206.095646.

Scheiner S. M. (1993). Genetics and Evolution of Phenotypic Plasticity. *Annual Review of Ecology and Systematics* **24**.1: 35– 68. DOI: 10.1146/annurev.es.24.110193.000343.

— (2013). The genetics of phenotypic plasticity. XII. Temporal and spatial heterogeneity. *Ecology and Evolution* **3**.13: 4596– 4609. DOI: 10.1002/ece3.792.

Scherrer D., Vitasse Y., Guisan A., Wohlgemuth T., and Lischke H. (2020). Competition and demography rather than dispersal limitation slow down upward shifts of trees' upper elevation limits in the Alps. *Journal of Ecology* **108**.6: 2416– 2430. DOI: 10.1111/1365-2745.13451.

Schlichting C. D. (2008). Hidden Reaction Norms, Cryptic Genetic Variation, and Evolvability. *Annals of the New York Academy of Sciences* **1133**.1: 187– 203. DOI: 10.1196/annals.1438.010.

Schluter D. (2000a). Ecological Character Displacement in Adaptive Radiation. *The American Naturalist* **156**.S4: S4– S16. DOI: 10.1086/303412.

— (2000b). *The Ecology of Adaptive Radiation*. Oxford University Press, UK. 302 pp.

Schmitt J., Stinchcombe J. R., Heschel M. S., and Huber H. (2003). The Adaptive Evolution of Plasticity: Phytochrome-Mediated Shade Avoidance Responses. *Integrative and Comparative Biology* **43**.3: 459– 469. DOI: 10.1093/icb/43.3.459.

Schuster W. S. F. and Mitton J. B. (2000). Paternity and gene dispersal in limber pine (Pinus flexilis James). *Heredity* **84**.3: 348– 361. DOI: 10.1046/j.1365-2540.2000.00684.x.

Scott A. D., Zimin A. V., Puiu D., Workman R., Britton M., Zaman S., Caballero M., Read A. C., Bogdanove A. J., Burns E., Wegrzyn J., Timp W., Salzberg S. L., and Neale D. B. (2020). A Reference Genome Sequence for Giant Sequoia. *G3 Genes|Genomes|Genetics* **10**.11: 3907– 3919. DOI: 10.1534/g3.120.401612.

Scotti I., González-Martínez S. C., Budde K. B., and Lalagüe H. (2016). Fifty years of genetic studies: what to make of the large amounts of variation found within populations? *Annals of Forest Science* **73**.1: 69– 75. DOI: 10.1007/s13595-015-0471-z.

Scotti-Saintagne C., Bodénès C., Barreneche T., Bertocchi E., Plomion C., and Kremer A. (2004). Detection of quantitative trait loci controlling bud burst and height growth in Quercus robur L. *Theoretical and Applied Genetics* **109**.8: 1648– 1659. DOI: 10.1007/s00122-004-1789-3.

Seidl R., Thom D., Kautz M., Martin-Benito D., Peltoniemi M., Vacchiano G., Wild J., Ascoli D., Petr M., Honkaniemi J., Lexer M. J., Trotsiuk V., Mairota P., Svoboda M., Fabrika M., Nagel T. A., and Reyer C. P. O. (2017). Forest disturbances under climate change. *Nature Climate Change* **7**.6: 395– 402. DOI: 10.1038/nclimate3303.

Sella G. and Barton N. H. (2019). Thinking about the evolution of complex traits in the era of genome-wide association studies. *Annual Review of Genomics and Human Genetics* **20**.

Serra-Varela M. J., Grivet D., Vincenot L., Broennimann O., Gonzalo-Jiménez J., and Zimmermann N. E. (2015). Does phylogeographical structure relate to climatic niche divergence? A test using maritime pine (Pinus pinaster Ait.) *Global Ecology and Biogeography* **24**.11: 1302– 1313. DOI: 10.1111/geb.12369.

Shaw R. G. (2018). From the Past to the Future: Considering the Value and Limits of Evolutionary Prediction. *The American Naturalist* **193**.1: 1– 10. DOI: 10.1086/700565.

Shaw R. G. and Etterson J. R. (2012). Rapid climate change and the rate of adaptation: insight from experimental quantitative genetics. *New Phytologist* **195**.4: 752– 765. DOI: 10.1111/j.1469-8137.2012.04230.x.

Sheth S. N. and Angert A. L. (2018). Demographic compensation does not rescue populations at a trailing range edge. *Proceedings of the National Academy of Sciences* **115**.10: 2413– 2418. DOI: 10.1073/pnas.1715899115.

Shi H., Kichaev G., and Pasaniuc B. (2016). Contrasting the Genetic Architecture of 30 Complex Traits from Summary Association Data. *The American Journal of Human Genetics* **99**.1: 139– 153. DOI: 10.1016/j.ajhg.2016.05.013.

Shoemaker J. S., Painter I. S., and Weir B. S. (1999). Bayesian statistics in genetics: a guide for the uninitiated. *Trends in Genetics* **15**.9: 354– 358. DOI: 10.1016/S0168-9525(99)01751-5.

Shoval O., Sheftel H., Shinar G., Hart Y., Ramote O., Mayo A., Dekel E., Kavanagh K., and Alon U. (2012). Evolutionary Trade-Offs, Pareto Optimality, and the Geometry of Phenotype Space. *Science* **336**.6085: 1157– 1160. DOI: 10.1126/science.1217405.

Sierra-de-Grado R., Pando V., Martínez-Zurimendi P., Peñalvo A., Báscones E., and Moulia B. (2008). Biomechanical differences in the stem straightening process among Pinus pinaster provenances. A new approach for early selection of stem straightness. *Tree Physiology* **28**.6: 835– 846. DOI: 10.1093/treephys/28.6.835.

Sierra-Lucero V., McKeand S., Huber D., Rockwood D., and White T. (2002). Performance Differences and Genetic Parameters for Four Coastal Provenances of Loblolly Pine in the Southeastern United States. *Forest Science* **48**.4: 732– 742. DOI: 10.1093/forestscience/48.4.732.

Simpson G. G. (1944). *Tempo and Mode in Evolution*. Columbia University Press, New York.

Siol M., Wright S. I., and Barrett S. C. H. (2010). The population genomics of plant adaptation. *New Phytologist* **188**.2: 313– 332. DOI: 10.1111/j.1469-8137.2010.03401.x.

Sittaro F., Paquette A., Messier C., and Nock C. A. (2017). Tree range expansion in eastern North America fails to keep pace with climate warming at northern range limits. *Global Change Biology* **23**.8: 3292– 3301. DOI: 10.1111/gcb.13622.

Slatkin M. (1985). Gene Flow in Natural Populations. *Annual Review of Ecology and Systematics* **16**: 393– 430.

Sohail M., Maier R. M., Ganna A., Bloemendal A., Martin A. R., Turchin M. C., Chiang C. W., Hirschhorn J., Daly M. J., Patterson N., Neale B., Mathieson I., Reich D., and Sunyaev S. R. (2019). Polygenic adaptation on height is overestimated due to uncorrected stratification in genome-wide association studies. *eLife* **8**: e39702. DOI: 10.7554/eLife.39702.

Sork V. L., Aitken S. N., Dyer R. J., Eckert A. J., Legendre P., and Neale D. B. (2013). Putting the landscape into the genomics of trees: approaches for understanding local adaptation and population responses to changing climate. *Tree Genetics & Genomes* **9**.4: 901– 911. DOI: 10.1007/s11295-013-0596-x.

Sork V. L. (2018). Genomic Studies of Local Adaptation in Natural Plant Populations. *Journal of Heredity* **109**.1: 3– 15. DOI: 10.1093/jhered/esx091.

Sork V. L., Stowe K. A., and Hochwender C. (1993). Evidence for Local Adaptation in Closely Adjacent Subpopulations of Northern Red Oak (Quercus rubra L.) Expressed as Resistance to Leaf Herbivores. *The American Naturalist* **142**.6: 928– 936. DOI: 10.1086/285581.

Soulé M. E. (1991). Conservation: Tactics for a Constant Crisis. *Science* **253**.5021: 744– 750. DOI: 10.1126/science.253.5021.744.

Spichtig M. and Kawecki T. J. (2004). The Maintenance (or Not) of Polygenic Variation by Soft Selection in Heterogeneous Environments. *The American Naturalist* **164**.1: 70– 84. DOI: 10.1086/421335.

Spitze K. (1993). Population structure in Daphnia obtusa: quantitative genetic and allozymic variation. *Genetics* **135**.2: 367– 374.

Stapley J., Reger J., Feulner P. G. D., Smadja C., Galindo J., Ekblom R., Bennison C., Ball A. D., Beckerman A. P., and Slate J. (2010). Adaptation genomics: the next generation. *Trends in Ecology & Evolution* **25**.12: 705– 712. DOI: 10.1016/j.tree.2010.09.002.

Steffen W., Persson Å., Deutsch L., Zalasiewicz J., Williams M., Richardson K., Crumley C., Crutzen P., Folke C., Gordon L., Molina M., Ramanathan V., Rockström J., Scheffer M., Schellnhuber H. J., and Svedin U. (2011). The Anthropocene: From Global Change to Planetary Stewardship. *AMBIO* **40**.7: 739. DOI: 10.1007/s13280-011-0185-x.

Stein A., Gerstner K., and Kreft H. (2014). Environmental heterogeneity as a universal driver of species richness across taxa, biomes and spatial scales. *Ecology Letters* **17**.7: 866– 880. DOI: 10.1111/ele.12277.

Stern A. J., Speidel L., Zaitlen N. A., and Nielsen R. (2020). Disentangling selection on genetically correlated polygenic traits using whole-genome genealogies. *bioRxiv*. DOI: 10.1101/2020.05.07.083402.

Stevens K. A., Wegrzyn J. L., Zimin A., Puiu D., Crepeau M., Cardeno C., Paul R., Gonzalez-Ibeas D., Koriabine M., Holtz-Morris A. E., Martínez-García P. J., Sezen U. U., Marçais G., Jermstad K., McGuire P. E., Loopstra C. A., Davis J. M., Eckert A., de Jong P., Yorke J. A., Salzberg S. L., Neale D. B., and Langley C. H. (2016). Sequence of the Sugar Pine Megagenome. *Genetics* **204**.4: 1613– 1626. DOI: 10.1534/genetics.116.193227.

Stock A. J., Campitelli B. E., and Stinchcombe J. R. (2014). Quantitative genetic variance and multivariate clines in the Ivyleaf morning glory, Ipomoea hederacea. *Philosophical Transactions of the Royal Society B: Biological Sciences* **369**.1649: 20130259. DOI: 10.1098/rstb.2013.0259.

Storfer A., Murphy M. A., Evans J. S., Goldberg C. S., Robinson S., Spear S. F., Dezzani R., Delmelle E., Vierling L., and Waits L. P. (2007). Putting the 'landscape' in landscape genetics. *Heredity* **98**.3: 128– 142. DOI: 10.1038/sj.hdy.6800917.

Storfer A., Patton A., and Fraik A. K. (2018). Navigating the Interface Between Landscape Genetics and Landscape Genomics. *Frontiers in Genetics* **9**: 68. DOI: 10.3389/fgene.2018.00068.

Stovall A. E. L., Shugart H., and Yang X. (2019). Tree height explains mortality risk during an intense drought. *Nature Communications* **10**.1: 1– 6. DOI: 10.1038/s41467-019-12380-6.

Supple M. A., Bragg J. G., Broadhurst L. M., Nicotra A. B., Byrne M., Andrew R. L., Widdup A., Aitken N. C., and Borevitz J. O. (2018). Landscape genomic prediction for restoration of a Eucalyptus foundation species under climate change. *eLife* **7**. Ed. by Kliebenstein D. J.: e31835. DOI: 10.7554/eLife.31835.

Svensson E. and Calsbeek R. (2012). *The Adaptive Landscape in Evolutionary Biology*. Oxford University Press, UK. 338 pp.

Tabor H. K., Risch N. J., and Myers R. M. (2002). Candidate-gene approaches for studying complex genetic traits: practical considerations. *Nature Reviews Genetics* **3**.5: 391– 397. DOI: 10.1038/nrg796.

Tam V., Patel N., Turcotte M., Bossé Y., Paré G., and Meyre D. (2019). Benefits and limitations of genome-wide association studies. *Nature Reviews Genetics* **20**.8: 467– 484. DOI: 10.1038/s41576-019-0127-1.

Tapias R., Climent J., Pardos J. A., and Gil L. (2004). Life histories of Mediterranean pines. *Plant Ecology* **171**.1: 53– 68. DOI: 10.1023/B:VEGE.0000029383.72609.f0.

Templeton A. R. (2006). *Population Genetics and Microevolutionary Theory*. John Wiley & Sons. 769 pp. ISBN: 978-1-118-50436-9.

Thibault E., Soolanayakanahally R., and Keller S. R. (2020). Latitudinal clines in bud flush phenology reflect genetic variation in chilling requirements in balsam poplar, Populus balsamifera. *American Journal of Botany* **107**.11: 1597– 1605. DOI: 10.1002/ajb2.1564.

Thompson J. N. (1998). Rapid evolution as an ecological process. *Trends in Ecology & Evolution* **13**.8: 329– 332. DOI: 10.1016/S0169-5347(98)01378-0.

Thomson C. E., Winney I. S., Salles O., and Pujol B. (2018). A guide to using a multiple-matrix animal model to disentangle genetic and nongenetic causes of phenotypic variance. *bioRxiv*. DOI: 10.1101/318451.

Tiffin P. and Ross-Ibarra J. (2014). Advances and limits of using population genetics to understand local adaptation. *Trends in Ecology & Evolution* **29**.12: 673– 680. DOI: 10.1016/j.tree.2014.10.004.

Tigano A. and Friesen V. L. (2016). Genomics of local adaptation with gene flow. *Molecular Ecology* **25**.10: 2144– 2164. DOI: 10.1111/mec.13606.

Triantis K. A., Mylonas M., Lika K., and Vardinoyannis K. (2003). A model for the species–area–habitat relationship. *Journal of Biogeography* **30**.1: 19– 27. DOI: 10.1046/j.1365-2699.2003.00805.x.

Tufto J. (2000). The Evolution of Plasticity and Nonplastic Spatial and Temporal Adaptations in the Presence of Imperfect Environmental Cues. *The American Naturalist* **156**.2: 121– 130. DOI: 10.1086/303381.

Urban M. C. (2015). Accelerating extinction risk from climate change. *Science* **348**.6234: 571–573. DOI: 10.1126/science.aaa4984.

Valladares F., Gianoli E., and Gómez J. M. (2014a). Ecological limits to plant phenotypic plasticity. *New Phytologist* **176**.4: 749– 763. DOI: 10.1111/j.1469-8137.2007.02275.x.

Valladares F., Matesanz S., Guilhaumon F., Araújo M. B., Balaguer L., Benito-Garzón M., Cornwell W., Gianoli E., Kleunen M. van, Naya D. E., Nicotra A. B., Poorter H., and Zavala M. A. (2014b). The effects of phenotypic plasticity and local adaptation on forecasts of species range shifts under climate change. *Ecology Letters* **17**.11: 1351– 1364. DOI: 10.1111/ele.12348.

Van Boheemen L. A. and Hodgins K. A. (2020). Rapid repeatable phenotypic and genomic adaptation following multiple introductions. *Molecular Ecology* **29**.21: 4102– 4117. DOI: 10.1111/mec.15429.

Van Heerwaarden B., Kellermann V., Schiffer M., Blacket M., Sgrò C. M., and Hoffmann A. A. (2009). Testing evolutionary hypotheses about species borders: patterns of genetic variation towards the southern borders of two rainforest Drosophila and a related habitat generalist. *Proceedings of the Royal Society B: Biological Sciences* **276**.1661: 1517– 1526. DOI: 10.1098/rspb.2008.1288.

Vanhove M., Pina-Martins F., Coelho A. C., Branquinho C., Costa A., Batista D., Príncipe A., Sousa P., Henriques A., Marques I., Belkadi B., Knowles L. L., and Paulo O. S. (2021). Using gradient Forest to predict climate response and adaptation in Cork oak. *Journal of Evolutionary Biology* **34**.6: 910– 923. DOI: 10.1111/jeb.13765.

VanRaden P. M. (2008). Efficient Methods to Compute Genomic Predictions. *Journal of Dairy Science* **91**.11: 4414– 4423. DOI: 10.3168/jds.2007-0980.

Vehtari A., Gelman A., and Gabry J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing* **27**.5: 1413– 1432. DOI: 10.1007/s11222-016-9696-4.

Via S. (1984). The Quantitative Genetics of Polyphagy in an Insect Herbivore. I. Genotype-Environment Interaction in Larval Performance on Different Host Plant Species. *Evolution* **38**.4: 881– 895. DOI: 10.2307/2408398.

Via S. and Lande R. (1985). Genotype-Environment Interaction and the Evolution of Phenotypic Plasticity. *Evolution* **39**.3: 505– 522. DOI: 10.1111/j.1558-5646.1985.tb00391.x.

Viñas R. A., Caudullo G., Oliveira S., and de Rigo D. (2016). Pinus pinaster in Europe: distribution, habitat, usage and threats.

Visscher P. M. and Goddard M. E. (2019). From R.A. Fisher's 1918 Paper to GWAS a Century Later. *Genetics* **211**.4: 1125– 1130. DOI: 10.1534/genetics.118.301594.

Visscher P. M., Wray N. R., Zhang Q., Sklar P., McCarthy M. I., Brown M. A., and Yang J. (2017). 10 Years of GWAS Discovery: Biology, Function, and Translation. *The American Journal of Human Genetics* **101**.1: 5– 22. DOI: 10.1016/j.ajhg.2017.06.005.

Visser M. E. (2008). Keeping up with a warming world; assessing the rate of adaptation to climate change. *Proceedings of the Royal Society B: Biological Sciences* **275**.1635: 649– 659. DOI: 10.1098/rspb.2007.0997.

Vitasse Y. V., Delzon S. D., Bresson C. C. B. C., Michalet R. M., and Kremer A. K. (2009). Altitudinal differentiation in growth and phenology among populations of temperate-zone tree species growing in a common garden. *Canadian Journal of Forest Research*. DOI: 10.1139/X09-054.

Vitousek P. M. (1992). Global environmental change: an introduction. *Annual review of Ecology and Systematics* **23**.1: 1– 14.

Vizcaíno-Palomar N., Fady B., Alía R., Raffin A., Mutke S., and Benito Garzón M. (2020). The legacy of climate variability over the last century on populations' phenotypic variation in tree height. *Science of The Total Environment* **749**: 141454. DOI: 10.1016/j.scitotenv.2020.141454.

Vizcaíno-Palomar N., Revuelta-Eugercios B., Zavala M. A., Alía R., and González-Martínez S. C. (2014). The Role of Population Origin and Microenvironment in Seedling Emergence and Early Survival in Mediterranean Maritime Pine (Pinus pinaster Aiton). *PLOS ONE* **9**.10: e109132. DOI: 10.1371/journal.pone.0109132.

Vogt G. (2015). Stochastic developmental variation, an epigenetic source of phenotypic diversity with far-reaching biological consequences. *Journal of Biosciences* **40**.1: 159– 204. DOI: 10.1007/s12038-015-9506-8.

Waddington C. H. (1952). Selection of the Genetic Basis for an Acquired Character. *Nature* **169**.4302: 625– 626. DOI: 10.1038/169625b0.

— (1953). Genetic Assimilation of an Acquired Character. *Evolution* **7**.2: 118– 126. DOI: 10.2307/2405747.

Waldvogel A.-M., Feldmeyer B., Rolshausen G., Exposito-Alonso M., Rellstab C., Kofler R., Mock T., Schmid K., Schmitt I., Bataillon T., Savolainen O., Bergland A., Flatt T., Guillaume F., and Pfenninger M. (2020). Evolutionary genomics can improve prediction of species' responses to climate change. *Evolution Letters* **4**.1: 4– 18. DOI: 10.1002/evl3.154.

Walsh B. and Blows M. W. (2009). Abundant Genetic Variation + Strong Selection = Multivariate Genetic Constraints: A Geometric View of Adaptation. *Annual Review of Ecology, Evolution, and Systematics* **40**.1: 41– 59. DOI: 10.1146/annurev.ecolsys.110308.120232.

Walsh B. and Lynch M. (2018). *Evolution and Selection of Quantitative Traits*. Oxford University Press. 1490 pp. ISBN: 978-0-19-256664-5.

Walther G.-R., Post E., Convey P., Menzel A., Parmesan C., Beebee T. J. C., Fromentin J.-M., Hoegh-Guldberg O., and Bairlein F. (2002). Ecological responses to recent climate change. *Nature* **416**.6879: 389– 395. DOI: 10.1038/416389a.

Wang T., Hamann A., Yanchuk A., O'neill G. A., and Aitken S. N. (2006). Use of response functions in selecting lodgepole pine populations for future climates. *Global Change Biology* **12**.12: 2404– 2416. DOI: 10.1111/j.1365-2486.2006.01271.x.

Wang T., O'Neill G. A., and Aitken S. N. (2010). Integrating environmental and genetic effects to predict responses of tree populations to climate. *Ecological Applications* **20**.1: 153– 163. DOI: 10.1890/08-2257.1.

Warren R. L., Keeling C. I., Yuen M. M. S., Raymond A., Taylor G. A., Vandervalk B. P., Mohamadi H., Paulino D., Chiu R., Jackman S. D., Robertson G., Yang C., Boyle B., Hoffmann M., Weigel D., Nelson D. R., Ritland C., Isabel N., Jaquish B., Yanchuk A., Bousquet J., Jones S. J. M., MacKay J., Birol I., and Bohlmann J. (2015). Improved white spruce (Picea glauca) genome assemblies and annotation of large gene families of conifer terpenoid and phenolic defense metabolism. *The Plant Journal* **83**.2: 189– 212. DOI: 10.1111/tpj.12886.

Weed A. S., Ayres M. P., and Hicke J. A. (2013). Consequences of climate change for biotic disturbances in North American forests. *Ecological Monographs* **83**.4: 441– 470. DOI: 10.1890/13-0160.1.

Weir B. S. and Cockerham C. C. (1984). Estimating F-Statistics for the Analysis of Population Structure. *Evolution* **38**.6: 1358– 1370. DOI: 10.2307/2408641.

West-Eberhard M. J. (2003). *Developmental Plasticity and Evolution*. Oxford University Press. 815 pp. ISBN: 978-0-19-802856-7.

Westneat D. F., Wright J., and Dingemanse N. J. (2015). The biology hidden inside residual within-individual phenotypic variation. *Biological Reviews* **90**.3: 729– 743. DOI: 10.1111/brv.12131.

Wiens J. J. (2016). Climate-Related Local Extinctions Are Already Widespread among Plant and Animal Species. *PLOS Biology* **14**.12: e2001104. DOI: 10.1371/journal.pbio.2001104.

Wolak M. E. and Reid J. M. (2017). Accounting for genetic differences among unknown parents in microevolutionary studies: how to include genetic groups in quantitative genetic animal models. *Journal of Animal Ecology* **86**.1: 7– 20. DOI: 10.1111/1365-2656.12597.

Wolf J. B. and Wade M. J. (2009). What are maternal effects (and what are they not)? *Philosophical Transactions of the Royal Society B: Biological Sciences* **364**.1520: 1107– 1115. DOI: 10.1098/rstb.2008.0238.

Woodall C. W., Oswalt C. M., Westfall J. A., Perry C. H., Nelson M. D., and Finley A. O. (2009). An indicator of tree migration in forests of the eastern United States. *Forest Ecology and Management* **257**.5: 1434– 1444. DOI: 10.1016/j.foreco.2008.12.013.

Wray N. R., Kemper K. E., Hayes B. J., Goddard M. E., and Visscher P. M. (2019). Complex Trait Prediction from Genome Data: Contrasting EBV in Livestock to PRS in Humans: Genomic Prediction. *Genetics* **211**.4: 1131– 1141. DOI: 10.1534/genetics.119.301859.

Wray N. R., Wijmenga C., Sullivan P. F., Yang J., and Visscher P. M. (2018). Common Disease Is More Complex Than Implied by the Core Gene Omnigenic Model. *Cell* **173**.7: 1573– 1580. DOI: 10.1016/j.cell.2018.05.051.

Wu H. X. and Ying C. C. (2004). Geographic pattern of local optimality in natural populations of lodgepole pine. *Forest Ecology and Management* **194**.1: 177– 198. DOI: 10.1016/j.foreco.2004.02.017.

Wund M. A. (2012). Assessing the Impacts of Phenotypic Plasticity on Evolution. *Integrative and Comparative Biology* **52**.1: 5– 15. DOI: 10.1093/icb/ics050.

Wyckoff P. H. and Clark J. S. (2002). The relationship between growth and mortality for seven co-occurring tree species in the southern Appalachian Mountains. *Journal of Ecology* **90**.4: 604– 615. DOI: 10.1046/j.1365-2745.2002.00691.x.

Yang J., Bakshi A., Zhu Z., Hemani G., Vinkhuyzen A. A. E., Lee S. H., Robinson M. R., Perry J. R. B., Nolte I. M., Vliet-Ostaptchouk J. V. van, Snieder H., Esko T., Milani L., Mägi R., Metspalu A., Hamsten A., Magnusson P. K. E., Pedersen N. L., Ingelsson E., Soranzo N., Keller M. C., Wray N. R., Goddard M. E., and Visscher P. M. (2015). Genetic variance estimation with imputed variants finds negligible missing heritability for human height and body mass index. *Nature Genetics* **47**.10: 1114– 1120. DOI: 10.1038/ng.3390.

Yeaman S., Chen Y., and Whitlock M. C. (2010). No Effect of Environmental Heterogeneity on the Maintenance of Genetic Variation in Wing Shape in Drosophila Melanogaster. *Evolution* **64**.12: 3398– 3408. DOI: 10.1111/j.1558-5646.2010.01075.x.

Yeaman S., Hodgins K. A., Lotterhos K. E., Suren H., Nadeau S., Degner J. C., Nurkowski K. A., Smets P., Wang T., Gray L. K., Liepe K. J., Hamann A., Holliday J. A., Whitlock M. C., Rieseberg L. H., and Aitken S. N. (2016). Convergent local adaptation to climate in distantly related conifers. *Science* **353**.6306: 1431– 1433. DOI: 10.1126/science.aaf7812.

Yeaman S. and Jarvis A. (2006). Regional heterogeneity and gene flow maintain variance in a quantitative trait within populations of lodgepole pine. *Proceedings of the Royal Society B: Biological Sciences* **273**.1594: 1587– 1593. DOI: 10.1098/rspb.2006.3498.

Yeaman S. and Otto S. P. (2011). Establishment and Maintenance of Adaptive Genetic Divergence Under Migration, Selection, and Drift. *Evolution* **65**.7: 2123– 2129. DOI: 10.1111/j.1558-5646.2011.01277.x.

Yengo L., Sidorenko J., Kemper K. E., Zheng Z., Wood A. R., Weedon M. N., Frayling T. M., Hirschhorn J., Yang J., Visscher P. M., and the GIANT Consortium (2018). Meta-analysis of genome-wide association studies for height and body mass index in 700000 individuals of European ancestry. *Human Molecular Genetics* **27**.20: 3641– 3649. DOI: 10.1093/hmg/ddy271.

Zas R., Merlo E., Díaz R., and Fernández López J. (2003). Stability across sites of Douglas-fir provenances in northern Spain. **10**.1: 71– 82.

Zas R., Sampedro L., Solla A., Vivas M., Lombardero M. J., Alía R., and Rozas V. (2020). Dendroecology in common gardens: Population differentiation and plasticity in resistance, recovery and resilience to extreme drought events in Pinus pinaster. *Agricultural and Forest Meteorology* **291**: 108060. DOI: 10.1016/j.agrformet.2020.108060.

Zhou Y., Zhang L., Liu J., Wu G., and Savolainen O. (2014). Climatic adaptation and ecological divergence between two closely related pine species in Southeast China. *Molecular Ecology* **23**.14: 3504– 3522. DOI: 10.1111/mec.12830.

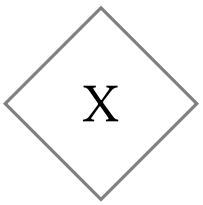
Zhu K., Woodall C. W., and Clark J. S. (2012). Failure to migrate: lack of tree range expansion in response to climate change. *Global Change Biology* **18**.3: 1042– 1052. DOI: 10.1111/j.1365-2486.2011.02571.x.

Zhu Y., Hogan J. A., Cai H., Xun Y., Jiang F., and Jin G. (2017). Biotic and abiotic drivers of the tree growth and mortality trade-off in an old-growth temperate forest. *Forest Ecology and Management* **404**: 354– 360. DOI: 10.1016/j.foreco.2017.09.004.

Zimin A., Stevens K. A., Crepeau M. W., Holtz-Morris A., Koriabine M., Marçais G., Puiu D., Roberts M., Wegrzyn J. L., de Jong P. J., Neale D. B., Salzberg S. L., Yorke J. A., and Langley C. H. (2014). Sequencing and Assembly of the 22-Gb Loblolly Pine Genome. *Genetics* **196**.3: 875– 890. DOI: 10.1534/genetics.113.159715.

Zonneveld B. J. M. (2012). Conifer genome sizes of 172 species, covering 64 of 67 genera, range from 8 to 72 picogram. *Nordic Journal of Botany* **30**.4: 490– 502. DOI: 10.1111/j.1756-1051.2012.01516.x.

Zuk O., Hechter E., Sunyaev S. R., and Lander E. S. (2012). The mystery of missing heritability: Genetic interactions create phantom heritability. *Proceedings of the National Academy of Sciences* **109**.4: 1193– 1198. DOI: 10.1073/pnas.1119675109.



Supplementary information - Chapter 1

1 The data

1.1 Phenotypic data

1.1.1 Details of the eight phenotypic traits

Traits	Common gardens	Dates of measurement	Tree age	Survival	Units	Trees	Populations	Clones	Transf.
Height	Portugal	October 2012	20	0.66	mm	2746	33	521	-
Height	Bordeaux	November 2013	25	0.97	mm	3238	33	430	-
Height	Bordeaux	November 2018	85	0.96	mm	3209	33	430	-
Height	Asturias	November 2012	21	0.96	mm	3973	33	522	-
Mean bud burst date	Bordeaux	2013, 2014, 2015, 2017	-	-	°C-day	3175	33	430	center
Mean duration of bud burst	Bordeaux	2014, 2015, 2017	-	-	°C-day	3187	33	430	center
Specific Leaf Area	Portugal	-	-	-	m^2/kg	2642	33	520	log
$\delta^{13}C$	Portugal	-	-	-	‰	1939	33	517	center

Table X.1. Information about the phenotypic traits. Tree age is in months. Survival is the proportion of survival in the common gardens at the measurement date. Some variables were log-transformed (*log*) or mean-centered (*center*) prior to analyses, which is indicated in the *Transf.* column.

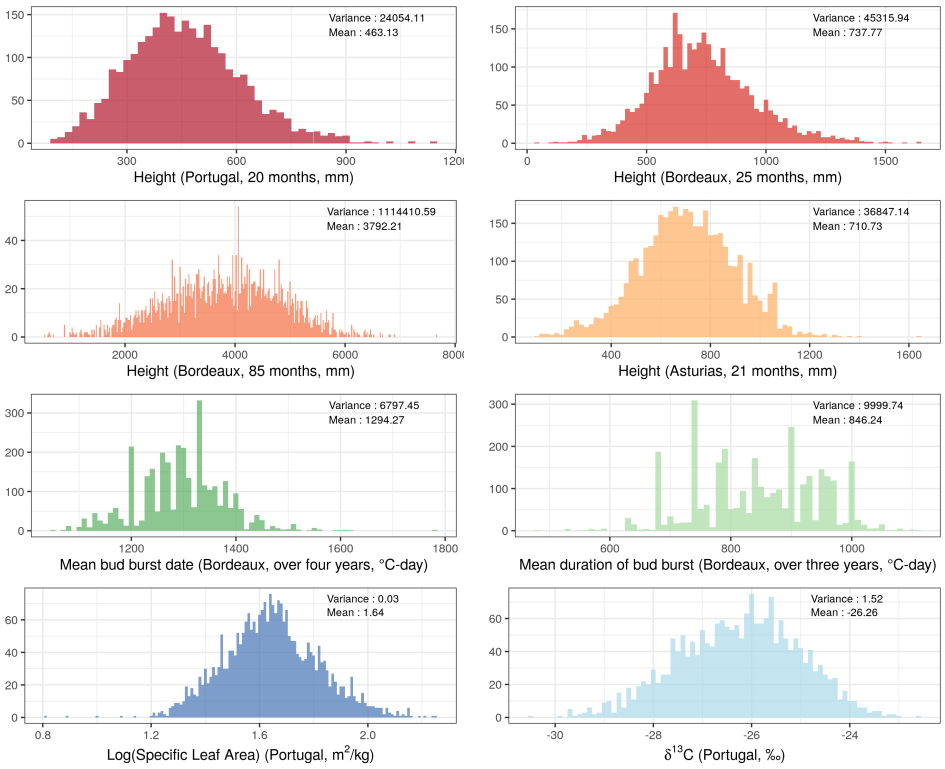


Figure X.1. Distribution of the phenotypic traits used in the study.

1.1.2 Population-specific distributions, means and variances

Pop.	Trait means								Trait variances							
	Ht (Portugal)	Ht (Bordeaux 2013)	Ht (Bordeaux 2018)	Ht (Asturias)	SLA	$\delta^{13}C$	meanBB	meanDBB	Ht (Portugal)	Ht (Bordeaux 2013)	Ht (Bordeaux 2018)	Ht (Asturias)	SLA	$\delta^{13}C$	meanBB	meanDBB
gray!6 ALT	525.56	760.19	4275.38	764.43	5.08	-25.94	1333.05	883.19	25044.03	25994.08	518190.0	43958.36	0.56	1.26	10179.89	11621.08
ARM	505.48	793.33	4450.16	818.89	5.19	-26.01	1282.95	859.95	14693.67	30154.84	416420.9	27155.20	0.53	0.75	5787.12	10512.04
gray!6 ARN	423.44	651.87	3302.00	652.35	5.19	-26.83	1309.37	861.13	20251.08	26534.21	487704.6	21763.14	0.96	1.68	1199.55	5285.84
BAY	394.29	618.36	3054.03	592.79	5.18	-26.49	1299.02	829.97	15606.98	24433.97	461502.5	20475.10	0.96	1.91	5347.63	7758.00
gray!6 BON	421.33	600.43	3273.83	702.61	5.37	-26.84	1197.00	752.85	15994.80	30225.90	1159572.0	26013.68	1.08	1.12	3843.38	7596.92
CAD	481.58	847.76	4631.79	781.60	5.34	-25.79	1342.52	898.14	32160.59	55641.88	894272.5	42048.76	0.85	1.11	13379.00	8728.85
gray!6 CAR	405.52	505.32	2321.30	584.58	4.81	-26.74	1282.36	813.30	18268.47	18912.40	396287.2	11029.61	0.58	1.22	9461.82	8783.43
CAS	501.67	753.01	4202.78	769.32	5.07	-25.90	1310.38	877.46	28617.67	35004.68	554102.0	30359.25	1.00	1.01	6072.29	12050.36
gray!6 CEN	484.82	663.51	3348.65	723.14	4.91	-26.32	1272.32	849.04	25174.51	31090.09	648606.5	26016.07	0.58	1.36	7633.28	15793.70
COC	406.47	587.13	2735.71	588.77	5.22	-26.47	1320.11	848.06	17647.00	25181.40	573367.0	25089.71	0.89	1.39	6714.06	7676.71
gray!6 COM	516.92	809.33	3220.71	756.13	4.86	-26.14	1260.71	836.87	22918.15	35020.23	342303.2	57311.18	0.53	0.83	4511.64	11303.29
CUE	402.29	587.38	2793.39	582.10	5.06	-26.35	1313.72	851.46	15749.06	29948.47	745261.0	19274.90	1.02	1.76	7290.54	8915.74
gray!6 HOU	549.58	841.95	4578.18	831.13	5.32	-25.79	1287.73	832.88	30575.88	36168.73	683021.5	37114.18	0.80	1.09	3998.32	8832.28
LAM	443.56	706.79	3824.82	736.38	5.37	-26.25	1322.51	893.70	18614.34	28385.84	503323.6	35455.80	0.90	1.44	4263.94	6633.86
gray!6 LEI	467.30	765.64	4156.02	743.20	5.12	-26.34	1317.84	880.28	23362.76	44161.14	799974.2	41372.76	0.81	1.29	8737.24	11448.55
MIM	447.76	783.87	4182.64	718.20	5.34	-26.05	1283.85	832.76	22395.23	51903.59	755436.3	33853.99	0.91	1.94	6008.16	8888.56
gray!6 OLB	445.60	700.35	3410.89	712.97	5.24	-26.43	1285.80	817.56	15336.79	23496.30	487763.2	24782.38	0.92	1.28	5575.52	9353.93
OLO	549.10	803.57	4390.16	828.55	5.41	-25.63	1287.50	829.00	22867.78	47499.69	534286.6	29013.66	0.84	1.08	4316.43	9240.09
gray!6 ORI	375.58	667.40	2935.64	639.17	5.40	-26.92	1239.57	797.85	15107.03	21736.00	391840.3	25020.48	1.10	0.94	4685.37	7988.44
PET	525.13	829.51	4471.39	781.10	5.29	-25.81	1306.26	864.40	24369.26	38187.87	665897.4	41792.11	0.67	1.53	3723.47	10308.61
gray!6 PIA	494.60	838.62	4163.58	783.88	5.22	-26.51	1201.04	793.13	28122.80	44180.50	558897.3	35430.62	0.84	1.39	3697.91	9221.80
PIE	435.18	630.77	3318.63	620.29	5.43	-26.26	1278.11	847.36	19269.41	23336.65	472940.1	30191.09	1.02	1.37	3240.16	7546.22
gray!6 PLE	468.99	770.15	4044.84	683.62	5.42	-25.92	1306.42	870.85	25399.24	46364.32	868825.2	33619.27	1.12	1.23	5312.36	7901.84
PUE	520.26	865.10	4460.59	775.65	4.98	-26.10	1353.32	902.72	33564.79	45797.49	841781.7	37674.17	0.72	1.11	5397.23	8249.97
gray!6 QUA	458.63	706.61	3265.79	689.08	4.92	-26.74	1304.61	862.73	20764.43	25415.59	679962.7	27766.85	0.61	1.53	5845.11	8570.03
SAC	436.84	779.80	4317.55	699.71	5.17	-25.69	1333.98	882.42	10276.24	26602.00	665431.4	40020.50	0.56	0.57	10387.92	11888.99
gray!6 SAL	396.90	582.90	2915.65	592.48	5.33	-26.74	1253.29	802.65	20116.52	15944.42	339030.8	19588.04	0.95	1.01	5040.26	6804.81
SEG	453.87	819.44	4484.53	728.92	5.12	-26.06	1320.39	878.22	19458.77	47514.06	767773.7	26366.13	0.81	1.07	6328.79	10553.06
gray!6 SIE	445.13	860.65	4266.96	755.56	5.42	-26.11	1337.77	887.32	28151.96	49975.12	431452.8	29270.25	0.88	1.34	5703.93	7581.65
STJ	535.06	893.84	4559.44	794.76	5.43	-25.92	1296.49	838.08	23677.59	45913.57	632962.2	37600.18	0.81	1.60	6258.52	10697.50
gray!6 TAM	339.09	570.29	2250.45	555.13	5.76	-27.82	1261.34	813.01	12214.55	16014.62	365171.0	23907.35	0.79	0.89	3427.87	7982.86
VAL	431.30	672.13	3393.73	656.09	5.13	-26.56	1258.81	825.76	16849.74	43523.72	860272.1	23268.04	0.92	1.06	6513.53	10003.81
gray!6 VER	503.62	813.65	4401.61	780.73	5.46	-25.75	1325.48	864.43	22141.52	44992.37	680372.1	34422.39	0.73	1.57	4422.45	8800.61

Table X.2. Population-specific means and variances of the eight phenotypic traits (from left to right): height in Portugal (October 2012), height in Bordeaux (France, November 2013), height in Bordeaux (France, November 2018), height in Asturias (Spain, November 2012), mean bud burst date in Bordeaux (over the years 2013, 2014, 2015 and 2017), mean duration of bud burst in Bordeaux (over the years 2014, 2015 and 2017), specific leaf area in Portugal and $\delta^{13}C$ in Portugal.

Figure X.2. Population-specific height distributions in Portugal (October 2012). The color gradient corresponds to the extreme minimum temperature in the population location over the period 1901-1950.

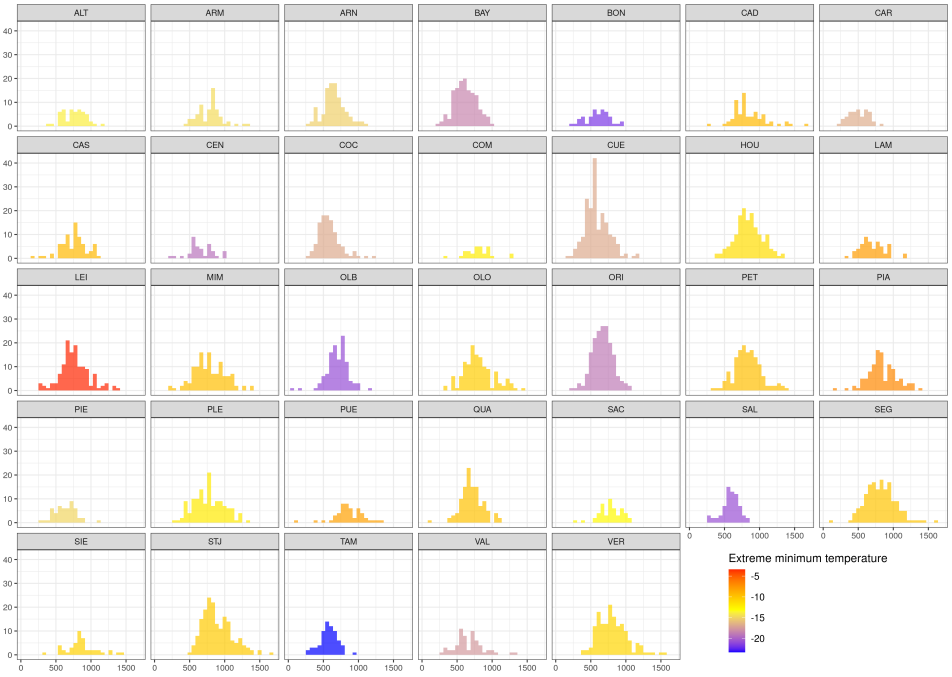


Figure X.3. Population-specific height distributions in Bordeaux (France, November 2013). The color gradient corresponds to the extreme minimum temperature in the population location over the period 1901-1950.

Figure X.4. Population-specific height distributions in Bordeaux (France, November 2018). The color gradient corresponds to the extreme minimum temperature in the population location over the period 1901-1950.

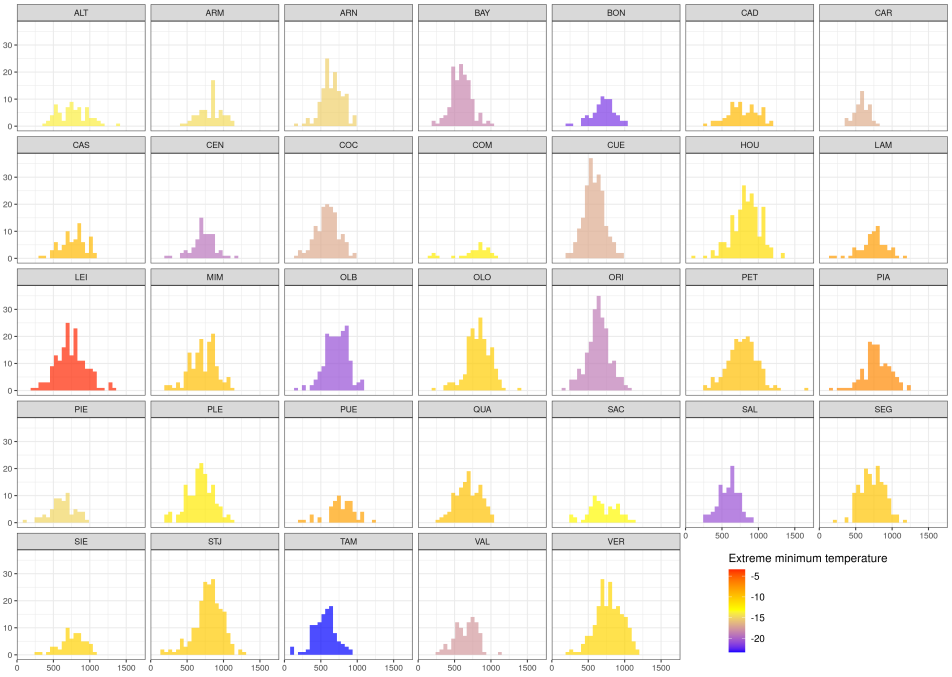


Figure X.5. Population-specific height distributions in Asturias (Spain, November 2012). The color gradient corresponds to the extreme minimum temperature in the population location over the period 1901-1950.

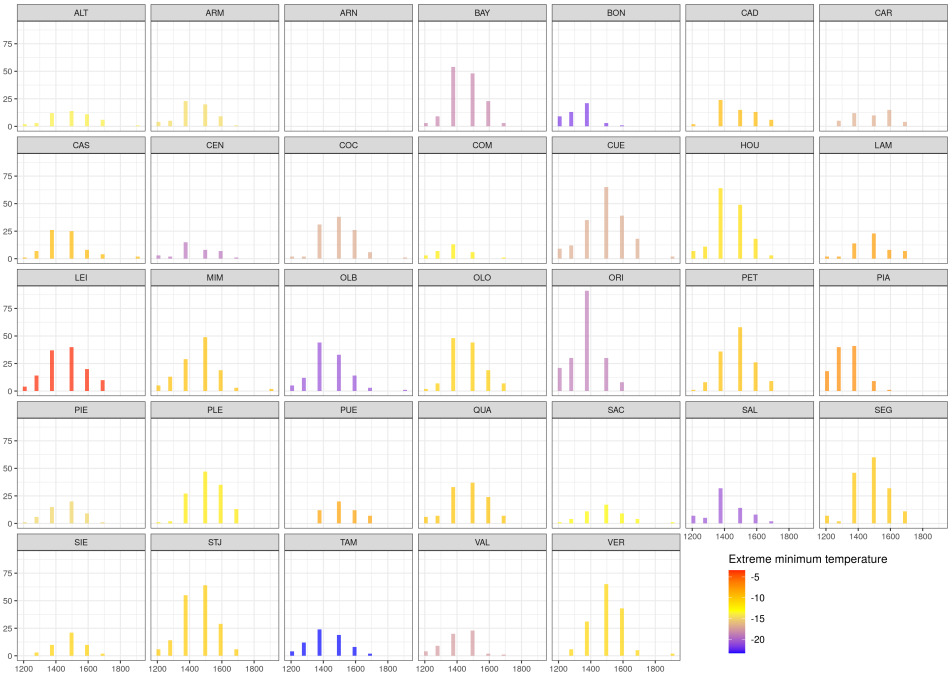


Figure X.6. Population-specific distributions of the mean bud burst date in Bordeaux (France) averaged over the years 2013, 2014, 2015 and 2017. The color gradient corresponds to the extreme minimum temperature in the population location over the period 1901-1950.

Figure X.7. Population-specific distributions of the mean duration of the bud burst date in Pieroton (France) averaged over the years 2014, 2015 and 2017. The color gradient corresponds to the extreme minimum temperature in the population location over the period 1901-1950.

Figure X.8. Population-specific distributions of the specific leaf area (SLA) in Portugal. The color gradient corresponds to the extreme minimum temperature in the population location over the period 1901-1950.

Figure X.9. Population-specific distributions of the isotope discrimination ($\delta^{13}C$) in Portugal. The color gradient corresponds to the extreme minimum temperature in the population location over the period 1901-1950.

1.2 Environment of the common gardens

Coordinates, soil and topographic variables	Units	Asturias	Bordeaux	Portugal
Latitude	degrees	43.42	44.75	40.11
Longitude	degrees	-6.54	-0.78	-7.48
Topographic ruggedness index	unitless	17.18	0.81	19.76
Depth available to roots	cm	30.00	70.00	70.00
Clay content in the topsoil (0-30 cm)	%	20.00	4.00	10.00
Sand content in the topsoil (0-30 cm)	%	49.00	87.00	75.00
Silt content in the topsoil (0-30 cm)	%	31.00	9.00	15.00

Table X.3. Values of the geographical coordinates, soil and topographic variables of the three common gardens.

Annual climatic variables	Units	Asturias	Bordeaux	Portugal
Mean Coldest Month Temperature (MCMT)	°C	5.80	5.60	4.70
Mean Warmest Month Temperature (MWMT)	°C	17.75	19.70	19.75
Temperature difference (MWMT-MCMT)	°C	11.95	14.10	15.05
Extreme minimum temperature (**EMT**)	°C	-12.50	-13.30	-14.30
Mean summer precipitation	mm	423.00	496.00	370.00
Mean spring precipitation	mm	242.00	226.00	271.00
Summer heat moisture index (**SHM**)	°C/mm	134.47	105.91	274.31

Table X.4. Values of the annual climatic variables in the three common gardens.

Monthly climatic variables	Units	Asturias	Bordeaux	Portugal
Minimum temperature - January	°C	2.50	2.00	1.50
Minimum temperature - February	°C	3.00	2.50	1.60
Minimum temperature - March	°C	4.20	3.80	3.10
Minimum temperature - April	°C	5.40	6.30	4.70
Minimum temperature - May	°C	7.70	9.20	7.30
Minimum temperature - June	°C	10.50	12.10	10.90
Minimum temperature - July	°C	12.30	13.90	13.10
Minimum temperature - August	°C	12.70	13.80	13.40
Minimum temperature - September	°C	11.30	11.80	11.30
Minimum temperature - October	°C	8.30	8.60	8.20
Minimum temperature - November	°C	5.40	4.80	4.20
Minimum temperature - December	°C	3.50	2.80	2.20
Maximum temperature - January	°C	9.10	9.20	7.90
Maximum temperature - February	°C	10.40	11.10	8.80
Maximum temperature - March	°C	12.30	13.90	11.10
Maximum temperature - April	°C	14.10	16.60	13.90
Maximum temperature - May	°C	16.70	20.00	17.20
Maximum temperature - June	°C	20.20	23.30	22.10
Maximum temperature - July	°C	22.50	25.50	25.70
Maximum temperature - August	°C	22.80	25.60	26.10
Maximum temperature - September	°C	20.90	23.30	22.00
Maximum temperature - October	°C	16.70	18.30	16.60
Maximum temperature - November	°C	12.20	12.80	11.20
Maximum temperature - December	°C	9.70	9.50	8.40
Total precipitation - January	mm	88.00	83.00	112.00
Total precipitation - February	mm	79.00	74.00	122.00
Total precipitation - March	mm	93.00	76.00	120.00
Total precipitation - April	mm	72.00	70.00	78.00
Total precipitation - May	mm	77.00	80.00	73.00
Total precipitation - June	mm	50.00	63.00	45.00
Total precipitation - July	mm	34.00	50.00	11.00
Total precipitation - August	mm	37.00	54.00	9.00
Total precipitation - September	mm	61.00	82.00	52.00
Total precipitation - October	mm	92.00	97.00	102.00
Total precipitation - November	mm	100.00	98.00	122.00
Total precipitation - December	mm	113.00	103.00	124.00
Hargreaves climatic moisture deficit - January	mm	17.68	17.88	18.17
Hargreaves climatic moisture deficit - February	mm	26.56	29.31	25.17
Hargreaves climatic moisture deficit - March	mm	46.40	57.06	44.75
Hargreaves climatic moisture deficit - April	mm	68.38	82.97	72.07
Hargreaves climatic moisture deficit - May	mm	92.24	115.64	101.41
Hargreaves climatic moisture deficit - June	mm	115.38	137.41	134.79
Hargreaves climatic moisture deficit - July	mm	130.36	151.65	166.13
Hargreaves climatic moisture deficit - August	mm	115.21	134.58	151.93
Hargreaves climatic moisture deficit - September	mm	79.08	92.08	92.73
Hargreaves climatic moisture deficit - October	mm	45.16	50.68	48.35
Hargreaves climatic moisture deficit - November	mm	21.58	23.63	23.31
Hargreaves climatic moisture deficit - December	mm	14.70	14.68	16.03

Table X.5. Values of the monthly climatic variables in the three common gardens.

1.3 Potential drivers of the within-population genetic variation

1.3.1 Population admixtures scores

Population	Longitude	Latitude	gpNA	gpC	gpCS	gpFA	gpIA	gpSES	mainGP	A	D	D_{fst}
CEN	-4.491	40.278	0.012	0.002	0.884	0.003	0.048	0.051	gpCS	0.116	0.031	0.011
ARN	-5.116	40.195	0.010	0.002	0.955	0.008	0.010	0.014	gpCS	0.045	0.014	0.005
ALT	-6.494	43.283	0.003	0.000	0.109	0.092	0.793	0.002	gpIA	0.207	0.061	0.019
SAL	-3.063	41.835	0.010	0.004	0.944	0.027	0.008	0.007	gpCS	0.056	0.016	0.006
COM	-3.954	36.834	0.240	0.028	0.126	0.011	0.039	0.557	gpSES	0.443	0.218	0.061
CAD	-6.418	43.540	0.002	0.001	0.050	0.010	0.935	0.002	gpIA	0.065	0.018	0.005
VAL	-4.311	40.516	0.012	0.003	0.940	0.007	0.013	0.025	gpCS	0.060	0.018	0.007
MIM	-1.303	44.134	0.004	0.002	0.025	0.951	0.012	0.006	gpFA	0.049	0.014	0.006
LEI	-8.957	39.783	0.004	0.003	0.511	0.007	0.473	0.002	gpCS	0.489	0.125	0.031
BAY	-2.877	41.523	0.003	0.004	0.967	0.013	0.009	0.004	gpCS	0.033	0.009	0.003
SIE	-6.493	43.528	0.003	0.001	0.076	0.015	0.903	0.001	gpIA	0.097	0.028	0.008
LAM	-6.219	43.559	0.002	0.001	0.004	0.051	0.942	0.000	gpIA	0.058	0.020	0.007
TAM	-5.017	33.600	0.932	0.000	0.027	0.000	0.039	0.001	gpNA	0.068	0.045	0.020
COC	-4.498	41.255	0.017	0.005	0.831	0.060	0.044	0.044	gpCS	0.169	0.044	0.015
OLO	-1.831	46.566	0.003	0.001	0.007	0.979	0.006	0.003	gpFA	0.021	0.007	0.003
STJ	-2.029	46.764	0.003	0.002	0.028	0.946	0.017	0.004	gpFA	0.054	0.015	0.006
CUE	-4.484	41.375	0.003	0.001	0.872	0.063	0.058	0.002	gpCS	0.128	0.030	0.009
PET	-1.300	44.064	0.003	0.001	0.021	0.966	0.004	0.004	gpFA	0.034	0.009	0.004
ORI	-2.351	37.531	0.248	0.005	0.028	0.001	0.010	0.708	gpSES	0.292	0.180	0.044
SEG	-8.450	42.817	0.003	0.001	0.149	0.013	0.831	0.003	gpIA	0.169	0.046	0.012
OLB	-0.623	40.173	0.080	0.009	0.777	0.006	0.003	0.125	gpCS	0.223	0.078	0.032
QUA	-0.359	38.972	0.092	0.006	0.499	0.005	0.015	0.382	gpCS	0.501	0.142	0.056
CAS	-6.983	43.501	0.001	0.000	0.005	0.001	0.992	0.001	gpIA	0.008	0.003	0.001
PLE	-2.344	47.781	0.005	0.001	0.064	0.920	0.005	0.005	gpFA	0.080	0.020	0.008
BON	-1.661	39.986	0.164	0.010	0.640	0.003	0.002	0.181	gpCS	0.360	0.139	0.057
HOU	-1.150	45.183	0.004	0.001	0.026	0.960	0.007	0.002	gpFA	0.040	0.011	0.004
VER	-1.091	45.552	0.003	0.002	0.018	0.972	0.004	0.001	gpFA	0.028	0.008	0.003
ARM	-6.458	43.305	0.005	0.007	0.022	0.006	0.959	0.001	gpIA	0.041	0.015	0.005
CAR	-4.277	41.172	0.001	0.001	0.904	0.060	0.023	0.011	gpCS	0.096	0.021	0.007
PIE	9.038	41.973	0.007	0.969	0.023	0.000	0.001	0.001	gpC	0.031	0.014	0.005
PUE	-6.631	43.548	0.002	0.000	0.021	0.001	0.973	0.002	gpIA	0.027	0.008	0.003
PIA	9.465	42.021	0.004	0.974	0.010	0.001	0.008	0.003	gpC	0.026	0.012	0.005
SAC	-8.364	42.118	0.004	0.001	0.291	0.003	0.699	0.002	gpIA	0.301	0.079	0.020

Table X.6. Population admixture scores for the 33 provenances. The columns gpNA, gpC, gpCS, gpFA, gpIA and gpSES contain the proportion of belonging to the gene pools of Northern Africa, Corsica, Central Spain, French Atlantic region, Iberian Atlantic region and south-eastern Spain, respectively. A and D are the two population admixture scores used in the study. D_{fst} is similar to D, as it was calculated by weighting the proportions of ancestry from foreign gene pools by the pairwise F_{ST} between the main and foreign gene pools, while D was calculated by weighting the proportions of ancestry from foreign gene pools by the sum of the allele frequency divergence of the main and foreign gene pool from the common ancestral one (obtained from Jaramillo-Correa et al. 2015). As D and D_{fst} were highly correlated, we did not keep D_{fst} in the following analyses.

1.3.2 Climate harshness and environmental heterogeneity indexes

Environmental variables	Units
Monthly minimum temperature (12 variables)	°C
Monthly maximum temperature (12 variables)	°C
Monthly total precipitation (12 variables)	mm
Monthly Hargreaves climatic moisture deficit (12 variables)	mm
Mean Warmest Month Temperature (MWMT)	°C
Mean Coldest Month Temperature (MCMT)	°C
Temperature difference (MWMT-MCMT)	°C
Mean summer precipitation	mm
Mean spring precipitation	mm
Extreme minimum temperature (**EMT**)	°C
Summer heat moisture index (**SHM**)	°C/mm
Topographic ruggedness index	unitless
Depth available to roots	cm
Clay content in the topsoil (0-30 cm)	%
Sand content in the topsoil (0-30 cm)	%
Silt content in the topsoil (0-30 cm)	%

Table X.7. Climatic, topographic and soil variables used to describe the environmental heterogeneity around the population location (all variables) and the climate harshness at the population location (SHM and EMT). The climatic variables were averaged over the period 1901-1950, except the extreme minimum temperature (EMT) which corresponds to the minimum temperature over the same period.

Forested areas: In Europe, the forested areas were extracted from Copernicus ("Forest Type 2015" at 100-m resolution), which describes the land cover in 2015. We kept raster cells attributed to either broadleaved forests (1), coniferous forests (2) or mixed forests (3). The other raster cells were considered as non-forested areas. In Morocco, the forested areas were extracted from the GeoNetwork: Land cover of Morocco - Globcover Regional, which describes the land cover in 2005. Raster cells with the following LCCCode were considered as forested areas:

- 0003 / 0004 (Mosaic cropland vegetation)
- 0004 // 0003 (Mosaic vegetation / cropland)
- 21446 // 21450-121340 / 21454 (Mosaic forest or shrubland/grassland)
- 21450 (Closed to open shrubland)
- 21454 // 21446 // 21450 (Mosaic grassland/forest or shrubland)
- 21496 // 21497-15048 (Closed to open broadleaved evergreen or semidecidous forest)
- 21496-121340 // 21497-129401 (Closed (>40%) broadleaved evergreen or semidecidous forest)
- 21497-121340 (Closed (>40%) broadleaved decidous forest)
- 21497-15045 (Closed to open (>15%) mixed broadleaved decidous and needleaved evergreen forest)
- 21499-121340 (Closed (>40%) needeleaved evergreen forest)
- 21518 (Closed to open broadleaved decidous shrubland)

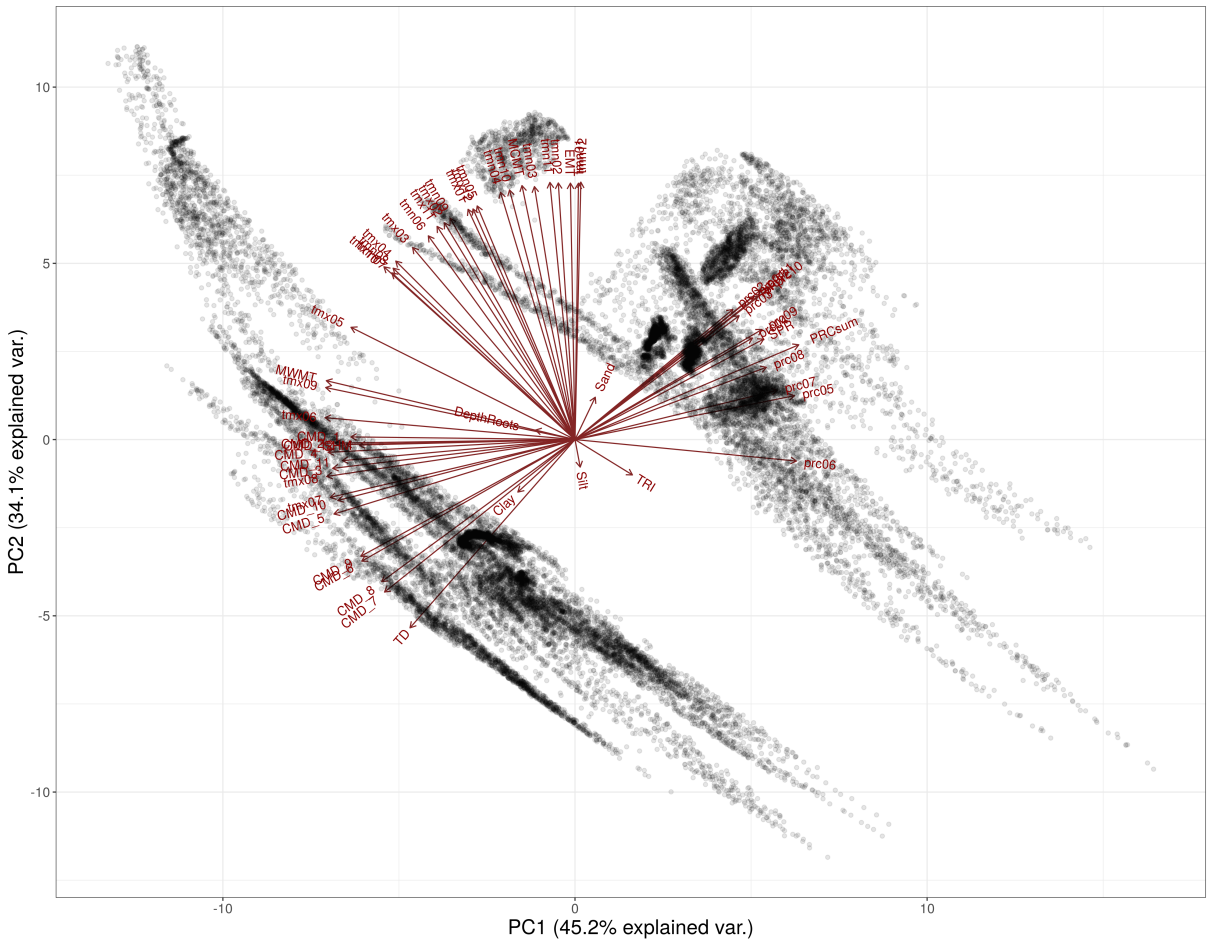


Figure X.10. Principal component analysis of all the climatic, topographic and soil variables in Table X.7.

1.3.3 Correlation among the potential drivers

	A	D	EH1[20km]	EH2[20km]	EH1[1.6km]	EH2[1.6km]	SHM	invEMT
A	-	-	-	-	-	-	-	-
D	**0.914**	-	-	-	-	-	-	-
EH1[20km]	-0.071	-0.013	-	-	-	-	-	-
EH2[20km]	0.197	0.349	**0.778**	-	-	-	-	-
EH1[1.6km]	0.375	0.400	0.561	0.698	-	-	-	-
EH2[1.6km]	0.420	0.565	0.333	**0.799**	0.686	-	-	-
SHM	0.381	0.536	0.335	0.541	0.275	0.519	-	-
invEMT	-0.009	0.126	0.197	0.082	0.218	0.044	0.406	-

Table X.8. Correlations among the potential underlying drivers of the within-population genetic variation. Correlations higher than 0.7 are in bold.

2 Model equation and priors

We modeled each trait y_{bpcr} such as:

$$y_{bpcr} \sim \mathcal{N}(\mu_{bpc}, \sigma_r^2)$$
$$\mu_{bpc} = \beta_0 + B_b + P_p + C_{c(p)}$$

where β_0 is the global intercept, B_b the block intercepts, P_p the population intercepts, $C_{c(p)}$ the clone intercepts and σ_r^2 the residual variance.

The prior of β_0 was weakly informative and centered around the mean of the observed values for the trait under considered, as follows:

$$\beta_0 \sim \mathcal{N}(\mu_y, 2)$$

The population and block intercepts, P_p and B_b were considered normally-distributed with variances σ_P^2 and σ_B^2, such as:

$$\begin{bmatrix} B_b \\ P_p \end{bmatrix} \sim \mathcal{N}\left(0, \begin{bmatrix} \sigma_B^2 \\ \sigma_P^2 \end{bmatrix}\right)$$

The clone intercepts $C_{c(p)}$ were considered to follow some population-specific normal distributions, such as:

$$C_{c(p)} \sim \mathcal{N}(0, \sigma_{C_p}^2)$$

where $\sigma_{C_p}^2$ are the population-specific variances among clones.
To partition the total variance, we parameterize our model so that only the total variance, σ_{tot}^2 has a prior, such that:

$$\begin{aligned}
\sigma_{tot}^2 &= \sigma_r^2 + \sigma_B^2 + \overline{\sigma_{C_p}^2} + \sigma_P^2 \\
\sigma_r &= \sigma_{tot} \times \sqrt{(\pi_r)} \\
\sigma_B &= \sigma_{tot} \times \sqrt{(\pi_B)} \\
\sigma_P &= \sigma_{tot} \times \sqrt{(\pi_P)} \\
\overline{\sigma_{C_p}} &= \sigma_{tot} \times \sqrt{(\pi_C)} \\
\sigma_{tot} &\sim \mathcal{S}^*(0, 1, 3)
\end{aligned}$$

where $\overline{\sigma_{C_p}}$ and $\overline{\sigma_{C_p}^2}$ are the mean of the population-specific among-clones standard deviations (σ_{C_p}) and variances ($\sigma_{C_p}^2$), respectively, and $\sum_l^4 \pi_l = 1$ (using the `simplex` function in `Stan`). The population-specific among-clones standard deviations σ_{C_p} follow a log-normal distribution with mean $\overline{\sigma_{C_p}}$ and variance σ_K^2, such as:

$$\begin{aligned}
\sigma_{C_p} &\sim \mathcal{LN}\left(\ln(\overline{\sigma_{C_p}}) - \frac{\sigma_K^2}{2} + \beta_X X_p, \sigma_K^2\right) \\
\sigma_K &\sim \exp(1)
\end{aligned}$$

with X_p the potential driver considered and β_x its associated coefficient.

Here we provide further explanation regarding the use of the log-normal distribution for σ_{C_p} in the model:

$$\sigma_{C_p} \sim \mathcal{LN}(\mu, \sigma_K^2) \Leftrightarrow \ln(\sigma_{C_p}) \sim \mathcal{N}(\mu, \sigma_K^2) \text{ with } \mu \text{ the median of } \sigma_{C_p}$$

$$\text{By definition: } \mathbb{E}(\sigma_{C_p}) = \exp\left(\mu + \frac{\sigma_K^2}{2}\right)$$

$$\text{We want: } \mathbb{E}(\sigma_{C_p}) = \exp\left(\mu + \frac{\sigma_K^2}{2}\right) = \sigma_{tot} \times \sqrt{(\pi_4)}$$

$$\text{Therefore: } \mu = \ln\left(\sigma_{tot} \times \sqrt{(\pi_4)}\right) - \frac{\sigma_K^2}{2}$$

3 Model accuracy on simulated data

We simulated data based on the real experimental design of two traits (height in Portugal at 20-month old and height in Bordeaux at 25-month old), which means that there were the same number of blocks, populations, clones per population and trees per clone as in the real experimental design. We ran 100 simulations, which are summarized in the tables below:

Parameter	True value	Mean standard error	Mean bias of the mean	Mean bias of the median	80% conf. int. coverage	95% conf. int. coverage
β_X	0.1	0.055	-0.002	-0.003	83	97
σ_K	0.1	0.067	0.020	0.014	92	99

Table X.9. Summary of the 100 model outputs ran on simulated data based on height in Bordeaux at 25-month old.

Parameter	True value	Mean standard error	Mean bias of the mean	Mean bias of the median	80% conf. int. coverage	95% conf. int. coverage
β_X	0.1	0.052	-0.005	-0.005	78	96
σ_K	0.1	0.065	0.015	0.009	95	98

Table X.10. Summary of the 100 model outputs ran on simulated data based on height in Portugal at 20-month old.

4 β_X interpretation

We have:

$$\sigma_{C_p} \sim \mathcal{LN}\left(\ln(\overline{\sigma_{C_p}}) - \frac{\sigma_K^2}{2} + \beta_X \tilde{X}_p, \sigma_K^2\right)$$

with $\tilde{X}_p = (X_p - \mu_{X_p})/\sigma_{X_p}$ (the explanatory variables were scaled before the analyses; μ_{X_p} is the mean of X_p and σ_{X_p} is its standard deviation).

By definition:

$$\ln(\sigma_{C_p}) \sim \mathcal{N}\left(\ln(\overline{\sigma_{C_p}}) - \frac{\sigma_K^2}{2} + \beta_X \tilde{X}_p, \sigma_K^2\right)$$

We want to calculate the percent of change in σ_{C_p} associated with a one-unit increase in $\tilde{X}_p$, that is a one-standard deviation increase in X_p. For that, we call σ_{new} the value of σ_{C_p} after increasing $\tilde{X}_p$ by one unit, and we have:

$$\begin{aligned}\ln(\sigma_{new}) &= \ln(\overline{\sigma_{C_p}}) - \frac{\sigma_K^2}{2} + \beta_X(\tilde{X}_p + 1)\\ &= \ln(\sigma_{C_p}) + \beta_X\end{aligned}$$

Therefore:

$$\begin{aligned}\ln(\sigma_{new}) - \ln(\sigma_{C_p}) &= \beta_X\\ \frac{\sigma_{new}}{\sigma_{C_p}} &= \exp(\beta_X)\\ 100 \times \left(\frac{\sigma_{new}}{\sigma_{C_p}} - 1\right) &= 100 \times (\exp(\beta_X) - 1)\\ 100 \times \left(\frac{\sigma_{new} - \sigma_{C_p}}{\sigma_{C_p}}\right) &= 100 \times (\exp(\beta_X) - 1)\end{aligned}$$

is the percent change in σ_{C_p} associated with a one-unit increase in $\tilde{X}_p$ (that is, a one-standard deviation increase in X_p). For instance, a one-standard deviation increase in the inverse of the extreme minimum temperature is associated, on average, with $100 \times (\exp(-0.395) - 1) = -32.6\%$ change in σ_{C_p} for height in Portugal, with $100 \times (\exp(-0.243) - 1) = -21.6\%$ change in σ_{C_p} for height in Bordeaux at 25-month old and with $100 \times (\exp(-0.197) - 1) = -17.9\%$ change in σ_{C_p} for height in Asturias. Similarly, a one-standard deviation increase in the summer heat moisture index is associated, on average, with $100 \times (\exp(-0.17) - 1) = -15.6\%$ change in σ_{C_p} for height in Bordeaux at 25-month old and with $100 \times (\exp(-0.272) - 1) = -23.8\%$ change in σ_{C_p} for height in Asturias.

5 Model outputs

5.1 $\mathcal{R}^2$ estimates

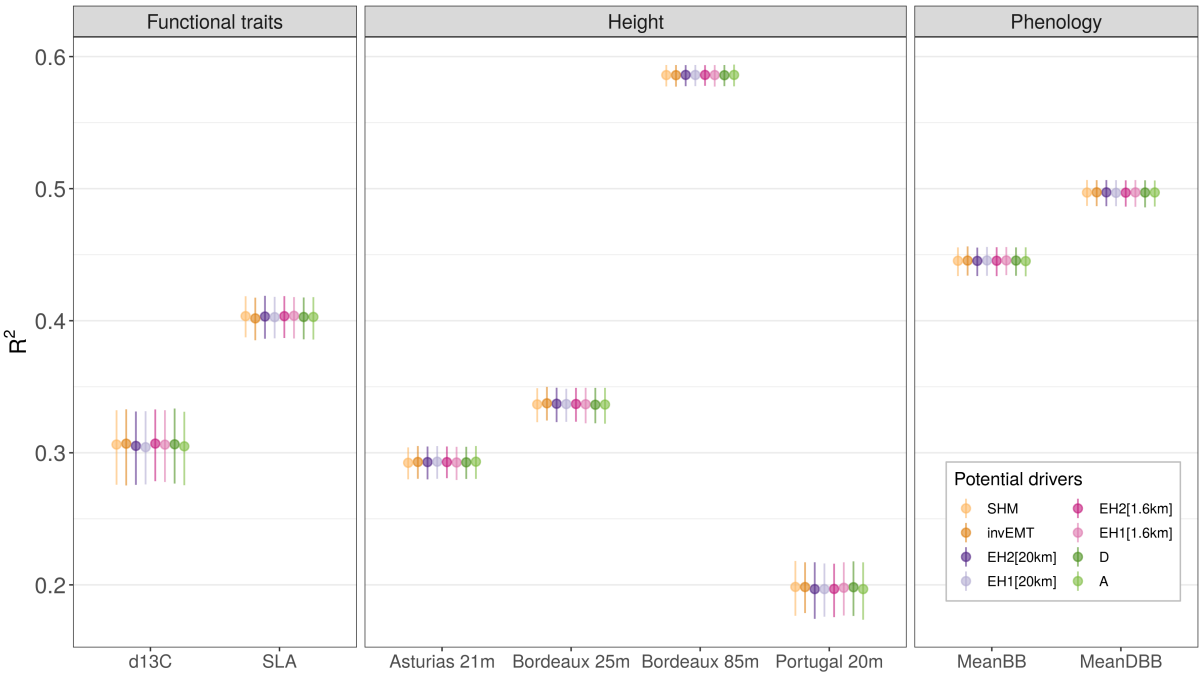


Figure X.11. Median and 95% intervals of the posterior distributions of the total variance explained by the models ($\mathcal{R}^2$).

5.2 β_X estimates for the eight potential drivers

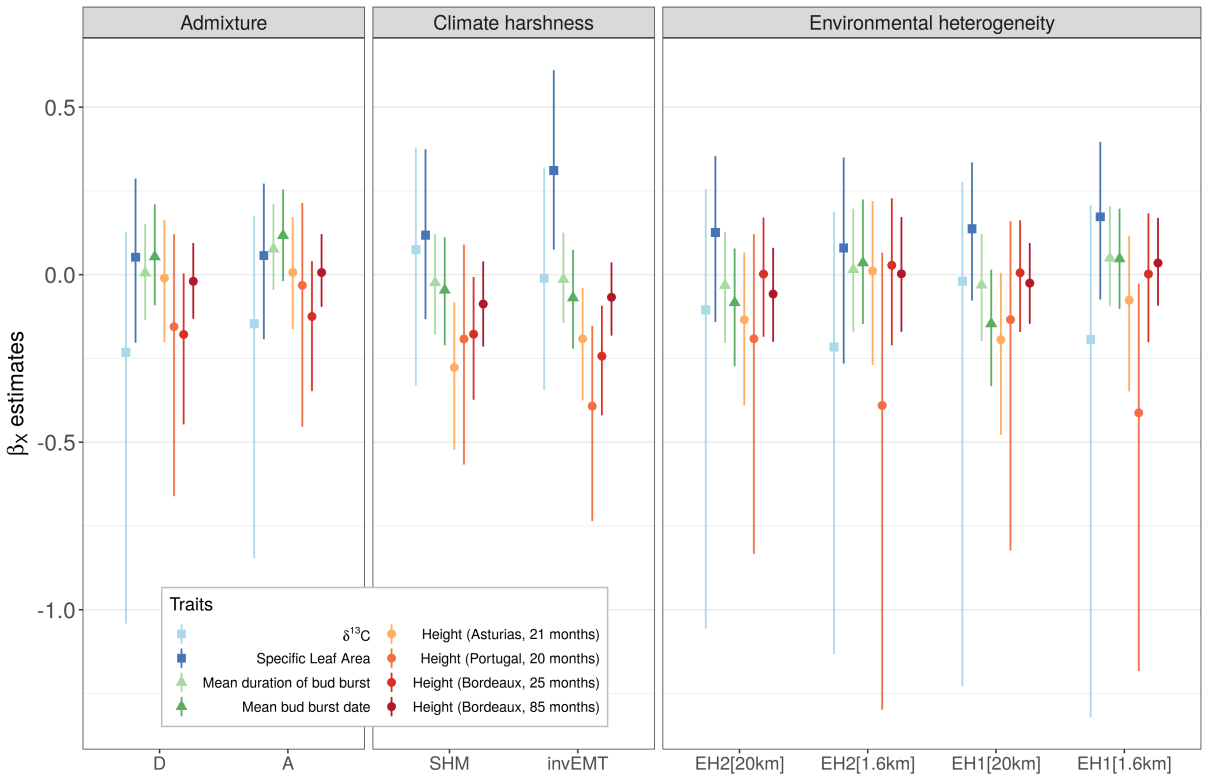


Figure X.12. Median and 95% intervals of the posterior distributions of β_X, the coefficient corresponding to the potential drivers of the within-population genetic variation.

5.3 σ_{C_p} estimates and variance partitioning

5.3.1 Height (Portugal, 20 months)

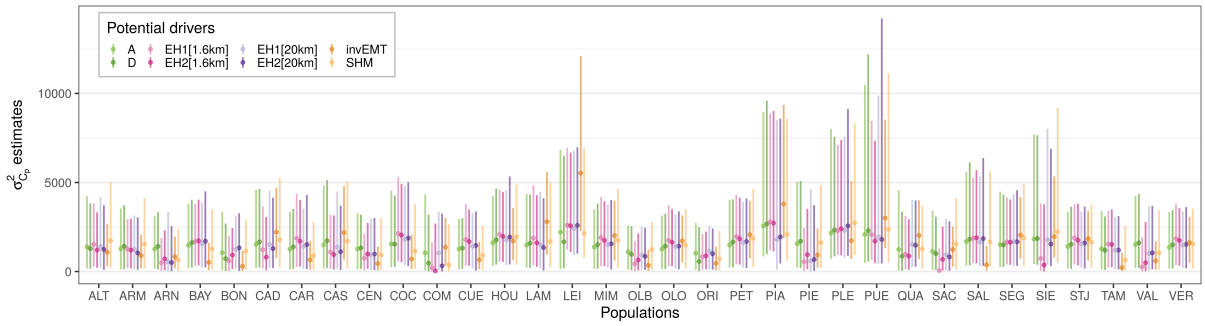


Figure X.13. Median and 95% intervals of the posterior distributions of $\sigma^2_{C_p}$, corresponding to the within-population total genetic variance (i.e. population-specific among clones variance).

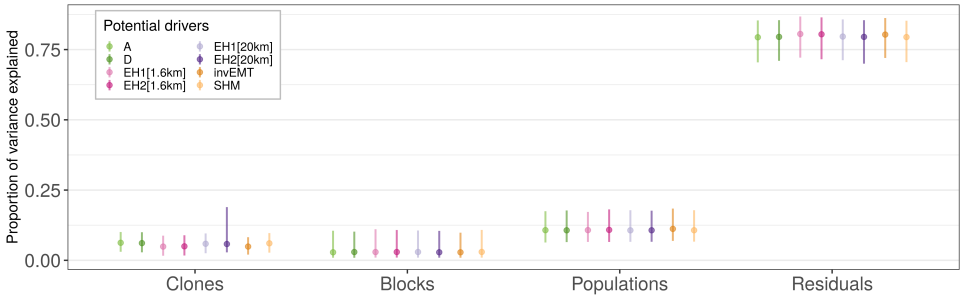


Figure X.14. Proportion of variance explained by the different components, namely the clones (π_C), the blocks (π_B), the populations (π_P) and the residuals (π_r).

5.3.2 Height (Bordeaux, 25 months)

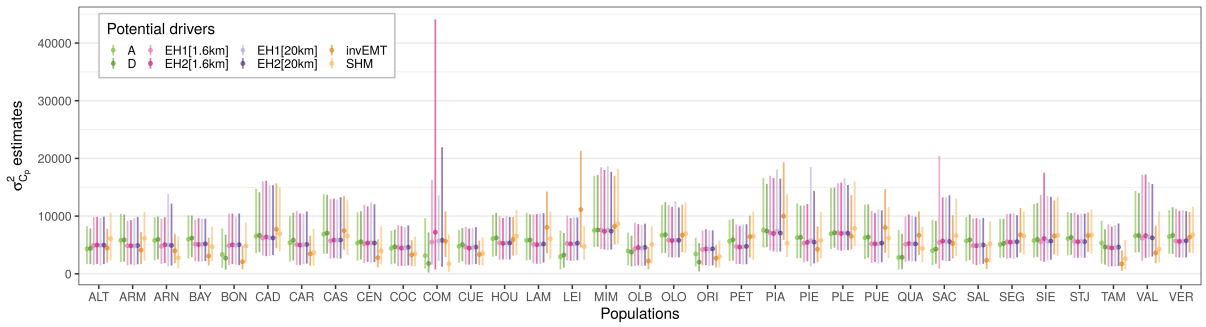


Figure X.15. Median and 95% intervals of the posterior distributions of $\sigma^2_{C_p}$, corresponding to the within-population total genetic variance (i.e. population-specific among clones variance).

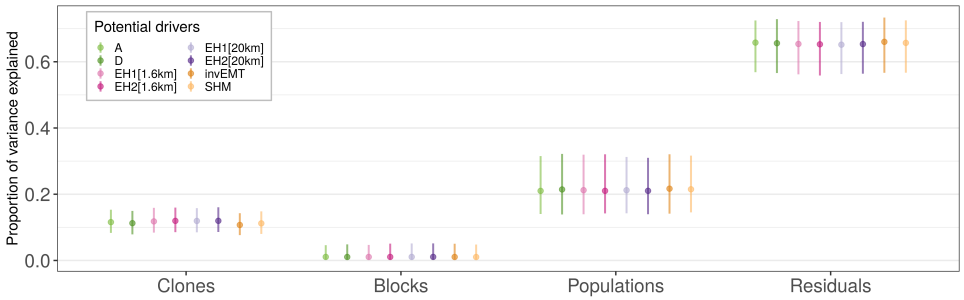


Figure X.16. Proportion of variance explained by the different components, namely the clones (π_C), the blocks (π_B), the populations (π_P) and the residuals (π_r).

5.3.3 Height (Bordeaux, 85 months)

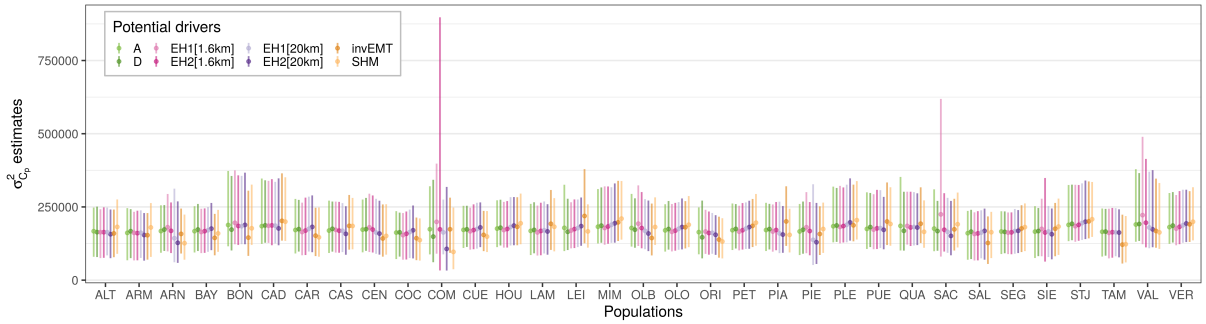


Figure X.17. Median and 95% intervals of the posterior distributions of $\sigma^2_{C_p}$, corresponding to the within-population total genetic variance (i.e. population-specific among clones variance).

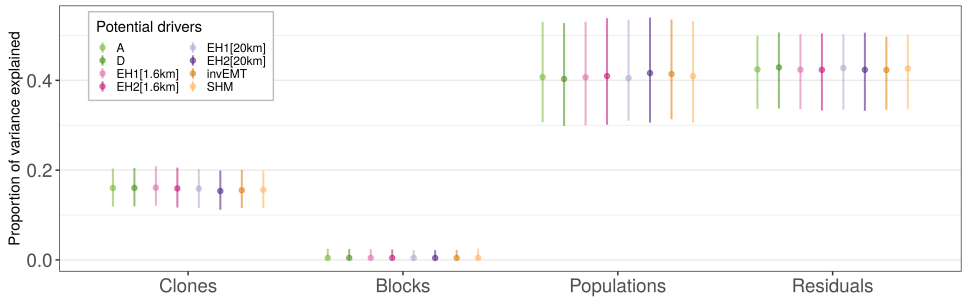


Figure X.18. Proportion of variance explained by the different components, namely the clones (π_C), the blocks (π_B), the populations (π_P) and the residuals (π_r).

5.3.4 Height (Asturias, 21 months)

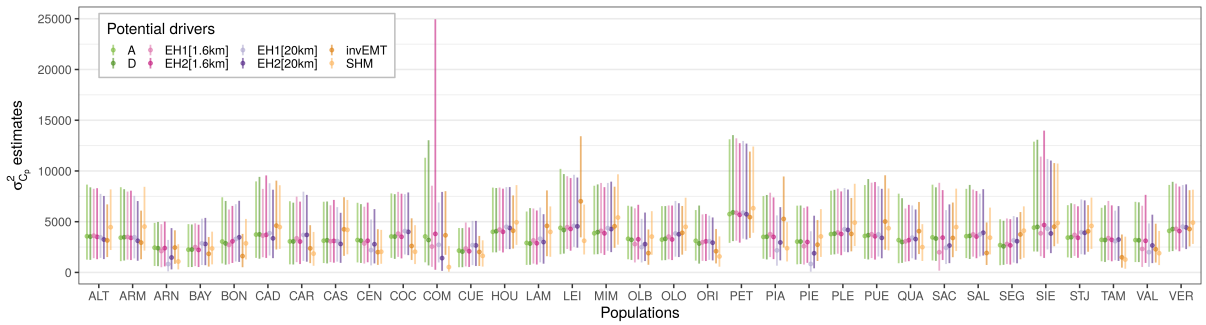


Figure X.19. Median and 95% intervals of the posterior distributions of $\sigma^2_{C_p}$, corresponding to the within-population total genetic variance (i.e. population-specific among clones variance).

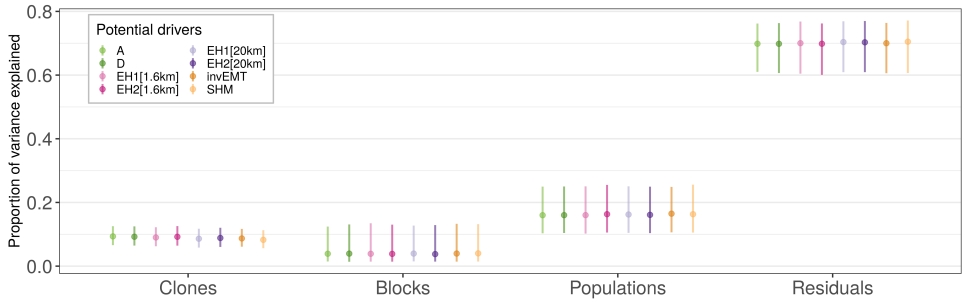


Figure X.20. Proportion of variance explained by the different components, namely the clones (π_C), the blocks (π_B), the populations (π_P) and the residuals (π_r).

5.3.5 Mean bud burst date (Bordeaux)

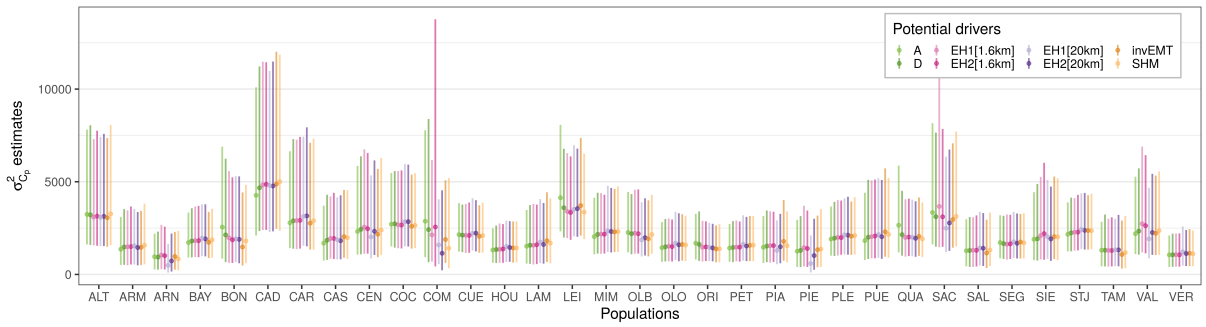


Figure X.21. Median and 95% intervals of the posterior distributions of $\sigma^2_{C_p}$, corresponding to the within-population total genetic variance (i.e. population-specific among clones variance).

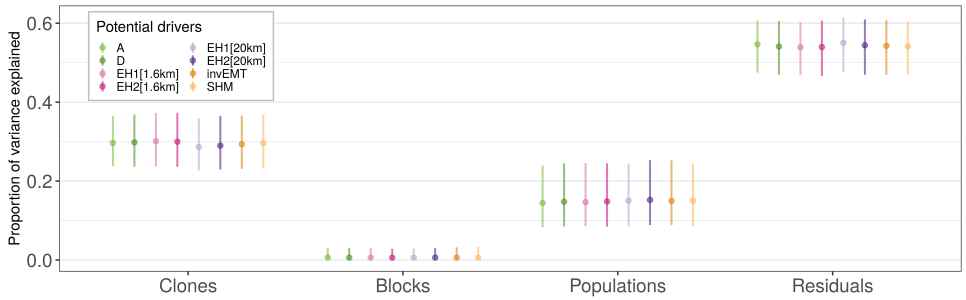


Figure X.22. Proportion of variance explained by the different components, namely the clones (π_C), the blocks (π_B), the populations (π_P) and the residuals (π_r).

5.3.6 Mean duration of bud burst (Bordeaux)

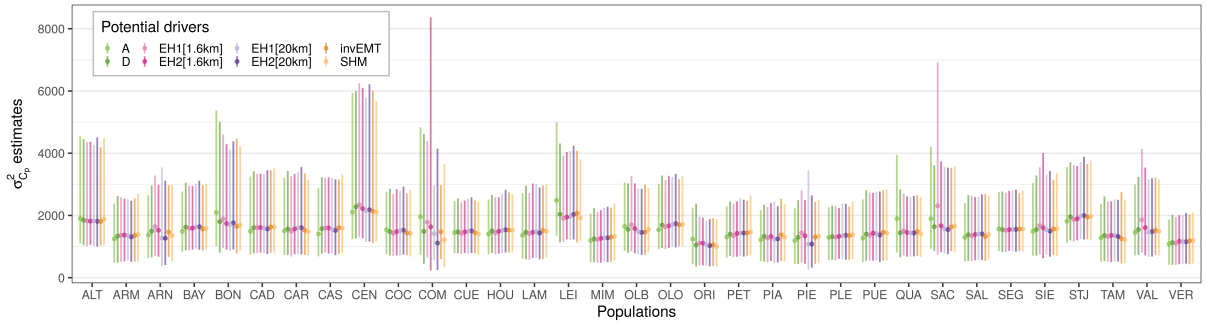


Figure X.23. Median and 95% intervals of the posterior distributions of $\sigma^2_{C_p}$, corresponding to the within-population total genetic variance (i.e. population-specific among clones variance).

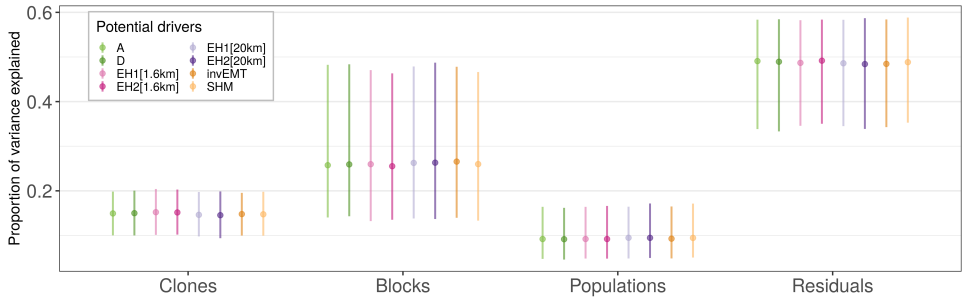


Figure X.24. Proportion of variance explained by the different components, namely the clones (π_C), the blocks (π_B), the populations (π_P) and the residuals (π_r).

5.3.7 Specific Leaf Area (Portugal)

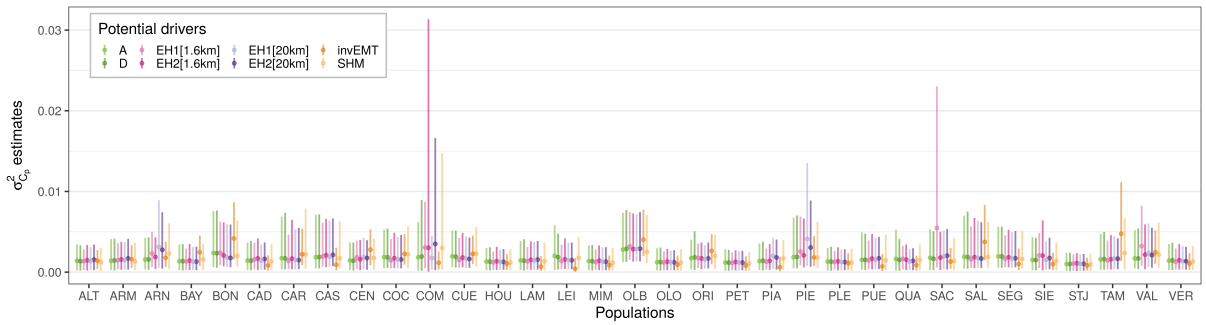


Figure X.25. Median and 95% intervals of the posterior distributions of $\sigma^2_{C_p}$, corresponding to the within-population total genetic variance (i.e. population-specific among clones variance).

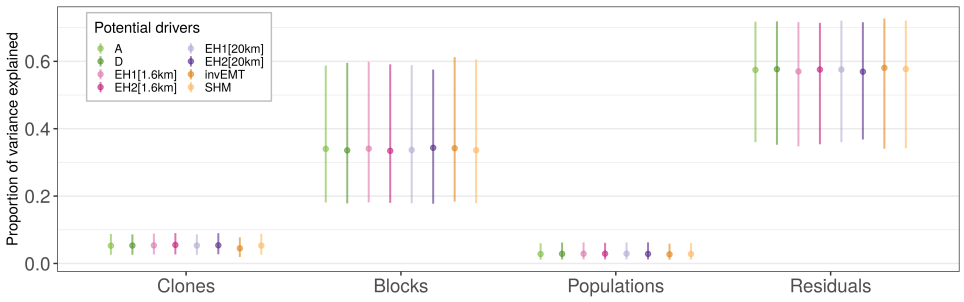


Figure X.26. Proportion of variance explained by the different components, namely the clones (π_C), the blocks (π_B), the populations (π_P) and the residuals (π_r).

5.3.8 δ^{13}C (Portugal)

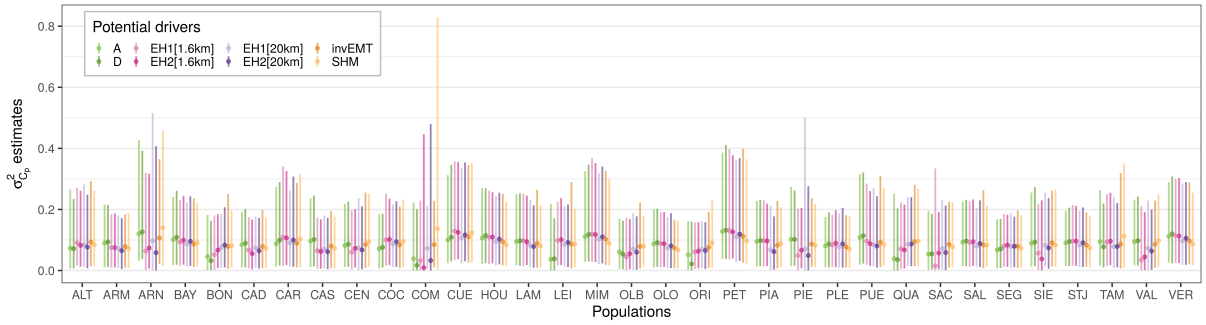


Figure X.27. Median and 95% intervals of the posterior distributions of $\sigma^2_{C_p}$, corresponding to the within-population total genetic variance (i.e. population-specific among clones variance).

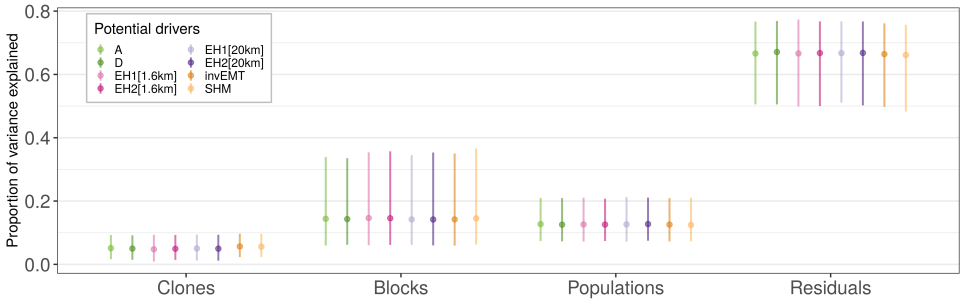


Figure X.28. Proportion of variance explained by the different components, namely the clones (π_C), the blocks (π_B), the populations (π_P) and the residuals (π_r).

5.4 Correlation between the number of clones per population and σ_{C_p}

Drivers	Ht (Portugal)	Ht (Bordeaux 2013)	Ht (Bordeaux 2018)	Ht (Asturias)	meanBB	meanDBB	SLA	$\delta^{13}C$
A	0.035	0.047	0.119	0.105	-0.217	-0.161	-0.119	0.152
D	0.038	0.068	0.149	0.122	-0.238	-0.189	-0.107	0.147
EH1[1.6km]	0.423	0.024	-0.283	0.317	-0.276	-0.381	-0.364	0.519
EH2[1.6km]	0.480	-0.166	0.087	0.066	-0.285	-0.262	-0.399	0.564
EH1[20km]	0.272	-0.042	0.350	0.321	-0.080	-0.093	-0.245	0.440
EH2[20km]	0.404	-0.035	0.497	0.433	-0.110	-0.006	-0.424	0.567
invEMT	0.109	0.185	0.230	0.126	-0.153	-0.131	-0.086	0.198
SHM	0.080	0.156	0.203	0.162	-0.186	-0.115	-0.208	-0.042

Table X.11. Pearson correlation coefficients between the estimates of the within-population genetic variation (i.e. σ_{C_p}) and the number of clones per population, for each combination of trait and potential driver of the within-population genetic variation. The eight phenotypic traits are (from left to right): height in Portugal (October 2012), height in Bordeaux (France, November 2013), height in Bordeaux (France, November 2018), height in Asturias (Spain, November 2012), mean bud burst date in Bordeaux (over the years 2013, 2014, 2015 and 2017), mean duration of bud burst in Bordeaux (over the years 2014, 2015 and 2017), specific leaf area in Portugal and the isotope discrimination of $\delta^{13}C$ in Portugal.

6 Climatic transfer distances

We estimated the association between climatic transfer distances for both EMT and SHM and the within-population genetic variation. The climatic transfer distances were calculated as the absolute difference between the climate in the location of origin of each population and the climate in the test site, for instance the climatic transfer distance for EMT between the population p and the common garden s was: equal to $abs(CTD_{EMT,s,p})$ with $CTD_{EMT,s,p} = EMT_p - EMT_s$.

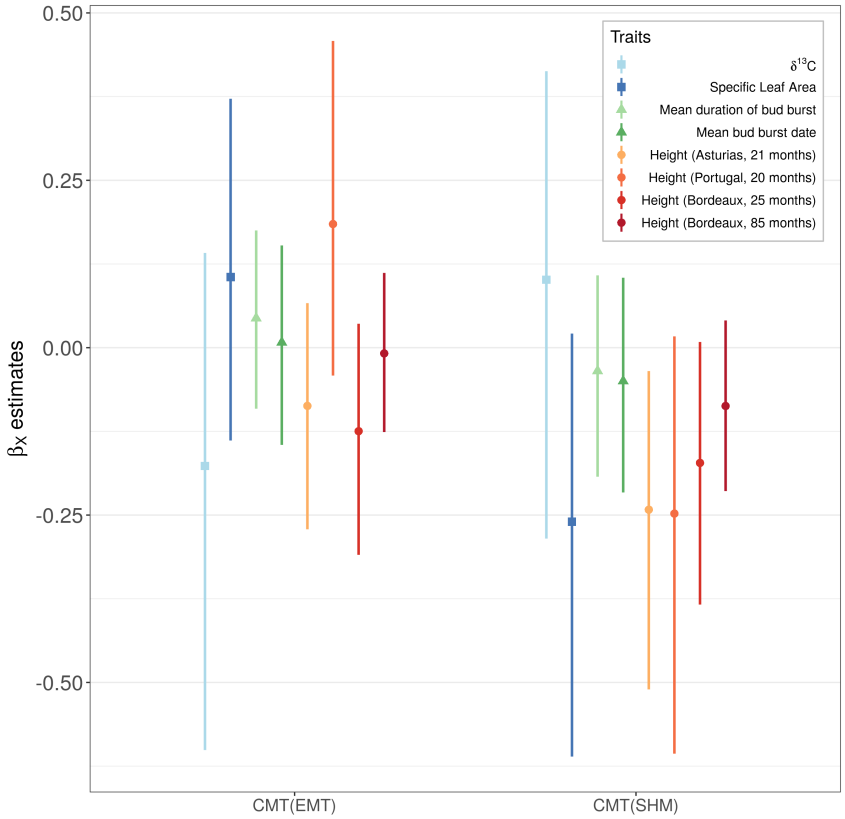


Figure X.29. Median and 95% intervals of the posterior distributions of β_X, the coefficient corresponding to the potential drivers (here the climatic transfer distances) of the within-population genetic variation.

We did not detect an association between climatic transfer distance and within-population genetic variation for SLA, as hypothesized in the discussion.

7 Genetic diversity

We estimated the association between genetic diversity (i.e. expected heterozygosity H_e) and within-population genetic variation. The expected heterozygosity H_e was extracted from Rodríguez-Quilón et al. (2016), in which H_e was calculated either on 12 nuclear microsatellites or on 266 SNP markers from the same populations as in our study (see Appendix S1 of Rodríguez-Quilón et al. 2016).

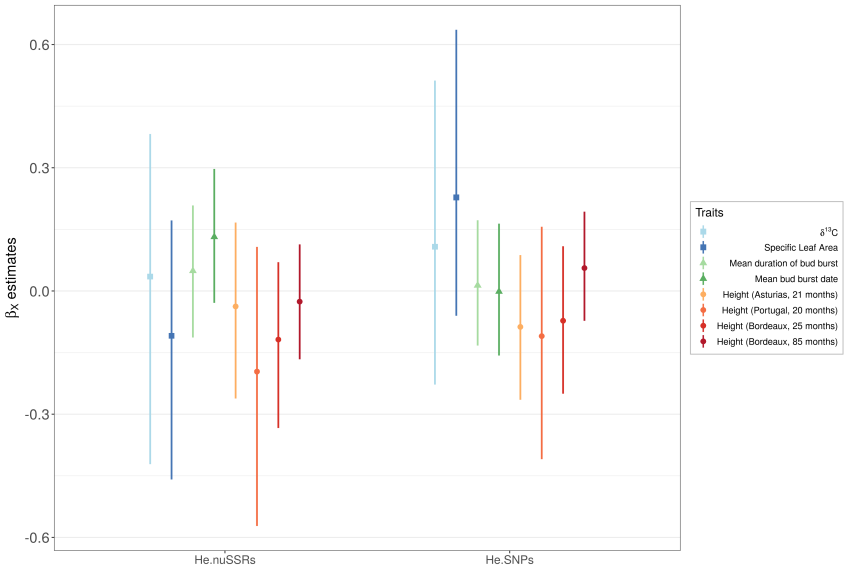


Figure X.30. Median and 95% intervals of the posterior distributions of β_X, the coefficient corresponding to the relationship between within-population quantitative genetic variation and genetic diversity, measured as the expected heretozygosity H_e estimated from either 12 nuclear microsatellites (He.nuSSRs) or 266 SNP markers (He.SNPs).

We did not detect an association between genetic diversity and within-population genetic variation for any trait, suggesting no influence of demographic processes affecting effective population size on within-population genetic variation.

8 Validation step

In the validation step, we performed exactly the same analyses (i.e. using the same model formula and code) as for the CLONAPIN height data, but on independent height data kindly provided by Ricardo Alia. This independent height data comes from a progeny test near Asturias, planted in 2005, and in which 23 provenances are shared with the CLONAPIN data (see Tables X.12 and X.13).

Populations	Mean	Variance	Number of families	Number of individuals
ALT	116.42373	1,697.8691	7	59
ARM	127.13793	2,275.1203	8	87
ARN	115.44615	1,723.3808	12	130
BAY	99.32948	1,503.3269	16	173
CAD	132.54000	2,112.1499	10	100
CAR	92.02128	1,415.1952	4	47
CAS	110.46970	2,974.8683	6	66
CEN	108.35000	1,003.4641	4	40
COC	109.85714	1,847.6504	8	77
CUE	98.58389	908.6635	14	149
LAM	118.09459	2,725.6211	7	74
LEI	114.00862	3,988.9825	12	116
MIM	115.20000	1,388.7236	9	90
ORI	91.91156	1,777.4784	15	147
PIA	123.42553	1,373.9454	4	47
PIE	106.27778	828.5654	2	18
PLE	122.73077	2,050.5288	9	104
PUE	119.90000	1,735.9694	5	50
SAL	97.21250	1,034.1695	8	80
SEG	117.18354	1,917.2846	16	158
SIE	116.38596	2,180.2412	5	57
TAM	86.90728	1,139.1780	16	151
VAL	106.48352	1,632.6747	8	91

Table X.12. Mean, variance, number of families and number of individuals in each population for the height measurements in the progeny test near Asturias when the trees were 3-year old.

Populations	Mean	Variance	Number of families	Number of individuals
ALT	343.4138	14,256.949	7	58
ARM	346.8605	15,262.263	8	86
ARN	315.4656	8,490.866	12	131
BAY	275.3810	9,417.399	16	168
CAD	372.6667	10,699.909	10	96
CAR	255.5778	8,827.659	4	45
CAS	312.1746	16,498.792	6	63
CEN	297.4250	9,311.892	4	40
COC	308.0789	9,849.060	8	76
CUE	277.6667	8,059.922	14	147
LAM	318.2676	16,332.885	7	71
LEI	331.9057	18,746.258	12	106
MIM	327.0444	8,190.200	9	90
ORI	262.3333	9,773.510	15	144
PIA	349.7234	6,867.813	4	47
PIE	295.5000	5,594.618	2	18
PLE	356.3010	11,389.154	9	103
PUE	336.8269	11,972.773	5	52
SAL	280.4872	6,440.773	8	78
SEG	338.6382	11,798.934	16	152
SIE	330.2500	12,265.718	5	56
TAM	272.7919	6,791.801	16	149
VAL	293.1000	10,550.788	8	90

Table X.13. Mean, variance, number of families and number of individuals in each population for the height measurements in the progeny test near Asturias when the trees were 6-year old.

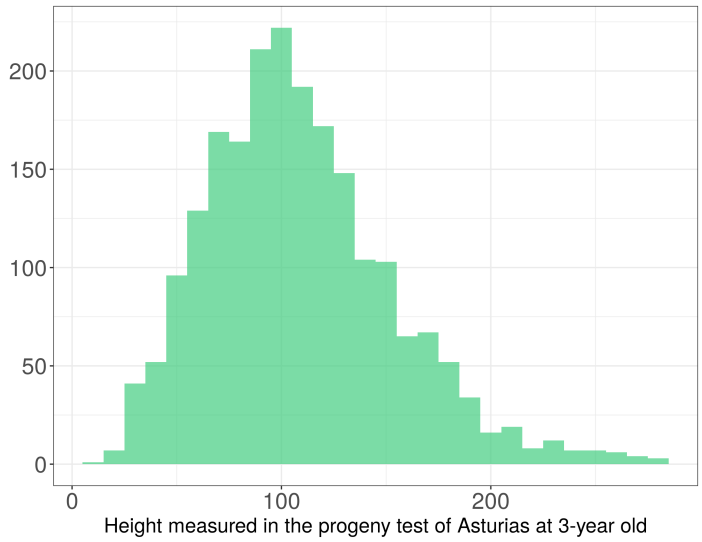


Figure X.31. Distribution of height measurements at 3-year old in the progeny test near Asturias (independent dataset used for the validation analysis).

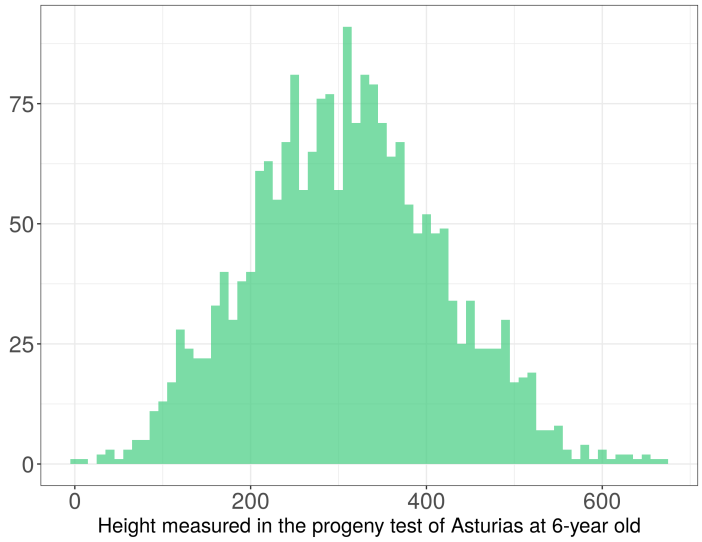


Figure X.32. Distribution of height measurements at 6-year old in the progeny test near Asturias (independent dataset used for the validation analysis).

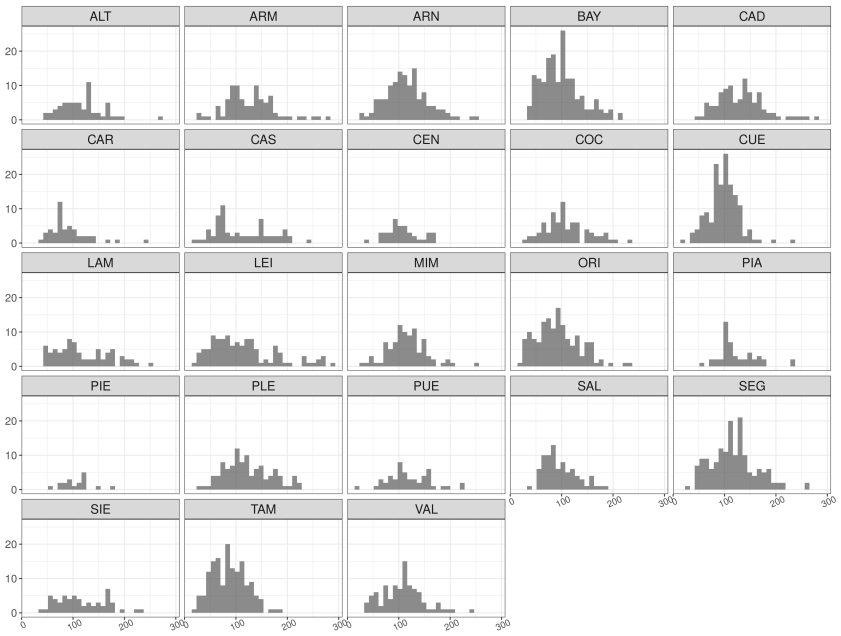


Figure X.33. Height distribution at 3-year old for the 23 populations shared between the CLONAPIN dataset and the independent dataset used in the validation analysis, i.e. a progeny test near Asturias.

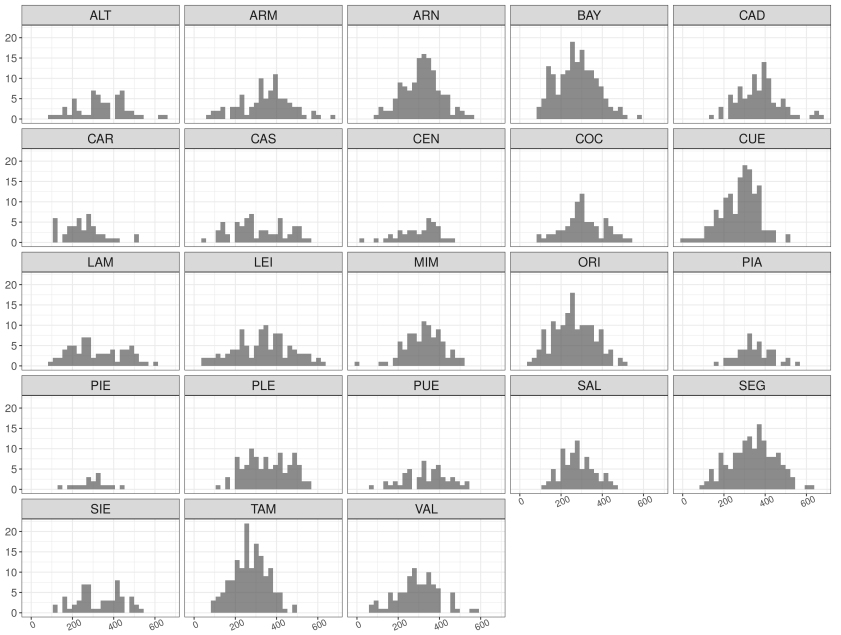


Figure X.34. Height distribution at 6-year old for the 23 populations shared between the CLONAPIN dataset and the independent dataset used in the validation analysis, i.e. a progeny test near Asturias.

8.2 Model equation and priors

We used the same mathematical model as the one used on CLONAPIN data (see section 2 in the Supplementary Information) but replacing clones by families.

We modeled each trait y_{bpfr} such as:

$$\begin{aligned} y_{bpfr} &\sim \mathcal{N}(\mu_{bpcf}, \sigma_r^2) \\ \mu_{bpf} &= \beta_0 + B_b + P_p + F_{f(p)} \end{aligned} \tag{8.1}$$

where β_0 is the global intercept, B_b the block intercepts, P_p the population intercepts, $F_{f(p)}$ the family intercepts and σ_r^2 the residual variance.
The prior of β_0 was weakly informative and centered around the mean of the observed values for the trait under considered, as follows:

$$\beta_0 \sim \mathcal{N}(\mu_y, 2)$$

The population and block intercepts, P_p and B_b were considered normally-distributed with variances σ_P^2 and σ_B^2, such as:

$$\begin{bmatrix} B_b \\ P_p \end{bmatrix} \sim \mathcal{N}\left(0, \begin{bmatrix} \sigma_B^2 \\ \sigma_P^2 \end{bmatrix}\right)$$

The family intercepts $F_{f(p)}$ were considered to follow some population-specific normal distributions, such as:

$$F_{f(p)} \sim \mathcal{N}(0, \sigma_{F_p}^2)$$

where $\sigma_{F_p}^2$ are the population-specific variances among families.
To partition the total variance, we parameterize our model so that only the total variance, σ_{tot}^2 has a prior, such that:

$$\begin{aligned} \sigma_{tot}^2 &= \sigma_r^2 + \sigma_B^2 + \overline{\sigma_{F_p}^2} + \sigma_P^2 \\ \sigma_r &= \sigma_{tot} \times \sqrt{(\pi_r)} \\ \sigma_B &= \sigma_{tot} \times \sqrt{(\pi_B)} \\ \sigma_P &= \sigma_{tot} \times \sqrt{(\pi_P)} \\ \overline{\sigma_{F_p}} &= \sigma_{tot} \times \sqrt{(\pi_F)} \\ \sigma_{tot} &\sim \mathcal{S}^*(0, 1, 3) \end{aligned} \tag{8.2}$$

where $\overline{\sigma_{F_p}}$ and $\overline{\sigma_{F_p}^2}$ are the mean of the population-specific among-families standard deviations (σ_{F_p}) and variances ($\sigma_{F_p}^2$), respectively, and $\sum_l^4 \pi_l = 1$ (using the `simplex` function in `Stan`).
The population-specific among-families standard deviations σ_{F_p} follow a log-normal distribution with mean $\overline{\sigma_{F_p}}$ and variance σ_K^2, such as:

$$\begin{aligned} \sigma_{F_p} &\sim \mathcal{LN}\left(\ln(\overline{\sigma_{F_p}}) - \frac{\sigma_K^2}{2} + \beta_X X_p, \sigma_K^2\right) \\ \sigma_K &\sim \exp(1) \end{aligned} \tag{8.3}$$

with X_p the potential driver considered and β_x its associated coefficient.

8.3 β_X estimates for the eight potential drivers

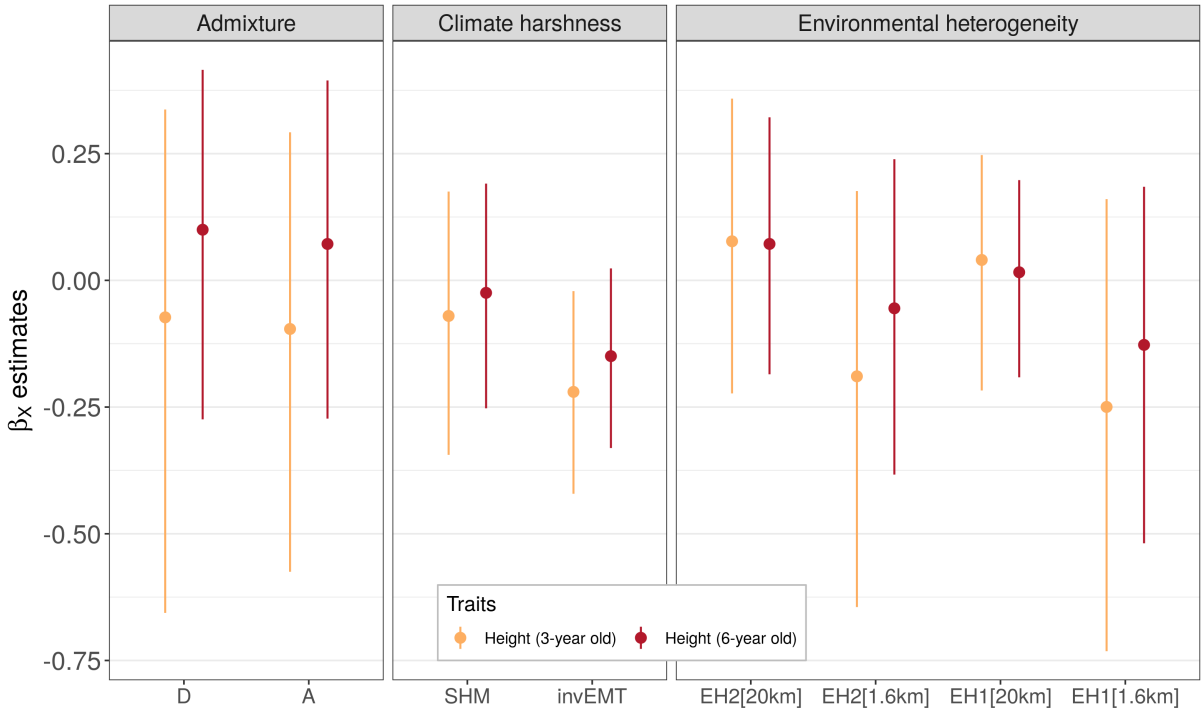


Figure X.35. Median and 95% intervals of the posterior distributions of β_X, the coefficient corresponding to the association between the eight potential drivers and the within-population additive genetic variation.

8.4 σ_{C_p} estimates and variance partitioning

8.4.1 Height at 3-year old

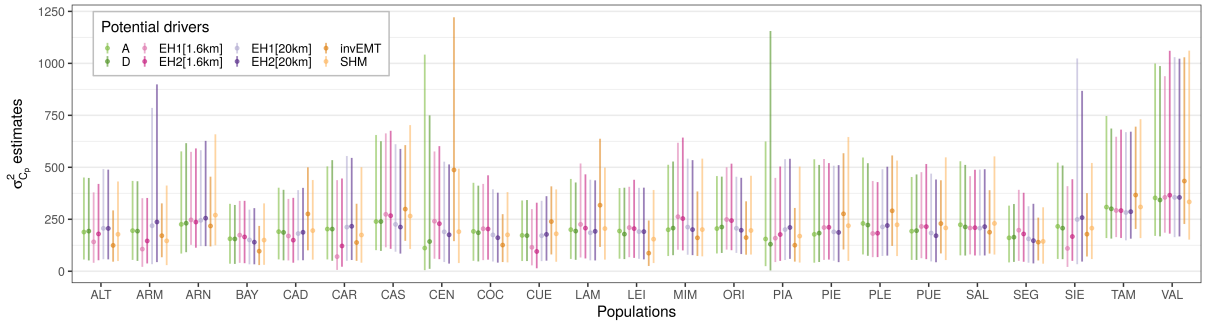


Figure X.36. Median and 95% intervals of the posterior distributions of $\sigma^2_{C_p}$, corresponding to the association between the within-population additive genetic variance (i.e. population-specific among families variance) and the eight potential drivers.

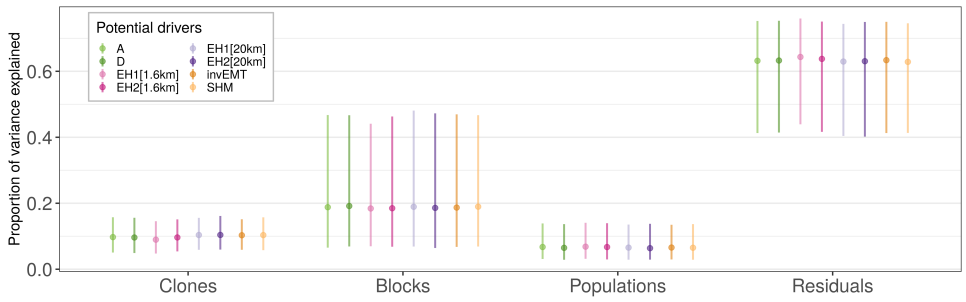


Figure X.37. Proportion of variance explained by the different components, namely the clones (π_C), the blocks (π_B), the populations (π_P) and the residuals (π_r).

8.4.2 Height at 6-year old

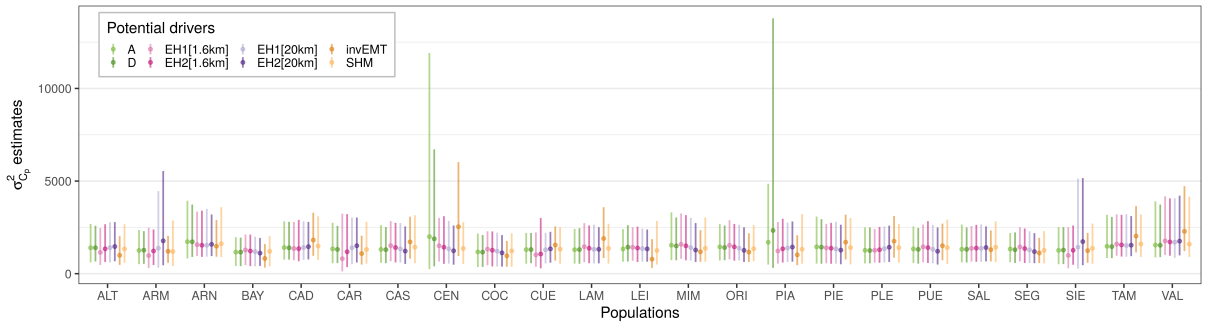


Figure X.38. Median and 95% intervals of the posterior distributions of $\sigma^2_{C_p}$, corresponding to the within-population total genetic variance (i.e. population-specific among clones variance).

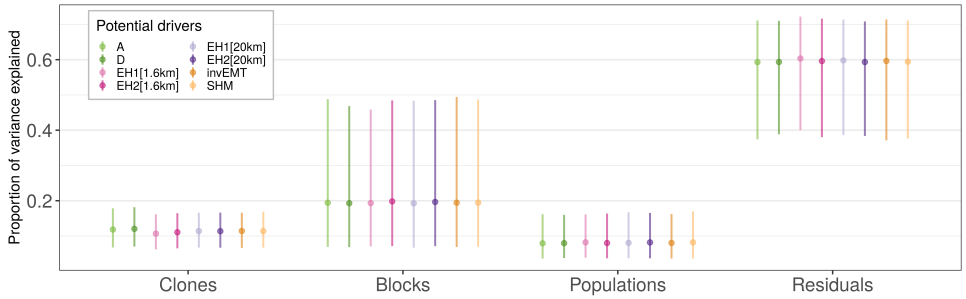


Figure X.39. Proportion of variance explained by the different components, namely the clones (π_C), the blocks (π_B), the populations (π_P) and the residuals (π_r).

9 Changes since preregistration

This study was pre-registered at the Center for Open Science (https://osf.io/knx6z/?view_only=41bb7b5cbf7241d0856e8b9e393cc795). Some changes have been made in the final manuscript compared to what was indicated in the pre-registration. There are listed below:

- There was a mistake in Table 2 of the pre-registration: we did not have the soil moisture index (SMI). Moreover, the notation for the summer heat moisture index has changed: instead of SumHMI, it is now noted as SHM.
- The initial number of clones and populations were 523 and 34 respectively. However, calculating the genetic variation in one population (from Madisouka) was impossible

as there was only one clone in that population. That's why there are 522 clones and 33 populations in the final manuscript.

- In the part *Statistical models* of the pre-registration, we specified: 'We will test three different models, from the simplest to the most complex, and we will keep the complex model if it converges and if the credible intervals are not too wide compared to the simpler models.' As the most complex worked well, we did not run the simpler models and we directly used the most complex model in the manuscript. In addition, there was a mistake in the pre-registration formula regarding the estimation of σ_{C_P} with the log-normal distribution, which was corrected in the final manuscript (see section *Model equation and priors* of the Supplementary Information).

- In the part *EnvironmentIndices* of the pre-registration, we indicated that we would calculate EHW, the environmental heterogeneity in a 20-km around each population location, as the variance of the PC1 scores weighted by the relative probability of gene flow from the surrounding region. However, the pollen dispersal kernels we used (from de-Lucas et al. 2008) are highly leptokurtic, which means that the probability of gene flow among trees located more than 200m from the GPS coordinates of the population is very low. Since the resolution of the climatic variables was only 1×1 km, we obtained implausible values for EHW and therefore decided not to use it. Instead, we calculated the variance of PC1 scores within a 1.6 km radius of the population locations. Furthermore, as the first two components of the PCA both explained a large part of the environmental variation (45.2% and 34.1%, respectively; X.10), we decided to calculate the environmental heterogeneity indices based on the PC1 and PC2 scores, resulting in four indices in the end: EH1[20km], EH2[20km], EH1[1.6km] and EH2[1.6km]. Last, we indicated that we would use the soil moisture index (SMI) as a measure of climate harshness but that was a mistake, we used the summer heat moisture index (SHM).

XI

Supplementary information - Chapter 2

1 Details about the experimental design

	All	Training P1	Test P1	Training P2	Test P2	Training P3	Test P3
All sites	33,121	24,840	8,281	27,349	5,772	26,172	6,949
Asturias	11,920	8,934	2,986	9,813	2,107	9,420	2,500
Bordeaux	6,473	4,836	1,637	5,285	1,188	5,063	1,410
Cáceres	340	249	91	297	43	272	68
Madrid	1,046	807	239	876	170	855	191
Portugal	13,342	10,014	3,328	11,078	2,264	10,562	2,780
Asturias - 10 months old	3,967	2,949	1,018	3,268	699	3133	834
Asturias - 21 months old	3,979	3,022	957	3,275	704	3,143	836
Asturias - 37 months old	3,974	2,963	1,011	3,270	704	3,144	830
Bordeaux - 25 months old	3,237	2,420	817	2,643	594	2,532	705
Bordeaux - 37 months old	3,236	2,416	820	2,642	594	2,531	705
Cáceres - 8 months old	340	249	91	297	43	272	68
Madrid - 13 months old	1,046	807	239	876	170	855	191
Portugal - 11 months old	4,152	3,102	1,050	3,442	710	3,266	886
Portugal - 15 months old	3,773	2,833	940	3,134	639	2,980	793
Portugal - 20 months old	2,752	2,078	674	2,288	464	2,192	560
Portugal - 27 months old	2,665	2,001	664	2,214	451	2,124	541

Table XI.1. Number of observations in the entire dataset (after filtering) and in each of the three partitions. In the P1 partition, the training dataset was obtained by randomly sampling 75% of the observations and the test dataset contains the remaining 25% observations. In the P2 partition, the training dataset was obtained by randomly sampling 28 provenances and the test dataset contains the remaining 6 provenances. The P3 partition corresponds to a non-random split between a training dataset of 28 provenances and a test dataset containing 6 provenances with at least one provenance from each under-represented gene pool (i.e. northern Africa, south-eastern Spain and Corsican gene pools).

Code	Name	Number of genotypes	Number of trees	Number of observations
ALT	Alto de la Llama	9	216	580
ARM	Armayán	8	213	547
ARN	Arenas de San Pedro	17	412	1083
BAY	Bayubas de Abajo	18	462	1181
BON	Boniches	9	221	594
CAD	Cadavedo	10	245	658
CAR	Carbonero el Mayor	6	156	398
CAS	Castropol	10	246	642
CEN	Cenicientos	9	207	561
COC	Coca	18	424	1114
COM	Cómpeta	4	109	272
CUE	Cuellar	28	680	1750
HOU	Hourtin	26	645	1669
LAM	Lamuño	9	216	563
LEI	Leiria	23	549	1439
MAD	Madisouka	1	19	54
MIM	Mimizan	18	445	1111
OLB	Olba	22	552	1476
OLO	Olonne sur Mer	24	563	1441
ORI	Oria	26	651	1720
PET	Petrocq	24	594	1496
PIA	Pinia	16	413	1046
PIE	Pineta	9	220	582
PLE	Pleucadec	20	480	1234
PUE	Puerto de Vega	8	198	497
QUA	Quatretonda	17	448	1156
SAC	San Cipriano de Ribaterme	9	208	499
SAL	San Leonardo	14	323	804
SEG	Sergude (Huerto Semillero)	21	536	1340
SIE	Sierra de Barcia	8	203	506
STJ	St-Jean des Monts	28	718	1824
TAM	Tamrabta	15	320	839
VAL	Valdemaqueda	12	286	750
VER	Le Verdon	27	663	1695

Table XI.2. Provenance information: provenance codes used in the study, provenance names, number of genotypes, trees and observations (an observation being a height-growth measurement in a given year on one individual) per provenance.

Provenance	NA	C	CS	FA	IA	SES
ALT	0.003	0.000	0.119	0.096	**0.780**	0.003
ARM	0.005	0.007	0.021	0.006	**0.959**	0.001
ARN	0.010	0.002	**0.958**	0.007	0.010	0.013
BAY	0.003	0.004	**0.966**	0.013	0.010	0.004
BON	0.152	0.010	**0.654**	0.003	0.002	0.179
CAD	0.002	0.001	0.053	0.010	**0.933**	0.002
CAR	0.001	0.001	**0.904**	0.060	0.022	0.011
CAS	0.001	0.000	0.005	0.001	**0.991**	0.001
CEN	0.013	0.002	**0.892**	0.003	0.041	0.050
COC	0.017	0.005	**0.826**	0.061	0.043	0.047
COM	0.239	0.028	0.127	0.011	0.039	**0.556**
CUE	0.003	0.001	**0.874**	0.063	0.056	0.002
HOU	0.004	0.001	0.026	**0.960**	0.007	0.002
LAM	0.002	0.001	0.003	0.050	**0.943**	0.000
LEI	0.004	0.003	**0.512**	0.007	0.472	0.002
MAD	**0.764**	0.001	0.000	0.002	0.000	0.233
MIM	0.004	0.002	0.024	**0.952**	0.013	0.005
OLB	0.080	0.010	**0.776**	0.006	0.003	0.126
OLO	0.004	0.001	0.007	**0.980**	0.006	0.003
ORI	0.249	0.005	0.027	0.001	0.009	**0.709**
PET	0.003	0.001	0.021	**0.966**	0.004	0.004
PIA	0.004	**0.974**	0.010	0.001	0.008	0.003
PIE	0.007	**0.970**	0.022	0.000	0.000	0.001
PLE	0.005	0.001	0.060	**0.924**	0.005	0.004
PUE	0.002	0.000	0.021	0.001	**0.974**	0.002
QUA	0.092	0.006	**0.499**	0.005	0.015	0.383
SAC	0.004	0.001	0.268	0.002	**0.723**	0.002
SAL	0.010	0.004	**0.944**	0.027	0.008	0.008
SEG	0.003	0.001	0.151	0.013	**0.829**	0.003
SIE	0.003	0.001	0.089	0.015	**0.891**	0.001
STJ	0.003	0.002	0.027	**0.947**	0.017	0.004
TAM	**0.937**	0.000	0.025	0.000	0.037	0.001
VAL	0.012	0.003	**0.943**	0.006	0.011	0.025
VER	0.003	0.002	0.019	**0.972**	0.003	0.001

Table XI.3. Mean proportion belonging to each gene pool for each provenance. For each provenance, the highest proportion belonging to a given gene pool is in bold. The gene pools come from: northern Africa (NA), Corsica (C), central Spain (CS), French Atlantic region (FA), Iberian Atlantic region (IA) and south-eastern Spain (SES).

2 Height-associated positive-effect alleles (PEAs)

2.1 Calculation of the counts of height-associated positive-effect alleles

This section complements the section 3.2 in the manuscript and we explain here in more details how we calculated the counts of global and regional height-associated positive-effect alleles. As already explained in the manuscript, for each of the four GWAS (a global GWAS and three regional GWAS), we selected the 350 SNPs with the highest absolute estimates of the posterior effect size (i.e. Rao-Blackwellized estimates), corresponding approximately to the estimated number of SNPs with non-zero effects on height in a previous study (i.e. the level of polygenicity; de Miguel et al. 2020). Then, for each selected allele, if the posterior effect size was negative, it was converted to positive values and the reference allele flipped to select only alleles that have a positive effect on height (positive-effect alleles; PEAs). Thus, we ended up with four groups of PEAs. One group had a global positive effect (i.e. range-wide effect) on height and was used to calculate the count of global PEAs that each sapling has ($gPEA_g$ variable). Trees with the same genotype had the same $gPEA_g$, such as $gPEA_g = \sum_{l=1}^{350} G_{lg}$, where $G_{lg} = \{0, 1, 2\}$ is the number of global PEAs that the genotype g has at the locus l. The three other groups of PEAs had a regional positive-effect on height (i.e. specific effect in a given geographical region/in a particular environment) and were used to calculate the number of regional PEAs that each tree has ($rPEA_{gr}$ variable). $rPEA_{gr}$ was calculated based on both the tree genotype g and the region r of its planting site, such as $rPEA_{gr} = \sum_{l=1}^{350} G_{lgr}$, where $G_{lgr} = \{0, 1, 2\}$ is the number of PEAs specific to the geographical region r that the genotype g has at the locus l (see Fig. XI.1 below for a summary diagram of how the counts of global and regional PEAs were obtained).

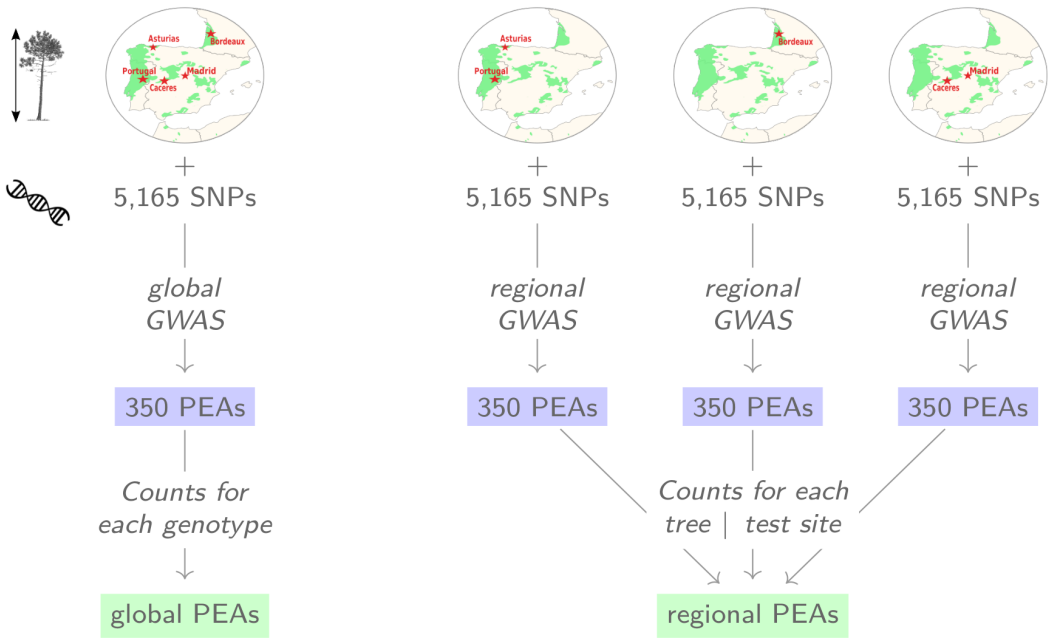


Figure XI.1. Schematic representation of the calculation of the global PEA counts ($gPEAg$ variable) and regional PEA counts ($rPEAgr$ variable).

2.2 Shared proportion of globally and regionally selected height-associated SNPs

A small proportion of regionally-selected height-associated SNPs was shared among the different regions: 20% (69 SNPs) shared between the French Atlantic (Bordeaux) and the Iberian Atlantic region (Asturias and Portugal), 12% (41 SNPs) shared between the Mediterranean (Cáceres and Madrid) and the French Atlantic region and 24% (83 SNPs) shared between the Iberian Atlantic and the Mediterranean region (Fig. XI.2). Interestingly, height-associated SNPs that were shared among different regions show consistently similar effects across regions (e.g. positive effects in two or more regions rather than antagonist effects): we did not detect a single SNP with an antagonistic effect (Fig. XI.2).

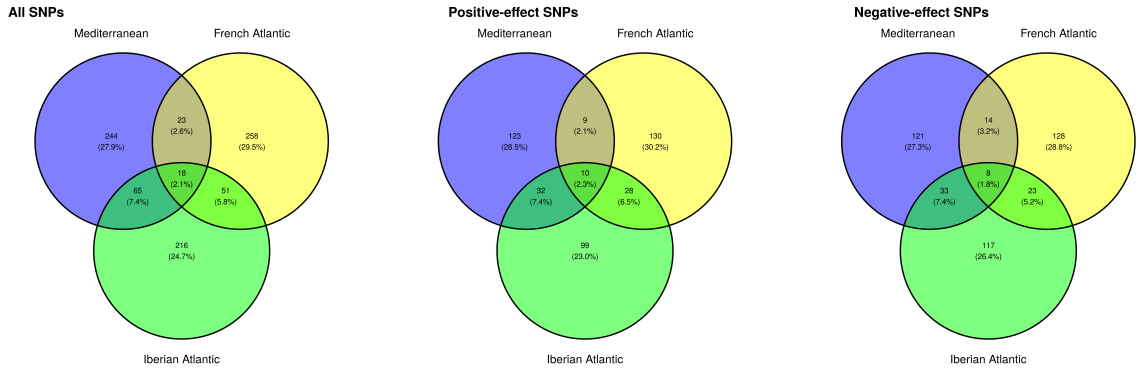


Figure XI.2. Number and proportion of height-associated alleles shared among regions: the French Atlantic region (Bordeaux), the Iberian Atlantic region (Asturias and Portugal) and the Mediterranean region (Cáceres and Madrid). This figure was obtained before transforming the height-associated negative-effect alleles in positive-effect alleles by flipping their reference allele.

82.9% of the globally-selected SNPs were at least selected once in a regional GWAS too (Fig. XI.3). Height-associated SNPs that were selected both globally and regionally (at least in one region) show consistently similar effects (i.e. either positive or negative, but not antagonistic effects): we did not find a single SNP with an antagonist effect when selected globally or regionally.

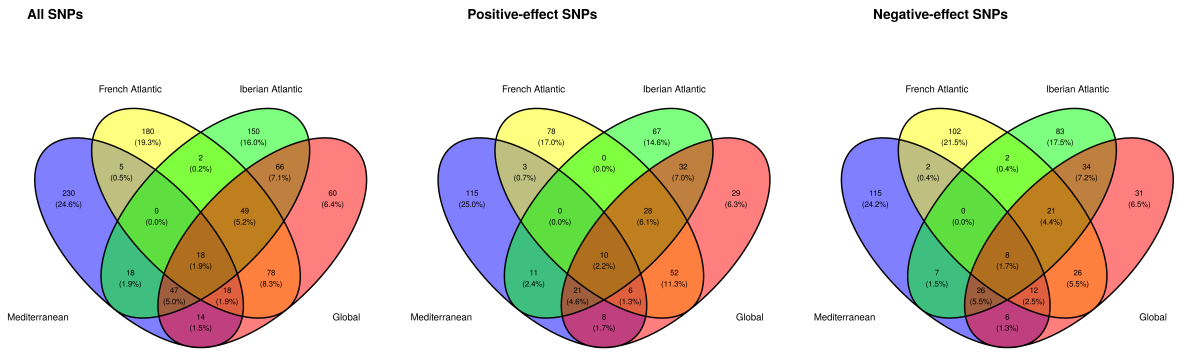


Figure XI.3. Number and proportion of alleles shared among the sets of 350 height-associated alleles selected globally and those selected regionally.

Importantly, we cannot exclude that the proportion of SNPs shared among regions or shared among globally-selected and regionally-selected SNPs is a function of the sample size. Indeed, the two regions with the lowest sample size (French Atlantic and Mediterranean regions) are also the ones that share the lowest number of height-associated SNPs (Fig. XI.2). Similarly, the

globally-selected height-associated SNPs share the highest proportion with regionally-selected height-associated SNPs from the Iberian Atlantic region (the region with the highest sample size) and the lowest proportion with regionally-selected height-associated SNPs from the Mediterranean region (the region with the lowest sample size) (Fig. XI.3).

3 Climatic data

3.1 In the test sites

We extracted monthly climatic data from the EuMedClim database at 1-km resolution (Fréjaville and Benito Garzón 2018). We calculated six variables that describe both extreme and average temperature and precipitation conditions in the test sites during the year preceding the measurements: the mean of monthly precipitation (*mean.pre*, mm), minimum of monthly minimum temperatures (*min.tmn*, °C), minimum of monthly precipitation during summer -June to September- (*min.presummer*, °C), the mean of monthly maximum temperatures (*mean.tmax*, °C), maximum of monthly precipitation (*max.pre*, mm), maximum of monthly maximum temperatures (*max.tmx*, °C). These variables had at most a correlation coefficient of 0.85 among each other (Fig. XI.4). Due to the unbalanced number of measurements among test sites (trees were measured only once in the hottest and driest sites, Cáceres and Madrid, as survival was very low), some of these variables were slightly correlated with tree age (at most with a correlation coefficient of 0.56 for the mean of the monthly precipitation; Fig. XI.4). We decided not to include soil variables (from the European Soil Database: https://esdac.jrc.ec.europa.eu/) in the analyses as they were highly correlated to some of the climatic variables. Likewise, we did not include variables related to water balance or evapotranspiration potential as they were highly correlated with temperature and precipitation variables.

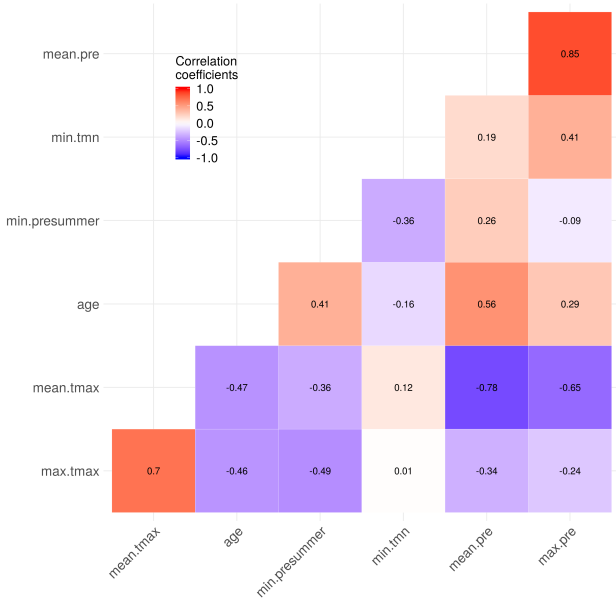


Figure XI.4. Correlation matrix between variables related to the climatic conditions in the test sites and tree age at the time of the measurements. The climatic variables, calculated over the year preceding the measurements, are: the mean of monthly precipitation (*mean.pre*, mm), minimum of monthly minimum temperatures (*min.tmn*, °C), minimum of monthly precipitation during summer -June to September- (*min.presummer*, °C), the mean of monthly maximum temperatures (*mean.tmax*, °C), maximum of monthly precipitation (*max.pre*, mm), maximum of monthly maximum temperatures (*max.tmx*, °C).

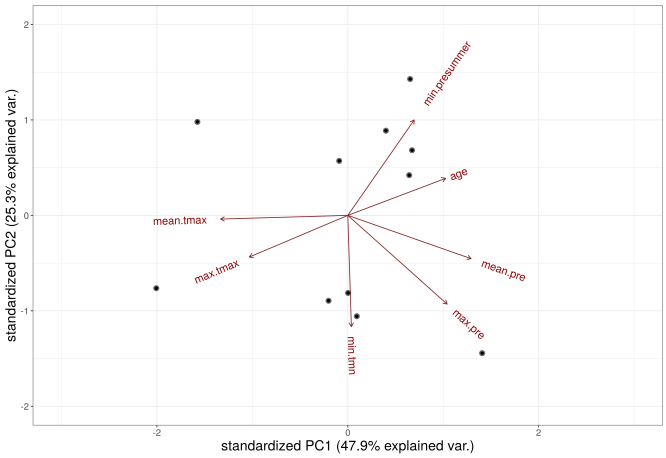


Figure XI.5. Principal component analysis of the variables related to the climatic conditions in the test sites and tree age at the time of the measurements. See Fig. XI.4 for the meaning of variable abbreviations.

The climatic similarity among test sites during the year preceding the measurements was described by the covariance matrix Ω. This covariance matrix was used to estimate the association between height-growth variation and the climatic similarity between test sites in *models M3* to *M6* (Table V.1), following Jarquín et al. (2014); see also a similar approach but using Euclidean distance matrices in Thomson et al. (2018).

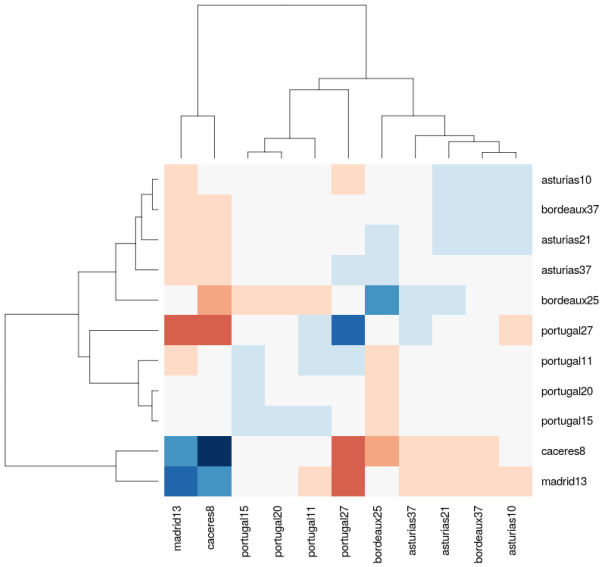


Figure XI.6. Heatmap of the covariance matrix Ω describing the climatic similarity among test sites during the year preceding the measurements. The labels correspond to the name of the test sites followed by the age of the trees at the date of the measurement (in months).

3.2 In the provenances

We extracted yearly data from the EuMedClim database at 1-km resolution (Fréjaville and Benito Garzón 2018). We calculated four variables that describe the mean temperature and precipitation in the provenance locations over the period from 1901 to 2009, representing the climate under which provenances have evolved: the average of the annual daily mean temperature (*mean.temp*, °C), the average of the maximum temperature of the warmest month

(*max.temp*, °C), the average of the annual precipitation (*mean.pre*, mm) and the average of the precipitation of the driest month (*min.pre*, mm). These variables had at most a correlation coefficient of 0.77 among each other (Fig. XI.7), and three of them were correlated to the population genetic structure, i.e. the gene pool assignment ($|\rho| \geq 0.6$). Indeed, the provenance proportion belonging to the French Atlantic gene pool was positively correlated (ρ=0.83) to the average of the precipitation during the driest month, whereas the proportion belonging to the Central Spain gene pool was negatively correlated (ρ=-0.68) to the average of the annual precipitation and positively correlated (ρ=0.6) to the average of the maximum temperature of the warmest month (Fig. XI.7). However, the confounding effect introduced by these correlations was mitigated by some provenances belonging to different gene pools but occurring in similar climates (i.e. French and Iberian Atlantic provenances), and some provenances occurring in different climates but belonging to the same gene pool (i.e. Corsican provenances). Soil variables from the European Soil Database (https://esdac.jrc.ec.europa.eu/) were not included in our study as they were highly correlated to some of the selected climatic variables.

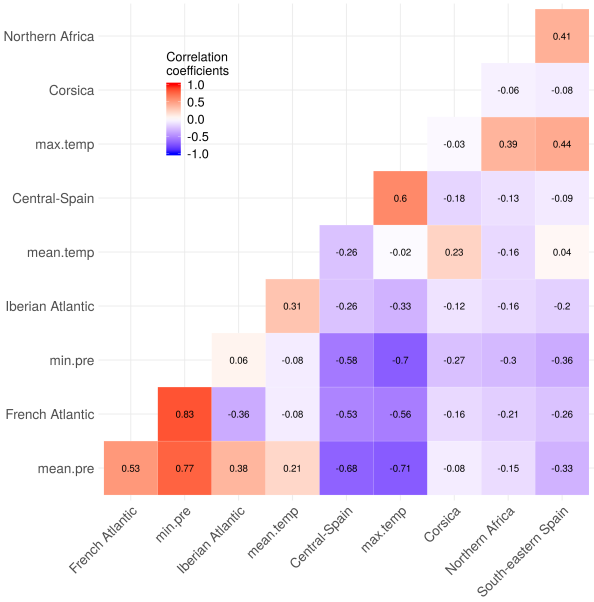


Figure XI.7. Correlation matrix of the variables related to the climatic conditions in the provenance locations and variables related to the population genetic structure (genotype proportion belonging to each gene pool). The genotypes belong to 6 distinct gene pools from: Northern Africa, Corsica, Central Spain, French Atlantic region, Iberian Atlantic region and south-eastern Spain. The climatic variables, calculated over the period from 1901 to 2009, are: the average of the annual daily mean temperature (*mean.temp*, °C), the average of the maximum temperature of the warmest month (*max.temp*, °C), the average of the annual precipitation (*mean.pre*, mm) and the average of the precipitation of the driest month (*min.pre*, mm).

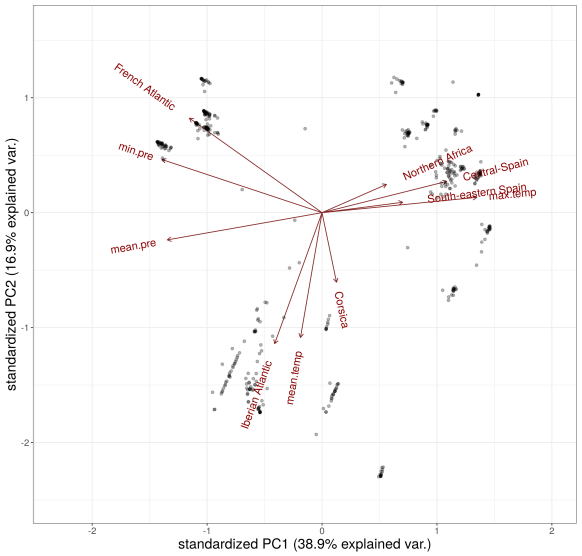


Figure XI.8. Principal component analysis of the variables related to the climatic conditions in the provenance locations and variables related to the population genetic structure (genotype proportion belonging to each gene pool). See Fig. XI.7 for the meaning of variable abbreviations.

The climatic similarity among provenances was described by the covariance matrix Φ. This covariance matrix was used to estimate the association between height-growth variation and the climatic similarity between provenances in *model M6*, following Jarquín et al. (2014); see also a similar approach but using Euclidean distance matrices in Thomson et al. (2018).

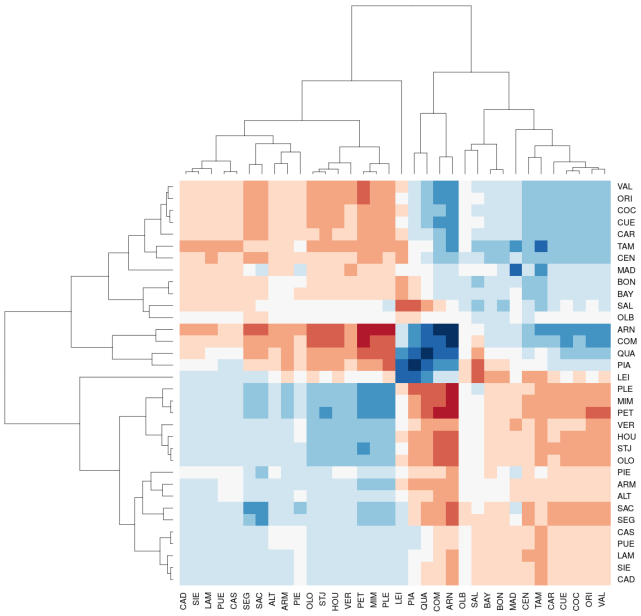


Figure XI.9. Heatmap of the covariance matrix Φ of the climatic variables in the provenances. Labels correspond to the provenance names.

4 Model equations and priors

4.1 Model equations

In this section, the complete equations of each model are specified. The priors are indicated in the next section.

4.1.1 Baseline *models M1* and *M2*: separating the genetic and plastic components

***Model M0*:**

$$\begin{aligned} \log(h_{isb}) &\sim \mathcal{N}(\mathbf{X}\beta + \mu_{sb}\,,\ \sigma^2) \\ \mathbf{X}\beta &= \beta_0 + \beta_{age}\mathrm{age}_i + \beta_{age2}\mathrm{age}_i^2 \\ \mu_{sb} &= S_s + B_{b(s)} \end{aligned}$$

***Model M1*:**

$$\begin{aligned} \log(h_{isbpg}) &\sim \mathcal{N}(\mathbf{X}\beta + \mu_{sbpg}\,,\ \sigma^2) \\ \mathbf{X}\beta &= \beta_0 + \beta_{age}\mathrm{age}_i + \beta_{age2}\mathrm{age}_i^2 \\ \mu_{sbpg} &= S_s + B_{b(s)} + P_p + G_{g(p)} \end{aligned}$$

Model M2:

$$\begin{aligned} \log(h_{isbpg}) &\sim \mathcal{N}(\mathbf{X}\beta + \mu_{sbpg}\,,\ \sigma^2) \\ \mathbf{X}\beta &= \beta_0 + \beta_{age}\mathrm{age}_i + \beta_{age2}\mathrm{age}_i^2 \\ \mu_{sbpg} &= S_s + B_{b(s)} + P_p + G_{g(p)} + S_sP_p \end{aligned}$$

where $\mathbf{X}$ is the 3-column design matrix and β is a vector including the intercept β_0 and the coefficients β_{age} and β_{age2} of the fixed effect variables (*age* and age^2, respectively). μ_{sbpg} is the vector of varying intercepts with the provenance intercepts P_p, the genotype intercepts $G_{g(p)}$, the site intercepts S_s, the block intercepts $B_{b(s)}$ and the interaction between the site and provenance intercepts S_sP_p.

4.1.2 Explanatory *models M3* to *M6*: potential drivers underlying height-growth variation

***Model M3*:**

$$\begin{aligned} \log(h_{isbpg}) &\sim \mathcal{N}(\mathbf{X}\beta + \mu_{sbpg}\,,\ \sigma^2) \\ \mathbf{X}\beta &= \beta_0 + \beta_{age}\mathrm{age}_i + \beta_{age2}\mathrm{age}_i^2 \\ \mu_{isbpg} &= S_s + B_{b(s)} + P_p + G_{g(p)} + cs_{is} \\ cs_{is} &\sim \mathcal{N}(0, \Omega\sigma^2_{cs_{is}}) \end{aligned}$$

***Models M4, M5* and *M6*:**

In *models M4* and *M5*, we hypothesized that the genetic component of height growth was influenced by the proportion belonging to each gene pool (proxy of the population demographic history and genetic drift). In *M4*, following Wolak and Reid (2017), gene pools j were allowed to vary in their mean relative contribution g_j on height growth as follows:

$$\log(h_{isbpg}) \sim \mathcal{N}(\mathbf{X}\beta + \mu_{sbpg}\,,\ \sigma^2)$$
$$\mathbf{X}\beta = \beta_0 + \beta_{age}\text{age}_i + \beta_{age2}\text{age}_i^2$$
$$\mu_{ijsbpg} = S_s + B_{b(s)} + P_p + G_{g(p)} + cs_{is} + \sum_{j=1}^{6} q_{gj} g_j$$
$$g_j \sim \mathcal{N}(0, \sigma^2_{g_j})$$

where q_{gj} corresponds to the proportion of each genotype g belonging to the gene pool j (as estimated in Jaramillo-Correa et al. 2015) and g_j is the mean relative contribution of gene pool j on height growth. In this model, trees from the same gene pool are considered to be unrelated, which is a reasonable assumption given the sampling scheme (see Materials & Methods).

M5 extends *M4* by allowing gene pools j to vary in their total genetic variance $\sigma^2_{A_j}$, following Muff et al. (2019). This involves replacing the genotype varying intercepts $G_{g(p)}$ in *M4* by partial genetic values a_{gj} standing for the relative contribution of gene pool j to the genetic value a_g of genotype g (approximated in *M4* by the genotype intercepts $G_{g(p)}$). Thus, *M5* can be expressed as *M4* but with μ_{ijsbpg} equal to:

$$\log(h_{isbpg}) \sim \mathcal{N}(\mathbf{X}\beta + \mu_{sbpg}\,,\ \sigma^2)$$
$$\mathbf{X}\beta = \beta_0 + \beta_{age}\text{age}_i + \beta_{age2}\text{age}_i^2$$
$$\mu_{ijsbpg} = S_s + B_{b(s)} + P_p + cs_{is} + \sum_{j=1}^{6} q_{gj} g_j + \sum_{j=1}^{6} a_{gj}$$
$$\mathbf{a}_j^{\intercal} = (a_{1j}, ..., a_{nj})^{\intercal} \sim \mathcal{N}(0, \sigma^2_{A_j}\,\mathbf{A_j})$$

with $\mathbf{A_j}$ the genomic relationship matrix specific to the gene pool j and $\sigma^2_{A_j}$, the total genetic variance in gene pool j. $\mathbf{A_j}$ matrices were calculated based on SNPs that did not show any association with height at range-wide geographical scales (see Muff et al. 2019 for details on $\mathbf{A_j}$ calculation). Using the modeled residual variance σ^2 and gene-pool specific total genetic variances $\sigma^2_{A_j}$, we calculated the gene-pool specific broad-sense heritability as: $H^2_j = \sigma^2_{A_j}/(\sigma^2_{A_j} + \sigma^2)$.

In *model M6*, we hypothesized that populations are genetically adapted to the climatic conditions in which they evolved. Thus, we aimed to quantify the association between height growth and the climatic similarity among provenances, while accounting also for the proportion belonging to each gene pool. We kept the genotype varying intercepts (like in *M1* to *M4*) but not the gene pool-specific total genetic variances (unlike *M5*). Thus, *M6* extends *M4* as:

$$\log(h_{isbpg}) \sim \mathcal{N}(\mathbf{X}\beta + \mu_{sbpg}\,,\ \sigma^2)$$
$$\mathbf{X}\beta = \beta_0 + \beta_{age}\text{age}_i + \beta_{age2}\text{age}_i^2$$
$$\mu_{ijsbpg} = S_s + B_{b(s)} + P_p + G_{g(p)} + cs_{is} + cp_p + \sum_{j=1}^{6} q_{gj} g_j$$
$$cp_p \sim \mathcal{N}(0, \Phi\,\sigma^2_{cp_p})$$

where Φ is the covariance matrix describing the climatic similarity between provenances p (Fig. XI.9) and cp_p are varying intercepts associated with each provenance p. In *M6*, the genetic component was partitioned among the regression on the climatic covariates (cp_p), the gene pool covariates (g_j), and the deviations related to the genotype ($G_{g(p)}$) and provenance (P_p)

effects (resulting, for instance, from adaptation to environmental variables not measured in our study).

4.1.3 Predictive *models M7* to *M12*: combining climatic and genomic information to improve predictions

Model M7:

$$\log(h_{isb}) \sim \mathcal{N}(\mathbf{X}\beta + \mu_{jsbpg}\,,\ \sigma^2)$$
$$\mathbf{X}\beta = \beta_0 + \beta_{age}\text{age}_i + \beta_{age2}\text{age}_i^2$$
$$\mu_{jsbpg} = S_s + B_{b(s)} + \sum_{j=1}^{6} q_{gj}g_j + \beta_{min.pre,s}min.pre_p + \beta_{max.temp,s}max.temp_p + \beta_{gPEA,s}gPEA_g$$

Model M8:

$$\log(h_{isbr}) \sim \mathcal{N}(\mathbf{X}\beta + \mu_{jsbpgr}\,,\ \sigma^2)$$
$$\mathbf{X}\beta = \beta_0 + \beta_{age}\text{age}_i + \beta_{age2}\text{age}_i^2$$
$$\mu_{jsbpgr} = S_s + B_{b(s)} + \sum_{j=1}^{6} q_{gj}g_j + \beta_{min.pre,s}min.pre_p + \beta_{max.temp,s}max.temp_p + \beta_{rPEA,s}rPEA_{gr}$$

Model M9:

$$\log(h_{isb}) \sim \mathcal{N}(\mathbf{X}\beta + \mu_{jsb}\,,\ \sigma^2)$$
$$\mathbf{X}\beta = \beta_0 + \beta_{age}\text{age}_i + \beta_{age2}\text{age}_i^2$$
$$\mu_{jsb} = S_s + B_{b(s)} + \sum_{j=1}^{6} q_{gj}g_j$$

Model M10:

$$\log(h_{isb}) \sim \mathcal{N}(\mathbf{X}\beta + \mu_{jsbp}\,,\ \sigma^2)$$
$$\mathbf{X}\beta = \beta_0 + \beta_{age}\text{age}_i + \beta_{age2}\text{age}_i^2$$
$$\mu_{jsbp} = S_s + B_{b(s)} + \beta_{min.pre,s}min.pre_p + \beta_{max.temp,s}max.temp_p$$

Model M11:

$$\log(h_{isbg}) \sim \mathcal{N}(\mathbf{X}\beta + \mu_{sbg}\,,\ \sigma^2)$$
$$\mathbf{X}\beta = \beta_0 + \beta_{age}\text{age}_i + \beta_{age2}\text{age}_i^2$$
$$\mu_{sbg} = S_s + B_{b(s)} + \beta_{gPEA,s}gPEA_g$$

Model M12:

$$\log(h_{isbgr}) \sim \mathcal{N}(\mathbf{X}\beta + \mu_{sbgr}\,,\ \sigma^2)$$
$$\mathbf{X}\beta = \beta_0 + \beta_{age}\text{age}_i + \beta_{age2}\text{age}_i^2$$
$$\mu_{sbgr} = S_s + B_{b(s)} + \beta_{rPEA,s}rPEA_{gr}$$

where $min.pre_p$ and $max.temp_p$ are the climatic variables in the provenance locations, $\beta_{min.pre,s}$ and $\beta_{max.temp,s}$ their site-specific slopes, $gPEA_g$ and $rPEA_{gr}$ the counts of global and regional PEAs and $\beta_{gPEA,s}$ and $\beta_{rPEA,s}$ their site-specific slopes.

4.2 Model priors

In **all models:**

$$\begin{bmatrix} S_s \\ B_{b(s)} \\ P_p \\ G_{g(p)} \\ S_s P_p \end{bmatrix} \sim \mathcal{N}\left(0, \begin{bmatrix} \sigma_S \\ \sigma_B \\ \sigma_P \\ \sigma_G \\ \sigma_{Inter} \end{bmatrix}\right)$$

$$(\sigma, \sigma_S, \sigma_B, \sigma_P, \sigma_G, \sigma_{Inter}, \sigma_{cs_{is}}, \sigma_{g_j}, \sigma_{A_j}, \sigma_{cp_p})^\intercal \sim \text{StudentT}(3, 0, 10)$$

$$\beta_0 \sim \mathcal{N}(0, 5)$$

$$\begin{bmatrix} \beta_{age} \\ \beta_{age2} \end{bmatrix} \sim \mathcal{N}(0, 1)$$

In ***model M7***:

$$\begin{bmatrix} S_s \\ \beta_{min.pre,s} \\ \beta_{max.temp,s} \\ \beta_{gPEA,s} \end{bmatrix} \sim \text{MVNormal}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \text{S}\right)$$

$$\text{S} = \begin{pmatrix} \sigma_S & 0 & 0 & 0 \\ 0 & \sigma_{\beta_{min.pre,s}} & 0 & 0 \\ 0 & 0 & \sigma_{\beta_{max.temp,s}} & 0 \\ 0 & 0 & 0 & \sigma_{\beta_{gPEA,s}} \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 & \rho \\ 1 & 1 & \rho & 1 \\ 1 & \rho & 1 & 1 \\ \rho & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} \sigma_S & 0 & 0 & 0 \\ 0 & \sigma_{\beta_{min.pre,s}} & 0 & 0 \\ 0 & 0 & \sigma_{\beta_{max.temp,s}} & 0 \\ 0 & 0 & 0 & \sigma_{\beta_{gPEA,s}} \end{pmatrix}$$

$$\begin{bmatrix} \sigma_S \\ \sigma_{\beta_{min.pre,s}} \\ \sigma_{\beta_{max.temp,s}} \\ \sigma_{\beta_{gPEA,s}} \end{bmatrix} \sim \text{StudentT}(3, 0, 10)$$

$$\begin{pmatrix} 1 & 1 & 1 & \rho \\ 1 & 1 & \rho & 1 \\ 1 & \rho & 1 & 1 \\ \rho & 1 & 1 & 1 \end{pmatrix} \sim \text{LKJcorr}(4)$$

where $\beta_{x,s}$ corresponds to $\beta_{min.pre,s}$ in M7 and $\beta_{max.temp,s}$ in M8.

In ***model M8***: same as M7 but replacing $\beta_{gPEA,s}$ and $\sigma_{\beta_{gPEA,s}}$ by $\beta_{rPEA,s}$ and $\sigma_{\beta_{rPEA,s}}$, respectively.

In ***models M9***, ***M10***, ***M11*** and ***M12***, same as M7 and M8.

5 Model comparison

5.1 Description of the different indices used to compare the models

Similarly to the in-sample proportion of the variance explained by each model m in each site s ($\mathcal{R}^2_{ms}|age$), we calculated the out-of-sample proportion of the variance predicted by each model m in each site s conditional on the age effect as follows:

$$prediction\ \mathcal{R}^2_{ms}|age = \frac{V_{\text{pred}_{ms}} - V_{\text{age}_{2s}}}{V_{\text{y}_s} - V_{\text{age}_{2s}}}$$

where $V_{\text{pred}_{ms}}$ is the variance of the modeled predictive means from model m in site s of the test dataset, $V_{\text{age}_{2s}}$ is the variance predicted by the age effect in the *model M2* in site s and V_{y_s} is the phenotypic variance in the site s of the test dataset. Estimates of *prediction* $\mathcal{R}^2_{ms}|age$ in the three partitions are reported in Table V.4.

We then calculated other indices that are not presented in the main manuscript as they are not necessary to support the main objectives of the paper, especially whether the models combining the climatic and genomic drivers of the genetic component can improve the prediction on new provenances. However, they are still useful to compare the goodness-of-fit and predictive ability of the models, that's why we report them here.

We calculated the total in-sample proportion of the variance explained by each model m such as:

$$\mathcal{R}^2_m = \frac{V_{\text{pred}_m}}{V_{\text{y}}}$$

where V_{pred_m} is the variance of the modeled predictive means from model m in the training dataset and V_{y} is the phenotypic variance in the training dataset. Similarly, we calculated the *prediction* $\mathcal{R}^2_m$ on the test dataset, that is the total out-of-sample proportion of variance predicted by each model m in the test dataset.

We calculated the in-sample proportion of the variance explained by the fixed effects of each model m such as:

$$\mathcal{R}^2_{m\,(\text{fix})} = \frac{V_{\text{pred}_m(\text{fix})}}{V_{\text{y}}}$$

where $V_{\text{pred}_m(\text{fix})}$ is the variance explained by the fixed effects of model m in the training dataset and V_{y} is the phenotypic variance in the training dataset. Similarly, we calculated the *prediction* $\mathcal{R}^2_m(fix)$ on the test dataset, that is the out-of-sample proportion of variance predicted by the fixed effects of each model m in the test dataset.

Last, we calculated the total in-sample proportion of the variance explained by each model m conditional on the age effect, such as:

$$\mathcal{R}^2_m|age = \frac{V_{\text{pred}_m} - V_{\text{age}_2}}{V_{\text{y}} - V_{\text{age}_2}}$$

where V_{age_2} is the variance explained by the age effect in *model M2*. Similarly, we calculated the *prediction* $\mathcal{R}^2_m|age$ on the test dataset, that is the total out-of-sample proportion of variance predicted by each model m conditional on the age effect in the test dataset.

To both evaluate the model goodness-of-fit and predictive ability, we also calculated the model **mean predictive error** of each model m (mean of observed minus predicted responses, PE_m) on the training and test datasets of the three partitions.

$\mathcal{R}^2_m$, $\mathcal{R}^2_{m\,(\text{fix})}$, $\mathcal{R}^2_m|age$, PE_m and their predictive equivalents are presented in Tables XI.4 (P1 partition), XI.9 (P2 partition) and XI.12 (P3 partition).

To assess the model predictive ability, we also calculated the $ELPD_{loo}$, which is the Bayesian leave-one-out estimate of the expected log pointwise predictive density (equation 4 in Vehtari et al. 2017). This is a method for estimating out-of-sample prediction accuracy of Bayesian models, which is asymptotically equal to WAIC (Vehtari et al. 2017) and has the great advantage that it can be estimated without refitting the model. $ELPD_{loo}$ estimates can be found in Table XI.6 for the P1 partition and its pairwise comparisons between models in Tables XI.7 (P1 partition) and XI.10 (P2 partition). $ELPD_{loo}$, like WAIC, provides various advantages over AIC and DIC, especially that it is not a point estimate and, on the contrary, has an entire posterior distribution (Vehtari et al. 2017). Moreover, calculating $ELPD_{loo}$ using Pareto-smoothed importance sampling as we did in the present study using the *loo* R package, leads to more robust estimates than with WAIC (e.g. in cases with weak priors or influential observations). Models with higher $ELPD_{loo}$ are expected to have a higher predictive ability for new observations. In other words, $ELPD_{loo}$ indicates which model best captures each left-out data point. Therefore, $ELPD_{loo}$ indicates whether models have good predictive ability for new observations, but not for new groups (e.g., new provenances in our case). To estimate to predictive ability on new provenances with the $ELPD_{loo}$, we would have had to divide the dataset into k partitions (e.g. 34 partitions and leaving one provenance out each time, 34 being the number of provenances; or 6 partitions and leaving ~6 provenances out each time) and run the models k times, which would have been very computationally heavy and was not feasible in our case given the computation time of some models (almost a week for *M5*). This is why we used the *prediction* $\mathcal{R}^2_{ms}|age$ instead of the $ELPD_{loo}$ to compare the models as it allowed us to calculate the variance explained and predicted by the models conditional on the age effect, and also to compare their predictive ability on new provenances without running the models again.

5.2 In-sample proportion of variance explained conditional on age

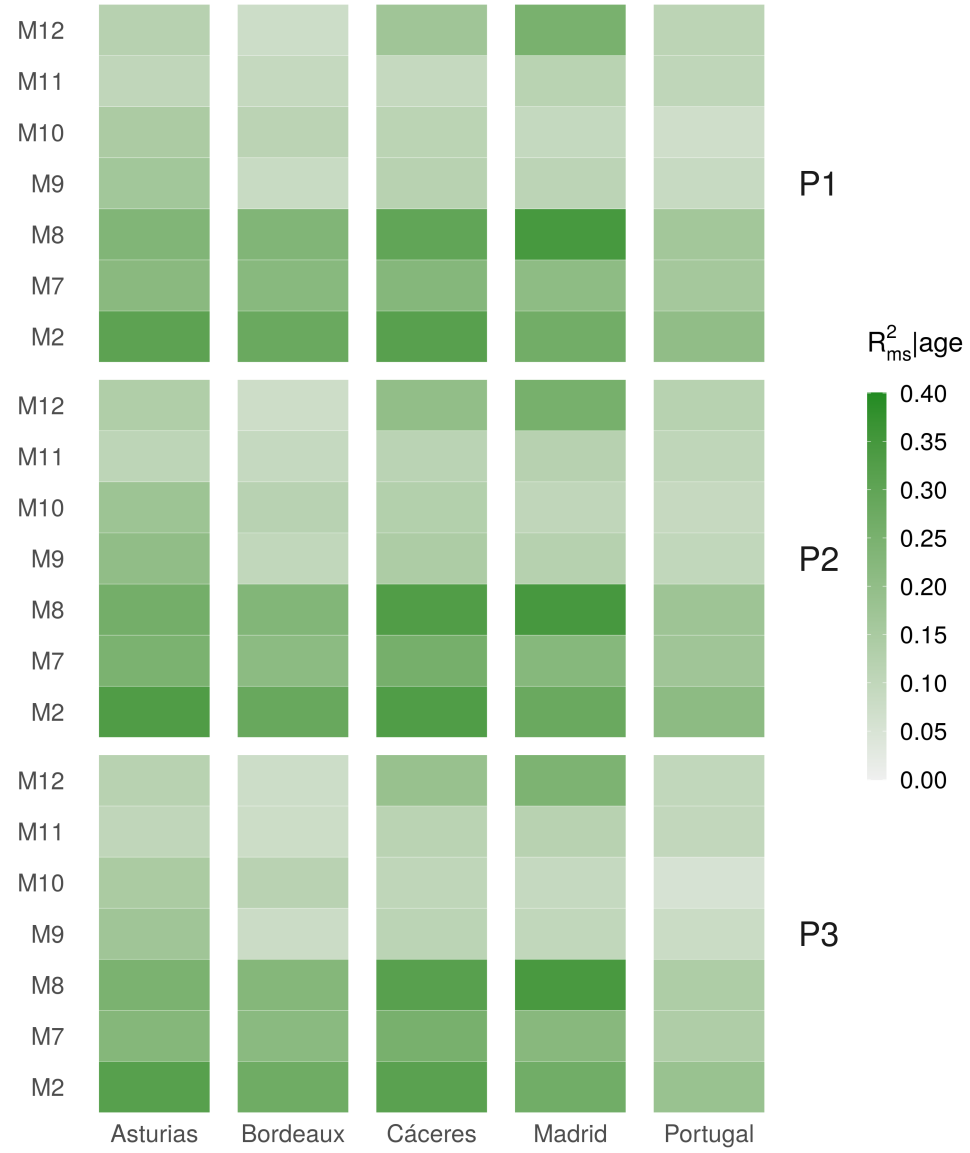


Figure XI.10. In-sample proportion of explained variance conditional on the age effect ($\mathcal{R}^2_{ms}|age$) in the training datasets (data used to fit the models) of the P1, P2 and P3 partitions. In the P1 partition, the training dataset was obtained by randomly sampling 75% of the observations and the test dataset contains the remaining 25% observations. In the P2 partition, the training dataset was obtained by randomly sampling 28 provenances and the test data set contains the remaining 6 provenances. The P3 partition corresponds to a non-random split between a training dataset of 28 provenances and a test dataset containing 6 provenances with at least one provenance from each under-represented gene pool. The exact values of the $\mathcal{R}^2_{ms}|age$ estimates and their associated credible intervals can be found in Tables XI.4 (P1 partition), XI.9 (P2 partition) and XI.12 (P3 partition).

5.3 P1 partition (random split of the observations)

5.3.1 Variance explained and predicted

Models	Explanatory part: training P1				Predictive part: test P1			
	$\mathcal{R}^2_m\|age$	$\mathcal{R}^2_m$	$\mathcal{R}^2_m(fix)$	PE_m	prediction $\mathcal{R}^2_m\|age$	prediction $\mathcal{R}^2_m$	prediction $\mathcal{R}^2_m(fix)$	PE_m
M0	0.463 [0.439-0.487]	0.768 [0.758-0.779]	0.575 [0.564-0.586]	0.267 [0.010-0.84]	0.462 [0.437-0.487]	0.773 [0.762-0.784]	0.584 [0.573-0.596]	0.269 [0.010-0.851]
M1	0.571 [0.548-0.594]	0.815 [0.805-0.825]	0.569 [0.559-0.578]	0.231 [0.008-0.754]	0.559 [0.535-0.583]	0.814 [0.804-0.824]	0.578 [0.568-0.588]	0.236 [0.008-0.769]
M2	0.578 [0.556-0.601]	0.818 [0.808-0.828]	0.568 [0.559-0.578]	0.229 [0.009-0.752]	0.567 [0.543-0.591]	0.817 [0.807-0.827]	0.577 [0.568-0.587]	0.235 [0.008-0.763]
M3	0.574 [0.552-0.597]	0.816 [0.807-0.826]	0.566 [0.506-0.621]	0.231 [0.009-0.747]	0.563 [0.538-0.587]	0.815 [0.805-0.825]	0.575 [0.514-0.631]	0.235 [0.009-0.763]
M4	0.575 [0.553-0.597]	0.817 [0.807-0.826]	0.568 [0.507-0.623]	0.230 [0.008-0.747]	0.563 [0.54-0.587]	0.815 [0.806-0.826]	0.577 [0.516-0.634]	0.235 [0.009-0.763]
M5	0.573 [0.551-0.595]	0.816 [0.806-0.825]	0.568 [0.508-0.622]	0.231 [0.008-0.751]	0.562 [0.539-0.585]	0.815 [0.805-0.825]	0.577 [0.516-0.632]	0.236 [0.009-0.767]
M6	0.575 [0.553-0.598]	0.817 [0.807-0.826]	0.567 [0.504-0.624]	0.230 [0.009-0.748]	0.563 [0.539-0.587]	0.816 [0.805-0.826]	0.576 [0.512-0.635]	0.235 [0.009-0.763]
M7	0.546 [0.523-0.569]	0.804 [0.794-0.814]	0.570 [0.560-0.58]	0.241 [0.009-0.788]	0.529 [0.506-0.553]	0.801 [0.792-0.811]	0.580 [0.570-0.590]	0.243 [0.008-0.790]
M8	0.554 [0.532-0.577]	0.807 [0.798-0.817]	0.570 [0.560-0.580]	0.238 [0.008-0.780]	0.535 [0.512-0.558]	0.804 [0.794-0.814]	0.580 [0.569-0.590]	0.239 [0.009-0.783]
M9	0.502 [0.478-0.525]	0.785 [0.775-0.795]	0.574 [0.564-0.585]	0.253 [0.010-0.822]	0.491 [0.467-0.516]	0.785 [0.775-0.796]	0.584 [0.573-0.595]	0.255 [0.009-0.826]
M10	0.497 [0.473-0.520]	0.783 [0.773-0.793]	0.574 [0.564-0.585]	0.255 [0.009-0.821]	0.485 [0.460-0.509]	0.783 [0.772-0.793]	0.584 [0.573-0.594]	0.257 [0.010-0.828]
M11	0.500 [0.477-0.523]	0.784 [0.774-0.794]	0.571 [0.561-0.581]	0.257 [0.010-0.812]	0.495 [0.471-0.519]	0.787 [0.777-0.797]	0.580 [0.570-0.591]	0.259 [0.009-0.831]
M12	0.506 [0.482-0.530]	0.787 [0.776-0.797]	0.571 [0.560-0.581]	0.255 [0.009-0.804]	0.498 [0.473-0.523]	0.788 [0.778-0.799]	0.580 [0.570-0.591]	0.257 [0.011-0.825]

Table XI.4. Summary of model performance in the P1 partition, in which the training dataset was obtained by randomly sampling 75% of the observations and the test dataset contains the remaining 25% observations. $\mathcal{R}^2_m|age$ and *prediction* $\mathcal{R}^2_m|age$ correspond to the proportion of variance explained and predicted by the models conditional on the age effect in the training (in-sample) and test (out-of-sample) datasets, respectively. $\mathcal{R}^2_m$ and $\mathcal{R}^2_m(fix)$ correspond to the proportion of explained variance in the training dataset by the entire model and by the fixed variables only, respectively. *Prediction* $\mathcal{R}^2_m$ and $\mathcal{R}^2_m(fix)$ correspond to the proportion of variance in the test dataset predicted by the entire model or by the fixed effects, respectively. PE_m corresponds to the mean predictive error (mean of observed minus predicted responses). For $\mathcal{R}^2$ and PE, the mean (over all iterations for $\mathcal{R}^2$ or over all observations for PE) and the 95% credible intervals are given. For more details on the calculation of each index, see section 5.1 of the Supplementary Information. See Table V.1 and main text for description of model the components.

M0 explained less variance (lower $\mathcal{R}^2_m$ and $\mathcal{R}^2_m|age$) than the other models as it did not account for the genetic component of height growth (Table XI.4). The models that accounted for the genetic component with varying intercepts for the provenances explained 81.5% of the variance (Table XI.4). The models combining genomic (PEAs and gene pools) and climatic drivers (*M7* and *M8*) explained more variance (higher $\mathcal{R}^2_m$ and $\mathcal{R}^2_m|age$) than models including separately each driver (*M9* to *M12*). $\mathcal{R}^2_m$ and *prediction* $\mathcal{R}^2_m$ were similar for all models, meaning that the models predicted well new observations (i.e. new observations but from the same sites and same provenances).

5.3.2 Variance partitioning conditional on the age effect

We examine here the partitioning of the variance conditional on the age effect, which provides insight into the genetic and plastic components of deviations from the mean height-growth trajectory (Table XI.5). It was not possible to extract the variance explained by the different drivers of the genetic and plastic components (i.e. for the plastic component, the variance associated with site intercepts and the intercepts associated with climatic similarity between sites) as they were confounded, and thus their explained variance could not be disentangled.

Models	Environment	Genetic	Genetic x Environment	Residuals
M0	45.5% [43.7 - 47.2]	-	-	54.6% [53.2 - 56.0]
M1	46.5% [44.9 - 48.0]	11.4% [10.7 - 12.1]	-	42.9% [41.9 - 43.9]
M2	46.8% [42.5 - 51.2]	11.0% [9.6 - 12.5]	1.5% [-1.0 - 4.3]	42.2% [41.3 - 43.1]

Table XI.5. Partitioning of the variance explained on the data used for sampling conditional on the age effect (i.e. the variance of the deviations from the mean height-growth trajectory).

5.3.3 Bayesian LOO estimate of the expected log predictive density ($ELPD_{loo}$)

Models	$ELPD_{loo}$
M0	-9216 [165]
M1	-6652 [182]
M2	-6487 [182]
M3	-6551 [184]
M4	-6549 [184]
M5	-6591 [184]
M6	-6550 [184]
M7	-7135 [179]
M8	-6926 [180]
M9	-8276 [172]
M10	-8405 [173]
M11	-8334 [169]
M12	-8171 [169]

Table XI.6. $ELPD_{loo}$ of models fitted on the training dataset of the P1 partition. The mean and the standard deviation (in brackets) are given.

	M1	M10	M11	M12	M2	M3	M4	M5	M6	M7	M8	M9
M0	**-2563.62 [71.79]**	**-810.25 [39.04]**	**-881.82 [41.73]**	**-1044.94 [45.56]**	**-2728.44 [73.82]**	**-2665.02 [73.19]**	**-2666.62 [73.22]**	**-2624.87 [73.01]**	**-2665.94 [73.18]**	**-2080.78 [63.39]**	**-2289.82 [66.96]**	**-939.27 [42.76]**
M1		**1753.37 [61.78]**	**1681.8 [61.36]**	**1518.68 [62.97]**	**-164.82 [18.35]**	**-101.4 [14.94]**	**-103 [14.96]**	-61.24 [17.18]	**-102.32 [14.92]**	**482.84 [40.42]**	**273.8 [45.04]**	**1624.35 [59.03]**
M10			-71.57 [58.87]	-234.69 [61.4]	**-1918.19 [64.12]**	**-1854.77 [63.52]**	**-1856.37 [63.39]**	**-1814.61 [63.14]**	**-1855.68 [63.31]**	**-1270.53 [51.27]**	**-1479.56 [55.64]**	**-129.01 [21.51]**
M11				**-163.12 [31.09]**	**-1846.62 [63.88]**	**-1783.2 [62.89]**	**-1784.8 [63.1]**	**-1743.04 [62.64]**	**-1784.12 [63.05]**	**-1198.96 [47.55]**	**-1408 [58.33]**	-57.45 [62.64]
M12					**-1683.5 [63.52]**	**-1620.08 [64.65]**	**-1621.68 [64.92]**	**-1579.93 [64.43]**	**-1621 [64.89]**	**-1035.84 [56.61]**	**-1244.88 [48.18]**	105.67 [64.94]
M2						63.42 [24.36]	61.82 [24.4]	**103.58 [25.8]**	62.51 [24.39]	**647.66 [44.45]**	**438.63 [45.28]**	**1789.18 [61.98]**
M3							-1.6 [1.24]	**40.16 [8.61]**	-0.91 [1.09]	**584.24 [43.19]**	**375.21 [47.81]**	**1725.76 [61.02]**
M4								**41.76 [8.52]**	0.68 [0.78]	**585.84 [43.06]**	**376.8 [47.78]**	**1727.35 [60.75]**
M5									**-41.07 [8.54]**	**544.09 [42.25]**	**335.05 [47]**	**1685.6 [60.47]**
M6										**585.16 [43.02]**	**376.12 [47.77]**	**1726.67 [60.72]**
M7											**-209.04 [35]**	**1141.51 [48.56]**
M8												**1350.55 [52.98]**

Table XI.7. $ELPD_{loo}$ differences among models fitted on the training dataset of the P1 partition. The $ELPD_{loo}$ difference among models corresponds to the $ELPD_{loo}$ of the row model minus the $ELPD_{loo}$ of the column model. Thus, a negative difference in $ELPD_{loo}$ means that the row model has a lower $ELPD_{loo}$ than the column model, and therefore a lower predictive ability on new observations. The standard error of the model differences is indicated between brackets. Two models are considered significantly different when their absolute $ELPD_{loo}$ difference is higher than four times the standard error (in bold).

5.3.4 Site-specific predicted variance conditional on the age effect

Models	Asturias	Bordeaux	Cáceres	Madrid	Portugal
M0	0.064 [-0.027-0.155]	0.002 [-0.02-0.024]	0.017 [0.004-0.043]	0.035 [0.015-0.063]	0.029 [0.007-0.051]
M1	0.341 [0.257-0.43]	0.215 [0.192-0.239]	0.227 [0.196-0.263]	0.204 [0.18-0.231]	0.161 [0.139-0.184]
M2	0.349 [0.264-0.436]	0.259 [0.229-0.289]	0.266 [0.201-0.342]	0.25 [0.208-0.297]	0.171 [0.148-0.194]
M3	0.242 [0.142-0.341]	0.362 [0.324-0.401]	0.227 [0.197-0.262]	0.203 [0.18-0.23]	0.168 [0.141-0.198]
M4	0.244 [0.144-0.346]	0.363 [0.324-0.403]	0.228 [0.197-0.265]	0.203 [0.179-0.23]	0.169 [0.14-0.198]
M5	0.246 [0.15-0.344]	0.357 [0.32-0.397]	0.227 [0.195-0.263]	0.2 [0.176-0.228]	0.166 [0.138-0.195]
M6	0.243 [0.144-0.342]	0.363 [0.324-0.402]	0.228 [0.197-0.265]	0.203 [0.179-0.231]	0.169 [0.14-0.199]
M7	0.251 [0.168-0.336]	0.186 [0.157-0.215]	0.184 [0.112-0.276]	0.158 [0.117-0.205]	0.122 [0.101-0.145]
M8	0.248 [0.163-0.332]	0.2 [0.171-0.23]	0.235 [0.147-0.344]	0.337 [0.269-0.414]	0.137 [0.114-0.159]
M9	0.167 [0.082-0.256]	0.071 [0.05-0.092]	0.093 [0.079-0.116]	0.093 [0.075-0.117]	0.063 [0.042-0.084]
M10	0.124 [0.037-0.211]	0.093 [0.067-0.12]	0.08 [0.029-0.148]	0.086 [0.052-0.126]	0.052 [0.031-0.074]
M11	0.129 [0.045-0.215]	0.074 [0.049-0.098]	0.071 [0.024-0.138]	0.102 [0.066-0.143]	0.087 [0.064-0.109]
M12	0.124 [0.04-0.21]	0.063 [0.038-0.088]	0.128 [0.061-0.219]	0.262 [0.196-0.335]	0.102 [0.079-0.124]

Table XI.8. Summary table of site-specific proportion of predicted variance conditional on the age effect (prediction $\mathcal{R}^2_{ms}|age$) in the test datasets (data not used to fit the models) of the P1 partition. The numbers shown are the mean and the 95% credible intervals.

5.4 P2 partition (random split of the provenances)

5.4.1 Variance explained and predicted

Models	Explanatory part: training P2				Predictive part: test P2			
	$\mathcal{R}^2_m\|age$	$\mathcal{R}^2_m$	$\mathcal{R}^2_m(fix)$	PE_m	prediction $\mathcal{R}^2_m\|age$	prediction $\mathcal{R}^2_m$	prediction $\mathcal{R}^2_m(fix)$	PE_m
M0	0.454 [0.431-0.478]	0.766 [0.756-0.776]	0.577 [0.567-0.587]	0.268 [0.011-0.839]	0.425 [0.403-0.447]	0.745 [0.736-0.755]	0.563 [0.553-0.573]	0.263 [0.010-0.881]
M1	0.571 [0.550-0.592]	0.816 [0.807-0.825]	0.571 [0.562-0.581]	0.230 [0.008-0.746]	0.428 [0.409-0.448]	0.747 [0.738-0.755]	0.557 [0.548-0.566]	0.264 [0.010-0.877]
M2	0.578 [0.558-0.599]	0.819 [0.810-0.828]	0.571 [0.562-0.58]	0.228 [0.009-0.743]	0.426 [0.374-0.482]	0.746 [0.723-0.770]	0.556 [0.548-0.565]	0.264 [0.010-0.870]
M7	0.546 [0.524-0.567]	0.805 [0.796-0.814]	0.574 [0.564-0.583]	0.239 [0.008-0.779]	0.502 [0.481-0.522]	0.779 [0.770-0.788]	0.559 [0.550-0.568]	0.247 [0.008-0.861]
M8	0.554 [0.532-0.575]	0.809 [0.799-0.818]	0.574 [0.564-0.583]	0.237 [0.009-0.771]	0.540 [0.518-0.562]	0.796 [0.786-0.806]	0.559 [0.550-0.569]	0.246 [0.008-0.844]
M9	0.505 [0.483-0.528]	0.788 [0.778-0.798]	0.578 [0.568-0.587]	0.251 [0.009-0.805]	0.479 [0.458-0.501]	0.769 [0.760-0.779]	0.563 [0.554-0.573]	0.264 [0.009-0.906]
M10	0.499 [0.476-0.522]	0.785 [0.775-0.795]	0.577 [0.567-0.587]	0.252 [0.009-0.808]	0.492 [0.470-0.514]	0.775 [0.765-0.785]	0.563 [0.553-0.573]	0.265 [0.009-0.905]
M11	0.493 [0.470-0.516]	0.783 [0.773-0.793]	0.573 [0.563-0.583]	0.257 [0.009-0.813]	0.471 [0.449-0.493]	0.766 [0.756-0.776]	0.559 [0.549-0.569]	0.258 [0.010-0.856]
M12	0.503 [0.480-0.526]	0.787 [0.777-0.797]	0.573 [0.563-0.583]	0.254 [0.010-0.802]	0.522 [0.499-0.545]	0.788 [0.778-0.798]	0.559 [0.549-0.568]	0.263 [0.010-0.833]

Table XI.9. Summary of model performance in the P2 partition, in which the training dataset was obtained by randomly sampling 28 provenances and the test data set contains the remaining 6 provenances. $\mathcal{R}^2_m|age$ and prediction $\mathcal{R}^2_m|age$ correspond to the proportion of variance explained and predicted by the models conditional on the age effect in the training (in-sample) and test (out-of-sample) datasets, respectively. $\mathcal{R}^2_m$ and $\mathcal{R}^2_m(fix)$ correspond to the proportion of explained variance in the training dataset by the entire model and by the fixed variables only, respectively. Prediction $\mathcal{R}^2_m$ and $\mathcal{R}^2_m(fix)$ correspond to the proportion of variance in the test dataset predicted by the entire model or by the fixed effects, respectively. PE_m corresponds to the mean predictive error (mean of observed minus predicted responses). For $\mathcal{R}^2$ and PE, the mean (over all iterations for $\mathcal{R}^2$ or over all observations for PE) and the 95% credible intervals are given. For more details on the calculation of each index, see section 5 of the Supplementary Information. See Table V.1 and main text for description of model the components.

5.4.2 Bayesian LOO estimate of the expected log predictive density ($ELPD_{loo}$)

	M1	M10	M11	M12	M2	M7	M8	M9
M0	**-3097.43 [80.4]**	**-1180.71 [47.7]**	**-1001.98 [45.13]**	**-1293 [50.8]**	**-3307.11 [82.58]**	**-2503.89 [70.49]**	**-2740.39 [74.19]**	**-1339.65 [50.55]**
M1		**1916.72 [66.85]**	**2095.45 [68.57]**	**1804.43 [69.11]**	**-209.68 [20.62]**	**593.55 [43.39]**	**357.04 [49.06]**	**1757.78 [63.01]**
M10			178.73 [65.95]	-112.29 [67.87]	**-2126.4 [69.47]**	**-1323.17 [54.04]**	**-1559.68 [58.5]**	**-158.94 [25.24]**
M11				**-291.02 [34.88]**	b	**-1501.91 [53.72]**	**-1738.41 [64.64]**	**-337.67 [69]**
M12					**-2014.11 [69.7]**	**-1210.89 [60.93]**	**-1447.39 [52.32]**	-46.65 [70.45]
M2						**803.22 [48.2]**	**566.72 [49.33]**	**1967.46 [66.09]**
M7							**-236.51 [35.87]**	**1164.23 [50.25]**
M8								**1400.74 [54.56]**

Table XI.10. $ELPD_{loo}$ differences among models fitted on the training dataset of the P2 partition. The $ELPD_{loo}$ difference among models corresponds to the $ELPD_{loo}$ of the row model minus the $ELPD_{loo}$ of the column model. Thus, a negative difference in $ELPD_{loo}$ means that the row model has a lower $ELPD_{loo}$ than the column model, and therefore a lower predictive ability on new observations. The standard error of the model differences is indicated between brackets. Two models are considered significantly different when their absolute $ELPD_{loo}$ difference is higher than four times the standard error (in bold). Detail about interpretation of $ELPD_{loo}$ differences can be found in section 5.3.3.

5.4.3 Site-specific predicted variance conditional on the age effect

Models	Asturias	Bordeaux	Cáceres	Madrid	Portugal
M0	0.052 [-0.024-0.128]	-0.001 [-0.022-0.021]	0.035 [0.007-0.088]	0.042 [0.017-0.076]	0.033 [0.013-0.053]
M2	0.005 [-0.061-0.074]	0.004 [-0.016-0.024]	0.035 [0.008-0.086]	0.034 [0.015-0.062]	0.017 [0.000-0.035]
M1	0.011 [-0.056-0.079]	0.004 [-0.015-0.023]	0.035 [0.008-0.082]	0.036 [0.016-0.063]	0.019 [0.002-0.036]
M7	0.149 [0.079-0.219]	0.140 [0.114-0.167]	0.247 [0.150-0.371]	0.229 [0.170-0.299]	0.111 [0.092-0.131]
M8	0.124 [0.054-0.196]	0.166 [0.137-0.195]	0.290 [0.182-0.422]	0.356 [0.281-0.439]	0.101 [0.082-0.121]
M9	0.166 [0.093-0.240]	0.101 [0.078-0.123]	0.162 [0.132-0.212]	0.139 [0.115-0.169]	0.076 [0.056-0.095]
M10	0.178 [0.100-0.252]	0.152 [0.120-0.185]	0.204 [0.090-0.358]	0.131 [0.085-0.187]	0.081 [0.061-0.101]
M11	0.100 [0.025-0.177]	0.100 [0.072-0.129]	0.189 [0.081-0.330]	0.209 [0.138-0.292]	0.103 [0.082-0.125]
M12	0.141 [0.065-0.217]	0.070 [0.044-0.095]	0.274 [0.150-0.434]	0.329 [0.253-0.413]	0.131 [0.109-0.152]

Table XI.11. Summary table of site-specific proportion of predicted variance conditional on the age effect (prediction $\mathcal{R}^2_{ms}|age$) in the test datasets (data not used to fit the models) of the P2 partition. The numbers shown are the mean and the 95% credible intervals.

5.5 P3 partition (non-random split of the provenances)

5.5.1 Variance explained and predicted

Models	**Explanatory part**: training P3				**Predictive part**: test P3			
	$\mathcal{R}^2_m\|age$	$\mathcal{R}^2_m$	$\mathcal{R}^2_m(fix)$	PE_m	prediction $\mathcal{R}^2_m\|age$	prediction $\mathcal{R}^2_m$	prediction $\mathcal{R}^2_m(fix)$	PE_m
M0	0.448 [0.425-0.472]	0.765 [0.755-0.775]	0.58 [0.569-0.591]	0.266 [0.01-0.842]	0.382 [0.362-0.402]	0.716 [0.706-0.725]	0.545 [0.535-0.555]	0.27 [0.009-0.853]
M1	0.557 [0.535-0.579]	0.811 [0.802-0.821]	0.575 [0.565-0.584]	0.232 [0.008-0.759]	0.383 [0.365-0.402]	0.716 [0.708-0.725]	0.54 [0.531-0.549]	0.271 [0.011-0.841]
M2	0.563 [0.54-0.584]	0.814 [0.804-0.823]	0.574 [0.564-0.584]	0.23 [0.008-0.753]	0.388 [0.343-0.438]	0.719 [0.698-0.741]	0.539 [0.531-0.549]	0.271 [0.011-0.842]
M7	0.53 [0.507-0.553]	0.8 [0.79-0.81]	0.577 [0.566-0.587]	0.241 [0.009-0.792]	0.439 [0.415-0.464]	0.742 [0.731-0.753]	0.542 [0.532-0.551]	0.242 [0.008-0.81]
M8	0.537 [0.514-0.56]	0.803 [0.793-0.813]	0.577 [0.567-0.587]	0.24 [0.009-0.784]	0.504 [0.478-0.532]	0.772 [0.76-0.785]	0.542 [0.533-0.551]	0.237 [0.009-0.78]
M9	0.487 [0.464-0.511]	0.782 [0.772-0.792]	0.58 [0.57-0.591]	0.253 [0.01-0.825]	0.406 [0.385-0.428]	0.727 [0.717-0.737]	0.546 [0.536-0.555]	0.286 [0.011-0.905]
M10	0.48 [0.457-0.503]	0.779 [0.769-0.789]	0.581 [0.571-0.591]	0.256 [0.01-0.827]	0.394 [0.373-0.414]	0.721 [0.712-0.731]	0.546 [0.536-0.555]	0.257 [0.01-0.807]
M11	0.484 [0.46-0.508]	0.781 [0.77-0.791]	0.576 [0.566-0.587]	0.257 [0.009-0.818]	0.404 [0.384-0.425]	0.726 [0.717-0.735]	0.542 [0.532-0.551]	0.258 [0.01-0.825]
M12	0.49 [0.466-0.513]	0.783 [0.773-0.793]	0.576 [0.566-0.587]	0.255 [0.01-0.809]	0.442 [0.421-0.462]	0.743 [0.734-0.753]	0.542 [0.532-0.552]	0.255 [0.009-0.819]

Table XI.12. Summary of model performance in the P3 partition, in which the training dataset contains 28 provenances and the test dataset contains the remaining 6 provenances. $\mathcal{R}^2_m|age$ and prediction $\mathcal{R}^2_m|age$ correspond to the proportion of variance explained and predicted by the models conditional on the age effect in the training (in-sample) and test (out-of-sample) datasets, respectively. $\mathcal{R}^2_m$ and $\mathcal{R}^2_m(fix)$ correspond to the proportion of explained variance in the training dataset by the entire model and by the fixed variables only, respectively. Prediction $\mathcal{R}^2_m$ and $\mathcal{R}^2_m(fix)$ correspond to the proportion of variance in the test dataset predicted by the entire model or by the fixed effects, respectively. PE_m corresponds to the mean predictive error (mean of observed minus predicted responses). For $\mathcal{R}^2$ and PE, the mean (over all iterations for $\mathcal{R}^2$ or over all observations for PE) and the 95% credible intervals are given. For more details on the calculation of each index, see section 5 of the Supplementary Information. See Table V.1 and main text for description of model the components.

5.5.2 Site-specific predicted variance conditional on the age effect

Models	Asturias	Bordeaux	Cáceres	Madrid	Portugal
M0	0.045 [-0.028-0.117]	0.000 [-0.024-0.025]	0.021 [0.005-0.048]	0.033 [0.014-0.059]	0.029 [0.010-0.048]
M1	0.010 [-0.054-0.074]	0.004 [-0.017-0.025]	0.018 [0.004-0.044]	0.024 [0.010-0.044]	0.016 [-0.001-0.034]
M2	0.006 [-0.057-0.071]	0.004 [-0.017-0.026]	0.019 [0.004-0.046]	0.023 [0.009-0.043]	0.015 [-0.002-0.032]
M7	0.188 [0.110-0.266]	0.246 [0.193-0.305]	0.271 [0.160-0.409]	0.205 [0.145-0.271]	0.124 [0.097-0.152]
M8	0.277 [0.197-0.360]	0.310 [0.254-0.371]	0.250 [0.152-0.369]	0.223 [0.176-0.276]	0.164 [0.134-0.197]
M9	0.107 [0.035-0.180]	0.061 [0.032-0.092]	0.105 [0.074-0.148]	0.096 [0.068-0.129]	0.062 [0.041-0.084]
M10	0.125 [0.056-0.195]	0.146 [0.113-0.18]	0.111 [0.040-0.209]	0.075 [0.043-0.114]	0.053 [0.034-0.073]
M11	0.061 [-0.010-0.130]	0.066 [0.04-0.092]	0.080 [0.033-0.146]	0.096 [0.063-0.134]	0.062 [0.042-0.081]
M12	0.061 [-0.007-0.132]	0.095 [0.066-0.126]	0.061 [0.032-0.102]	0.121 [0.091-0.154]	0.060 [0.041-0.080]

Table XI.13. Summary table of site-specific proportion of predicted variance conditional on the age effect (prediction $\mathcal{R}^2_{ms}|age$) in the test datasets (data not used to fit the models) of the P3 partition. The numbers shown are the mean and the 95% credible intervals.

6 Posterior distributions and parameter interpretation

6.1 P1 partition (random split of the observations)

6.1.1 Baseline *models M0, M1* and *M2*: separating the genetic and plastic components

In the baseline model *M0*, only the plastic component is included (via the site and block intercepts). The genetic component is not considered (no intercepts for the provenances and the genotypes). *M0* was performed to compare the gain in explanatory and predictive power of models that account for the genetic component, compared to the model *M0* that does not.

MODEL M0

Parameter	Median	SD	InfCI	SupCI
σ^2_S	0.105	0.370	0.027	0.979
σ^2_B	0.002	0.001	0.001	0.004
σ^2	0.123	0.001	0.121	0.125
β_0	6.273	0.176	5.942	6.664
β_{age}	0.601	0.003	0.596	0.607
β_{age2}	-0.153	0.003	-0.159	-0.147

Table XI.14. Parameter estimates of the varying-intercept variances (σ^2_S, σ^2_B), the global variance σ^2, the global intercept β_0 and the slopes associated with the age effect (β_{age} and β_{age2}) in *M0*. SD corresponds to the standard deviation and InfCI and SupCI correspond to the lower and upper bounds of the 0.95 credible interval.

MODEL M1

In *M1*, the plastic component was mainly attributed to the variance σ^2_S between sites (median of 0.108), while the variance σ^2_B between blocks was almost null (median of 0.002) (Table XI.15). Some sites showed heights deviating strongly from the global mean: Madrid, where trees grew the least (median of -0.376), and Asturias where they grew particularly well (median of 0.272) (Fig. V.3 & Table XI.16). The genetic component was equally attributed to the variance between provenances σ^2_P and genotypes σ^2_G, with a median of 0.013 and 0.012, respectively (Table XI.15).

Parameter	Median	SD	InfCI	SupCI
σ^2_P	0.013	0.004	0.008	0.023
σ^2_G	0.012	0.001	0.010	0.013
σ^2_S	0.108	0.444	0.028	1.046
σ^2_B	0.002	0.001	0.001	0.003
σ^2	0.098	0.001	0.096	0.100
β_0	6.247	0.193	5.842	6.643
β_{age}	0.598	0.003	0.592	0.603
β_{age2}	-0.150	0.003	-0.155	-0.145

Table XI.15. Parameter estimates of the varying-intercept variances (σ^2_S, σ^2_B, σ^2_P and σ^2_G), the global variance σ^2, the global intercept β_0 and the slopes associated with the age effect (β_{age} and β_{age2}) in *M1*.

Parameter	Median	SD	InfCI	SupCI
$S_{Asturias}$	0.292	0.193	-0.103	0.690
$S_{Bordeaux}$	0.151	0.193	-0.254	0.555
$S_{Caceres}$	0.041	0.194	-0.366	0.443
S_{Madrid}	-0.376	0.193	-0.780	0.019
$S_{Portugal}$	-0.110	0.193	-0.513	0.288

Table XI.16. Parameter estimates of the site intercepts S_s in *M1*. SD corresponds to the standard deviation and InfCI and SupCI correspond to the lower and upper bounds of the 0.95 credible interval.

Parameter	Median	SD	InfCI	SupCI
P_{ALT}	0.089	0.041	0.009	0.168
P_{ARM}	0.121	0.043	0.039	0.205
P_{ARN}	-0.093	0.032	-0.158	-0.031
P_{BAY}	-0.135	0.033	-0.196	-0.072
P_{BON}	-0.040	0.042	-0.121	0.041
P_{CAD}	0.061	0.040	-0.016	0.142
P_{CAR}	-0.162	0.047	-0.254	-0.071
P_{CAS}	0.054	0.041	-0.026	0.132
P_{CEN}	0.011	0.042	-0.073	0.093
P_{COC}	-0.154	0.034	-0.222	-0.088
P_{COM}	0.076	0.054	-0.033	0.181
P_{CUE}	-0.171	0.030	-0.233	-0.113
P_{HOU}	0.162	0.030	0.102	0.220
P_{LAM}	0.014	0.040	-0.064	0.093
P_{LEI}	0.038	0.032	-0.023	0.101
P_{MAD}	0.021	0.083	-0.146	0.181
P_{MIM}	0.005	0.034	-0.061	0.071
P_{OLB}	0.047	0.031	-0.013	0.108
P_{OLO}	0.141	0.032	0.078	0.204
P_{ORI}	-0.121	0.030	-0.180	-0.062
P_{PET}	0.123	0.030	0.065	0.182
P_{PIA}	0.089	0.034	0.023	0.154
P_{PIE}	-0.078	0.042	-0.163	0.002
P_{PLE}	0.000	0.032	-0.062	0.063
P_{PUE}	0.073	0.044	-0.011	0.161
P_{QUA}	0.020	0.034	-0.046	0.089
P_{SAC}	-0.034	0.042	-0.115	0.047
P_{SAL}	-0.171	0.036	-0.244	-0.100
P_{SEG}	0.007	0.031	-0.052	0.070
P_{SIE}	0.002	0.042	-0.078	0.086
P_{STJ}	0.172	0.029	0.117	0.231
P_{TAM}	-0.226	0.036	-0.297	-0.153
P_{VAL}	-0.047	0.038	-0.120	0.027
P_{VER}	0.106	0.030	0.047	0.165

Table XI.17. Parameter estimates of the provenance intercepts P_p in *M1*. SD corresponds to the standard deviation and InfCI and SupCI correspond to the lower and upper bounds of the 0.95 credible interval.

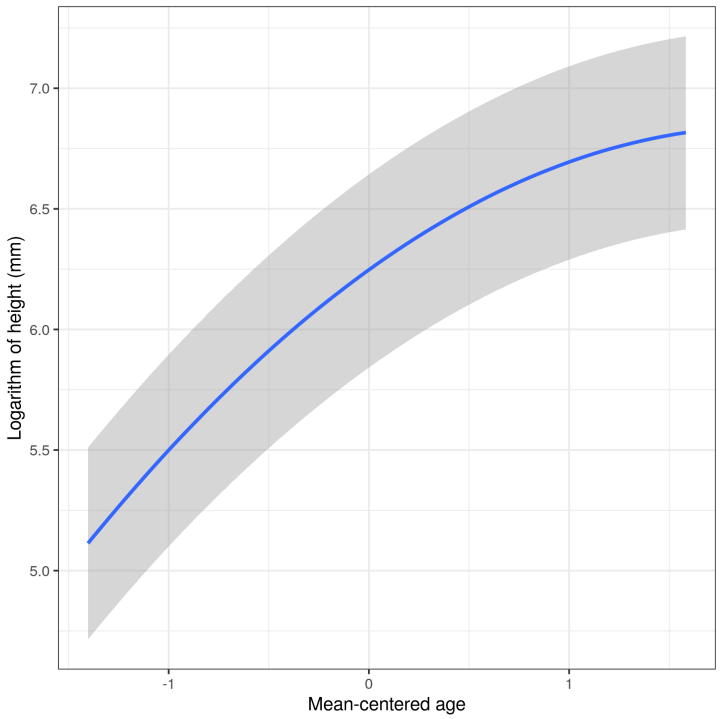


Figure XI.11. Relationship between mean-centered age and height on the log scale in *M1*.

MODEL M2

In *M2*, parameter estimates were similar to *M1*. The variance σ^2_{Inter} of the provenance-by-site interaction was much smaller than the variances among provenances or genotypes (median of 0.004 and 95% CIs: 0.003-0.006; Table XI.18).

Parameter	Median	SD	InfCI	SupCI
σ^2_P	0.011	0.004	0.007	0.022
σ^2_G	0.012	0.001	0.010	0.014
σ^2_{Inter}	0.004	0.001	0.003	0.006
σ^2_S	0.113	1.525	0.030	1.423
σ^2_B	0.002	0.001	0.001	0.003
σ^2	0.096	0.001	0.095	0.098
β_0	6.231	0.285	5.753	6.670
β_{age}	0.597	0.003	0.592	0.602
β_{age2}	-0.150	0.002	-0.155	-0.145

Table XI.18. Parameter estimates of the varying-intercept variances (σ^2_S, σ^2_B, σ^2_P, σ^2_G and σ^2_{Inter}), the global variance σ, the global intercept β_0 and the slopes associated with the age effect (β_{age} and β_{age2}). SD corresponds to the standard deviation and InfCI and SupCI correspond to the lower and upper bounds of the 0.95 credible interval.

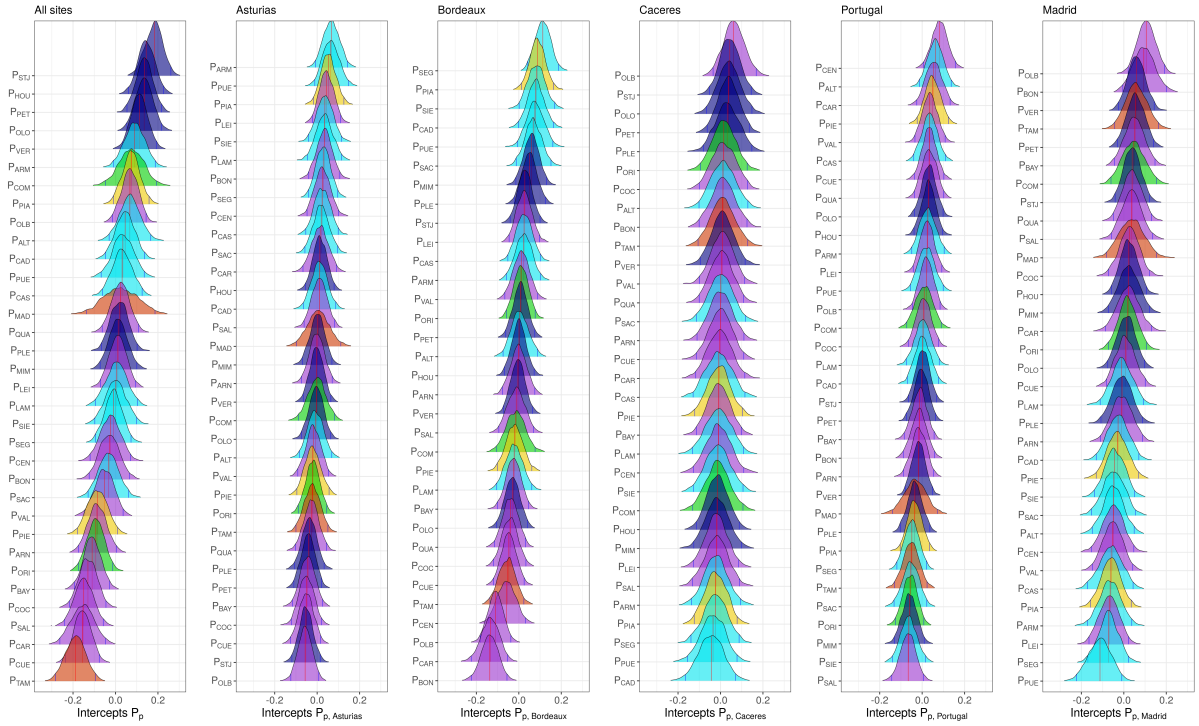


Figure XI.12. Posterior distributions of the provenance intercepts across all sites (P_p) and the site-specific intercepts of each provenance ($P_{p,s}$) from *model M2*. The colors correspond to the gene pool from which each population mainly belongs: gene pool in Northern Africa in orange, gene pool in Corsica in yellow, gene pool in Central Spain in purple, gene pool in the French Atlantic region in marine blue, gene pool in the Iberian Atlantic region in sky blue and gene pool in South-Eastern Spain in green.

Provenance varying intercepts are shown in Fig. V.3 (global intercepts across all sites from *M1*) and Fig. XI.12 (site-specific intercepts from *M2*). Most provenances from the French Atlantic gene pool (e.g. STJ, HOU, PET, OLO and VER) were taller on average than other provenances in all sites. Most provenances from the Iberian Atlantic gene pool (e.g. ARM or PUE) were among the tallest in Asturias and Bordeaux but the shortest in Madrid and Cáceres. In contrast, most provenances from the Central Spain gene pool were among the shortest, especially in Asturias and Bordeaux. The two provenances from the Corsican gene pool showed highly contrasted intercepts. The provenance Pinia (PIA) was taller than other provenances on average, especially in Asturias and Bordeaux. In contrast, the second Corsican provenance (Pineta, PIE), was shorter than other provenances on average but grew particularly well in Portugal. Provenances belonging to the south-eastern Spain gene pool (i.e. ORI and COM) also showed contrasted intercepts (trees from ORI were shorter than those from COM on average), but did not show strong differences between sites. Lastly, regarding provenances belonging to the northern African gene pool, the TAM provenance had the lowest growth in our dataset and was more likely to be taller in Cáceres and Madrid (the harsh Mediterranean sites) than in the other sites.

6.1.2 *Models M3 and M3bis*: potential drivers underlying the plastic component of height-growth variation

The plastic component of height-growth in *M3* (but also in all subsequent models until *M6*; Table V.1) was only marginally associated with the climatic similarity among sites ($\sigma^2_{cs_{is}}$ with a median of 0.023 in *M3*), compared to the variance associated with site intercepts (σ^2_S with a median of 0.126 in *M3*; Table XI.19). However, estimates of the intercepts (S_s and cs_{is}) and variances (σ^2_S and $\sigma^2_{cs_{is}}$) were uncertain (Tables XI.19 to XI.21), and the median and the credible interval of σ^2_S increased from *M1* to *M3*, suggesting that *M3* may hardly separate between σ^2_S and $\sigma^2_{cs_{is}}$ (Tables XI.15 & XI.19). To check this, we ran a supplementary model identical to *M3* but without the site intercepts S_s (see *model M3bis* below). In this model, the variance related to the climatic similarity among sites was nearly equal and as uncertain as in *M3*, suggesting that our variance estimation of the plastic component in *M3* was robust (Table XI.23). However, the posterior distributions of the intercepts in *M3bis* were different from *M3*: height growth was positively associated with the climatic conditions in Bordeaux and Asturias, and negatively with the climatic conditions in Madrid and Cáceres, the two Mediterranean sites, and to a lesser extent in Portugal (Table XI.24).

MODEL M3

Parameter	Median	SD	InfCI	SupCI
σ^2_G	0.012	0.001	0.010	0.013
σ^2_P	0.013	0.004	0.008	0.023
σ^2_S	0.126	1.694	0.027	1.895
$\sigma^2_{cs_{is}}$	0.023	0.314	0.003	0.274
σ^2_B	0.002	0.001	0.001	0.003
σ^2	0.097	0.001	0.096	0.099
β_0	6.295	0.315	5.825	6.808
β_{age}	0.604	0.013	0.577	0.625
β_{age2}	-0.186	0.016	-0.223	-0.157

Table XI.19. Parameter estimates of the varying-intercept variances (σ^2_S, σ^2_B, σ^2_P, σ^2_G and $\sigma^2_{cs_{is}}$), the global variance σ^2, the global intercept β_0 and the slopes associated with the age effect (β_{age} and β_{age2}). SD corresponds to the standard deviation and InfCI and SupCI correspond to the lower and upper bounds of the 0.95 credible interval.

Parameter	Median	SD	InfCI	SupCI
$S_{Asturias}$	0.283	0.326	-0.248	0.809
$S_{Bordeaux}$	0.164	0.341	-0.382	0.726
$S_{Caceres}$	0.086	0.341	-0.520	0.596
S_{Madrid}	-0.349	0.321	-0.902	0.092
$S_{Portugal}$	-0.162	0.320	-0.717	0.291

Table XI.20. Parameter estimates of the site intercepts S_s in *model M3*. SD corresponds to the standard deviation and InfCI and SupCI correspond to the lower and upper bounds of the 0.95 credible interval.

MODEL *M3bis*

Parameter	Median	SD	InfCI	SupCI
$cs_{1,Asturias}$	0.012	0.076	-0.139	0.168
$cs_{2,Asturias}$	-0.027	0.085	-0.198	0.141
$cs_{3,Asturias}$	0.010	0.040	-0.047	0.114
$cs_{1,Bordeaux}$	-0.109	0.126	-0.353	0.160
$cs_{2,Bordeaux}$	0.043	0.084	-0.096	0.246
$cs_{1,Caceres}$	-0.001	0.145	-0.347	0.242
$cs_{1,Madrid}$	-0.035	0.064	-0.166	0.095
$cs_{1,Portugal}$	0.028	0.055	-0.089	0.132
$cs_{2,Portugal}$	0.065	0.051	-0.040	0.165
$cs_{3,Portugal}$	0.033	0.048	-0.072	0.123
$cs_{4,Portugal}$	-0.019	0.061	-0.142	0.103

Table XI.21. Parameter estimates of the cs_{is} intercepts related to climatic similarity between test sites during the year preceding the measurements. As the saplings were measured 2 to 4 times in Asturias, Bordeaux and Portugal, the numbers 1 to 4 correspond to each measurement, in the temporal order in which they were done. For example, the intercept $cs_{1,Asturias}$ corresponds to the first measurement taken in Asturias, when the saplings were 10 month old. See the table XI.1 for the sapling age at each measurement. SD corresponds to the standard deviation and InfCI and SupCI correspond to the lower and upper bounds of the 0.95 credible interval.

Models	**Explanatory part**: training P1				**Predictive part**: test P1			
	$\mathcal{R}^2_m\|age$	$\mathcal{R}^2_m$	$\mathcal{R}^2_m(fix)$	$\mathbf{PE}_m$	prediction $\mathcal{R}^2_m\|age$	prediction $\mathcal{R}^2_m$	prediction $\mathcal{R}^2_m(fix)$	$\mathbf{PE}_m$
M3	0.574 [0.552-0.597]	0.816 [0.807-0.826]	0.566 [0.506-0.621]	0.231 [0.009-0.747]	0.563 [0.538-0.587]	0.815 [0.805-0.825]	0.575 [0.514-0.631]	0.235 [0.009-0.763]
M13	0.575 [0.553-0.597]	0.816 [0.807-0.826]	0.634 [0.594-0.673]	0.23 [0.009-0.746]	0.563 [0.54-0.587]	0.816 [0.806-0.825]	0.644 [0.604-0.685]	0.235 [0.009-0.767]

Table XI.22. Comparing the performance of *M3* and *M3bis* in the P1 partition, in which the training dataset was obtained by randomly sampling 75% of the observations and the test dataset contains the remaining 25% observations. $\mathcal{R}^2_m|age$ and *prediction* $\mathcal{R}^2_m|age$ correspond to the proportion of variance explained and predicted by the models conditional on the age effect in the training (in-sample) and test (out-of-sample) datasets, respectively. $\mathcal{R}^2_m$ and $\mathcal{R}^2_m(fix)$ correspond to the proportion of explained variance in the training dataset by the entire model and by the fixed variables only, respectively. *Prediction* $\mathcal{R}^2_m$ and $\mathcal{R}^2_m(fix)$ correspond to the proportion of variance in the test dataset predicted by the entire model or by the fixed effects, respectively. PE_m corresponds to the mean predictive error (mean of observed minus predicted responses). For $\mathcal{R}^2$ and PE, the mean (over all iterations for $\mathcal{R}^2$ or over all observations for PE) and the 95% credible intervals are given. For more details on the calculation of each index, see section 5 of the Supplementary Information.

Parameter	Median	SD	InfCI	SupCI
σ^2_B	0.029	0.009	0.017	0.053
σ^2_P	0.013	0.004	0.008	0.022
σ^2_G	0.012	0.001	0.010	0.013
$\sigma^2_{cs_{is}}$	0.025	0.083	0.006	0.229
σ^2	0.097	0.001	0.096	0.099
β_0	6.284	0.036	6.213	6.356
β_{age}	0.630	0.008	0.614	0.645
β_{age2}	-0.153	0.010	-0.172	-0.134

Table XI.23. Parameter estimates of the varying-intercept variances (σ^2_B, σ^2_P, σ^2_G and $\sigma^2_{cs_{is}}$), the global variance σ^2, the global intercept β_0 and the slopes associated with the age effect (β_{age} and β_{age2}). SD corresponds to the standard deviation and InfCI and SupCI correspond to the lower and upper bounds of the 0.95 credible interval.

Parameter	Median	SD	InfCI	SupCI
$cs_{1,Asturias}$	0.133	0.026	0.080	0.180
$cs_{2,Asturias}$	0.117	0.030	0.054	0.172
$cs_{3,Asturias}$	0.026	0.020	-0.009	0.069
$cs_{1,Bordeaux}$	0.085	0.042	-0.003	0.162
$cs_{2,Bordeaux}$	0.134	0.031	0.076	0.195
$cs_{1,Caceres}$	-0.158	0.053	-0.264	-0.057
$cs_{1,Madrid}$	-0.140	0.030	-0.198	-0.082
$cs_{1,Portugal}$	-0.047	0.022	-0.088	-0.004
$cs_{2,Portugal}$	-0.007	0.023	-0.051	0.041
$cs_{3,Portugal}$	-0.028	0.020	-0.066	0.011
$cs_{4,Portugal}$	-0.114	0.028	-0.168	-0.056

Table XI.24. Parameter estimates of the cs_{is} intercepts related to climatic similarity between test sites during the year preceding the measurements. SD corresponds to the standard deviation and InfCI and SupCI correspond to the lower and upper bounds of the 0.95 credible interval.

6.1.3 *Models M4 to M6*: potential drivers underlying the genetic component of height-growth variation

In *M4*, the variance $\sigma^2_{g_j}$ between gene pools was as important as the variance σ^2_G between genotypes, but had a very wide posterior distribution (Table XI.25). From *M3* to *M4*, adding the gene pool intercepts g_j resulted in a decreased variance between provenances (from a median of 0.013 to a median of 0.006, Tables XI.19 & XI.25). This indicates redundant information between gene pools and provenances. Genotypes belonging to the French Atlantic gene pool, and to a lesser extent to the Iberian Atlantic gene pool, were on average taller than genotypes belonging to the northern Africa gene pool, and to a lesser extent to the Central Spain gene pool (Fig. XI.14).

MODEL M4

Parameter	Median	SD	InfCI	SupCI
$\sigma^2_{g_j}$	0.013	0.044	0.002	0.093
σ^2_P	0.006	0.003	0.003	0.013
σ^2_G	0.012	0.001	0.010	0.013
σ^2_S	0.127	1.365	0.026	2.125
$\sigma^2_{cs_{is}}$	0.022	0.187	0.003	0.267
σ^2_B	0.002	0.001	0.001	0.003
σ^2	0.097	0.001	0.095	0.099
β_0	6.280	0.269	5.704	6.811
β_{age}	0.605	0.012	0.577	0.626
β_{age2}	-0.185	0.016	-0.221	-0.156

Table XI.25. Parameter estimates of the varying-intercept variances (σ^2_S, σ^2_B, σ^2_P, σ^2_G, $\sigma^2_{cs_{is}}$ and $\sigma^2_{g_j}$), the global variance σ^2, the global intercept β_0 and the slopes associated with the age effect (β_{age} and β_{age2}). SD corresponds to the standard deviation and InfCI and SupCI correspond to the lower and upper bounds of the 0.95 credible interval.

MODEL M5: gene pool-specific heritabilities, the telltale of distinct adaptive histories

Heritability variation across populations or gene pools can inform on differences in the drivers underlying their adaptive histories, such as evolutionary mechanisms (e.g. capacity of dispersion, selection strength, migration) and local environmental constraints (e.g. environmental heterogeneity). Using CLONAPIN data from all sites except Bordeaux, Rodríguez-Quilón et al. (2016) showed that populations from the Mediterranean gene pools had a higher heritability than those from the Atlantic gene pools. Nonetheless, their heritability estimates did not consider population admixture, a notable feature in maritime pine (Fig. V.1). In our study, we applied the recent methodology proposed by Muff et al. (2019) to calculate gene pool-specific total genetic variance and broad-sense heritabilities in a single model that accounts for population admixture. We showed that the total genetic variance of the Iberian Atlantic gene pool had a probability higher than 0.95 of being lower than that of the Corsican and south-eastern Spain gene pools, and a probability higher than 0.90 of being lower than that of the Central Spain gene pool (Table XI.27; see also Fig. XI.13), despite their geographical proximity. The total genetic variance of the French Atlantic gene pool had a probability higher than 0.90 of being lower than that of the Corsican and south-eastern Spain gene pools (Table XI.27; see also Fig. XI.13). However, this should be taken with caution as the total variance of the gene pools from south-eastern Spain, Corsica and northern Africa had wide posterior distributions, which is probably due to the small number of genotypes from these gene pools (see Table XI.3). In line with the genetic variance estimates, the medians of the gene-pool specific estimates of heritability H_j^2 varied between 0.104 in the Iberian Atlantic gene-pool (95% CIs: 0.065-0.146) and 0.223 in the south-eastern Spain gene pool (95% CIs: 0.093-0.363) (Table XI.28 and Fig. XI.13A). Interestingly, provenances that showed the highest broad-sense heritabilities (i.e. provenances from the Corsican, south-eastern Spain, and central Spain gene pools) are also the ones facing more contrasted climates. Indeed, Corsica and south-eastern Spain are mountainous areas, with strong environmental heterogeneity at small spatial scales. Central Spain is less contrasted spatially but experiences a high continentality, and thus strong daily and annually climatic variation. Noticeably, genotypes displayed genetic values mostly determined by the dominant gene pool from which they belong to, but also, to a lesser extent, by other gene pools (Fig. XI.13B). Taken together, these results may suggest that gene pools from regions with high environmental heterogeneity (in space and/or time) have also higher heritability. Indeed, theoretical and empirical works have proposed that high levels of adaptive genetic variance can be maintained in regions of high environmental heterogeneity (mainly spatial, but to a lesser extent also temporal) and some degree of gene flow between populations (McDonald and Yeaman 2018, Yeaman and Jarvis 2006, Yeaman and Otto 2011). Further research involving multiple adaptive traits and more detailed environmental data would be needed to confirm this hypothesis.

Parameter	Median	SD	InfCI	SupCI
$\sigma^2_{A_{NA}}$	0.015	0.010	0.005	0.041
$\sigma^2_{A_C}$	0.025	0.010	0.013	0.050
$\sigma^2_{A_{CS}}$	0.017	0.003	0.013	0.024
$\sigma^2_{A_{FA}}$	0.015	0.002	0.012	0.020
$\sigma^2_{A_{IA}}$	0.011	0.003	0.007	0.017
$\sigma^2_{A_{SES}}$	0.027	0.012	0.011	0.059
$\sigma^2_{g_j}$	0.004	0.031	0.000	0.064
σ^2_P	0.011	0.004	0.006	0.021
σ^2_S	0.121	0.565	0.026	1.259
$\sigma^2_{cs_{is}}$	0.024	0.157	0.003	0.262
σ^2_B	0.002	0.001	0.001	0.003
σ^2	0.098	0.001	0.096	0.099
β_0	6.289	0.220	5.819	6.708
β_{age}	0.605	0.012	0.578	0.626
β_{age2}	-0.185	0.016	-0.221	-0.158

Table XI.26. Parameter estimates of the varying-intercept variances (σ^2_S, σ^2_B, σ^2_P, σ^2_G, $\sigma^2_{cs_{is}}$ and σ_{A_j}), the gene-pool specific total genetic variances (σ_{A_j}), the global variance σ^2, the global intercept β_0 and the slopes associated with the age effect (β_{age} and β_{age2}). SD corresponds to the standard deviation and InfCI and SupCI correspond to the lower and upper bounds of the 0.95 credible interval.

Hypothesis	Estimate	Est.Error	CI.Lower	CI.Upper	Evid.Ratio	Post.Prob	Star
$\sigma^2_{A_{NA}} - \sigma^2_{A_C} < 0$	-0.035	0.043	-0.104	0.035	4.291	0.811	
$\sigma^2_{A_{NA}} - \sigma^2_{A_{CS}} < 0$	-0.007	0.035	-0.057	0.056	1.656	0.623	
$\sigma^2_{A_{NA}} - \sigma^2_{A_{FA}} < 0$	0.001	0.035	-0.049	0.065	1.127	0.530	
$\sigma^2_{A_{NA}} - \sigma^2_{A_{IA}} < 0$	0.019	0.036	-0.033	0.084	0.450	0.310	
$\sigma^2_{A_{NA}} - \sigma^2_{A_{SES}} < 0$	-0.042	0.052	-0.121	0.046	4.396	0.815	
$\sigma^2_{A_C} - \sigma^2_{A_{NA}} < 0$	0.035	0.043	-0.035	0.104	0.233	0.189	
$\sigma^2_{A_C} - \sigma^2_{A_{CS}} < 0$	0.029	0.030	-0.015	0.081	0.183	0.155	
$\sigma^2_{A_C} - \sigma^2_{A_{FA}} < 0$	0.037	0.029	-0.006	0.089	0.091	0.084	
$\sigma^2_{A_C} - \sigma^2_{A_{IA}} < 0$	0.055	0.030	0.009	0.109	0.022	0.022	
$\sigma^2_{A_C} - \sigma^2_{A_{SES}} < 0$	-0.006	0.044	-0.080	0.065	1.239	0.553	
$\sigma^2_{A_{CS}} - \sigma^2_{A_{NA}} < 0$	0.007	0.035	-0.056	0.057	0.604	0.377	
$\sigma^2_{A_{CS}} - \sigma^2_{A_C} < 0$	-0.029	0.030	-0.081	0.015	5.459	0.845	
$\sigma^2_{A_{CS}} - \sigma^2_{A_{FA}} < 0$	0.008	0.013	-0.013	0.030	0.384	0.277	
$\sigma^2_{A_{CS}} - \sigma^2_{A_{IA}} < 0$	0.026	0.017	-0.001	0.053	0.064	0.060	
$\sigma^2_{A_{CS}} - \sigma^2_{A_{SES}} < 0$	-0.035	0.038	-0.103	0.022	4.859	0.829	
$\sigma^2_{A_{FA}} - \sigma^2_{A_{NA}} < 0$	-0.001	0.035	-0.065	0.049	0.887	0.470	
$\sigma^2_{A_{FA}} - \sigma^2_{A_C} < 0$	-0.037	0.029	-0.089	0.006	10.952	0.916	*
$\sigma^2_{A_{FA}} - \sigma^2_{A_{CS}} < 0$	-0.008	0.013	-0.030	0.013	2.606	0.723	
$\sigma^2_{A_{FA}} - \sigma^2_{A_{IA}} < 0$	0.018	0.014	-0.006	0.041	0.117	0.105	
$\sigma^2_{A_{FA}} - \sigma^2_{A_{SES}} < 0$	-0.043	0.035	-0.107	0.010	9.435	0.904	*
$\sigma^2_{A_{IA}} - \sigma^2_{A_{NA}} < 0$	-0.019	0.036	-0.084	0.033	2.221	0.690	
$\sigma^2_{A_{IA}} - \sigma^2_{A_C} < 0$	-0.055	0.030	-0.109	-0.009	44.455	0.978	**
$\sigma^2_{A_{IA}} - \sigma^2_{A_{CS}} < 0$	-0.026	0.017	-0.053	0.001	15.667	0.940	*
$\sigma^2_{A_{IA}} - \sigma^2_{A_{FA}} < 0$	-0.018	0.014	-0.041	0.006	8.524	0.895	
$\sigma^2_{A_{IA}} - \sigma^2_{A_{SES}} < 0$	-0.061	0.036	-0.125	-0.006	29.928	0.968	**
$\sigma^2_{A_{SES}} - \sigma^2_{A_{NA}} < 0$	0.042	0.052	-0.046	0.121	0.227	0.185	
$\sigma^2_{A_{SES}} - \sigma^2_{A_C} < 0$	0.006	0.044	-0.065	0.080	0.807	0.447	
$\sigma^2_{A_{SES}} - \sigma^2_{A_{CS}} < 0$	0.035	0.038	-0.022	0.103	0.206	0.171	
$\sigma^2_{A_{SES}} - \sigma^2_{A_{FA}} < 0$	0.043	0.035	-0.010	0.107	0.106	0.096	
$\sigma^2_{A_{SES}} - \sigma^2_{A_{IA}} < 0$	0.061	0.036	0.006	0.125	0.033	0.032	

Table XI.27. One-sided hypothesis testing on the probability that the gene pool-specific total genetic variances in *M5* are different, using the function 'hypothesis' from the 'brms' package (Bürkner 2017). 'Est.Error' is the standard deviation of the estimated difference between two genetic variances ('Estimate'). The 'CI.Lower' and 'CI.Upper' are the lower and upper bounds of the 95% credible interval, respectively. 'Evid.Ratio' is the evidence ratio of each hypothesis, i.e. the posterior probability ('Post.Prob') under the hypothesis against its alternative. For instance, the evidence ratio of the hypothesis $\sigma^2_{A_{NA}} - \sigma^2_{A_C} < 0$ is the ratio of the posterior probability of $\sigma^2_{A_{NA}} - \sigma^2_{A_C} < 0$ and the posterior probability of $\sigma^2_{A_{NA}} - \sigma^2_{A_C} > 0$. The * and ** in the 'Star' column indicate hypotheses with a posterior probability higher than 0.90 and 0.95, respectively.

Using the estimates of the gene-pool specific total genetic variances from *M5*, the gene-pool specific heritabilities were calculated such as:

$$H_j^2 = \frac{\sigma_{A_j}^2}{\sigma_{A_j}^2 + \sigma^2}$$

where σ^2 is the modeled residual variance and $\sigma_{A_j}^2$ the gene-pool specific total genetic variances.

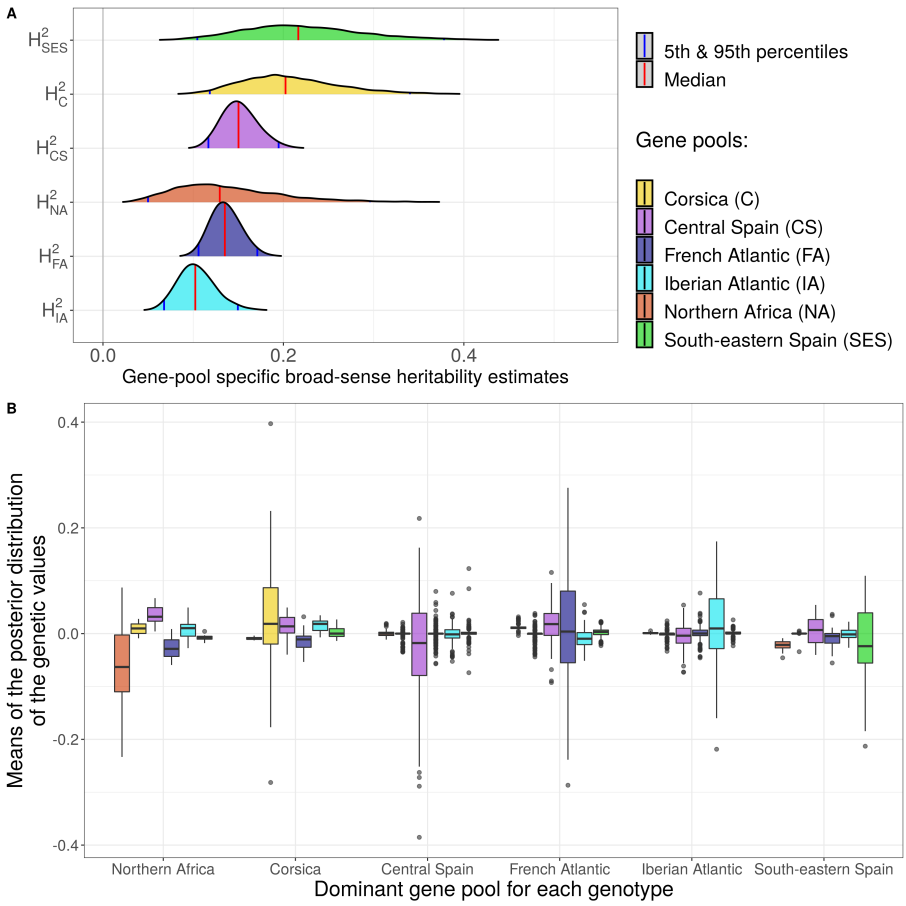


Figure XI.13. Posterior distribution of H_j^2, the gene pool-specific heritabilities obtained from *model M5* (A). Boxplot of he gene pool partial genetic values for genotypes grouped according to the main gene pool to which they belong (B). The colors represent the gene pool from which the partial genetic values were estimated. The lower, middle and upper parts of the hinges correspond to the 25%, 50% and 75% quantiles, respectively. The lower/upper whiskers extend from the hinge to the smallest/largest value no further than 1.5 time the IQR from the hinge (where IQR is the distance between the first and third quartiles).

Parameter	Median	SD	InfCI	SupCI
H^2_{NA}	0.141	0.064	0.034	0.266
H^2_{C}	0.210	0.057	0.107	0.322
H^2_{CS}	0.152	0.020	0.114	0.190
H^2_{FA}	0.136	0.017	0.105	0.169
H^2_{IA}	0.104	0.021	0.065	0.146
H^2_{SES}	0.223	0.070	0.093	0.363

Table XI.28. Gene pool-specific heritability estimates H_j^2. The gene pools come from: Northern Africa (NA), Corsica (C), Central Spain (CS), French Atlantic region (FA), Iberian Atlantic region (IA) and south-eastern Spain (SES). SD corresponds to the standard deviation and InfCI and SupCI correspond to the lower and upper bounds of the 0.95 credible interval.

MODEL M6

In *M6*, the variance $\sigma^2_{cp_p}$ associated with the provenance climate-of-origin had a higher median than the variances attributed to either gene pools or provenances ($\sigma^2_{g_j}$ and σ^2_P; Table XI.29). Including the provenance climate-of-origin in *M6* resulted in a decreased variance of the provenances and gene pools, suggesting confounded effects (Tables XI.25 & XI.29). Once these confounded effects are taking into account, trees from climatic regions neighboring the Atlantic Ocean were generally the tallest (e.g. CAD, SIE, PUE, LAM and CAS in northwestern Spain; all provenances along the French Atlantic coast; Fig. XI.14). Interestingly, the Leiria (LEI) provenance, which has a strong Iberian Atlantic component (Table XI.3) and had the highest climate intercept estimate (similar to that of the French Atlantic provenances), was not among the tallest provenances (Fig. XI.14). Also, the Corsican provenances showed contrasted climate intercepts, with a positive influence on height growth for Pinia (PIA) but not for Pineta (PIE), which could explain their striking differences in height-growth patterns (Fig. V.3). Finally, the four provenances from south-eastern Spain and northern Africa gene pools showed all negative climate intercepts (Fig. XI.14).

Parameter	Median	SD	InfCI	SupCI
$\sigma^2_{g_j}$	0.003	0.018	0.000	0.050
σ^2_P	0.005	0.002	0.003	0.011
$\sigma^2_{cp_p}$	0.009	0.291	0.000	0.244
σ^2_G	0.012	0.001	0.010	0.014
σ^2_S	0.128	0.761	0.026	1.436
$\sigma^2_{cs_{is}}$	0.023	0.111	0.003	0.297
σ^2_B	0.002	0.001	0.001	0.003
σ^2	0.097	0.001	0.096	0.099
β_0	6.280	0.235	5.777	6.723
β_{age}	0.605	0.013	0.576	0.627
β_{age2}	-0.185	0.018	-0.224	-0.156

Table XI.29. Parameter estimates of the varying-intercept variances (σ^2_S, σ^2_B, σ^2_P, σ^2_G, $\sigma^2_{cs_{is}}$ and $\sigma^2_{cp_p}$), the gene-pool specific total genetic variances (σ_{A_j}), the global variance σ^2, the global intercept β_0 and the slopes associated with the age effect (β_{age} and β_{age2}). SD corresponds to the standard deviation and InfCI and SupCI correspond to the lower and upper bounds of the 0.95 credible interval.

Summary figure of models *models M3, M5* and *M6*

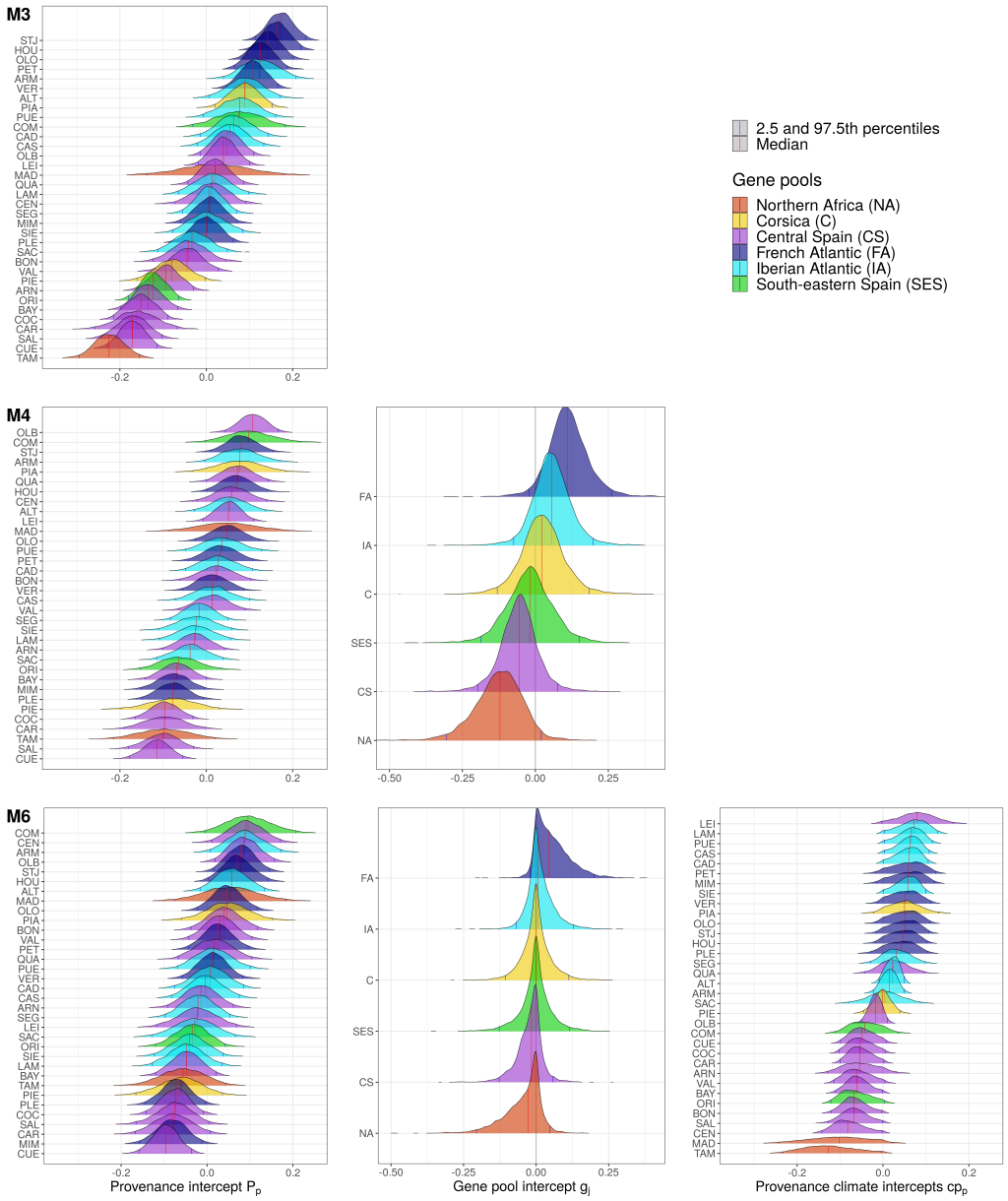


Figure XI.14. Posterior distributions of the provenance intercepts P_p, gene pool intercepts g_j and the cp_p intercepts related to the provenance climate. The first row corresponds to *M3* including only P_p. The second row corresponds to *M4* including both P_p and g_j. The third row corresponds to *M6* including P_p, g_j and cp_p. In *M3*, *M4* and *M6*, the plastic component is included in the same way via the intercepts S_S and cs_{is}. Only the genetic component changes between these models. Provenances were colored based on the main gene pool they belong to.

6.1.4 Predictive *models M7* and *M8*: combining climatic and genomic drivers

MODEL M7

Term	Median	SD	InfCI	SupCI
σ^2_B	0.002	0.001	0.001	0.003
$\sigma^2_{g_j}$	0.027	0.100	0.008	0.192
σ^2_S	0.172	0.992	0.036	2.252
$\sigma^2_{\beta_{max.temp,s}}$	0.001	0.005	0.000	0.009
$\sigma^2_{\beta_{min.pre,s}}$	0.001	0.003	0.000	0.007
$\sigma^2_{\beta_{gPEA,s}}$	0.018	0.048	0.005	0.135
σ^2	0.104	0.001	0.102	0.106
β_0	6.303	0.403	5.518	7.147
β_{age}	0.599	0.003	0.593	0.604
β_{age2}	-0.151	0.003	-0.156	-0.146

Table XI.30. Parameter estimates from model *M7* fitted on the P1 partition. SD corresponds to the standard deviation and InfCI and SupCI correspond to the lower and upper bounds of the 0.95 credible interval.

MODEL M8

Term	Median	SD	InfCI	SupCI
σ^2_B	0.002	0.001	0.001	0.003
$\sigma^2_{g_j}$	0.033	0.078	0.009	0.228
σ^2_S	0.185	1.548	0.041	2.349
$\sigma^2_{\beta_{max.temp,s}}$	0.003	0.009	0.001	0.020
$\sigma^2_{\beta_{min.pre,s}}$	0.002	0.005	0.000	0.014
$\sigma^2_{\beta_{rPEA,s}}$	0.032	0.133	0.010	0.202
σ^2	0.102	0.001	0.100	0.104
β_0	6.333	0.353	5.631	7.034
β_{age}	0.599	0.003	0.593	0.604
β_{age2}	-0.151	0.003	-0.156	-0.146

Table XI.31. Parameter estimates from model *M8* fitted on the P1 partition. SD corresponds to the standard deviation and InfCI and SupCI correspond to the lower and upper bounds of the 0.95 credible interval.

Summary figure of *models M7* and *M8*

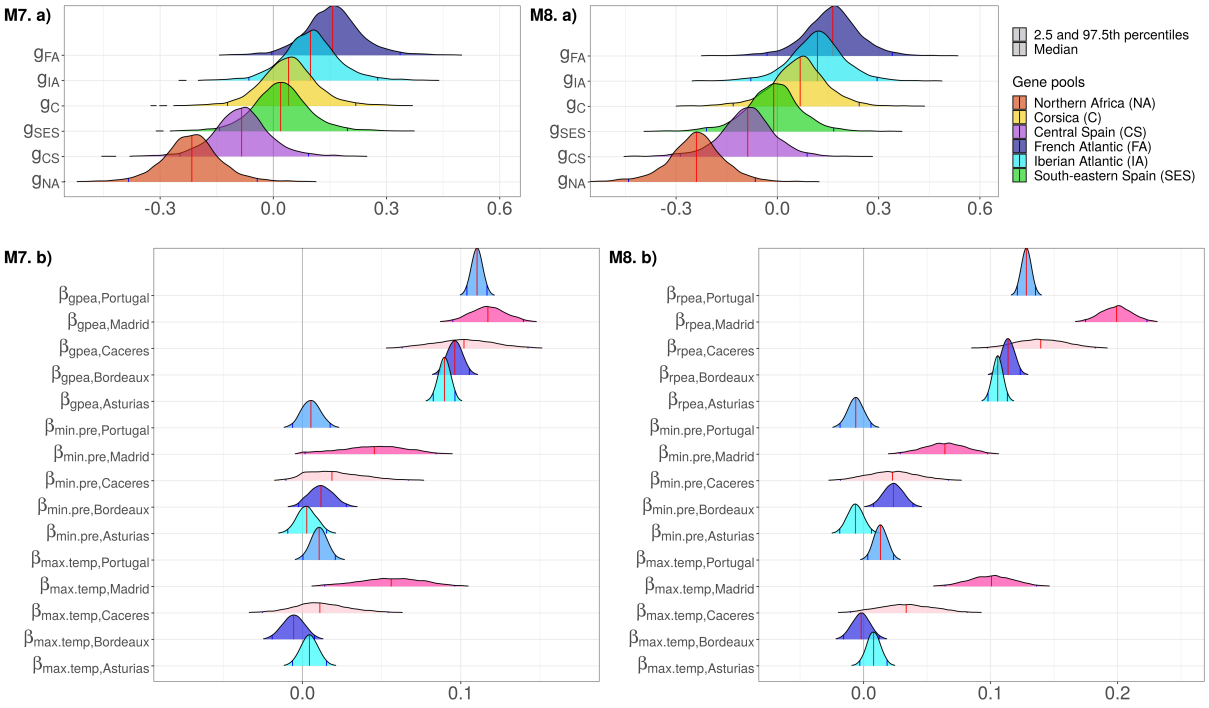


Figure XI.15. Posterior distributions of parameters from models *M7* and *M8* fitted on the P1 partition. In panels a), the parameter estimates g_j correspond to the effect of the gene pools j. In panels b), the parameter estimates correspond to the site-specific effects of gPEAs in *M7* ($\beta_{gPEA,s}$), rPEAs in *M8* ($\beta_{rPEA,s}$), the minimum precipitation during the driest month ($\beta_{min.pre,s}$) and the maximum temperature of the warmest month ($\beta_{max.temp,s}$). Parameters specific to the same site are colored the same to aid visualization.

6.1.5 Predictive *models M9* to *M12*: including separately climatic and genomic drivers

MODEL M9

Parameter	Median	SD	InfCI	SupCI
σ_{g_j}	0.024	0.077	0.007	0.170
σ^2_S	0.107	0.921	0.027	1.346
σ^2_B	0.002	0.001	0.001	0.003
σ^2	0.114	0.001	0.112	0.116
β_0	6.233	0.237	5.708	6.637
β_{age}	0.601	0.003	0.596	0.607
β_{age2}	-0.153	0.003	-0.158	-0.148

Table XI.32. Parameter estimates from model *M9* fitted on the P1 partition. SD corresponds to the standard deviation and InfCI and SupCI correspond to the lower and upper bounds of the 0.95 credible interval.

MODEL M10

Term	Median	SD	InfCI	SupCI
σ^2_B	0.002	0.001	0.001	0.003
σ^2_S	0.173	0.909	0.034	2.404
$\sigma^2_{\beta_{max.temp,s}}$	0.005	0.032	0.001	0.033
$\sigma^2_{\beta_{min.pre,s}}$	0.006	0.063	0.001	0.046
σ^2	0.115	0.001	0.113	0.117
β_0	6.202	0.399	5.237	6.869
β_{age}	0.601	0.003	0.596	0.607
β_{age^2}	-0.153	0.003	-0.159	-0.148

Table XI.33. Parameter estimates from model *M10* fitted on the P1 partition. SD corresponds to the standard deviation and InfCI and SupCI correspond to the lower and upper bounds of the 0.95 credible interval.

MODEL M11

Term	Median	SD	InfCI	SupCI
σ^2_B	0.002	0.001	0.001	0.003
σ^2_S	0.249	4.243	0.040	7.779
$\sigma^2_{\beta_{gPEA,s}}$	0.013	0.056	0.004	0.092
σ^2	0.114	0.001	0.112	0.117
β_0	6.322	0.763	4.768	7.742
β_{age}	0.599	0.003	0.593	0.604
β_{age2}	-0.151	0.003	-0.156	-0.146

Table XI.34. Parameter estimates from model *M11* fitted on the P1 partition. SD corresponds to the standard deviation and InfCI and SupCI correspond to the lower and upper bounds of the 0.95 credible interval.

MODEL M12

Term	Median	SD	InfCI	SupCI
σ^2_B	0.002	0.001	0.001	0.003
σ^2_S	0.332	1.572	0.044	4.542
$\sigma^2_{\beta_{rPEA,s}}$	0.023	0.047	0.007	0.152
σ^2	0.113	0.001	0.111	0.115
β_0	6.597	0.499	5.735	7.663
β_{age}	0.599	0.003	0.593	0.604
β_{age2}	-0.151	0.003	-0.156	-0.146

Table XI.35. Parameter estimates from model *M12* fitted on the P1 partition. SD corresponds to the standard deviation and InfCI and SupCI correspond to the lower and upper bounds of the 0.95 credible interval.

Summary figure of *models M9* to *M12*

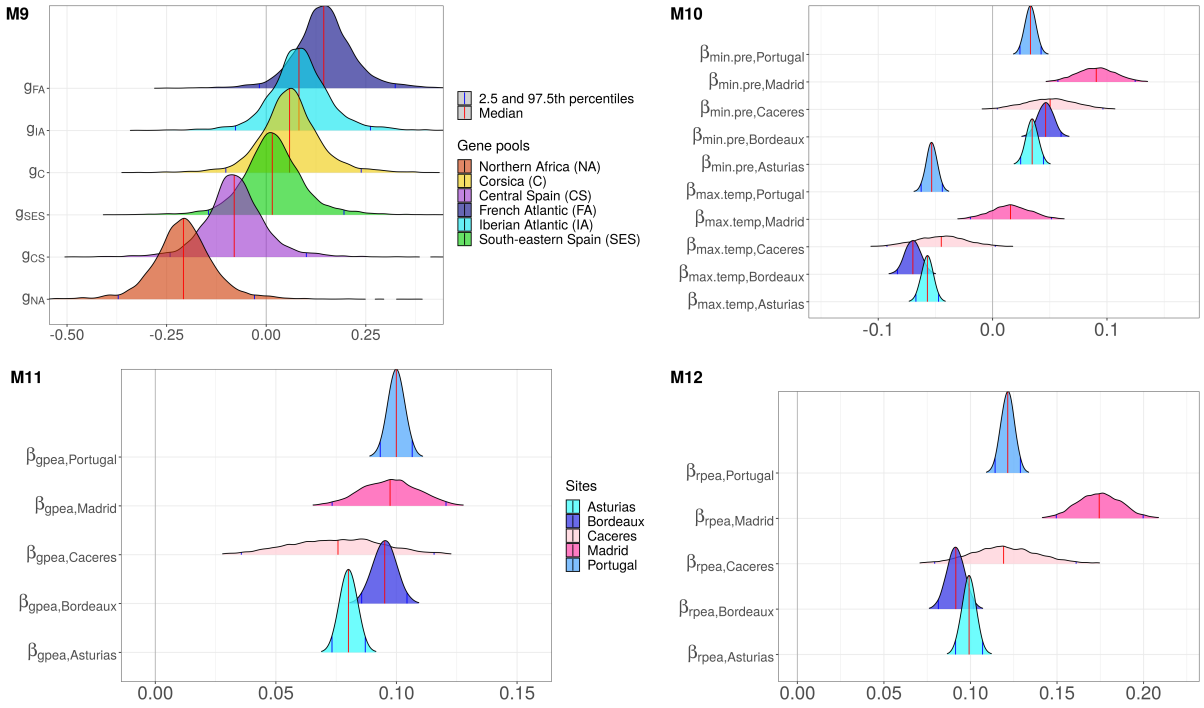


Figure XI.16. Posterior distributions of parameters from models *M9* and *M12* fitted on P1. For *M9*, the parameter estimates g_j correspond to the effect of the gene pools j. For *M10*, the parameter estimates correspond to the site-specific effects of the minimum precipitation during the driest month ($\beta_{min.pre,s}$) and the maximum temperature of the warmest month ($\beta_{max.temp,s}$). For *M11*, the parameter estimates correspond to the site-specific effect of the gPEAs ($\beta_{gPEA,s}$). For *M12*, parameter estimates correspond to the site-specific effects of the rPEAs ($\beta_{rPEA,s}$).

6.2 P2 partition (random split of the provenances)

The unknown provenances of the P2 test dataset were chosen randomly among the 34 provenances of our study.

6.2.1 *M0*, *M1* and *M2*

Parameter	Median	SD	InfCI	SupCI
σ^2_S	0.101	0.527	0.025	1.165
σ^2_B	0.002	0.001	0.001	0.004
σ^2	0.124	0.001	0.121	0.126
β_0	6.278	0.210	5.840	6.631
β_{age}	0.602	0.003	0.597	0.607
β_{age2}	-0.153	0.003	-0.158	-0.148

Table XI.36. Parameter estimates from model M0 fitted on the P2 partition. SD corresponds to the standard deviation and InfCI and SupCI correspond to the lower and upper bounds of the 0.95 credible interval.

Parameter	Median	SD	InfCI	SupCI
σ^2_P	0.016	0.005	0.009	0.029
σ^2_G	0.011	0.001	0.010	0.013
σ^2_S	0.110	0.540	0.028	1.032
σ^2_B	0.002	0.001	0.001	0.003
σ^2	0.097	0.001	0.095	0.099
β_0	6.240	0.190	5.866	6.627
β_{age}	0.599	0.002	0.594	0.603
β_{age2}	-0.150	0.002	-0.155	-0.146

Table XI.37. Parameter estimates from model *M1* fitted on the P2 partition. SD corresponds to the standard deviation and InfCI and SupCI correspond to the lower and upper bounds of the 0.95 credible interval.

Parameter	Median	SD	InfCI	SupCI
σ^2_P	0.014	0.005	0.008	0.027
σ^2_G	0.011	0.001	0.010	0.013
σ^2_{Inter}	0.005	0.001	0.003	0.007
σ^2_S	0.123	0.584	0.030	1.532
σ^2_B	0.002	0.001	0.001	0.003
σ^2	0.095	0.001	0.094	0.097
β_0	6.229	0.226	5.777	6.686
β_{age}	0.598	0.002	0.593	0.603
β_{age2}	-0.150	0.002	-0.155	-0.145

Table XI.38. Parameter estimates from model *M2* fitted on the P2 partition. SD corresponds to the standard deviation and InfCI and SupCI correspond to the lower and upper bounds of the 0.95 credible interval.

6.2.2 Predictive *models M7* and *M8*: combining climatic and genomic drivers

Term	Median	SD	InfCI	SupCI
σ^2_B	0.002	0.001	0.001	0.004
$\sigma^2_{g_j}$	0.027	0.076	0.008	0.195
σ^2_S	0.157	0.837	0.036	1.926
$\sigma^2_{\beta_{max.temp,s}}$	0.001	0.004	0.000	0.007
$\sigma^2_{\beta_{min.pre,s}}$	0.001	0.004	0.000	0.007
$\sigma^2_{\beta_{gPEA,s}}$	0.019	0.055	0.005	0.136
σ^2	0.103	0.001	0.101	0.104
β_0	6.310	0.345	5.624	7.068
β_{age}	0.600	0.003	0.595	0.605
β_{age2}	-0.151	0.002	-0.156	-0.147

Table XI.39. Parameter estimates from model *M7* fitted on the P2 partition. SD corresponds to the standard deviation and InfCI and SupCI correspond to the lower and upper bounds of the 0.95 credible interval.

Term	Median	SD	InfCI	SupCI
σ^2_B	0.002	0.001	0.001	0.003
$\sigma^2_{g_j}$	0.032	0.075	0.010	0.224
σ^2_S	0.190	1.195	0.041	2.366
$\sigma^2_{\beta_{max.temp,s}}$	0.003	0.006	0.001	0.019
$\sigma^2_{\beta_{min.pre,s}}$	0.002	0.004	0.000	0.014
$\sigma^2_{\beta_{rPEA,s}}$	0.032	0.085	0.010	0.241
σ^2	0.101	0.001	0.099	0.103
β_0	6.337	0.366	5.602	7.106
β_{age}	0.600	0.003	0.595	0.605
β_{age2}	-0.151	0.002	-0.156	-0.147

Table XI.40. Parameter estimates from model *M8* fitted on the P2 partition. SD corresponds to the standard deviation and InfCI and SupCI correspond to the lower and upper bounds of the 0.95 credible interval.

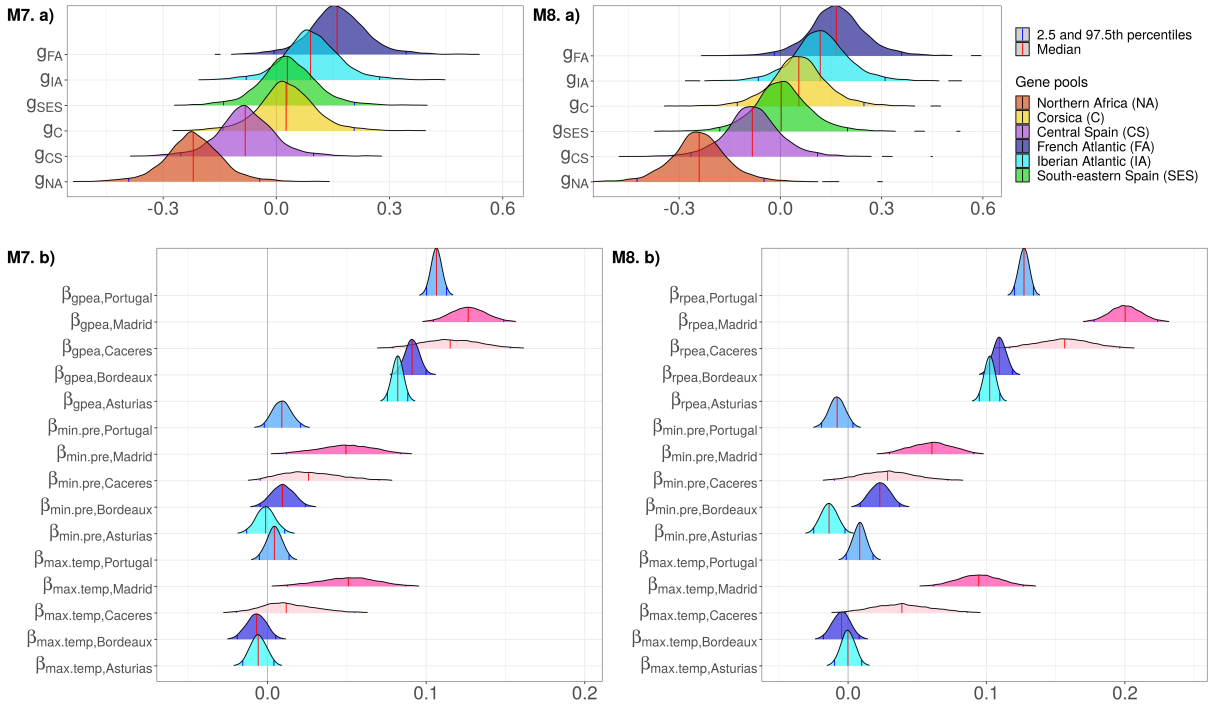


Figure XI.17. Posterior distributions of parameters from models *M7* and *M8* fit on P2. In panels a), the parameter estimates g_j correspond to the effect of the gene pools j. In panels b), the parameter estimates correspond to the site-specific effects of the PEAs ($\beta_{gPEA,s}$ for the global PEAs in *M7* and $\beta_{rPEA,s}$ for the regional PEAs in *M8*), the minimum precipitation during the driest month ($\beta_{min.pre,s}$) and the maximum temperature of the warmest month ($\beta_{max.temp,s}$). Parameters specific to the same test site are colored the same to aid visualization.

6.2.3 Predictive *models M9* to *M12*: including separately climatic and genomic drivers

Parameter	Median	SD	InfCI	SupCI
$\sigma^2_{g_j}$	0.027	0.083	0.008	0.199
σ^2_S	0.108	0.622	0.027	1.019
σ^2_B	0.002	0.001	0.001	0.004
σ^2	0.112	0.001	0.110	0.114
β_0	6.230	0.217	5.802	6.647
β_{age}	0.602	0.003	0.597	0.608
β_{age2}	-0.153	0.003	-0.158	-0.148

Table XI.41. Parameter estimates from model *M9* fitted on the P2 partition. SD corresponds to the standard deviation and InfCI and SupCI correspond to the lower and upper bounds of the 0.95 credible interval.

Parameter	Median	SD	InfCI	SupCI
σ^2_B	0.002	0.001	0.001	0.004
σ^2_S	0.183	1.789	0.036	3.006
$\sigma^2_{\beta_{max.temp,s}}$	0.005	0.011	0.002	0.035
$\sigma^2_{\beta_{min.pre,s}}$	0.006	0.018	0.002	0.041
σ^2	0.113	0.001	0.111	0.115
β_0	6.239	0.456	5.194	7.075
β_{age}	0.602	0.003	0.597	0.607
β_{age2}	-0.153	0.003	-0.158	-0.148

Table XI.42. Parameter estimates from model *M10* fitted on the P2 partition. SD corresponds to the standard deviation and InfCI and SupCI correspond to the lower and upper bounds of the 0.95 credible interval.

Term	Median	SD	InfCI	SupCI
σ^2_B	0.002	0.001	0.001	0.004
σ^2_S	0.265	2.482	0.038	5.597
$\sigma^2_{\beta_{gPEA,s}}$	0.013	0.031	0.004	0.085
σ^2	0.115	0.001	0.113	0.117
β_0	6.374	0.677	5.164	7.956
β_{age}	0.600	0.003	0.594	0.605
β_{age2}	-0.151	0.003	-0.156	-0.146

Table XI.43. Parameter estimates from model *M11* fitted on the P2 partition. SD corresponds to the standard deviation and InfCI and SupCI correspond to the lower and upper bounds of the 0.95 credible interval.

Term	Median	SD	InfCI	SupCI
σ^2_B	0.002	0.001	0.001	0.003
σ^2_S	0.332	2.743	0.042	4.656
$\sigma^2_{\beta_{rPEA,s}}$	0.026	0.083	0.008	0.177
σ^2	0.112	0.001	0.110	0.114
β_0	6.556	0.532	5.620	7.714
β_{age}	0.599	0.003	0.594	0.605
β_{age2}	-0.151	0.003	-0.156	-0.146

Table XI.44. Parameter estimates from model *M12* fitted on the P1 partition. SD corresponds to the standard deviation and InfCI and SupCI correspond to the lower and upper bounds of the 0.95 credible interval.

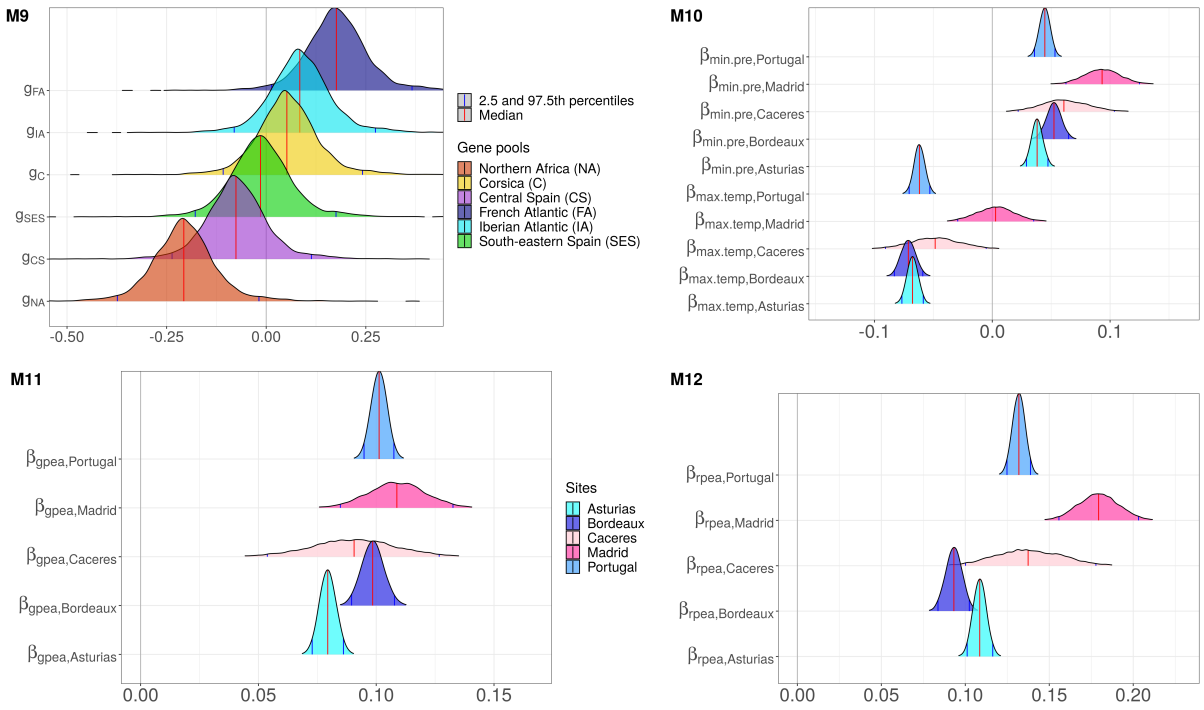


Figure XI.18. Posterior distributions of parameters from models *M9* and *M12* fitted on P2. For *M9*, the parameter estimates g_j correspond to the effect of the gene pools j. For *M10*, the parameter estimates correspond to the site-specific effects of the minimum precipitation during the driest month ($\beta_{min.pre,s}$) and the maximum temperature of the warmest month ($\beta_{max.temp,s}$). For *M11*, the parameter estimates correspond to the site-specific effect of the gPEAs ($\beta_{gPEA,s}$). For *M12*, parameter estimates correspond to the site-specific effects of the rPEAs ($\beta_{rPEA,s}$).

6.3 P3 partition (non-random split of the provenances)

Evaluation of model performance on new provenances was replicated on six other provenances to assess the robustness of the results. In the P3 partition, the new provenances were not totally randomly selected to ensure that each under-represented gene pool in our study was represented by at least one provenance. Thus, one provenance was randomly selected from the two provenances belonging mainly to the northern Africa gene pool. The same was done for the gene pools from south-eastern Spain and Corsica. The last three provenances were randomly selected from the three remaining gene pools.

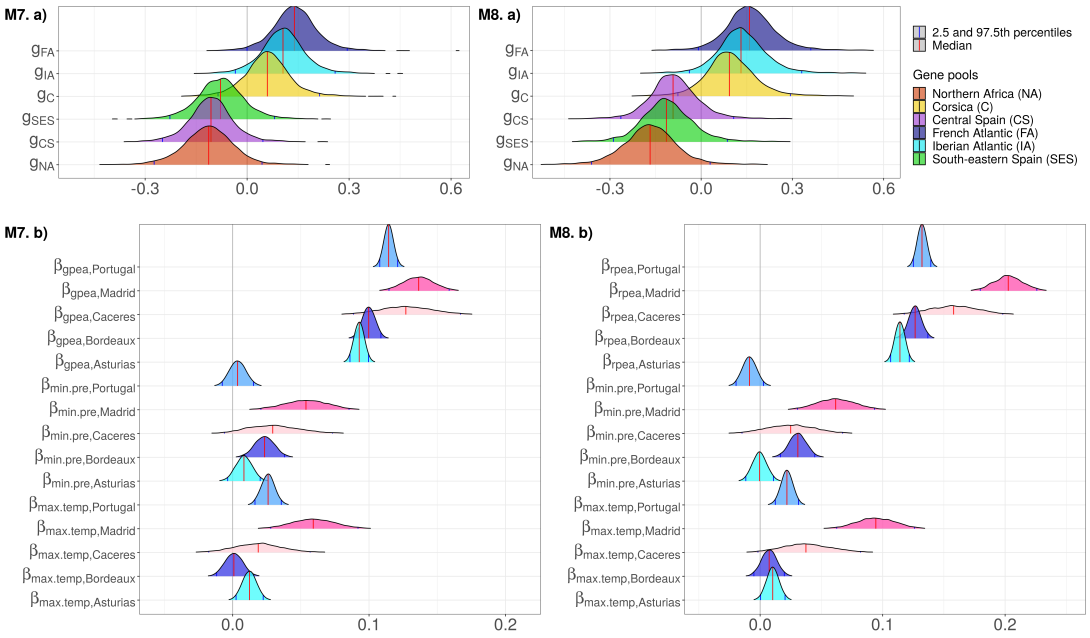


Figure XI.19. Posterior distributions of parameters from models *M7* and *M8* fit on P3. In panels a), the parameter estimates g_j correspond to the effect of the gene pools j. In panels b), the parameter estimates correspond to the site-specific effects of gPEAs in *M7* ($\beta_{gPEA,s}$, rPEAs in *M8* $\beta_{rPEA,s}$, the minimum precipitation during the driest month ($\beta_{min.pre,s}$) and the maximum temperature of the warmest month ($\beta_{max.temp,s}$). Parameters specific to the same site are colored the same to aid visualization.

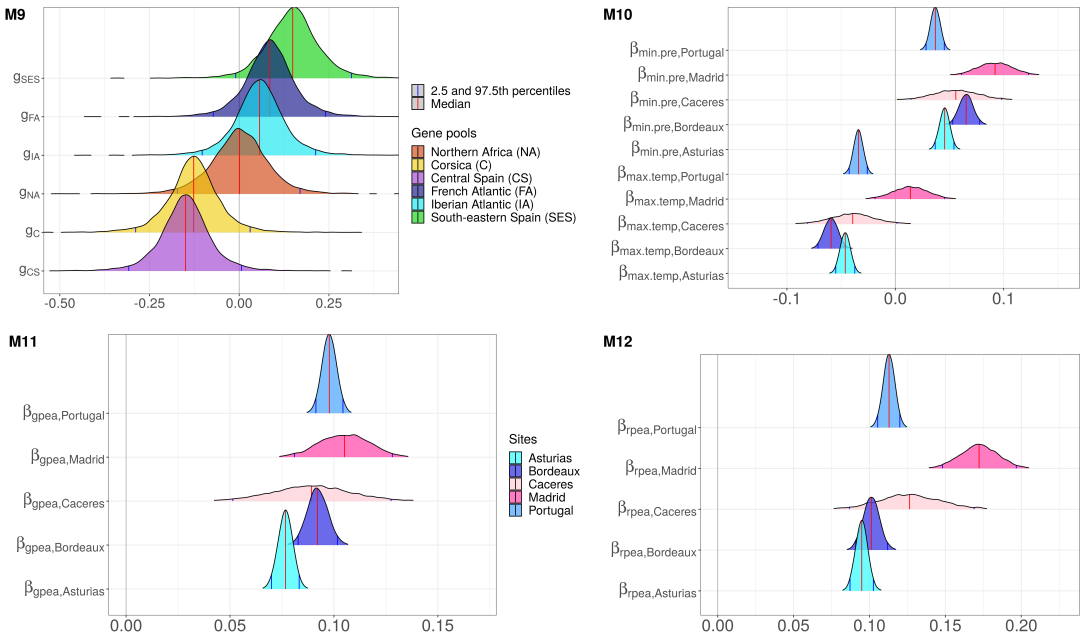


Figure XI.20. Posterior distributions of parameters from models *M9* and *M12* fit on P3. For *M9*, the parameter estimates g_j correspond to the effect of the gene pools j. For *M10*, the parameter estimates correspond to the site-specific effects of the minimum precipitation during the driest month ($\beta_{min.pre,s}$) and the maximum temperature of the warmest month ($\beta_{max.temp,s}$). For *M11*, the parameter estimates correspond to the site-specific effect of the gPEAs ($\beta_{gPEA,s}$). For *M12*, parameter estimates correspond to the site-specific effects of the rPEAs ($\beta_{rPEA,s}$).

6.4 Interpretation of the PEAs coefficients

Let's take model *M12* as an example. Here is the equation of *M12* (see equation 4.1.3 in the Supplementary Information):

$$\log(h_{isbr}) \sim \mathcal{N}(\mathbf{X}\beta + S_s + B_{b(s)} + \beta_{rPEA,s} rPEA_{gr}\ ,\ \sigma^2)$$

with $rPEA_{gr}$ is the scaled explanatory variable (i.e. the counts of regionally-selected positive-effect alleles). Let's call $r\tilde{PEA}_{gr}$ the explanatory variable before being scaled, that is $rPEA_{gr} = (r\tilde{PEA}_{gr} - \mu_{r\tilde{PEA}_{gr}})/\sigma_{r\tilde{PEA}_{gr}}$, where $\mu_{r\tilde{PEA}_{gr}}$ is the mean of $r\tilde{PEA}_{gr}$ and $\sigma_{r\tilde{PEA}_{gr}}$ is its standard deviation.

We want to calculate the percent of change in height associated with a one-unit increase in $rPEA_{gr}$, that is a one-standard deviation increase in $r\tilde{PEA}_{gr}$. For that, we call h_{new} the value of h_{isbr} after increasing $rPEA_{gr}$ by one unit, and we have:

$$\begin{aligned}\log(h_{new}) &= \log(\mathbf{X}\beta + S_s + B_{b(s)} + \beta_{rPEA,s}(rPEA_{gr} + 1)\\ &= \log(h_{isbr}) + \beta_{rPEA,s}\end{aligned}$$

Therefore:

$$\begin{aligned}\log(h_{new}) - \log(h_{isbr}) &= \beta_{rPEA,s}\\ \frac{h_{new}}{h_{isbr}} &= \exp(\beta_{rPEA,s})\\ 100 \times \left(\frac{h_{new}}{h_{isbr}} - 1\right) &= 100 \times \left(\exp(\beta_{rPEA,s}) - 1\right)\\ 100 \times \left(\frac{h_{new} - h_{isbr}}{h_{isbr}}\right) &= 100 \times (\exp(\beta_{rPEA,s}) - 1)\end{aligned}$$

is the percent change in h_{isbr} associated with a one-unit increase in $rPEA_{gr}$ (that is, a one-standard deviation increase in $r\tilde{PEA}_{gr}$). For instance, a one-standard deviation increase in the counts of rPEAs is associated, on average, with $100 \times (\exp(0.174) - 1) = 19\%$ change in height in Madrid, with $100 \times (\exp(0.120) - 1) = 12.7\%$ change in height in Cáceres, with $100 \times (\exp(0.092) - 1) = 9.6\%$ change in height in Bordeaux, with $100 \times (\exp(0.099) - 1) = 10.4\%$ change in height in Asturias and with $100 \times (\exp(0.122) - 1) = 13.0\%$ change in height in Portugal.

7 $Q_{ST} - F_{ST}$ analysis

To determine whether height growth shows footprints of adaptive differentiation, we performed a $Q_{ST} - F_{ST}$ analysis. We used the global F_{ST} estimate calculated in de Miguel et al. (2020) on the same data as our study (i.e. the 5,165 SNPs from the Illumina Infinium SNP array). de Miguel et al. (2020) used the *diveRsity* R package and 1,000 bootstrap iterations across loci to estimate the 95% confidence interval of the global F_{ST}. They obtained a F_{ST} of 0.112 (95% confidence interval: 0.090 - 0.141).

To calculate the Q_{ST}, we used the following formula from Spitze (1993):

$$Q_{ST} = \frac{\sigma_P^2}{\sigma_P^2 + 2\sigma_G^2}$$

where σ_P^2 is the variance among provenances, and σ_G^2 is the variance among clones (.i.e. genotypes) within provenances.

The median estimate of the Q_{ST} was 0.358 (95% confidence interval: 0.251-0.506). Quantitative (Q_{ST}) and molecular (F_{ST}) genetic differentiation among provenances were considered significantly different as their posterior distributions had non-overlapping 95% confidence intervals, which therefore suggests that there is adaptive differentiation in height growth in our study.

8 Distribution of heights in Cáceres and Madrid.

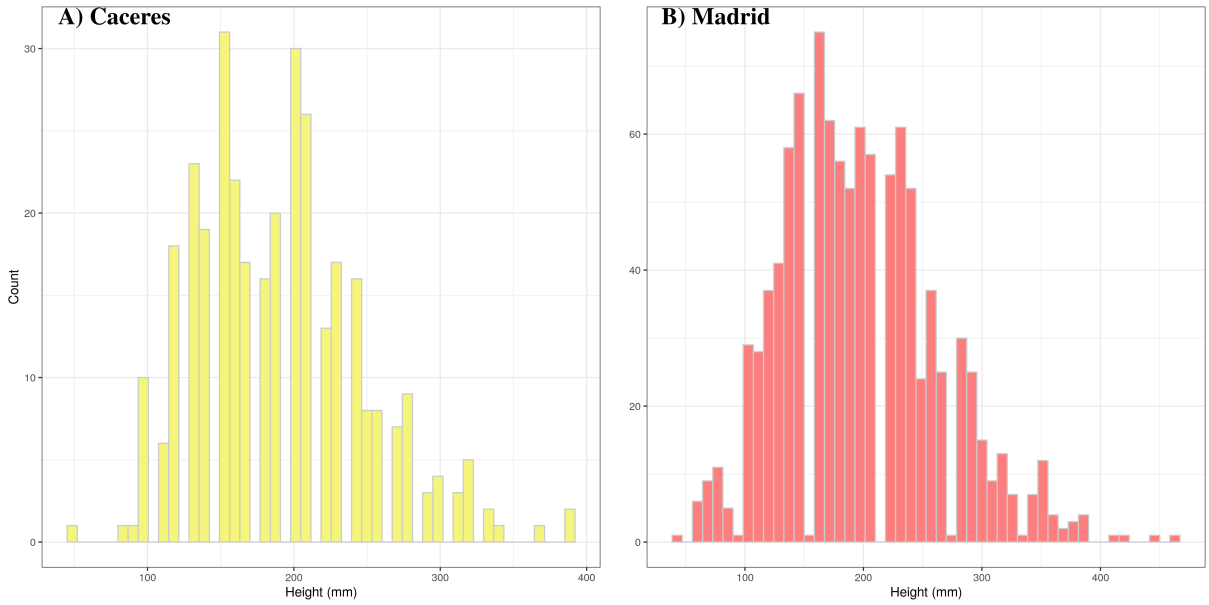
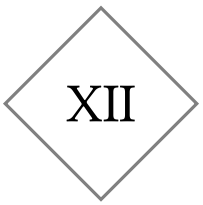


Figure XI.21. Distribution of height measurements (in mm) in Cáceres and Madrid.



Supplementary information - Chapter 3

1 Materials & methods

1.1 Population information

Code	Population	Number of trees
ALT	Alto de la Llama	8
ARM	Armayán	7
ARN	Arenas de San Pedro	14
BAY	Bayubas de Abajo	13
BON	Boniches	8
CAD	Cadavedo	8
CAR	Carbonero el Mayor	5
CAS	Castropol	8
CEN	Cenicientos	9
COC	Coca	15
COM	Cómpeta	3
CUE	Cuellar	23
HOU	Hourtin	25
LAM	Lamuño	9
LEI	Leiria	20
MAD	Madisouka	1
MIM	Mimizan	17
OLB	Olba	20
OLO	Olonne sur Mer	23
ORI	Oria	22
PET	Petrocq	22
PIA	Pinia	12
PIE	Pineta	9
PLE	Pleucadec	16
PUE	Puerto de Vega	7
QUA	Quatretonda	16
SAC	San Cipriano de Ribaterme	8
SAL	San Leonardo	10
SEG	Sergude (Huerto Semillero)	19
SIE	Sierra de Barcia	7
STJ	St-Jean des Monts	24
TAM	Tamrabta	12
VAL	Valdemaqueda	10
VER	Le Verdon	24

Table XII.1. Population information: population codes used in the study, population names and number of trees sampled in each population.

1.2 Climatic, soil, topographic and fire-related data

The eight environmental variables used in this study are:

- Climatic variables, which correspond to mean values over the period 1970-2000:

  - *bio5*, the maximum temperature of the warmest month (°C)
  - *bio6*, the minimum temperature of the coldest month (°C)
  - *bio12*, the annual precipitation (mm)
  - *bio15*, the precipitation seasonality (coefficient of variation)
- Soil-related variable:
  - *depth_roots*, the depth available to roots (cm)
  - *water_top*, the total available water content (mm)
- One topographic variable: *TRI* the topographic ruggedness index
- One fire-related variable: *BurnedArea* the average of the monthly burned area from June 1995 to December 2014 (hectares)

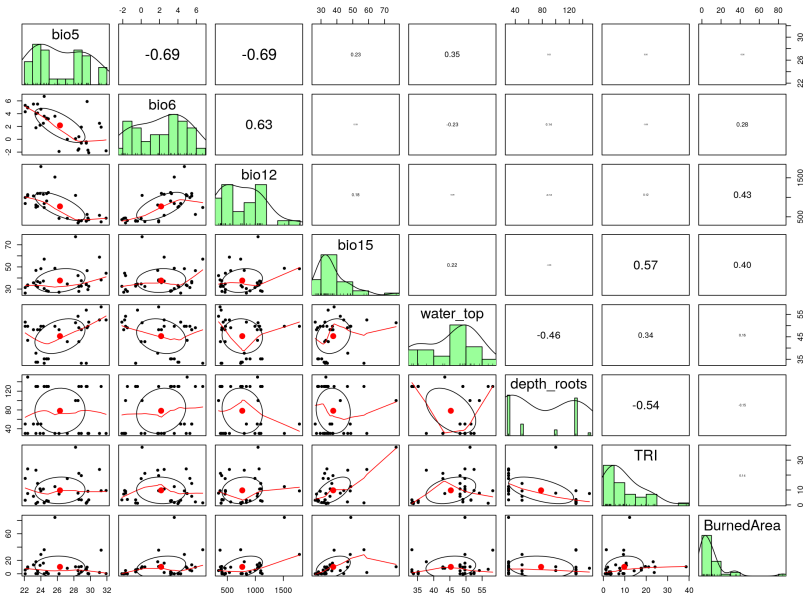
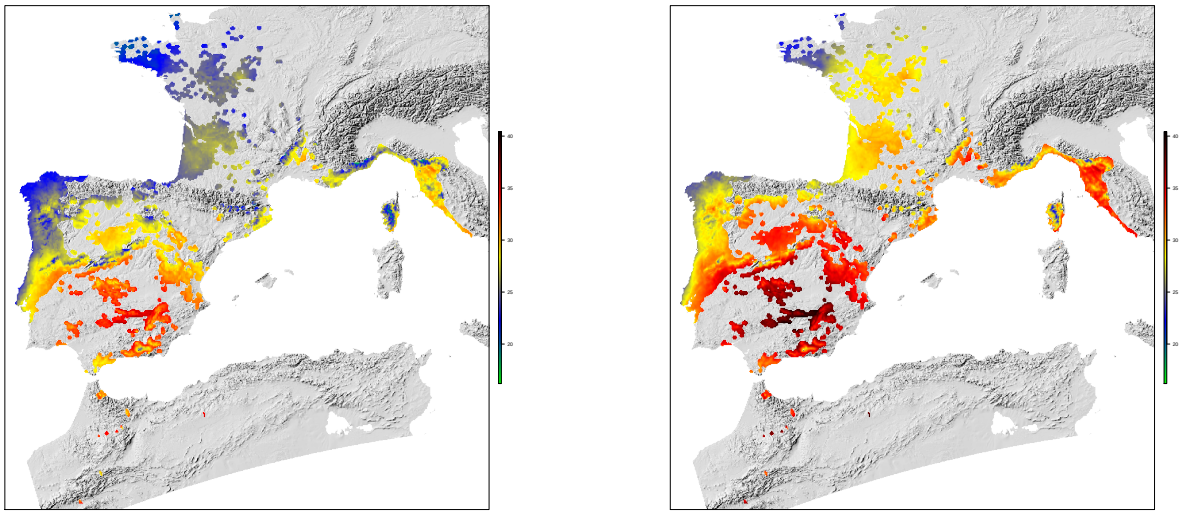
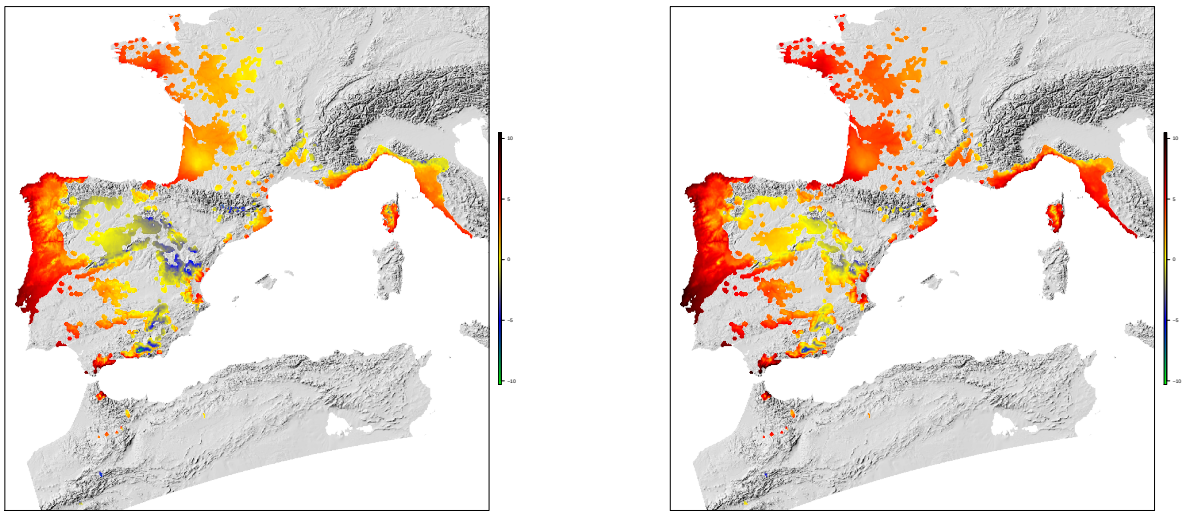


Figure XII.1. Distribution and correlation of the environmental covariates. This plot was generated with the function *pairs.panels* of the R package *psych*. Bivariate scatter plots are shown below the diagonal, histograms on the diagonal, and Pearson correlation coefficients above the diagonal.



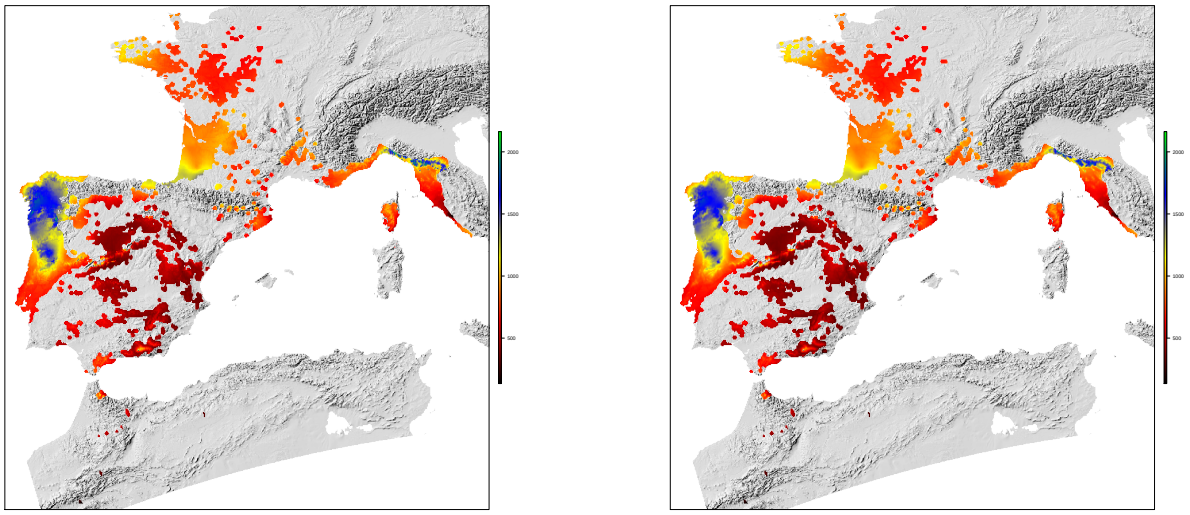
(a) Current climate. **(b)** Future climate (scenario SSP3-7.0).

Figure XII.2. Spatial variation in *bio5*, the **maximum temperature of the warmest month** (°C).



(a) Current climate. **(b)** Future climate (scenario SSP3-7.0).

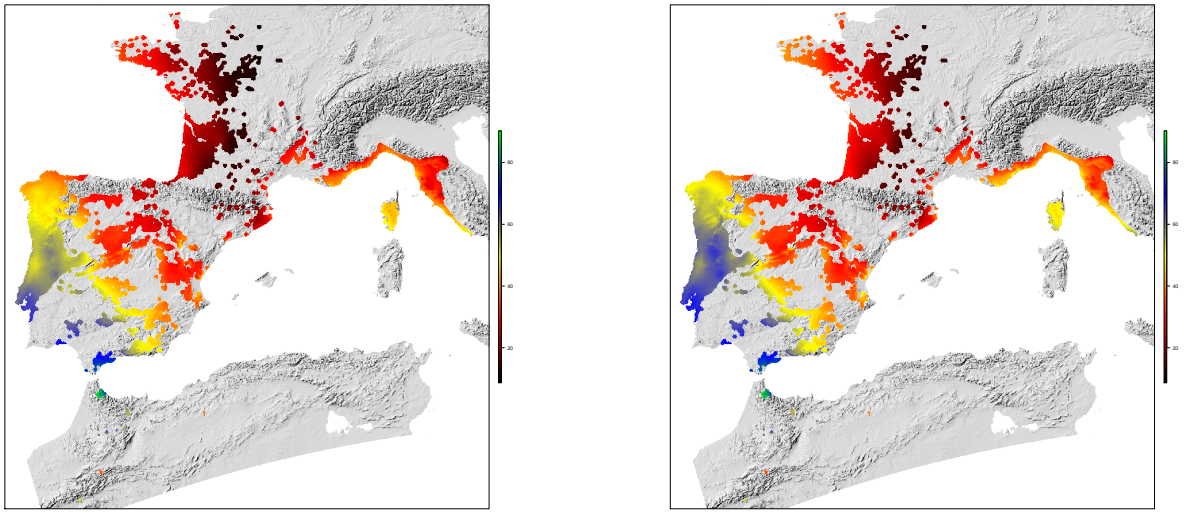
Figure XII.3. Spatial variation *bio6*, the **minimum temperature of the coldest month** (°C).



(a) Current climate.

(b) Future climate (scenario SSP3-7.0).

Figure XII.4. Spatial variation in *bio12*, the **annual precipitation** (mm).



(a) Current climate.

(b) Future climate (scenario SSP3-7.0).

Figure XII.5. Spatial variation in *bio15*, the **precipitation seasonality** (coefficient of variation).

To describe the potential future climate, we used the averaged predictions (over the period 2041-2060 at 2.5 arc-minutes spatial resolution) of nine global climate models (GCMs) from the WorldClim database (Fick and Hijmans 2017): BCC-CSM2-MR, CNRM-CM6-1, CNRM-ESM2-1, CanESM5, GFDL-ESM4, IPSL-CM6A-LR, MIROC-ES2L, MIROC6, MRI-ESM2-0. For the shared socio-economic pathway SSP3-7.0 (moderately alarming), the nine GCMs were used. For the shared socio-economic pathway SSP5-8.5 (strongly alarming), all GCMs except GFDL-ESM4 were used.

1.3 Validating genomic offset predictions in common gardens

In this part, we aimed to estimate the association of population performance in common gardens with (i) the predicted genomic offset of the populations in the new environmental

conditions of the common gardens and (ii) the climatic transfer distance of the populations (i.e. absolute difference between the climate in the location of origin of each population and the climate in the common garden).

As a measure of population performance in the common gardens, we used height data from five common gardens and mortality data from two common gardens. Commons gardens where height was measured were planted in different environments, i.e three under favorable conditions (i.e. the Atlantic region with mild winters, no severe cold events, high annual rainfall and relatively wet summers) and two in harsh environments (i.e. the Mediterranean region with high temperatures and an intense summer drought). Mortality was measured in the two common gardens under harsh environments in which a severe summer drought exacerbated by clay soils killed 92% and 72% of the trees. Details about the number of height and mortality observations per population in each common garden are given in Tables XII.2 and XII.3.

Pop.	Cáceres (8 months)			Madrid (13 months)		
	Nb of dead trees	Total nb of trees	Dead proportion	Nb of dead trees	Total nb of trees	Dead proportion
ALT	69	72	95.83	52	72	72.22
ARM	56	64	87.50	49	64	76.56
ARN	129	136	94.85	107	136	78.68
BAY	132	144	91.67	113	144	78.47
BON	62	72	86.11	48	72	66.67
CAD	74	80	92.50	62	80	77.50
CAR	44	48	91.67	36	48	75.00
CAS	78	80	97.50	66	80	82.50
CEN	66	72	91.67	50	72	69.44
COC	137	144	95.14	118	144	81.94
COM	28	32	87.50	21	32	65.62
CUE	211	224	94.20	186	224	83.04
HOU	186	208	89.42	145	208	69.71
LAM	68	72	94.44	57	72	79.17
LEI	162	184	88.04	147	184	79.89
MAD	8	8	100.00	5	8	62.50
MIM	135	144	93.75	116	144	80.56
OLB	149	176	84.66	113	176	64.20
OLO	173	192	90.10	145	192	75.52
ORI	196	208	94.23	162	208	77.88
PET	171	192	89.06	136	192	70.83
PIA	111	128	86.72	91	128	71.09
PIE	67	72	93.06	50	72	69.44
PLE	155	160	96.88	131	160	81.88
PUE	58	64	90.62	50	64	78.12
QUA	123	136	90.44	83	136	61.03
SAC	65	72	90.28	56	72	77.78
SAL	104	112	92.86	85	112	75.89
SEG	154	168	91.67	135	168	80.36
SIE	59	64	92.19	40	64	62.50
STJ	190	224	84.82	154	224	68.75
TAM	114	120	95.00	107	120	89.17
VAL	87	96	90.62	69	96	71.88
VER	197	216	91.20	153	216	70.83

Table XII.2. Number of dead trees, total number of trees and proportion of dead trees in the 34 populations used in the validation step in two common gardens: Cáceres at 8-month old and Madrid at 13-month old.

Pop.	Asturias (10 months)		Asturias (37 months)		Bordeaux (25 months)		Bordeaux (85 months)		Cáceres (8 months)		Madrid (13 months)		Portugal (11 months)		Portugal (27 months)	
	Count	Height	Count	Height	Count	Height	Count	Height	Count	Height	Count	Height	Count	Height	Count	Height
ALT	70	310.43	69	1366.09	52	760.19	52	4275.38	3	180.00	20	175.75	72	242.71	53	709.62
ARM	63	353.33	63	1491.59	63	793.33	63	4450.16	8	157.50	15	172.00	64	239.92	41	710.98
ARN	131	251.37	131	1150.69	107	651.87	105	3302.00	7	154.29	29	171.90	136	189.34	90	601.56
BAY	134	239.85	136	1059.56	140	618.36	139	3054.03	10	147.00	31	180.32	144	183.96	88	553.64
BON	67	287.01	69	1217.39	47	600.43	47	3273.83	9	184.44	24	211.04	72	201.94	58	613.79
CAD	75	306.40	75	1310.93	67	847.76	67	4631.79	5	142.00	18	180.56	80	228.38	58	687.07
CAR	48	240.62	48	1015.62	47	505.32	46	2321.30	3	146.67	12	156.67	48	194.17	27	581.85
CAS	76	307.76	76	1329.61	73	753.01	72	4202.78	2	165.00	14	156.43	80	218.00	49	716.12
CEN	68	282.35	70	1268.86	37	663.51	37	3348.65	6	155.00	22	167.27	72	235.49	54	686.85
COC	137	229.85	138	1063.70	108	587.13	105	2735.71	7	165.71	26	165.77	143	180.45	90	556.11
COM	28	341.43	32	1260.62	30	809.33	28	3220.71	4	182.50	11	249.09	32	223.28	22	703.18
CUE	216	229.68	215	991.07	187	587.38	183	2793.39	12	170.83	38	164.61	223	186.64	122	556.31
HOU	197	338.43	197	1444.92	154	841.95	154	4578.18	20	216.50	63	223.57	208	257.19	142	740.35
LAM	68	308.97	69	1328.26	56	706.79	56	3824.82	3	133.33	15	178.00	72	209.79	45	620.22
LEI	175	314.80	170	1318.06	133	765.64	133	4156.02	18	177.78	37	177.57	184	217.09	120	661.83
MAD	8	322.50	8	1242.50	8	572.50	8	2687.50	-	-	3	316.67	8	285.00	6	576.67
MIM	137	288.32	137	1299.20	124	783.87	121	4182.64	9	164.44	28	215.71	144	198.26	77	624.55
OLB	174	293.05	173	1228.15	113	700.35	112	3410.89	24	227.08	63	241.35	176	248.07	149	615.70
OLO	175	331.60	169	1401.95	129	803.57	128	4390.16	18	224.44	47	223.94	192	262.97	119	743.28
ORI	202	238.17	204	1150.34	181	667.40	181	2935.64	11	166.36	46	177.93	208	178.29	141	547.09
PET	178	325.56	181	1345.64	144	829.51	144	4471.39	20	213.50	56	233.66	191	235.34	117	713.76
PIA	121	325.21	120	1430.58	109	838.62	109	4163.58	17	189.41	37	180.68	127	212.20	85	712.00
PIE	68	250.74	69	1136.38	52	630.77	51	3318.63	5	154.00	22	160.91	71	194.58	55	629.27
PLE	152	281.71	152	1231.64	130	770.15	128	4044.84	5	228.00	29	193.45	160	201.22	89	646.29
PUE	62	312.90	63	1396.51	51	865.10	51	4460.59	5	142.00	14	150.00	64	209.38	34	770.00
QUA	131	286.34	131	1221.60	115	706.61	114	3265.79	12	189.17	53	201.04	136	225.96	100	637.40
SAC	67	300.30	71	1238.31	50	779.80	49	4317.55	5	206.00	16	156.88	47	173.94	37	622.43
SAL	107	241.96	105	1066.00	69	582.90	69	2915.65	8	151.25	27	177.96	112	158.93	57	559.47
SEG	159	289.75	156	1289.49	161	819.44	161	4484.53	13	151.54	33	151.82	168	187.62	91	651.65
SIE	64	293.91	62	1302.90	46	860.65	46	4266.96	5	174.00	24	181.04	64	193.12	35	643.71
STJ	205	329.66	209	1394.11	177	893.84	177	4559.44	32	232.81	70	241.64	224	261.58	154	732.01
TAM	109	225.69	112	991.61	69	570.29	67	2250.45	6	151.67	13	177.69	119	165.34	60	502.00
VAL	92	260.65	91	1179.57	61	672.13	59	3393.73	9	180.00	27	163.52	96	200.89	67	612.54
VER	205	331.12	205	1361.05	156	813.65	155	4401.61	19	221.58	63	233.97	215	244.33	133	668.12

Table XII.3. Number of trees and mean height of the 34 populations used in the validation step in five common gardens. Height measurements were taken in Asturias at 10 and 37-month old, in Bordeaux at 25 and 85-month old, Cáceres at 8-month old, Madrid at 13-month old and Portugal at 11 and 27-month old.

For each of the eight combinations of the four allele sets (reference SNPs and the three candidate SNP sets) and the two models used to estimate the current gene-environment relationships (i.e. GDM and GF), we predicted genomic offset of the 34 populations when transplanted in the common gardens based on the environmental differences (i.e. climatic, soil and topographic differences) between the location of the population and the common garden. Differences in burned area were not accounted for in genomic offset calculation as fire could not have influenced the population performance (i.e. height and mortality) in the common gardens, and therefore the burned area value was fixed to its value at the population location.

We also calculated the climatic transfer distance of each pair of population and common garden for five climatic covariates: *bio1* (the annual daily mean temperature, °C), *bio5* (the maximum temperature of the warmest month, °C), *bio6* (the minimum temperature of the coldest month, °C), *bio12* (the annual precipitation, mm) and *bio15* (the precipitation seasonality, coefficient of variation).

We first evaluated whether populations that grow the less in common gardens were those with the highest predicted genomic offset (or the highest climatic transfer distance). For that, in each of the five common gardens independently, we first estimated BLUPs for height with the following model:

$$\begin{aligned} H_{ip} &\sim \mathcal{N}(\mu_p, \sigma_r^2) \\ \mu_p &= \beta_0 + P_p + B_b \end{aligned} \tag{1.1}$$

with H_{ip} the height in the individual i in the population p, β_0 the global intercept, σ_r^2 the residual variance, P_p and B_b the population and block varying intercepts, respectively. This model was performed with the R package *brms* and we used to following priors:

$$
\begin{aligned}
\beta_0 &\sim \mathcal{N}(0, 10) \\
\begin{bmatrix} P_p \\ B_b \end{bmatrix} &\sim \mathcal{N}\left(0, \begin{bmatrix} \sigma_B^2 \\ \sigma_P^2 \end{bmatrix}\right) \\
(\sigma_r, \sigma_B, \sigma_P)^\intercal &\sim \text{Cauchy}(0, 10)
\end{aligned}
\tag{1.2}
$$

We extracted the mean mP_p of the posterior distributions of P_p (i.e. the BLUPs) to estimate the association between mP_p and the genomic offset or the climatic transfer distance, as follows:

$$
\begin{aligned}
mP_p &\sim \mathcal{N}(\mu_p, \sigma_r^2) \\
\mu_p &= \beta_0 + \beta_{X1} X_p + \beta_{X2} X_p^2
\end{aligned}
\tag{1.3}
$$

with β_0 the global intercept, σ_r^2 the residual variance and X_p the value of the genomic offset or climatic transfer distance for the population p. We included a quadratic term for X_p to allow for potential nonlinearity in the response, following Fitzpatrick et al. (2021). We used the following weakly informative priors:

$$
\begin{aligned}
\begin{bmatrix} \beta_0 \\ \beta_{X_1} \\ \beta_{X_2} \end{bmatrix} &\sim \mathcal{N}(0, 1) \\
\sigma_r &\sim \text{Exponential}(1)
\end{aligned}
\tag{1.4}
$$

Secondly, we evaluated whether populations that died more in common gardens were those with the highest predicted genomic offset (or the highest climatic transfer distance) with the following model:

$$
\begin{aligned}
a_p &\sim \text{Binomial}(N_p, p_p) \\
\text{logit}(p_p) &= \beta_0 + \beta_H H_p + \beta_{X1} X_p
\end{aligned}
\tag{1.5}
$$

with a_p the count of individual that died in the population p, N_p the total number of individuals in the population p (=number of individuals that were initially planted in the common garden), p_p is the estimated probability of mortality in the population p, X_p is the genomic offset or climatic transfer distance for the population p and H_p is the BLUPs for height of the population p (population varying intercepts calculated across all common gardens in the model 1 of Archambeau et al. 2021a). We included H_p as a covariate in the model to account for height differences before planting, as smaller trees had a higher mortality probability than taller trees. We used the following weakly informative priors:

$$
\begin{bmatrix} \beta_0 \\ \beta_H \\ \beta_{X_1} \end{bmatrix} \sim \mathcal{N}(0, 5)
\tag{1.6}
$$

We compared the proportion of variance explained (i.e. R^2, a measure of the model goodness-of-fit) of the height models and the predictive ability of both mortality and height models was evaluated with the leave-one-out cross-validation (LOOCV) procedure from the R package *loo*.

1.4 Validating genomic offset predictions in natural populations

In this part, we aimed to estimate the relationship between predicted genomic offset and mortality rates in natural populations across maritime pine range. We used mortality data from the French and Spanish National Forest Inventories (NFI) harmonized in Changenet et al. 2021.

The French data relies on temporary plots sampled between 2005 and 2014 while the Spanish data relies on permanent plots sampled during the second (from 1986 to 1996) and third NFIs (from 1997 to 2008). A tree was recorded as dead if its death was dated less than 5 years ago in the French NFI, or if it was alive in the second inventory but dead in the third one in the Spanish NFI.

We modeled the proportion p_i of maritime pines that died in the plot i during the census interval Δ_i with a complementary log-log link as follows:

$$\begin{aligned} m_i &\sim \text{Binomial}(N_i, p_i) \\ \log(-\log(1 - p_i)) &= \beta_{0,c} + \beta_{C,c} C_i + \beta_{GO,c} GO_i + \log(\Delta_i) \end{aligned} \tag{1.7}$$

with N_i the total number of maritime pines in the plot i, m_i the number of maritime pines that died during the census interval Δ_i in the plot i, C_i the basal area of all tree species confounded in the plot i (to account for the competition between trees) and GO_i the genomic offset predicted in the plot i. As the French and Spanish inventories present noticeable methodological differences that may bias the estimations, we estimated country-specific coefficients: the country-specific intercepts $\beta_{0,c}$ and the country-specific slopes $\beta_{C,c}$ and $\beta_{GO,c}$. We used the complementary log-log link jointly with the logarithm of the census interval Δ_i for the plot i to account for the different census intervals between inventories. We used the following weakly informative priors:

$$\begin{bmatrix} \beta_{0,c} \\ \beta_{C,c} \\ \beta_{GO,c} \end{bmatrix} \sim \mathcal{N}(0, 1) \tag{1.8}$$

The present model was performed for each of the sixteen combinations of the four allele sets (i.e. reference SNPs and the three candidate SNP sets), the two models used to estimate the current gene-environment relationships (i.e. GDM and GF) and the two scenarios of future climates (i.e. SSP3-7.0 and SSP5-8.5).

2 Results

2.1 Candidate SNPs identification

In BayPass, the candidate SNPs were selected based on a 5 dB threshold for the median Bayes Factor calculated over 5 independent runs, resulting in the identification of 26 candidates (1 associated with the maximum temperature of the warmest month, and 25 with the minimum temperature of the coldest month). For RDA, candidate SNPs were selected using a threshold of three standard deviations to identify outliers in the distribution of the SNP loadings on each significant RDA axis. The **common candidates** are the candidate SNPs selected by both GEA methods. The **candidates under expected strong selection** are the RDA candidates that show a strong association with at least one covariate, i.e. with $\beta_{RDA} > 0.3$, and all the BayPass candidates. The **merged candidates** are candidates selected by at least one of the two GEA methods.

	BayPass								RDA	TOTAL
	bio5	*bio6*	*bio12*	*bio15*	*depth_roots*	*water_top*	*TRI*	*BurnedArea*		
Common candidates	1	7	0	0	0	0	0	0	8	**8**
Candidates under expected strong selection	1	25	0	0	0	0	0	0	61	**79**
Merged candidates	1	25	0	0	0	0	0	0	352	**370**

Table XII.4. Candidate SNPs identified by each genotype–environment association (GEA) method.

β_{RDA} **threshold**	*bio5*	*bio6*	*bio12*	*bio15*	*water_top*	*depth_roots*	*TRI*	*BurnedArea*
0.10	190	175	154	202	153	136	179	118
0.30	17	11	19	10	13	3	4	6

Table XII.5. Number of candidate SNPs with β_{RDA} > 0.1 or 0.3 for each environmental covariate.

SNP sets	Mean	Median
Common candidates	0.07	0.01
Candidates under expected strong selection	0.03	0.01
Merged candidates	0.01	0.00
Reference SNPs	0.01	0.00

Table XII.6. Mean and median linkage disequilibrium in each SNP set.

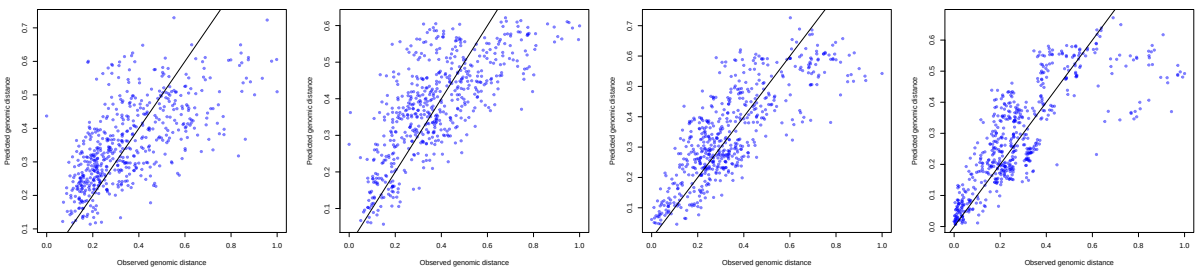
2.2 GDM and GF performance and covariate importance

In both the GDM and GF analyses, the maximum height of each fitted I-spline (GDM) or turnover function (GF) informs on the magnitude of genomic change along the gradient of the covariate considered, and therefore on the relative importance of that covariate in contributing to the genomic turnover while holding all other covariates constant.

2.2.1 GDM models

		Model performance				Covariate relative importance								
	Nb SNPs	*DevExp*	R^2_{CV9}	R^2_{CV6}	R^2_{CV2}	*Geographic*	*bio5*	*bio6*	*bio12*	*bio15*	*water_top*	*depth_roots*	*TRI*	*BurnedArea*
Common candidates	8	41.34	23.43 [3.00]	22.30 [3.70]	19.36 [6.15]	1.00	0.12	0.80	0.00	0.00	0.05	0.05	0.57	0.00
Candidates under expected strong selection	79	50.43	37.98 [2.59]	37.37 [2.98]	32.88 [5.95]	1.00	0.20	0.78	0.36	0.00	0.07	0.05	0.06	0.00
Merged candidates	370	62.56	50.02 [3.01]	49.09 [3.76]	40.18 [8.30]	1.00	0.02	0.27	0.10	0.09	0.04	0.02	0.40	0.00
Reference SNPs	9,817	63.58	50.45 [2.74]	48.21 [5.03]	43.09 [8.84]	1.00	0.00	0.05	0.08	0.10	0.03	0.00	0.30	0.00

Table XII.7. Number of SNPs in each set, model performance and covariate relative importance in the GDM analysis. *DevExp* is the percentage of deviance explained by the model. R^2_{CV9}, R^2_{CV6} and R^2_{CV2} are the mean and standard deviation in brackets of the coefficient of determination of 9-fold, 6-fold and 2-fold cross-validations repeated across 100 independent samples. The other columns correspond to the relative importance (scaled between 0 and 1) of the covariates in the GDM models. The names and units of the environmental covariates are given in section 1.2 of the Supplementary Information.



(a) Common candidates. **(b)** Candidates under expected strong selection. **(c)** Merged candidates. **(d)** Reference SNPs.

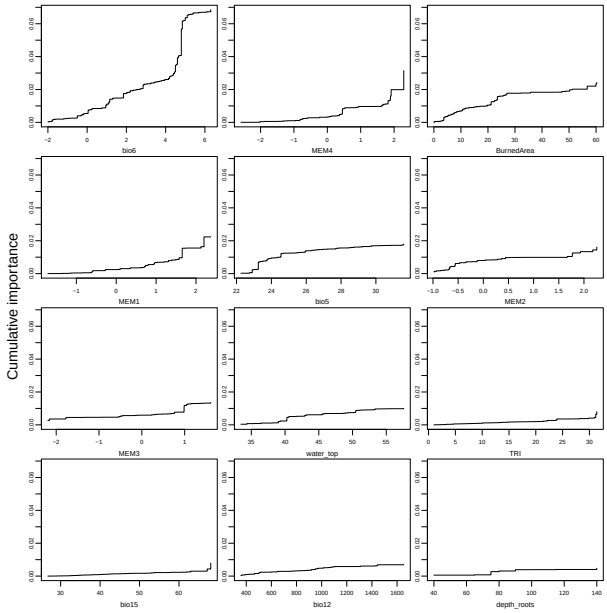
Figure XII.6. Predicted genomic distance versus observed genomic distance. The dots represent the population pairs and the line indicates where observations and predictions match.

2.2.2 GF models

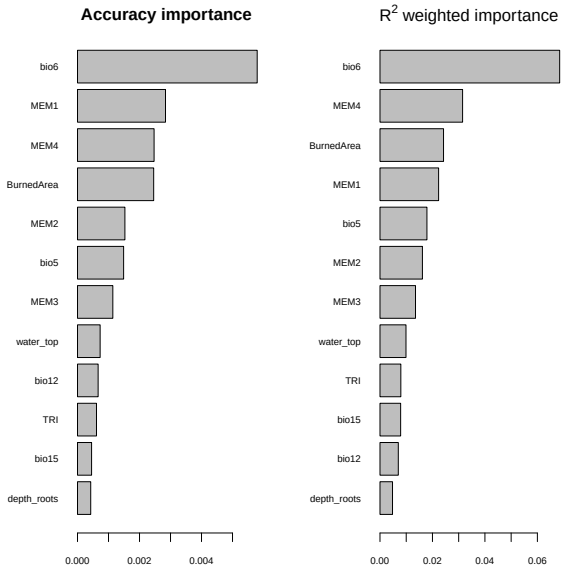
	Nb SNPs poly > 5	mean $\mathcal{R}^2$	minimum $\mathcal{R}^2$	maximum $\mathcal{R}^2$
Common candidates	8	0.23	0.00	0.51
Candidates under expected strong selection	78	0.37	0.02	0.69
Merged candidates	348	0.25	0.01	0.70
Reference SNPs	9650	0.29	0.00	0.88

Table XII.8. Number of SNPs that were polymorphic in more than five populations and were thus used in the GF analysis, and mean, minimum and maximum $\mathcal{R}^2$ across all SNPs in the GF models.

For each GF model performed on the four sets of SNPs, the figures below correspond to (a) the turnover functions for each environmental covariate and the first four Moran's eigen vectors (presented in order of importance), and (b) the overall importance of each covariate (i.e. mean accuracy importance and mean importance weighted by SNP $\mathcal{R}^2$). The names and units of the environmental covariates are given in section 1.2 of the Supplementary Information.

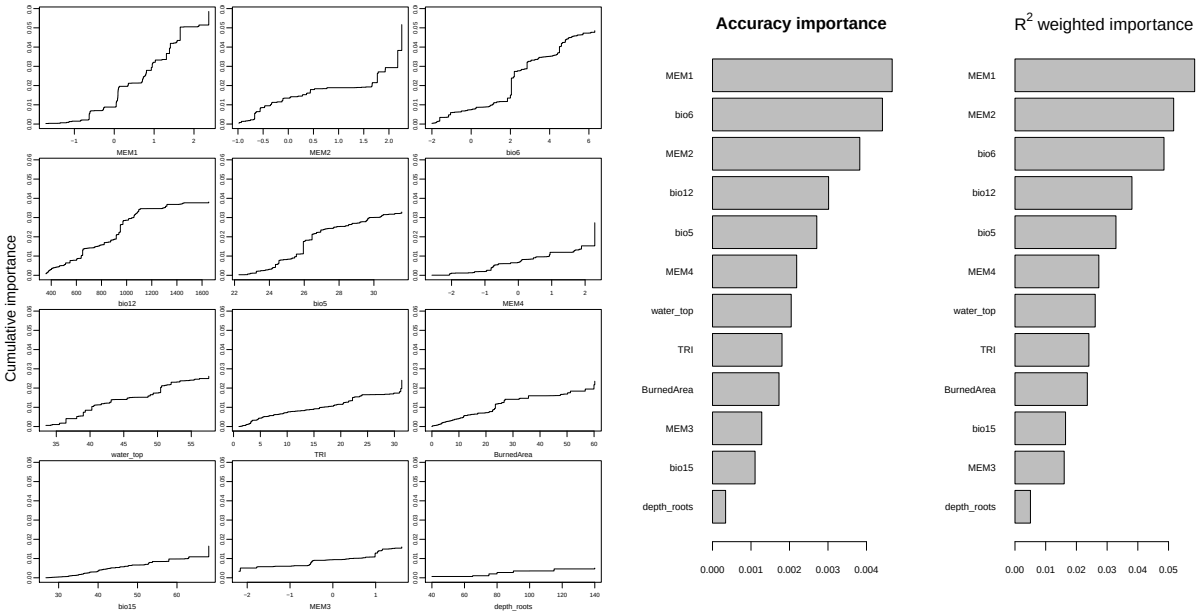


(a) Turnover functions.



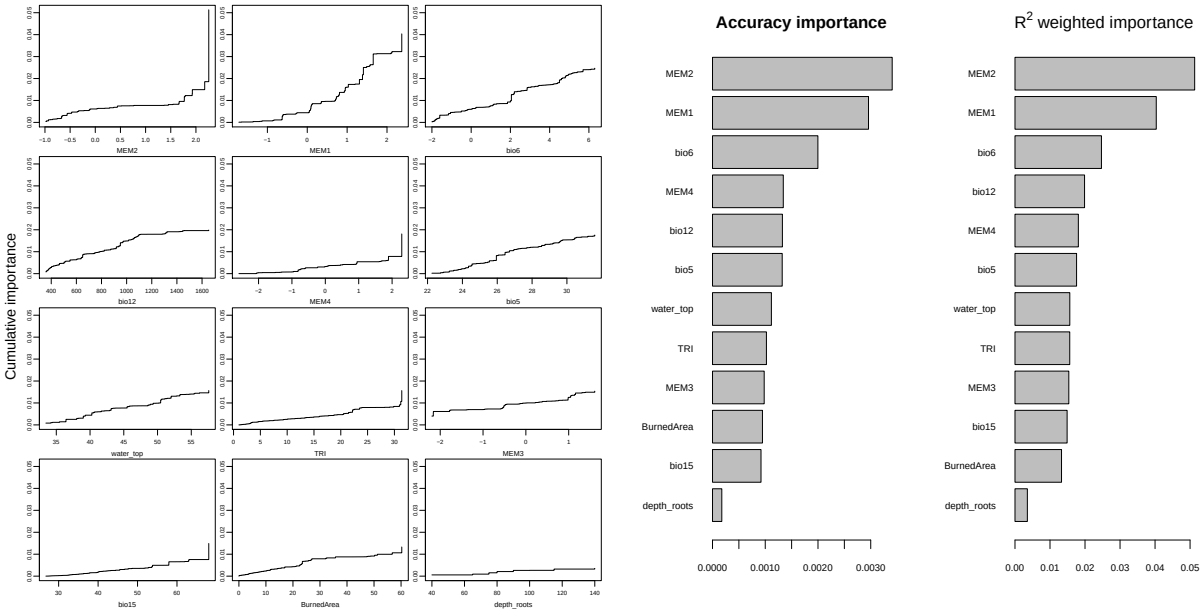
(b) Overall importance of the covariates.

Figure XII.7. Common SNP candidates.



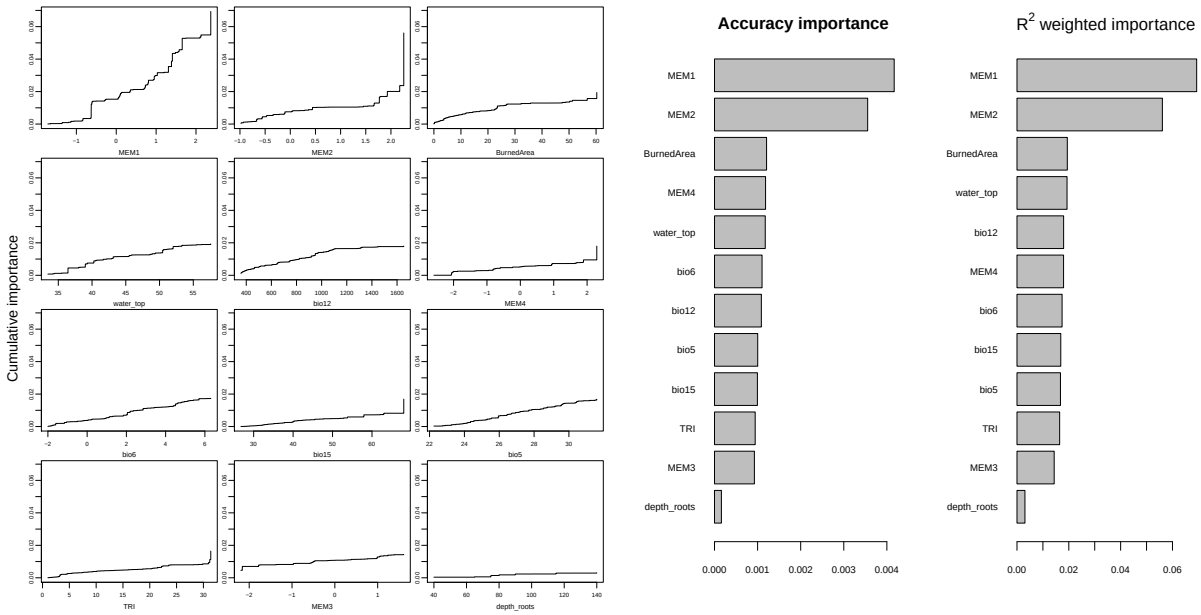
(a) Turnover functions. **(b)** Overall importance of the covariates.

Figure XII.8. Candidates under expected strong selection.



(a) Turnover functions. **(b)** Overall importance of the covariates.

Figure XII.9. Merged candidates.



(a) Turnover functions.

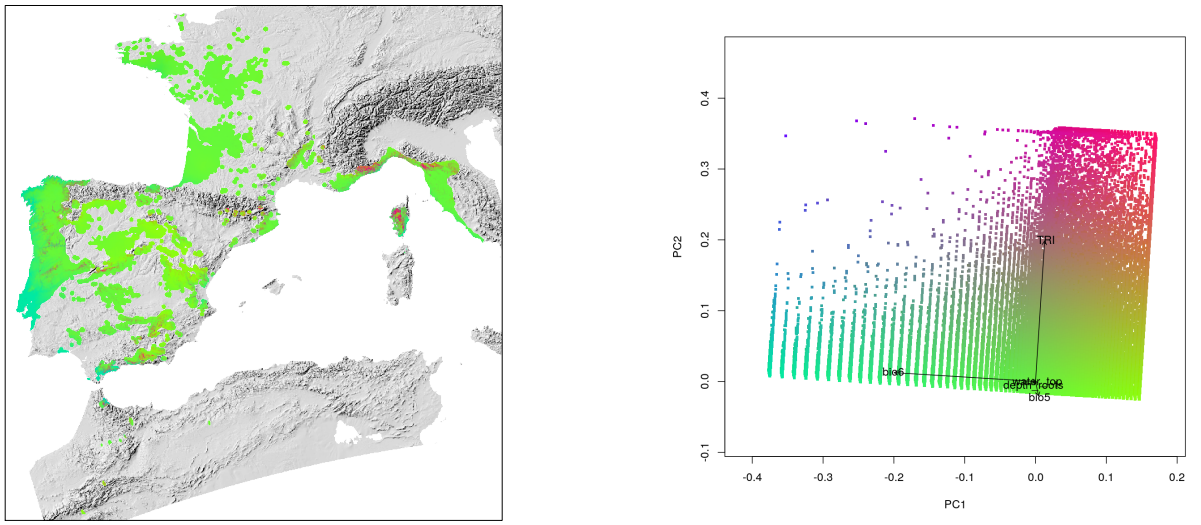
(b) Overall importance of the covariates.

Figure XII.10. Reference SNPs.

2.3 Predicted spatial variation in current and future genomic composition

For the projections of the genomic composition under future climates, only projections based on the moderately alarming scenario SSP3-7.0 are shown as those based on the strongly alarming scenario SSP5-8.5 are very similar.

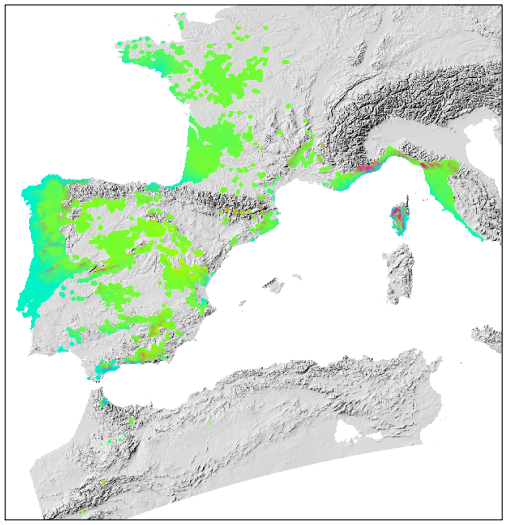
2.3.1 Common candidates



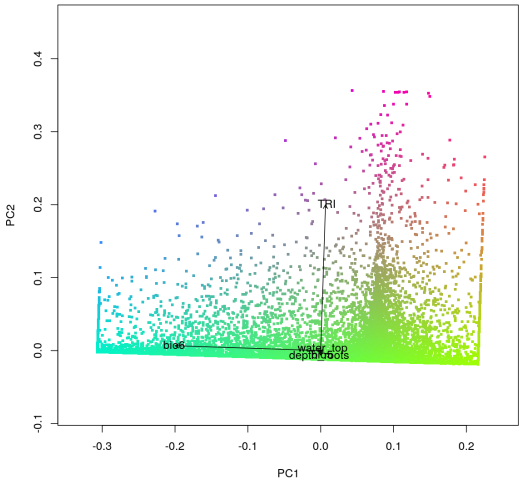
(a) Predicted spatial variation in genomic composition.

(b) PCA of the predicted variation in genomic composition

Figure XII.11. Genomic composition under **current climates** from **GDM models**.

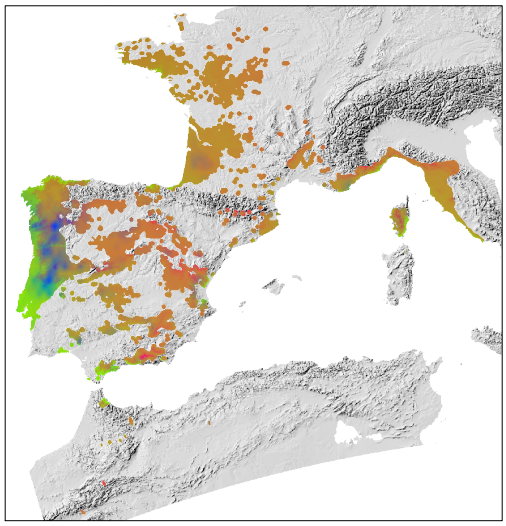


(a) Predicted spatial variation in genomic composition.

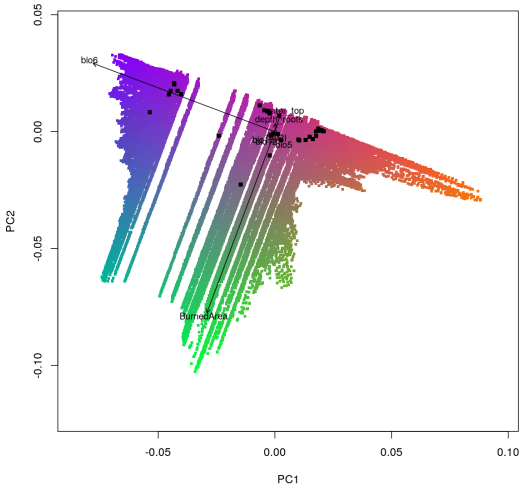


(b) PCA of the predicted variation in genomic composition

Figure XII.12. Genomic composition under **future climates** (scenario SSP3-7.0) from **GDM models**.

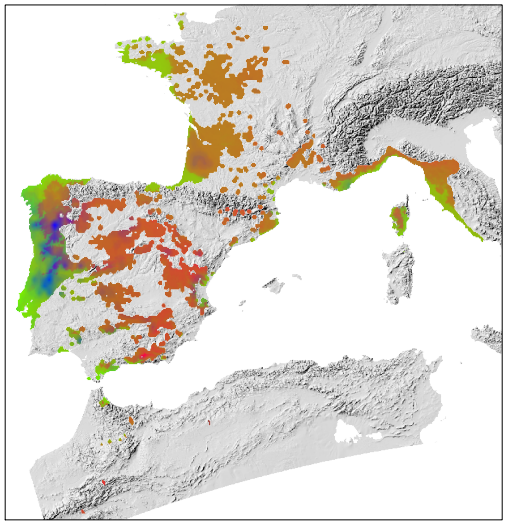


(a) Predicted spatial variation in genomic composition.

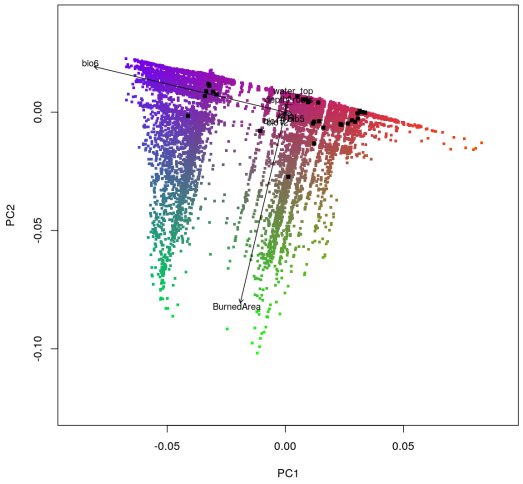


(b) PCA of the predicted variation in genomic composition

Figure XII.13. Genomic composition under **current climates** from **GF models**.



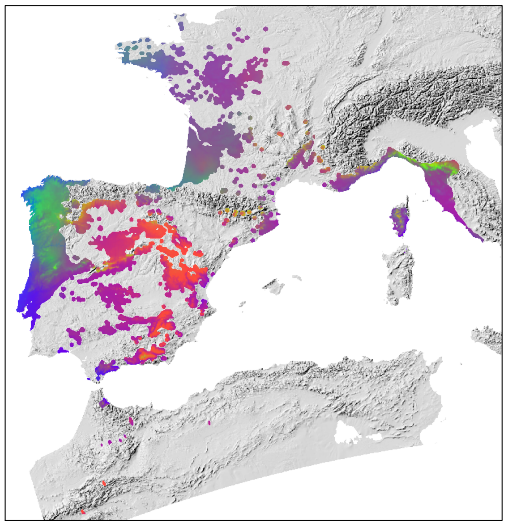
(a) Predicted spatial variation in genomic composition.



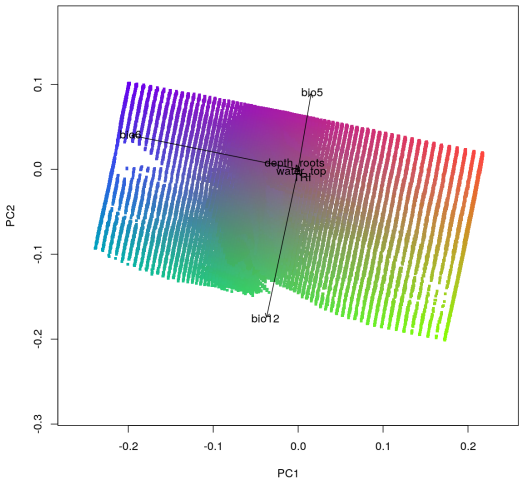
(b) PCA of the predicted variation in genomic composition

Figure XII.14. Genomic composition under **future climates** (scenario SSP3-7.0) from **GF models**.

2.3.2 Candidates under expected strong selection

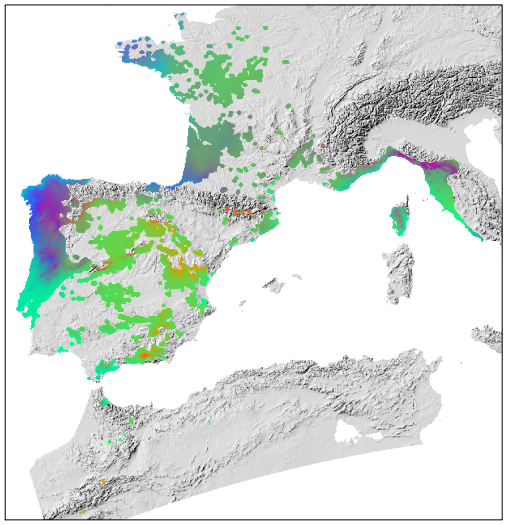


(a) Predicted spatial variation in genomic composition.

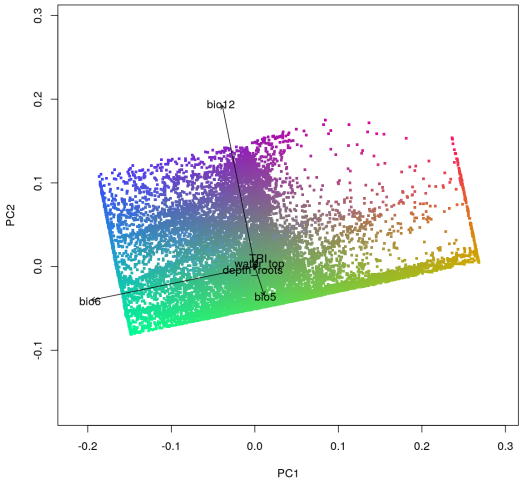


(b) PCA of the predicted variation in genomic composition

Figure XII.15. Genomic composition under **current climates** from **GDM models**.

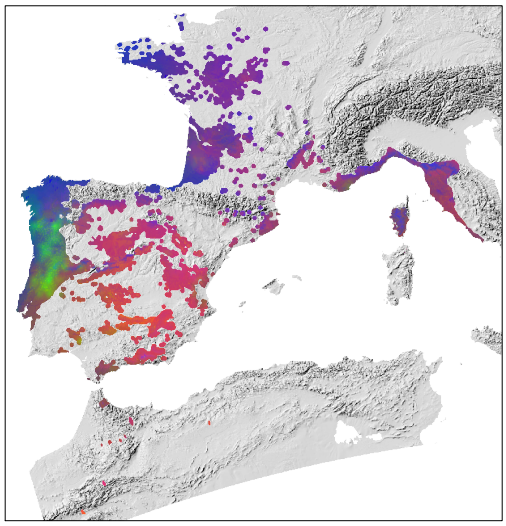


(a) Predicted spatial variation in genomic composition.

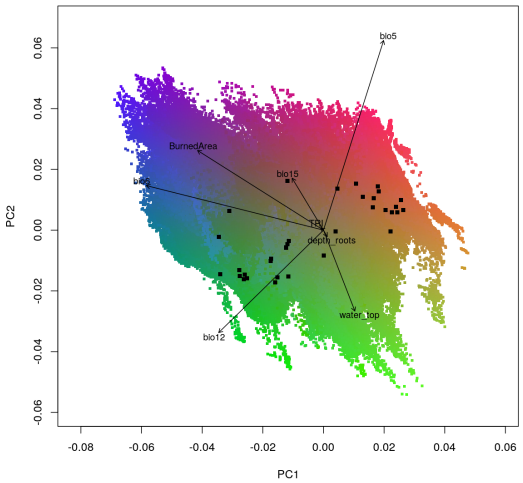


(b) PCA of the predicted variation in genomic composition

Figure XII.16. Genomic composition under **future climates** (scenario SSP3-7.0) from **GDM models**.

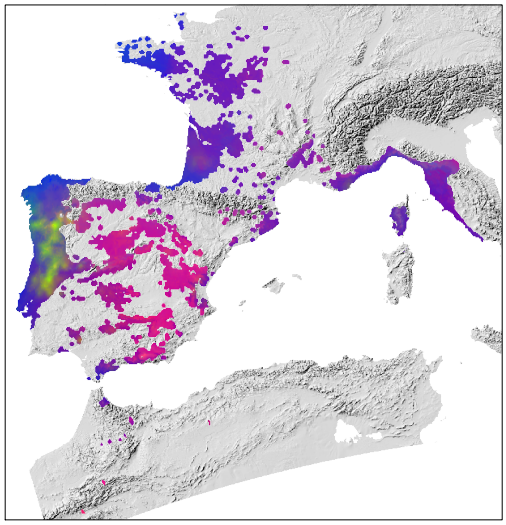


(a) Predicted spatial variation in genomic composition.

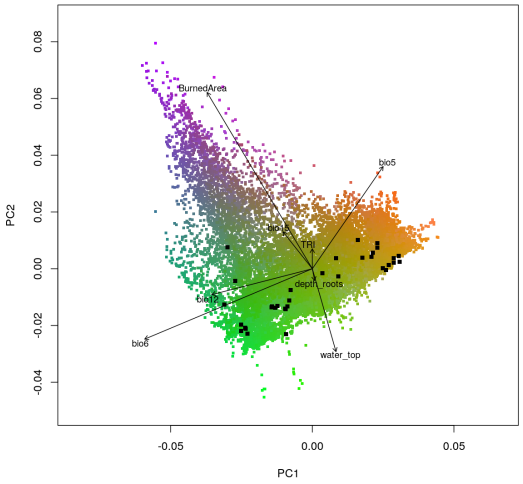


(b) PCA of the predicted variation in genomic composition

Figure XII.17. Genomic composition under **current climates** from **GF models**.



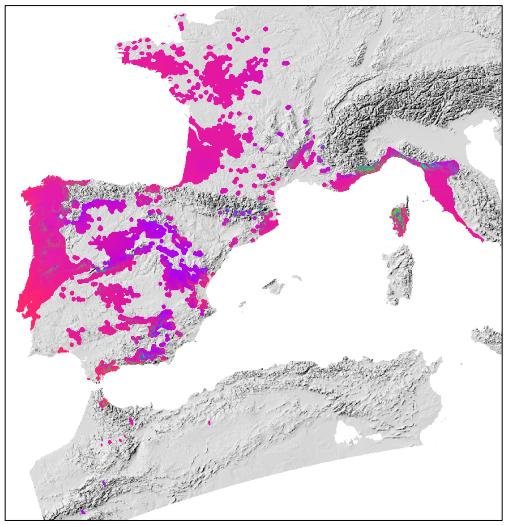
(a) Predicted spatial variation in genomic composition.



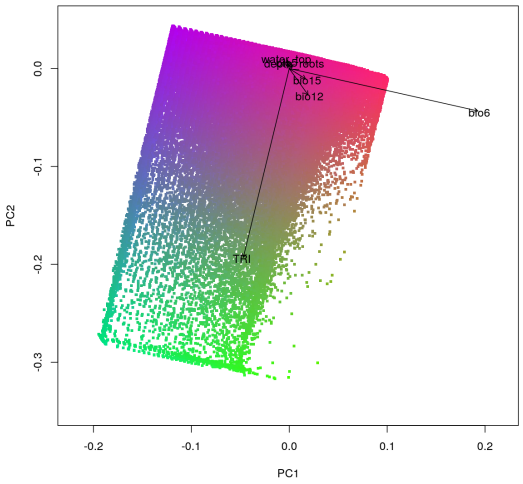
(b) PCA of the predicted variation in genomic composition

Figure XII.18. Genomic composition under **future climates** (scenario SSP3-7.0) from **GF models**.

2.3.3 Merged candidates

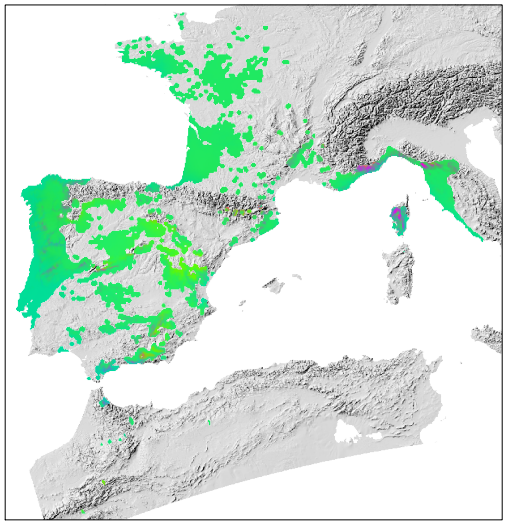


(a) Predicted spatial variation in genomic composition.

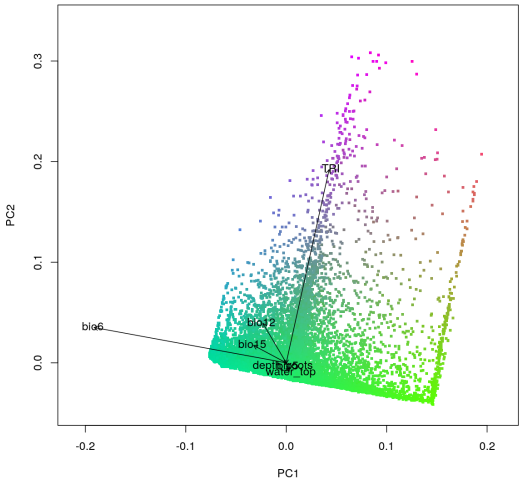


(b) PCA of the predicted variation in genomic composition

Figure XII.19. Genomic composition under **current climates** from **GDM models**.

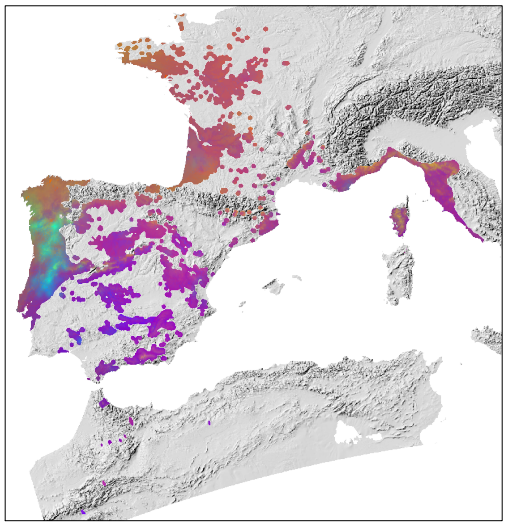


(a) Predicted spatial variation in genomic composition.

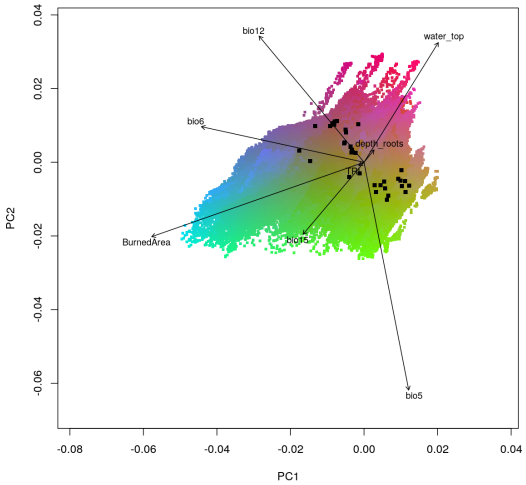


(b) PCA of the predicted variation in genomic composition

Figure XII.20. Genomic composition under **future climates** (scenario SSP3-7.0) from **GDM models**.

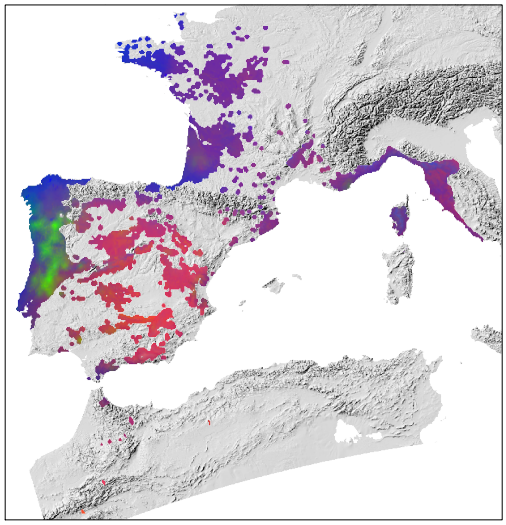


(a) Predicted spatial variation in genomic composition.

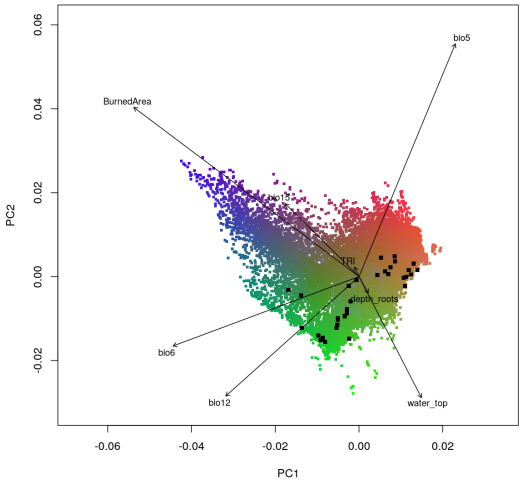


(b) PCA of the predicted variation in genomic composition

Figure XII.21. Genomic composition under **current climates** from **GF models**.



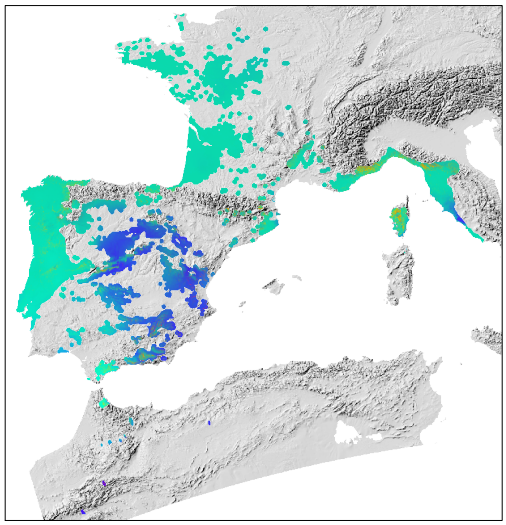
(a) Predicted spatial variation in genomic composition.



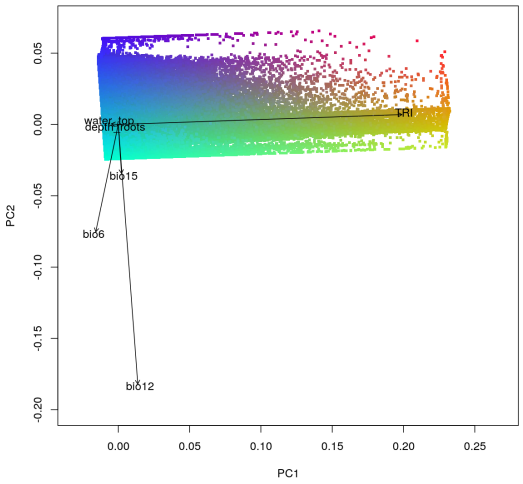
(b) PCA of the predicted variation in genomic composition

Figure XII.22. Genomic composition under **future climates** (scenario SSP3-7.0) from **GF models**.

2.3.4 Reference SNPs

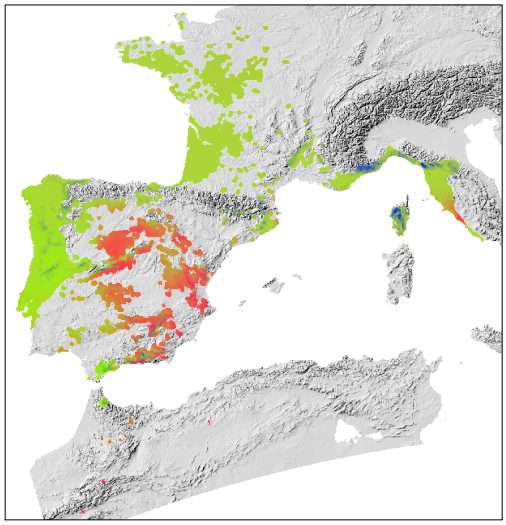


(a) Predicted spatial variation in genomic composition.

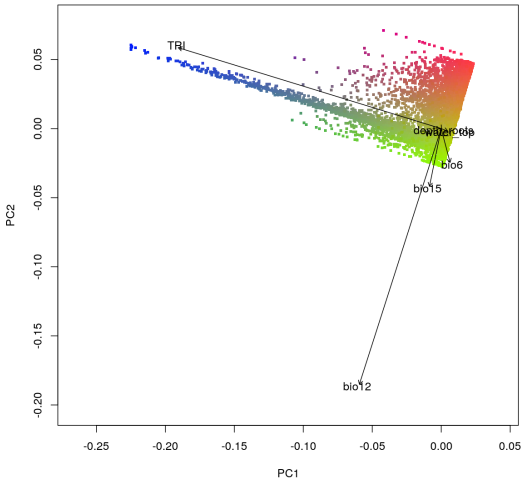


(b) PCA of the predicted variation in genomic composition

Figure XII.23. Genomic composition under **current climates** from **GDM models**.

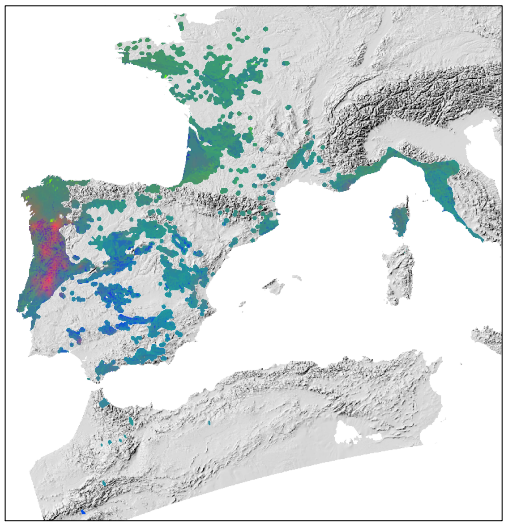


(a) Predicted spatial variation in genomic composition.

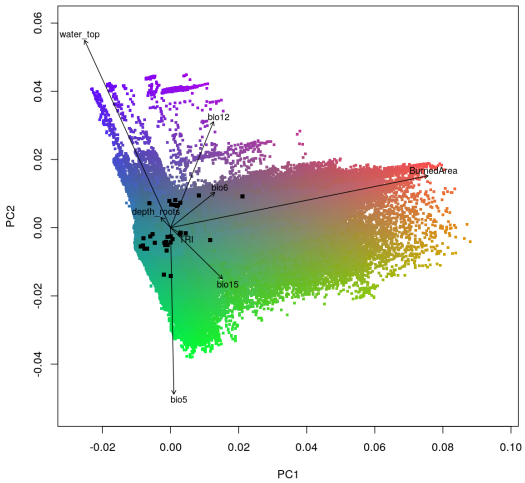


(b) PCA of the predicted variation in genomic composition

Figure XII.24. Genomic composition under **future climates** (scenario SSP3-7.0) from **GDM models**.

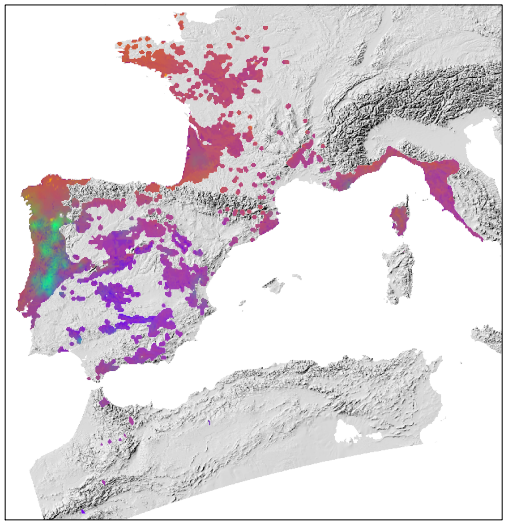


(a) Predicted spatial variation in genomic composition.

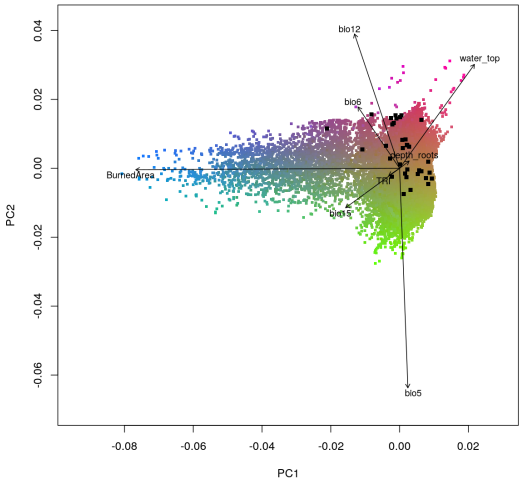


(b) PCA of the predicted variation in genomic composition

Figure XII.25. Genomic composition under **current climates** from **GF models**.



(a) Predicted spatial variation in genomic composition.



(b) PCA of the predicted variation in genomic composition

Figure XII.26. Genomic composition under **future climates** (scenario SSP3-7.0) from **GF models**.

2.4 Predicted spatial variation in genomic offset

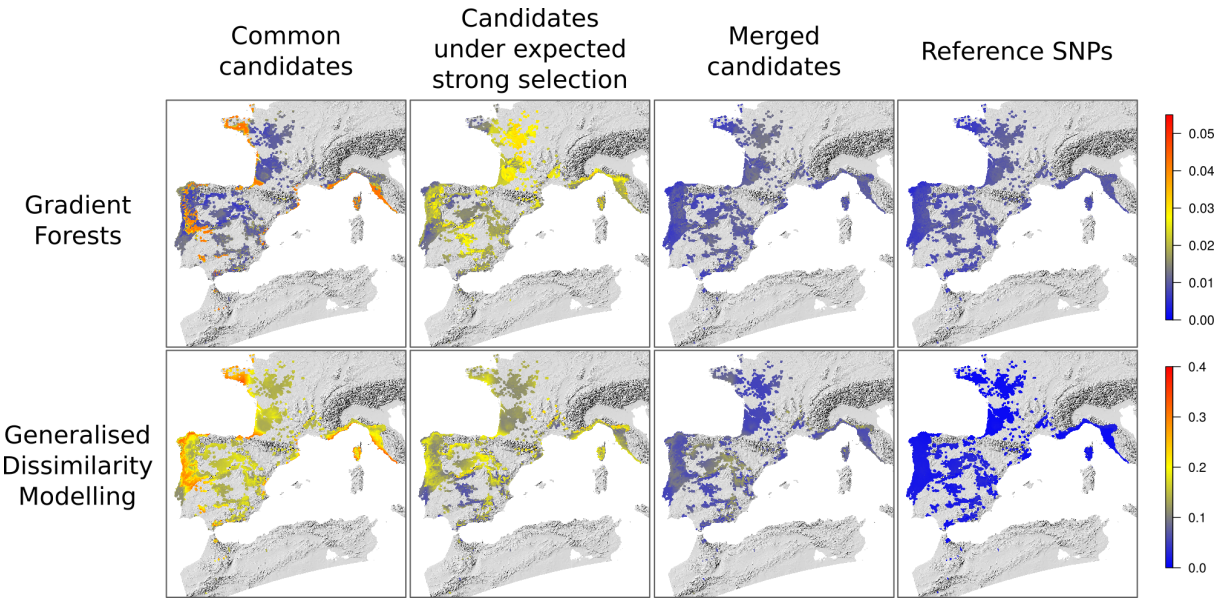


Figure XII.27. Predicted spatial variation in genomic offset for each combination of modelling approaches (i.e. Gradient Forests or Generalised Dissimilarity Modelling) and sets of SNPs (i.e. three sets of candidate SNPs and the reference SNPs) under the future climate scenario SSP5-8.5 (strongly alarming). See Figure XII.29 for the same predictions but visualized with different scales so that the spatial variation in genomic offset for the merged candidates and reference SNPs are visible.

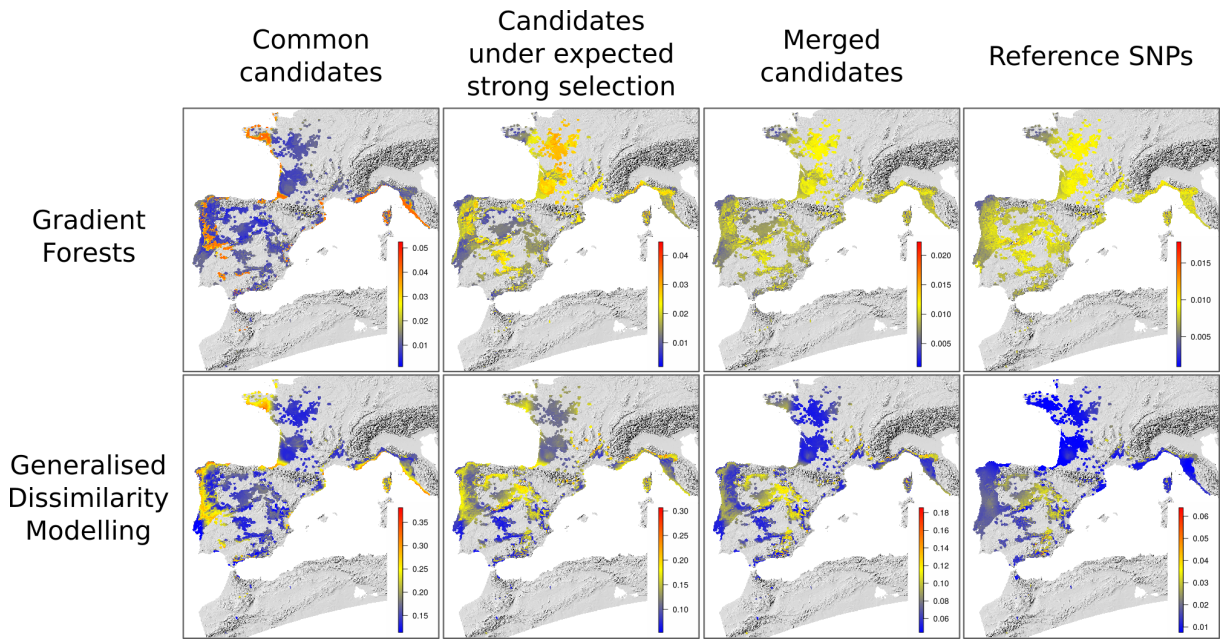


Figure XII.28. Same spatial predictions in genomic offset as in Figure VI.3 (future climate scenario SSP3-7.0) but visualized with different scales for each combination of modelling approaches (i.e. Gradient Forests or Generalised Dissimilarity Modelling) and sets of SNPs (i.e. three sets of candidate SNPs and the reference SNPs).

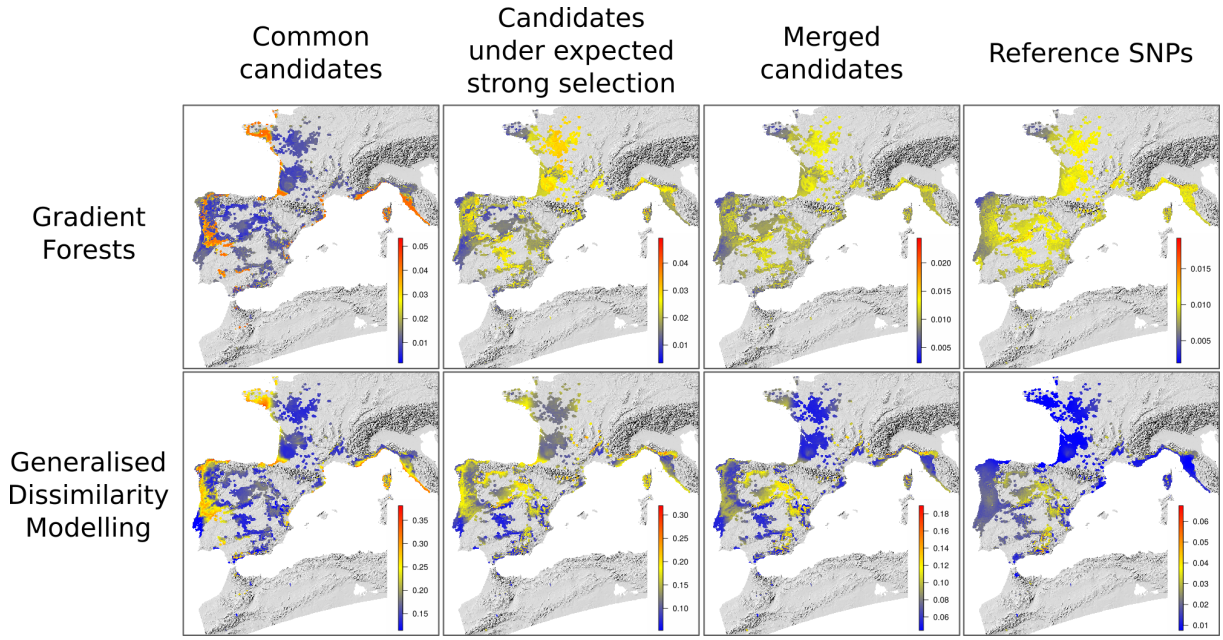


Figure XII.29. Same spatial predictions in genomic offset as in Figure XII.27 (future climate scenario SSP5-8.5) but visualized with different scales for each combination of modelling approaches (i.e. Gradient Forests or Generalised Dissimilarity Modelling) and sets of SNPs (i.e. three sets of candidate SNPs and the reference SNPs).

2.5 Validation in common gardens

In the validation part in the common gardens, we compared thirteen height and mortality models: five with a climatic transfer distance as covariate, four with a GDM-based predicted genomic offset and four with a GF-based predicted genomic offset. The five climatic variables used to calculated the climatic transfer distance tested were: the annual daily mean temperature (*bio1*; in °C), the maximum temperature of the warmest month (*bio5*; C°), the minimum temperature of the coldest month (*bio6*; °C), the annual precipitation (*bio12*; mm) and the precipitation seasonality (*bio15*; coefficient of variation). The GDM and GF-based genomic

offset correspond to the genomic offset predicted for each set of SNPs: the common candidates (Com), the candidates under expected strong selection (Mid), the merged candidates (Mer) and the reference SNPs (Ref).

2.5.1 Height models in five common gardens

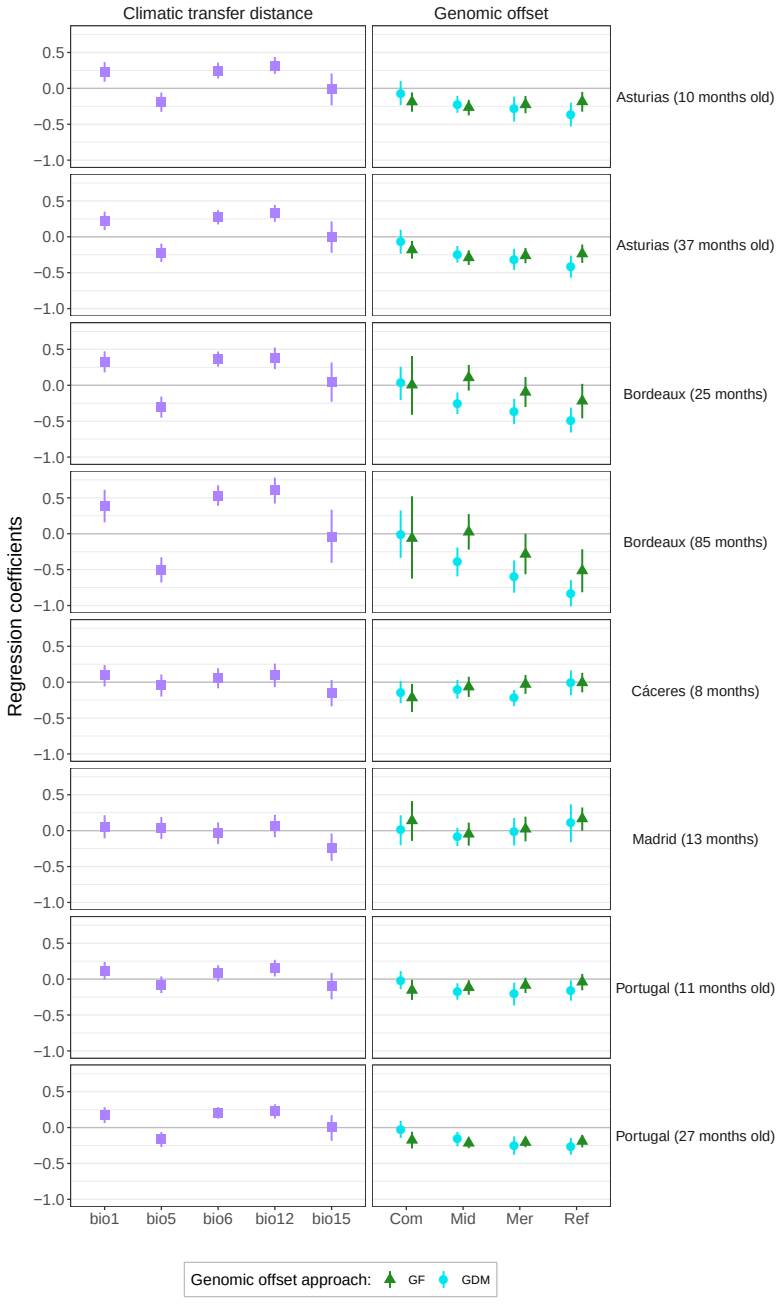


Figure XII.30. Regression coefficients corresponding to the linear association between tree height in five common gardens and the climatic transfer distances (left panels) or the predicted genomic offset (right panels). The climatic transfer distances were calculated based on five climatic variables: the annual daily mean temperature (*bio1*; in °C), the maximum temperature of the warmest month (*bio5*; C°), the minimum temperature of the coldest month (*bio6*; °C), the annual precipitation (*bio12*; mm) and the precipitation seasonality (*bio15*; coefficient of variation). Genomic offset predictions were obtained for each combination of two modelling approaches (GDM or GF) and four sets of SNPs, i.e. the common candidates (Com), the candidates under expected strong selection (Mid), the merged candidates (Mer) and the reference SNPs (Ref). Tree height in the common gardens was measured at 10 and 37-month old in Asturias, 25 and 85-month old in Bordeaux, 8-month old in Cáceres, 13-month old in Madrid and 11 and 27-month old in Portugal. Regression coefficients correspond to β_{X1} estimates in equation 1.3 in the Supplementary Information.

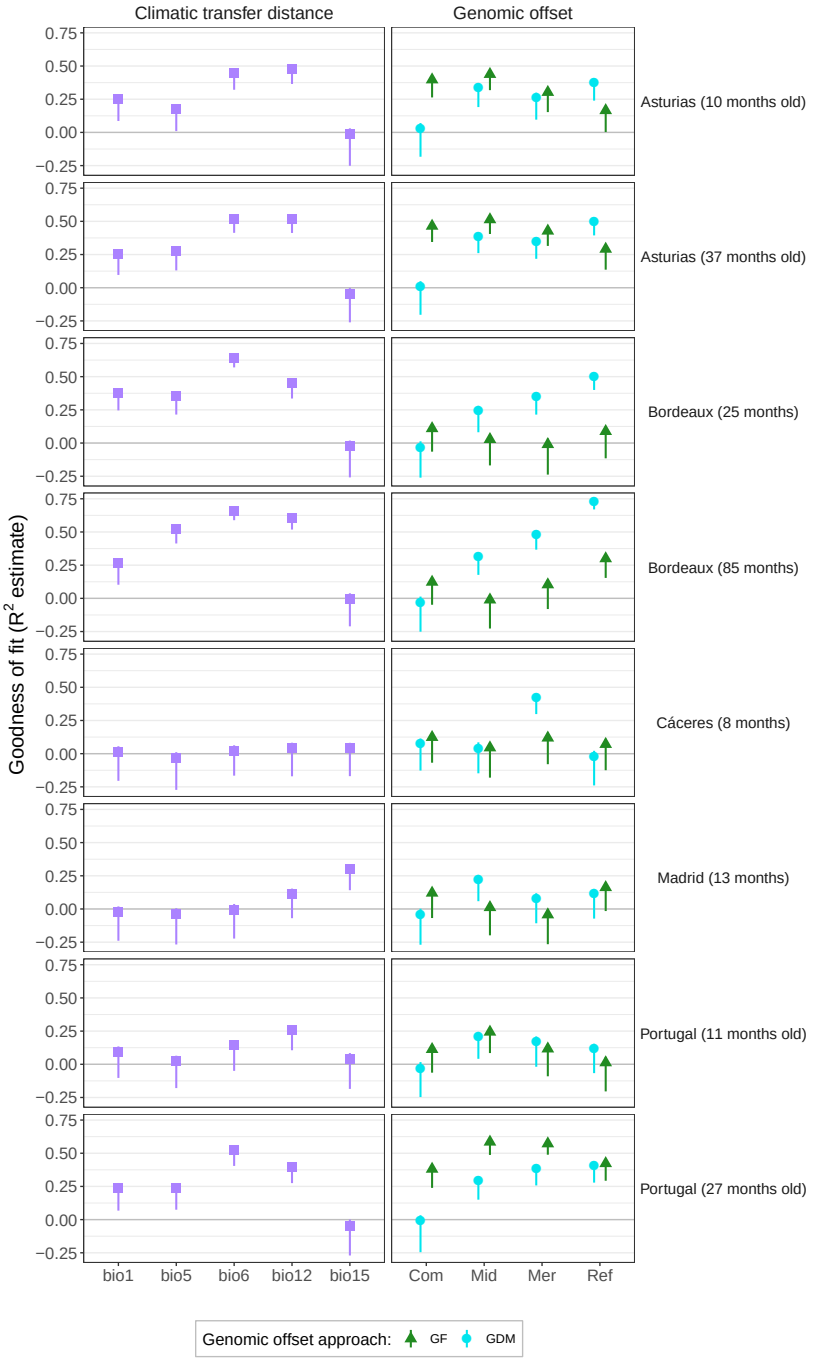


Figure XII.31. Proportion of variance explained ($\mathcal{R}^2$ estimate) of the models estimating the association between tree height in five common gardens and the climatic transfer distances (left panels) or the predicted genomic offset (right panels). The climatic transfer distances were calculated based on five climatic variables: the annual daily mean temperature (*bio1*; in °C), the maximum temperature of the warmest month (*bio5*; C°), the minimum temperature of the coldest month (*bio6*; °C), the annual precipitation (*bio12*; mm) and the precipitation seasonality (*bio15*; coefficient of variation). Genomic offset predictions were obtained for each combination of two modelling approaches (GDM or GF) and four sets of SNPs, i.e. the common candidates (Com), the candidates under expected strong selection (Mid), the merged candidates (Mer) and the reference SNPs (Ref). Tree height in the common gardens was measured at 10 and 37-month old in Asturias, 25 and 85-month old in Bordeaux, 8-month old in Cáceres, 13-month old in Madrid and 11 and 27-month old in Portugal.

Modelling approach	Climatic covariate/ set of SNPs	ELPD difference	SE difference
Climatic transfer distance	*bio12* - Annual precipitation (mm)	0.00	0.00
Climatic transfer distance	*bio6* - Min T° of the coldest month (°C)	-1.08	3.87
GF-based genomic offset	Candidates under expected strong selection	-1.48	2.33
GF-based genomic offset	Common candidates	-2.57	3.37
GDM-based genomic offset	Reference candidates	-3.42	2.86
GDM-based genomic offset	Candidates under expected strong selection	-4.24	3.01
GF-based genomic offset	Merged candidates	-5.15	2.95
Climatic transfer distance	*bio1* - Annual daily mean T° (°C)	-6.17	4.77
GDM-based genomic offset	Merged candidates	-6.49	3.81
Climatic transfer distance	*bio5* - Max T° of the warmest month (°C)	-7.87	2.13
GF-based genomic offset	Reference candidates	-8.34	3.56
GDM-based genomic offset	Common candidates	-10.32	3.68
Climatic transfer distance	*bio15* - Precipitation seasonality (coeff of variation)	-11.52	3.94

Table XII.9. Differences in expected predictive accuracy among the different models in the common garden in Asturias at 10-month old, estimated by making pairwise comparisons between each model and the model with the largest ELPD (model in the first row). The ELPD is the theoretical expected log pointwise predictive density for a new dataset (see equation 1 of Vehtari et al. 2017) and is estimated with the R package *loo*, which gives the Bayesian LOO estimate of the expected log pointwise predictive density (equation 4 of Vehtari et al. 2017). The predictive accuracy of two models can be robustly considered different if their ELPD difference is higher than 4 and is higher than four standard errors of the difference ('SE difference').

Modelling approach	Climatic covariate/ set of SNPs	ELPD difference	SE difference
Climatic transfer distance	*bio6* - Min T° of the coldest month (°C)	0.00	0.00
Climatic transfer distance	*bio12* - Annual precipitation (mm)	-0.35	4.35
GF-based genomic offset	Candidates under expected strong selection	-0.45	3.23
GDM-based genomic offset	Reference candidates	-0.94	3.71
GF-based genomic offset	Common candidates	-1.99	3.88
GF-based genomic offset	Merged candidates	-2.99	3.63
GDM-based genomic offset	Candidates under expected strong selection	-4.10	4.25
GDM-based genomic offset	Merged candidates	-5.43	4.41
Climatic transfer distance	*bio5* - Max T° of the warmest month (°C)	-6.68	4.14
GF-based genomic offset	Reference candidates	-6.88	4.17
Climatic transfer distance	*bio1* - Annual daily mean T° (°C)	-7.43	2.78
GDM-based genomic offset	Common candidates	-11.90	3.41
Climatic transfer distance	*bio15* - Precipitation seasonality (coeff of variation)	-12.65	3.47

Table XII.10. Differences in expected predictive accuracy among the different models in the common garden in Asturias at 37-month old. See legend of Table XII.9.

Modelling approach	Climatic covariate/ set of SNPs	ELPD difference	SE difference
Climatic transfer distance	*bio6* - Min T° of the coldest month (°C)	0.00	0.00
GDM-based genomic offset	Reference candidates	-6.26	6.80
Climatic transfer distance	*bio12* - Annual precipitation (mm)	-7.29	5.10
Climatic transfer distance	*bio1* - Annual daily mean T° (°C)	-9.19	4.11
Climatic transfer distance	*bio5* - Max T° of the warmest month (°C)	-9.90	4.40
GDM-based genomic offset	Merged candidates	-10.14	5.24
GDM-based genomic offset	Candidates under expected strong selection	-11.94	3.97
GF-based genomic offset	Common candidates	-14.66	3.75
GF-based genomic offset	Reference candidates	-15.34	4.41
GF-based genomic offset	Candidates under expected strong selection	-16.41	4.18
Climatic transfer distance	*bio15* - Precipitation seasonality (coeff of variation)	-17.50	4.15
GDM-based genomic offset	Common candidates	-17.55	4.21
GF-based genomic offset	Merged candidates	-17.92	4.46

Table XII.11. Differences in expected predictive accuracy among the different models in the common garden in Bordeaux at 25-month old. See legend of Table XII.9.

Modelling approach	Climatic covariate/ set of SNPs	ELPD difference	SE difference
GDM-based genomic offset	Reference candidates	0.00	0.00
Climatic transfer distance	*bio6* - Min T° of the coldest month (°C)	-3.06	6.24
Climatic transfer distance	*bio12* - Annual precipitation (mm)	-6.61	8.27
Climatic transfer distance	*bio5* - Max T° of the warmest month (°C)	-8.94	6.55
GDM-based genomic offset	Merged candidates	-10.69	3.11
GDM-based genomic offset	Candidates under expected strong selection	-14.47	5.15
GF-based genomic offset	Reference candidates	-15.12	5.85
Climatic transfer distance	*bio1* - Annual daily mean T° (°C)	-16.45	5.95
GF-based genomic offset	Common candidates	-18.75	6.11
GF-based genomic offset	Merged candidates	-19.43	5.31
Climatic transfer distance	*bio15* - Precipitation seasonality (coeff of variation)	-21.20	5.54
GF-based genomic offset	Candidates under expected strong selection	-21.36	5.82
GDM-based genomic offset	Common candidates	-21.90	5.75

Table XII.12. Differences in expected predictive accuracy among the different models in the common garden in Bordeaux at 85-month old. See legend of Table XII.9.

Modelling approach	Climatic covariate/ set of SNPs	ELPD difference	SE difference
GDM-based genomic offset	Merged candidates	0.00	0.00
GF-based genomic offset	Merged candidates	-5.80	4.25
GF-based genomic offset	Common candidates	-5.94	3.14
GDM-based genomic offset	Common candidates	-6.71	3.23
GF-based genomic offset	Reference candidates	-6.90	4.32
GDM-based genomic offset	Candidates under expected strong selection	-7.15	3.37
Climatic transfer distance	*bio15* - Precipitation seasonality (coeff of variation)	-7.28	4.20
Climatic transfer distance	*bio12* - Annual precipitation (mm)	-7.31	3.51
GF-based genomic offset	Candidates under expected strong selection	-7.35	4.31
Climatic transfer distance	*bio1* - Annual daily mean T° (°C)	-7.77	3.75
GDM-based genomic offset	Reference candidates	-7.88	3.95
Climatic transfer distance	*bio6* - Min T° of the coldest month (°C)	-7.90	3.53
Climatic transfer distance	*bio5* - Max T° of the warmest month (°C)	-8.57	3.84

Table XII.13. Differences in expected predictive accuracy among the different models in the common garden in Cáceres at 8-month old. See legend of Table XII.9.

Modelling approach	Climatic covariate/ set of SNPs	ELPD difference	SE difference
Climatic transfer distance	*bio15* - Precipitation seasonality (coeff of variation)	0.00	0.00
GDM-based genomic offset	Candidates under expected strong selection	-0.96	2.72
GF-based genomic offset	Reference candidates	-2.93	2.76
GF-based genomic offset	Common candidates	-3.35	3.53
Climatic transfer distance	*bio12* - Annual precipitation (mm)	-3.54	2.59
GDM-based genomic offset	Reference candidates	-3.69	2.18
GDM-based genomic offset	Merged candidates	-4.27	1.85
GF-based genomic offset	Candidates under expected strong selection	-5.17	2.30
Climatic transfer distance	*bio1* - Annual daily mean T° (°C)	-5.68	2.44
Climatic transfer distance	*bio6* - Min T° of the coldest month (°C)	-5.71	2.38
Climatic transfer distance	*bio5* - Max T° of the warmest month (°C)	-6.12	2.35
GF-based genomic offset	Merged candidates	-6.16	2.27
GDM-based genomic offset	Common candidates	-6.48	2.96

Table XII.14. Differences in expected predictive accuracy among the different models in the common garden in Madrid at 13-month old. See legend of Table XII.9.

Modelling approach	Climatic covariate/ set of SNPs	ELPD difference	SE difference
Climatic transfer distance	*bio12* - Annual precipitation (mm)	0.00	0.00
GF-based genomic offset	Candidates under expected strong selection	-0.57	3.20
GDM-based genomic offset	Merged candidates	-1.74	2.26
GDM-based genomic offset	Candidates under expected strong selection	-2.00	3.28
Climatic transfer distance	*bio6* - Min T° of the coldest month (°C)	-2.42	2.86
Climatic transfer distance	*bio1* - Annual daily mean T° (°C)	-2.54	2.66
GF-based genomic offset	Common candidates	-2.71	3.40
GDM-based genomic offset	Reference candidates	-2.81	2.01
GF-based genomic offset	Merged candidates	-3.30	2.61
Climatic transfer distance	*bio5* - Max T° of the warmest month (°C)	-4.12	1.89
Climatic transfer distance	*bio15* - Precipitation seasonality (coeff of variation)	-4.14	2.88
GDM-based genomic offset	Common candidates	-5.11	2.81
GF-based genomic offset	Reference candidates	-5.30	3.31

Table XII.15. Differences in expected predictive accuracy among the different models in the common garden in Portugal at 11-month old. See legend of Table XII.9.

Modelling approach	Climatic covariate/ set of SNPs	ELPD difference	SE difference
GF-based genomic offset	Candidates under expected strong selection	0.00	0.00
GF-based genomic offset	Merged candidates	-0.55	1.68
Climatic transfer distance	*bio6* - Min T° of the coldest month (°C)	-2.25	2.24
GF-based genomic offset	Reference candidates	-5.81	2.50
Climatic transfer distance	*bio12* - Annual precipitation (mm)	-6.52	3.16
GDM-based genomic offset	Reference candidates	-6.98	3.23
GDM-based genomic offset	Merged candidates	-7.07	3.06
GF-based genomic offset	Common candidates	-7.16	3.35
GDM-based genomic offset	Candidates under expected strong selection	-9.11	2.92
Climatic transfer distance	*bio1* - Annual daily mean T° (°C)	-10.26	3.79
Climatic transfer distance	*bio5* - Max T° of the warmest month (°C)	-10.60	4.50
GDM-based genomic offset	Common candidates	-15.19	3.96
Climatic transfer distance	*bio15* - Precipitation seasonality (coeff of variation)	-15.77	4.19

Table XII.16. Differences in expected predictive accuracy among the different models in the common garden in Portugal at 27-month old. See legend of Table XII.9.

2.5.2 Mortality models in two common gardens

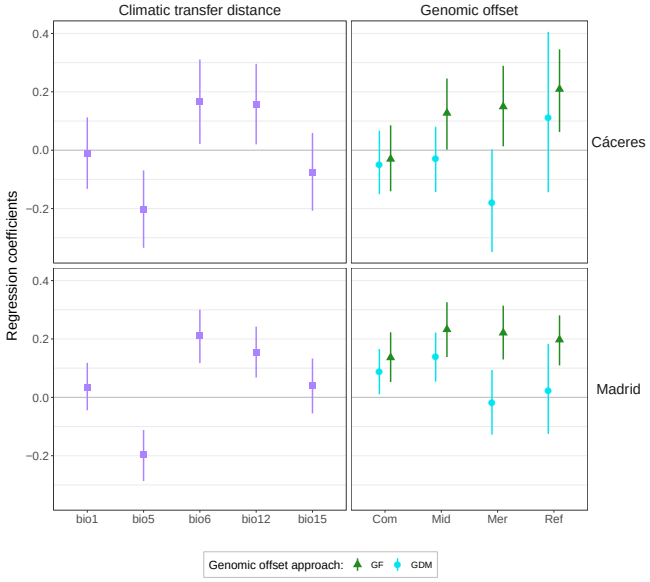


Figure XII.32. Regression coefficients corresponding to the linear association between tree mortality in two common gardens (under harsh conditions) and the climatic transfer distances (left panels) or the predicted genomic offset (right panels). See legend of Figure VI.4 for more details. Tree mortality was measured at 8-month old in Cáceres and 13-month old in Madrid. Regression coefficients correspond to β_{X1} estimates in equation 1.5 in the Supplementary Information.

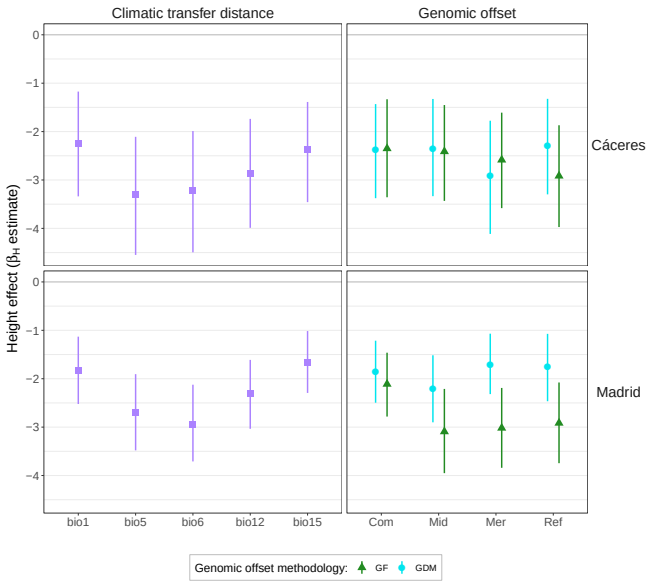


Figure XII.33. Regression coefficients describing the linear association between tree mortality in two common gardens (under harsh conditions) and the mean population height across all common gardens (i.e. BLUPs used as a proxy of tree height at the planting date) for each model considered (same models as in Figure XII.32). These regression coefficients correspond to the β_H estimates in equation 1.5 of the Supplementary Information.

Modelling approach	Climatic covariate/ set of SNPs	ELPD difference	SE difference
GF-based genomic offset	Candidates under expected strong selection	0.00	0.00
GF-based genomic offset	Merged candidates	-0.95	0.88
Climatic transfer distance	*bio6* - Min T° of the coldest month (°C)	-2.04	5.06
Climatic transfer distance	*bio5* - Max T° of the warmest month (°C)	-2.51	2.57
GF-based genomic offset	Reference candidates	-2.53	2.00
Climatic transfer distance	*bio12* - Annual precipitation (mm)	-6.42	4.54
GF-based genomic offset	Common candidates	-7.64	4.94
GDM-based genomic offset	Candidates under expected strong selection	-7.98	4.62
GDM-based genomic offset	Common candidates	-10.63	5.75
Climatic transfer distance	*bio15* - Precipitation seasonality (coeff of variation)	-12.52	7.98
Climatic transfer distance	*bio1* - Annual daily mean T° (°C)	-13.53	9.39
GDM-based genomic offset	Reference candidates	-14.03	7.62
GDM-based genomic offset	Merged candidates	-14.10	8.99

Table XII.17. Differences in expected predictive accuracy among the different models in the common garden in Madrid. See legend of Table XII.9.

Modelling approach	Climatic covariate/ set of SNPs	ELPD difference	SE difference
Climatic transfer distance	*bio5* - Max T° of the warmest month (°C)	0.00	0.00
GF-based genomic offset	Reference candidates	-0.23	2.02
GF-based genomic offset	Merged candidates	-2.05	2.77
GF-based genomic offset	Candidates under expected strong selection	-2.19	2.94
Climatic transfer distance	*bio6* - Min T° of the coldest month (°C)	-2.58	2.89
Climatic transfer distance	*bio12* - Annual precipitation (mm)	-2.64	3.36
GDM-based genomic offset	Merged candidates	-2.95	2.88
Climatic transfer distance	*bio15* - Precipitation seasonality (coeff of variation)	-3.87	3.44
GDM-based genomic offset	Reference candidates	-4.00	3.45
GDM-based genomic offset	Common candidates	-4.10	3.55
GDM-based genomic offset	Candidates under expected strong selection	-4.55	3.50
GF-based genomic offset	Common candidates	-4.64	3.49
Climatic transfer distance	*bio1* - Annual daily mean T° (°C)	-5.01	3.69

Table XII.18. Differences in expected predictive accuracy among the different models in the common garden in Cáceres. See legend of Table XII.9.